ADVANCES IN COMMUNICATIONS AND MEDIA RESEARCH

VOLUME 8

Contributors

Louis Rhéaume
Yves Rabeau
Francisco V. Cipolla-Ficarra
Vladica Tintor
Vlade Milićević
Milan Janković
Jovan Radunović
Hanlong Fu
Stephen E. Blythe
C. K. Cheung
Aditi Dubey
Kohei Arai
Lipur Sugiyanta
Kirsty Young
Adel Beldi
Zhiyuan Luo
Agnieszka Matysiak Szóstek
Panos Markopoulos
Berry Eggen
Hamed Rasouli
Alagan Anpalagan

Anthony V. Stavros
Editor

NOVA

ADVANCES IN COMMUNICATIONS AND MEDIA RESEARCH

VOLUME 8

ADVANCES IN COMMUNICATIONS AND MEDIA RESEARCH

Additional books in this series can be found on Nova's website
under the Series tab.

Additional E-books in this series can be found on Nova's website
under the E-book tab.

ADVANCES IN COMMUNICATIONS AND MEDIA RESEARCH

VOLUME 8

ANTHONY V. STAVROS
EDITOR

Nova Science Publishers, Inc.
New York

For permission to use material from this book please contact us:
Telephone 631-231-7269; Fax 631-231-8175
Web Site: http://www.novapublishers.com

LIBRARY OF CONGRESS CATALOGING-IN-PUBLICATION DATA

ISBN: 978-1-61324-902-4 (eBook)
ISSN: 2159-1695

Published by Nova Science Publishers, Inc. † New York

CONTENTS

PREFACE

In a society predicated on information, the media has a pervasive presence. From government policy to leisure television, the information age touches us all. The papers collected in this book constitute some of today's leading analyses of the information industry. Together, these essays represent a needed foundation for understanding the present state and future development of the mass media. Current trends in communications as well as media impact on public opinion are studied and reported on.

Chapter 1 - A confluence of technological leaps in devices, networks, and applications is setting the stage for wireless to change our lives the way the personal computer (PCs) did in the 1980s and the Internet in the late 1990s. There is a profound shift from wireline communications to wireless communications through cellphones, smartphones, Wi-Fi, WiMax, LTE, WLAN, and new promising technologies.Since the mid-90s, the telecom industry has gone into a period of sustained disruptive innovations which, combined with deregulation, led to a lot of turbulence and a sometimes difficult redefinition of business models. A true Schumpeterian wave of innovation enhanced by competition leads to creation of wealth as an unprecedented investment boom occurs in telecommunications sustained by overly optimistic and sometimes fraudulent forecasts of Internet traffic. But it also led to wealth destruction when the financial bubble ended in a stock market crash in 2000, whereas several telecom companies went bust, weakened by debts, substantial overcapacity, and a loss of market power. It then became clear that *traffic growth did not necessarily translate into revenue growth*. As overcapacity eased in the 2000s and telecom companies painfully restructured, the wave of innovation went on. Particularly, VoIP definitely made long-distance wireline service and dial-up ISP commodities. The telephony industry is now moving from a carriage sector toward a service sector as information technologies are at the heart of all business models in the world economy.

Chapter 2 - In the authors' research work they focus on the users who have a greater difficulty in understanding the functionality of mobile phones with the triad of communicability, usability and usefulness. It is a detailed study which starts with the models of the 90s and the first decade of the new millennium, and tries to underline the area where the formal and factual sciences interact. To this end an unprecedented evaluation methodology is established where there is a technological and communicative intersection. The participation of real users allows us to put into use a series of heuristic research techniques to face these problems from several points of view of interactive and communicational design. In this work the authors make a special analysis of the subject of the choice of the random samples inside descriptive statistics, since they are working with it to cut down costs in the methods and evaluation of the quality in the interactive systems, where the mobile multimedia phones are included, for example. Additionally, the obtained results allow not only to know the weak and strong points of the current multimedia mobile phones but the authors set up a little vademecum where the sociological and technological aspects of the users in the coming years are summed up.

Chapter 3 - This paper aims at illustrating the process of liberalisation within the telecom market in the Republic of Serbia from 2005 to 2010. During that period, the telecom market was primarily a state monopoly market. As such, this market was unregulated, characterised by the abuse of the dominant position, the operation of broadcasters without a broadcasting license, low investment rates, high service prices, relatively low service quality and, by and large, a low intensity of telecom service use. Relying on the Telecommunication Law, (adopted in 2005 and harmonised with the 1998 EU Directives), closely monitoring the development trends of new technologies and taking into account the experience and the best practices of regulatory authorities from the developed countries of the EU, the national regulatory authority managed, in a relatively short period of time, to secure the legal and economic framework for the liberalisation and opening up of the telecom market – firstly within the market segment of broadcasting, then mobile telephony, Internet services, cable distribution systems, and finally, broadband access. The shared efforts of telecom operators, the regulator and the state have resulted in the telecom market now being regulated and characterised by the predictability of business environment, high investment rates, high quality services and lower prices, market entry of foreign operators, higher degree of availability and use of telecom services, and lastly, higher revenues for operators and higher tax rates for the state. All of the aforementioned would have been virtually impossible had it not been for the appropriate regulatory framework, which had been efficiently implemented within the telecom sector.Furthermore, this paper will focus on each telecom market segment separately. Within the domain of fixed networks, the socialist heritage is evident in the low prices of fixed telephony, which are still below the price threshold. Consequently, as these prices considerably lag behind those of the neighbouring countries, there has not been much interest in the initiation of the liberalisation and privatisation process within this telecom market segment. Given the circumstances, one license has been granted to an operator for fixed broadband network services with national coverage and one license for the provision of services via CDMA technology. The liberalisation of the mobile telephony market segment was completed in 2006. This particular year was especially favourable as UTMS, a new technology in this field, had created the conditions for the introduction of new services and significantly increased the number of users. Moreover, these circumstances led to the issuance of new licenses and the privatisation of one part of the state property within this market segment. Three mobile telephony operators have introduced new, up-to-date and user-driven services and established absolute market competition where national traffic service prices rated among the lowest in Europe. Prior to the implementation of the new regulation, the Internet market segment had been undeveloped and had a relatively small number of users. The new regulation ensured the commercial provision of lower bit-rate Internet service in the 2.4 GHz band, whereas the issuance of authorisations for international interconnection regulated the issue of wholesale Internet. A considerable decrease in prices of international traffic services became evident following the issuance of authorisations for VoIP services. Moreover, the technical conditions for the roll-out of next generation broadband networks were fulfilled as well.Taking into account the regulatory, technological, as well as economic aspects, the development of broadcasting and cable distribution systems will be analysed under separate heading within this paper.Additionally, the economic effects of market liberalisation, reflected primarily in the transparency of the market, its openness and attractiveness for new investors, along with the scope of realized investments which constitute some of the largest investments in the Serbian economy, the increase in the number of users, the constant enhancement of the quality of services, as well as the price-cuts are likely to secure the conditions for a competitive business environment. All these trends which appear in the Serbian telecoms market are monitored in the comparative overviews of the key telecom market indicators. The concluding remarks of the paper offer an insightful illustration of the further development of the telecom market, drawing special attention to the challenges that may lie ahead, especially within the context of the new Electronic Communications Law, which has been harmonised with the 2003 and 2007 EU Directives.

Chapter 4 - The Telecommunications Act of 1996 is arguably the single most important piece of legislation since the Communications Act of 1934, one that affects the telecommunications industry, consumers, and ultimately the balance of political power. The 1996 Act was designed to usher in competition to telephony and cable by breaking down the cross-entry barriers that were put in place by the Communication Act of 1984. Yet immediately after the passage of the 1996 Act, telecommunication industries witnessed a deluge of mega mergers and acquisitions. This unprecedented merger wave resulted in a handful of conglomerates dominating industries that were previously separated by telecommunications regulations. The present chapter assesses the structural impact of the 1996 Act by analyzing the most recent industry trends and statistics of media ownership. The authors discuss, in particular, the implications of the 1996 Act for the National Broadband Plan.

Chapter 5 - Jordan's economy has been growing in recent years due in no small part to a most-favorable trade treaty with the United States. E-commerce emerged after the enactment of the Electronic Transactions Law ("ETL") in 2001. The ETL covers most E-commerce transactions, but is inapplicable to wills, real property deeds, powers of attorney, termination of insurance contracts and utility services, court documents and securities. The statute recognizes the enforceability and admissibility of electronic documents and signatures. The electronic form may be used to comply with statutory requirements pertinent to writing, originality, retention and signing. The ETL contains E-contract rules concerning attribution, acknowledgement of receipt, and the time and place that an electronic message can be assumed to have been sent and received. In an unusual but admirable move, the statute includes rules regarding the transfer of electronic notes and electronic funds transfers. The ETL contains third-generation E-signature rules; the utilization of stringent security procedures distinguishes secure E-signatures from insecure ones. The ETL lists several computer crimes: fraudulent use of a certificate; submission of false information to a Certification Authority ("CA"); offenses by CA's; and commission of other crimes using electronic means. Is the ETL up-to-date according to current trends in international E-commerce law? Not quite. The most significant of the recommended amendments are to: (1) eliminate the list of exclusions and recognize the legal validity of electronic wills; (2) add E-contract rules pertinent to automated and carriage contracts; (3) add consumer protections for E-commerce buyers; (4) establish Information Technology tribunals as a court-of-first-instance for E-commerce disputes; (5) claim "long arm" jurisdiction over foreign parties in E-commerce transactions; (6) designate the Jordanian Post Office as a licensed CA; and (7) add several new computer crimes.

Chapter 6 - The increasing public engagement in Internet communication and new media, especially in Web 2.0 technologies, has generated new interest regarding the effects of technological mediation in the area of language learning. The changes brought about due to the multimodal nature and stylistic variations of new media technology have resulted in increased theoretical reflection on language within a mediated communication reality (Rowe & Ryss, 2009). These theories imply the need for today's pedagogy to move beyond the traditional written and oral texts and to embrace digital text formats. Knobel & Wilber (2009) affirm that digital formats don't mesh well with traditional language learning practices such as book reports, comprehension questions, reading tasks, and weekly spelling tests. Thus, the Internet era with its Web 2.0 media tools, questions the dominant ideas around language and language learning. This in turns prompts the need for language learning to be rethought in the new media context.Web 2.0 refers to web applications that facilitate a user-centred design and interactive information sharing. Most "digital natives" (Prensky, 2001) or learners, who grew up using computers, are well-versed with Web 2.0 tools such as blogs, wikis, podcasts, and social networking sites. Unlike other applications, since the focal point of web 2.0 tools is the user, this paper aims to shed light into the potential of Web 2.0 tools to enhance the language ability of students, and to expand or broaden the range of language learning opportunities. The growth and development of Web 2.0 tools, which has been a result of the rapid advancement in new media and Internet technologies, can contribute to language learning by promoting interaction, exposure, feedback, reflection, empowerment and learner autonomy. They also provide an environment for language learners to reflect, comment, question and review their progress even outside the classroom in a more authentic environment (Pinkman, 2005). To conclude, the paper argues that the user-friendliness of Web 2.0 applications, when combined with appropriate guidelines, training and adequate knowledge, can enable teachers or "digital immigrants" (Prensky, 2001) to integrate Web 2.0 tools to assist students' language learning in the classroom. Web 2.0, thus, offers many opportunities for language learners and teachers. This not only has tremendous potential for the enhancement and improvement of language learning, but it is also more conducive for meeting the expanding language needs of the Internet era.

Chapter 7 - Wireless ad-hoc mesh network is a special kind of network, where all of the nodes move in time. Node is intended to help relaying packets of neighboring nodes using multi-hop routing mechanism in order to solve problem of dead communication. Wireless mesh network which engages broadcasting and contains multiple hops become increasingly vulnerable to problems such as routing problem and rapid increasing of overhead packets. During this progress, the delay on account of multi hop characteristics and redundant packets caused by communication nature potentially existed during communication. Typically, delay will increase in linearity with number of hops. There is a certain minimum level of delay that will be experienced due to the time it takes to transmit a packet through a link. Topology development holds a significant point prior to the data transmission. Without improved topology development protocol, this problem can decrease network's performance in overall data transmission. The authors analyze the delay performance of a multi-hop wireless network with a dynamic route between each source and final destination pair. There are fluctuate interference constraints on the set of links that impose a fundamental delay performance of any instant network topology. At first, the authors present a similar Link State

Routing network simulation to derive such referential lower bounds. The authors conduct extensive simulation studies to suggest that the average delay of multi-hop transmission policy can be made lower compared to the referential bound by using appropriate functions of network metrics. This chapter provides a broadcast framework that engages various network metrics and at the same time maintaining connectivity of nodes (mobile terminals). The framework captures the essential features of the wireless network metrics, i.e. bandwidth, throughput, network buffer, direction, and round trip time. This research is useful since, in many cases, it find that the throughput is the most important parameter in reduction of delay transmission. This result is confirmed with another composite simulation result. Most of network hop delay is impacted with this composite metric, particularly in delay minimization on the longer hops. The reduction achievement on average delay by this algorithm is 0.577% and the total average delay reduction for this simulated network is 0.683%. This research will be further designed primarily for achieving maximum throughput in the multiple wireless network area.

Chapter 8 - The use of popular online social networking sites (eg. MySpace, Facebook) has resulted in an abundance of authentic texts where words, graphics, pictures and music are combined by users to project a desired image of oneself. This Chapter proposes that these online profiles can be utilised for instructional purposes to develop students' digital literacy. This can be achieved by using social semiotic multimodal analysis to deconstruct online social networking profiles. The process of applying multimodal analysis to online social networking profiles develops student understanding of the impact of text, graphics and sound on the readers' interpretation which, in turn, facilitates students becoming critical readers and producers of online content. This Chapter presents an analytical framework that educators can use with students to deconstruct online content, with an example drawn from MySpace.

Chapter 9 - Since the 90 major changes have occurred in the telecommunications sector in Europe and particularly in France. Until this period, a governmental administration was in charge of managing post and telecommunications services in France. An owned corporation was created to manage telecommunications services and a separation between the two activities was occurred. The opening of the telecom market to competition in France has accelerated changes in strategic, organizational and human patterns of the company. The French Telecommunication Corporation (FTC) aimed to increase profits in a context characterized with a high competitors' aggressiveness. These factors have led the company to change its strategy, organization and the customer relationship management approach. In the 2000s, the board of directors decided to implement a major customer relationship management (CRM) software in commercial divisions of the company. In "business" branch, the project started in 2001 was the occasion to change internal processes, management methods and organization of the local agencies.

Chapter 10 - The Internet is a globally distributed network of networks supporting many applications and services. In addition to traditional applications (e.g. Email and FTP), new Internet applications such as multimedia streaming, Internet telephony, games and peer-to-peer file sharing become popular. Classification of network applications responsible for the generation of the traffic flows is important for network management tasks such as monitoring trends of the applications in operational networks and effective network planning and design. In particular, accurate and fast classification of network traffic based on statistical flow characteristics can offer substantial benefits to a number of key areas in Quality of Service support and service differentiation, enforcement of security policies, traffic engineering and support for Service Level Agreements for many network operators and service providers. The requirements for such classification of Internet traffic are increasing rapidly due to availability of high capacity communications links, heterogeneity of Internet traffic types and continued evolution of applications such as tunnelling and end-to-end encryption. Accurate and fast traffic classification requires algorithmic capability, in particular, machine learning algorithms. The idea of using machine learning techniques for Internet traffic classification is not new. Many machine learning algorithms have been successfully used to classify network traffic flows with good performance, but without information about the reliability in classifications. However, not knowing the confidence of predictions makes it difficult to measure and control the risk of error using a decision rule. Modern network resource management systems are becoming increasingly complex and as such require high quality, reliable predictions to optimiseperformance. Therefore, introducing more reliable algorithms for network resource management will result in higher quality of service for network users. In this chapter, the authors consider the problem of reliable Internet traffic classification. The authors present two recently developed machine learning algorithms, namely the Conformal Predictor and Venn Probability Machine, for making reliable decisions under uncertainty and achieving performance guarantees. Experiments on publicly available real Internet traffic datasets show these two algorithms work well. Performance comparison is also made between these two algorithms. Some guidelines on how to choose between these algorithms are provided and discussed.

Chapter 11 - As there is no generally accepted convention regarding which behaviours for negotiating communication initiation are considered as socially acceptable, people tend to

develop their own conventions that define the best practices for different situations and communities. Since system users rely on representations of their presence and actions to understand each others' degree of availability, finding ways to share relevant contextual information is necessary to help creating more refined and socially acceptable practices. However, current technologies tend to focus on provision of mechanisms that relate to functional rather than to social requirements and often disrupt the exchange of cues regarding the context communicators are in. For example, phones give the means to contact others anytime anywhere, yet do not provide any indications regarding whether the recipient is able to accept an incoming communication. As a result people are provided neither with sufficient information ab out the context, in which their communicators operate nor with means helping them to manage their communications in a way that is efficient and yet complies with social conventions similar to those used in face-to-face encounters. The goal of the study presented in this article is to investigate ways to ascertain mutual awareness about the recipient's availability status in Instant Messaging applications. The authors argue that availability indication alone is insufficient to leverage social behaviours in mediated communication and that it is crucial to introduce mechanisms stimulating mutual awareness regarding the communicative needs of communicators. To reach this goal a prototype named DoNTBother was implemented, which was evaluated in a web design company for a period of three weeks. The contributions of this research include the quantitative and qualitative measurements assessing the proposed solutions and a set of implications that promise to inform the design of future mechanisms supporting the attainment of mutual awareness in Instant Messaging applications.

Chapter 12 - Wireless relaying can be viewed as a new dimension in resource exploitation in radio communication systems. Relaying is a technique that uses certain terminals, either fixed or mobile, which have good communication links with the destination terminal to act as relay nodes for those who do not have. By applying this technique in a cellular system for example, either the throughput or network coverage can be increased. Multiple antenna schemes provide enormous performance gain in wireless systems by increasing the capacity of the network. Using space-time codes in the relaying scheme provides the diversity benefits of multiple antenna techniques. It has been suggested to use space-time block coding techniques that were proposed for multiple-antenna systems in a wireless relaying system. The idea of making the relay nodes act like multiple antennas is an interesting idea and has been studied extensively in the literature. Space-time coded relaying protocols can provide a higher throughput gain compared to the repetition-based relaying protocols. In this chapter, different relaying techniques are categorized and reviewed. Amplify-and-forward and decodeand- forward are the two main categories based on the relay functioning. Relaying can also be categorized in terms of wireless access protocol that is used in accessing the radio channel. Time-division, frequency-division are the legacy access protocols and space-division is a relative new one. The focus of this chapter is mainly on spacetime coded cooperation applied to wireless relays. Following a detailed discussion on distributed space-time coded relaying, some selected results are provided to show the advantages of space-time coded cooperation through wireless relays.

In: Advances in Communications,Volume 8
Editor: Anthony V. Stavros
ISBN: 978-1-61324-794-5

Chapter 1

RETHINKING NORTH AMERICAN TELEPHONY BUSINESS MODELS IN THE AGE OF TURBULENCE

***Louis Rhéaume*[1] *and Yves Rabeau*[2]**
TELUQ-UQAM, Université du Québec a Montréal, Canada

Keywords: telecommunications, business models, strategy.

1. INTRODUCTION

A confluence of technological leaps in devices, networks, and applications is setting the stage for wireless to change our lives the way the personal computer (PCs) did in the 1980s and the Internet in the late 1990s. There is a profound shift from wireline communications to wireless communications through cellphones, smartphones, Wi-Fi, WiMax, LTE, WLAN, and new promising technologies.

Since the mid-90s, the telecom industry has gone into a period of sustained disruptive innovations which, combined with deregulation, led to a lot of turbulence and a sometimes difficult redefinition of business models. A true Schumpeterian wave of innovation enhanced by competition leads to creation of wealth as an unprecedented investment boom occurs in telecommunications sustained by overly optimistic and sometimes fraudulent forecasts of Internet traffic. But it also led to wealth destruction when the financial bubble ended in a stock market crash in 2000, whereas several telecom companies went bust, weakened by debts, substantial overcapacity, and a loss of market power. It then became clear that *traffic growth did not necessarily translate into revenue growth*. As overcapacity eased in the 2000s and telecom companies painfully restructured, the wave of innovation went on. Particularly, VoIP definitely made long-distance wireline service and dial-up ISP commodities. The telephony industry is now moving from a carriage sector toward a service sector as information technologies are at the heart of all business models in the world economy.

[1] E-mail Address: rheaume.louis@teluq.uqam.ca
[2] E-mail Address: yrabeau@uqam.ca

Through M&A transactions, the former “Baby Bells” in America with their scale and huge markets have been doing relatively well in a highly competitive market. Owning the “last mile” of the network that runs into homes and offices remains a competitive advantage, although the rapid development of wireless technology is a threat to all former incumbents. In Europe, the former monopoly incumbents remain major players in the telecom market. With the recent turbulence of 2008-2010, telephone operators must update their current business model as technology continues to evolve and new services are offered. Telcos must also cope with regulation changes such as the new regulations with net neutrality which may postpone the plans of incumbents like Verizon and others to launch fibber optic networks.

The wireless technology is currently a major driving force in the telecom market as many players are offering new products and a whole array of services. New 4G networks, LTE, and WiMax, and the rapidly growing use of smartphones are impacting business models in all industries. As more industries use smartphones in the workplace, demand for downloadable apps is raising, letting all kinds of information technology firms offer new programs. New services are also offered to consumers, and new business competitors like Google getting revenues from advertising is a significant disruption in the wireless market.

Incumbents are still using the one stop shop model, offering a combination of product portfolio with fixed telephony and wireless. However, niche players are a threat by offering highly specialized products that cannot be offered by the former incumbent. The recent popularity of social networks and fixed data plans are also driving the traffic of Internet mobile networks. Technology standards in telecom tend to become more open, but some proprietary standards such as Apple and Research in Motion are successful in wireless.

As technology become more complex, “open innovation” models with coopetition is the rule. Telcos will have to sail in completely new waters if they want to survive to technological changes.

The first section deals with the wave of disruptive innovations in the communications industry over the last few years and stresses the ubiquity of IT in all aspects of the economy and society. The next section focuses on the search for profitable business models for telcos and providers of other communication services. Regulation in the US and Canada is analysed in the next section. Then, we turn to planning telecom business models with Net neutrality rules. The final section pinpoints, with short comments, the coming trends in the communications industry.

2. DISRUPTIVE INNOVATION AND NEW BUSINESS AND SOCIAL TRENDS

In this section, we show how disruptive innovations like the Internet and high speed transmission of information have completely changed the market of former telco incumbents and their competitors. Telcos must now deal with a complex world driven by incremental innovations, changes in business models in the whole economy, and new social trends set by individuals as they invade the Web.

The digitalization of information, the bandwidth revolution, and the emergence of the Internet are major innovations that have transformed the telecommunications industry as well as business practices all across the economy and the lives of citizens in general. In addition, the introduction of competition in the telecommunications sector of North-America and Europe has magnified the impact of innovations on the industry and made it more complex.

2.1.1. Major innovations

On the wireless side, the availability of new smartphones on 3G+ and 4G networks are critical. Some operators such as AT&T made alliances to obtain a temporary first-mover advantage with the sales exclusivity of the iPhone 3G with Apple. High-speed Mobile Internet, through WiFi, enables new powerful applications such as Mobile TV. The delivery of content through applications, proprietary content, or acquisition of content through alliances suggest that an optimal mix must be implemented by operators. Mobile app downloads jumped from 300 million in 2009 to 5 billion in 2010.

LTE and WiMax networks represent the fourth generation of wireless. LTE-Advanced offers 1 Gbit/s speed on peak downloads and 500 Mbit/s speed on peak uploads. Mobile WiMax offers peak data rates of 128 Mbit/s downlink and 56 Mbit/s uplink over 20 MHz wide channels.

A major innovation that started in 2004[1] was the introduction of mass-market VoIP services that utilize existing broadband Internet access, by which subscribers place and receive telephone calls in much the same manner as they would via the public switched telephone network (PSTN). VoIP proliferation has reached a new level. The leader, Skype, is now offering its service through several smartphones and with video calling (with iPhone, iPad and iTouch). In the first half of 2010, Skype users made 88.4 billion minutes of Skype-to-Skype calls, and approximately 40% of these were video calls. Google Talk is also a major player with rapid growth in VoIP.

In the wireline side, fiber optic represents an opportunity for telecom operators to own the “last mile” of the network that runs into homes and offices. It remains a competitive advantage, although the rapid development of wireless technology is a threat to all former incumbents. Since 2000, the prices for fiber-optic communications has dropped considerably. The price for rolling out fiber to the home has currently become more cost-effective than that of rolling out a copper based network. Prices have dropped to $850 per subscriber in the US and lower in countries like The Netherlands, where digging costs are low. The per-channel light signals propagating in the fibber have been modulated at rates as high as 111 gigabits per second by NTT[23] although 10 or 40 Gbit/s is typical in deployed systems.

The technology adoption of some new communication innovations have increased exponentially in recent years. The following graph shows that while it took around 75 years for the wireline phone to reach 50 million customers, it took only 12 years for the mobile phone, 5 years for the iPod, 4 years for e-mail, and only 3 years for Skype’s VoIP software. It can be explained by the fact that several communication networks are already in place and new innovations can leverage existing infrastructure. Another reason is the fact that

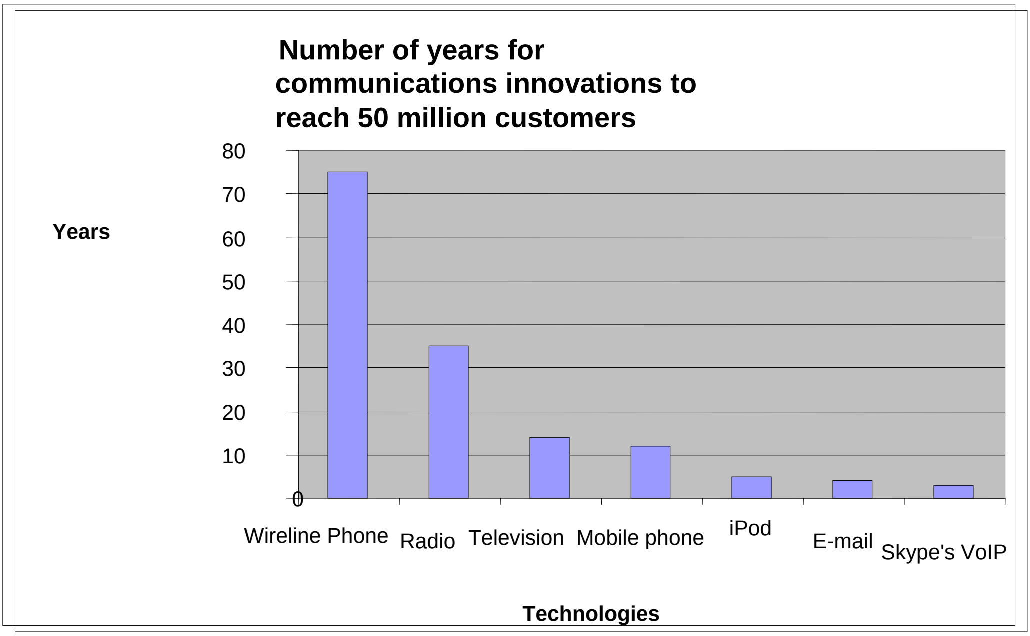
---

[1] “VOIP on the Verge”. Telecommuncations Online. November 1, 2004. Retrieved 2009-01-21.

[2] Tbps over a Single Optical Fiber: Successful Demonstration of World's Largest Capacity - 140 digital high-definition movies transmitted in one second. NTT Press Release. September 29, 2006 and by NSN.

[3] M. S. Alfiad, et al. (2008). "111 Gb/s POLMUX-RZ-DQPSK Transmission over 1140 km of SSMF with 10.7 Gb/s NRZ-OOK Neighbours". *Proceedings ECOC 2008*: pp. Mo.4.E.2.

communication consumers are more "technology educated" than before. The fastest growing tech firm at the convergence of social networks and mobile commerce is Groupon, which has reached 2 billion in revenue in its two years of existence.



Source: debitel

Figure 1.

2.1.2. New Business practices

With the rapid rise and fall of the dot-com companies in 2000, only a few original "pure plays" like Amazon, Yahoo!, and eBay survived. Then, a major change in the management paradigm occurred as real economy corporations started to invade the Net. Instead of using information technology (IT) to increase the productivity of a company specific function (like managing the payroll), IT was then beginning to be at the heart of business models as e-commerce was spreading in the whole economy. This was the beginning of the "digital economy"[1] era.

Unbundling the firm[2] was a first step toward the new digital business model. Outsourcing and "just in time supply chain" was, for several corporations, business practices that increased productivity and reduced costs. Then, one of the first versions of these new business models, called "fully integrated value chain" (FIVC), became a strategy to re-assemble an unbundled value chain to deliver the best products at the most competitive price to customers. It is an outward looking strategy aimed at optimizing the supply chain and implies a process

[1] Don Tapscott was among the first to write about the "digital economy" in "The Digital Economy", McGraw-Hill, 1995. See also: Don Tapscott, "Grown up Digital", McGraw-Hill, 2009.

[2] John Hagel III and Marc Singer, « Unbundling the Corporation », Harvard Business Review, March/April 1999.

of global cooperation which calls for electronic exchanges of complete information between all members of an FIVC; each firm knows exactly who is doing what and when. The FIVC deepens the relationship between buyers and vendors and secures full after sales services that can be delivered to a large extent via electronic exchange. Dell was a pioneer in adopting this business model followed by the like of Cisco, Intel, GE, and many others.

This model was further developed with the arrival of the so-called electronic manufacturing service providers (EMS). The EMS Providers' clients are generally large corporations in possession of an internationally recognised brand. To remain competitive in the new environment, these companies outsource the manufacturing of their products or part of their products to EMS providers as they prefer to focus on high value added activities such as marketing, financing, fundamental research, and product development in order to satisfy their clients' needs. EMS Providers are present in different countries and have a world-wide network of suppliers. Although EMS Providers are not involved in fundamental R&D, they do some specific R&D for large corporations aimed at designing, standardising, and integrating product components. Flextronics and Solectron are classic examples of EMS providers. Chart 1 illustrates this business model.

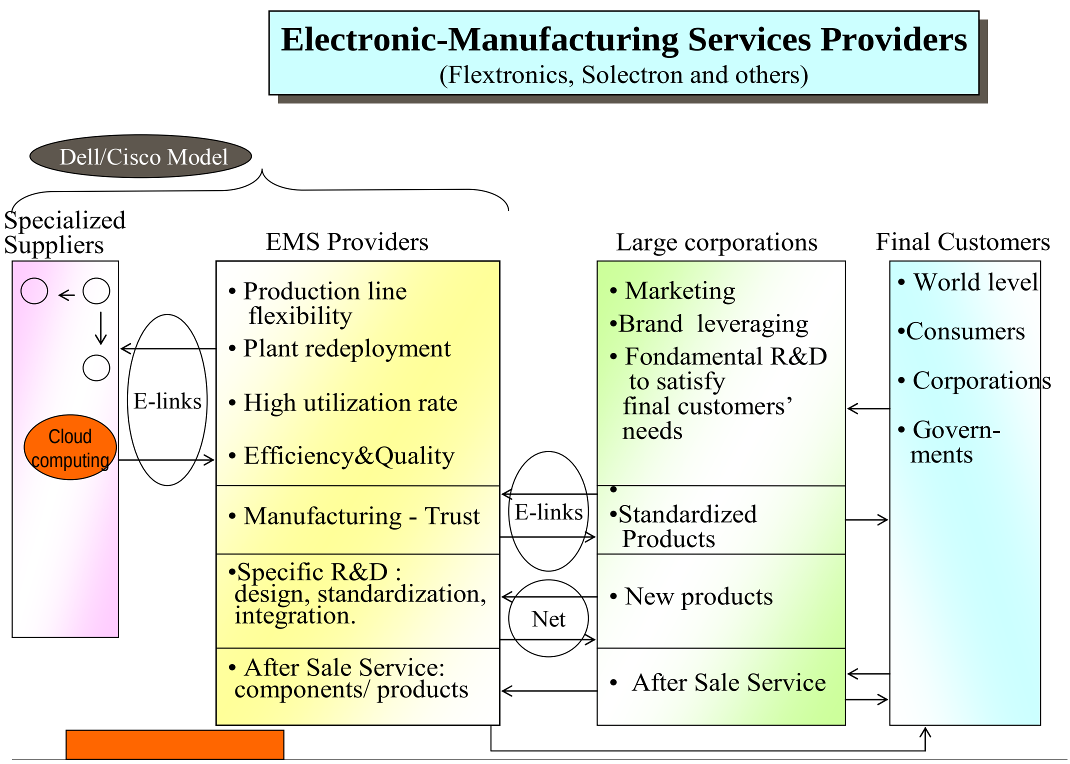


Chart 1.

With the growing complexity of electronic exchanges and the amount of data to be stored for day to day operations, firms like Flextronics and Siemens[1] are outsourcing several tasks to software providers offering “cloud computing” services. These firms can provide large storage capacity, various web apps, and other services to their clients. Furthermore, pioneer corporations like Dell, Ford, Jet Blue, and others have incorporated social media networks into their business model and many others are likely to follow that trend in the near future[2] as also demonstrated by a very recent McKinsey survey[3], thus extending the use of electronic exchanges in doing business.

In goods producing industries and also for some activities like the media in the service industry, most companies of all sizes operating globally or regionally have adopted a variant of the stylized business model depicted in Chart 1 and its extensions such as cloud computing and social networks. In fact, networked companies are everywhere in the world economy. This is part of the new world in which the telcos are now operating.

The service industry is also very web exchange intensive. In particular, well before the Internet era, the financial sector was relying on electronic exchange for its day to day transactions. The Internet has intensified the IT component of that sector’s business model. Financial institution clients of all sorts can have access to their accounts and make on-line transactions. All the media businesses are now on-line. The health sector, already IT intensive, is a growing user of electronic exchanges as complex services can be delivered on-line. Access to government services online for all taxpayers is largely available.

Internet companies or so-called “pure plays” are generating huge Internet traffic. A survivor from the dot-com crash such as e-Bay has a large ecosystem with several suppliers of services and a customer base of around 233 million users all around the world. By buying Critical Path Software in December 2010, E-Bay is taking the mobile bend in order to follow its customer base.

Among the Internet firms, Google is a remarkable example of network effects that can be brought on by electronic exchanges and deserves an examination of its business model to show what opportunities and threats net companies can create for telecom service providers. Google offers several Web services like on-line search, e-mail, social networking, and content of all sorts. It is also a supplier of web based software apps. The company has made strategic acquisitions in the recent past like You Tube and the on-line advertising company DoubleClick.

Google has a large ecosystem where all members can contribute to generate new innovations. It is an open-innovation business model where more than 130 million unique visitors[1] search daily for information. Doing so, they reveal their interests and are exposed to targeted advertising. Visitors can also contribute to the enhancement of the web site with

---

[1] R. King, “Flextronics, Siemens lead « Big Shift » to Cloud Computing”, Business Week, Special Report, December 6th, 2010.

[2] David Armano, “Six Social Media Trends for 2111” December 6th, 2010, cited in Harvard Business Review Weekly Hot List, December 13th, 2010.

[3] Jacques Bughin and Michael Chui, “The Rise of the Networked Enterprise: Web 2.0 finds its Payday”, McKinsey Global Institute, December 2010. Also in a 2009 McKinsey global survey, “How companies are benefiting from web 2.0”, September 2009, it is indicated that companies increased their market knowledge and make marketing more effective by observing social networks.

ideas for new services. Even if Google has made a few major acquisitions, it is not a vertically integrated company. For example, it did not buy important media companies and prefers to deal with several of them and get as much content as possible in order to attract as many viewers as possible to maximize advertising revenues. Media and other companies and also individuals are content providers from all around the world for Google. Over one million companies and other parties deliver advertising content to search identified users, generating a considerable revenue stream. It can tap on a large community of contributors to enhance its services and also innovate with new products. Innovations come from software vendors and Google's own staff as well as the open source community and from any other creators who want to contribute to Google's offerings. This open innovation system enhances the tools and technology at the heart of Google's business model.

Since it is still depending heavily on web search advertising, Google is trying to diversify its revenues sources[2]. So far, Google's Ad network represents 97% of the 2009 revenues. It faces serious competitors like Apple or Facebook[3] which are not making their customers' data available to Google. In 2010, by offering the innovative Social Box, Facebook is proposing to its members to use this platform as a personal communications hub[4]. If this offer is a success, it could bring more advertising on the Facebook network rather than competing directly against Google. But Google may get increasing ad revenues as more advertising moves to the Web. Google is also moving into other territories as it tries to get content from TV networks to supply GoogleTV services[5].

The innovation strategy must support the business model and not lose time and precious resources outside the center of gravity of the firm. At this end, Google is using an aggressive mix of Buy and Build innovations[6]. It bought Android for around $50M in 2005.In buying innovation, integration is the key. The founder of Android stayed with Google and was the champion of the development of Android's platform as an open-source operating system. As an open platform for mobile web connectivity, Android can be used by telecom operators as well as phone makers. As more people adopt Android, Google can expose potential customers to its services like online apps and bolster its presence in the mobile web. Smart phones can be used for e-commerce and also as another means for Google to deliver advertising, particularly local advertising since it is possible to localize people with mobile phones.

Google's mobile strategy is not to charge for the operating system itself, but the company profits from mobile ads displayed on Android phones. In 2010, mobile ads represent around $1 billion in revenue for Google. This revenue comes mainly from one of its latest

---

[1] For a comprehensive analysis of Google's open innovation ecosystem, see: B. Iyer and T. Davenport, "Reverse engineering Google's innovation machine", Harvard Business Review, April 2008.

[2] See: "How long will Google's magic last?, The Economist, December 4th, 2010.

[3] According to Experian Hitwise cited in the article "How long will Google's magic last", The Economist, December 4th , 2010, Facebook has surpassed Google in 2010 in terms of Web traffic in America.

[4] Om Malik, "Meet the New, New Facebook", from GigaOM, November 16th, 2010, reported by Business Week on line.

[5] In "Searching for the Future of Television", MIT technology Review, January/February, 2011, Richard D.Hof explores the future of television where TV sets are transformed into computer that be connected to the Internet and where the Net will have swallowed all information transmission.

[6] Google is the leader in the number of acquisitions for 2009-2010 in communications industries and made over $8 billion in acquisitions since its creation in 1998.

acquisition: AdMob. While mobile ad networks are very appealing and growing very fast, Google had to pay $750M for AdMob, a mobile-advertising start-up in November 2009[1]. Timing and integration is almost everything in acquisitions and the payoff on this investment will be more long term. In fact, the future strategic positioning of Google resides strongly on the premise that there will be more new Mobile Internet users than online internet users in the world soon. In fact, Android is now the fastest growing platform for high-end smartphones, a tough opponent for Apple, RIM, and Nokia.

With its developing technology combined with its formidable data bank on web users, Google is now in position to offer what telecom companies would like to offer as a value added service, namely infomediation service[2].

2.1.3. Individuals and households

For individuals, the Internet really took of in the Web 2.0 era where they took control of a growing part of the content diffused on the Net, thanks to new platforms like MySpace, You Tube, and more recently Facebook. In addition, many households are doing transactions over the Internet as most companies, institutions, and governments have a Web site offering transactions facilities. The spectacular growth of Facebook is another demonstration of the capacity of the Internet to generate network effects at world level and therefore creating huge traffic on the Web. The growing use of mobile smart phones and tablets is another major trend as mobile High Speed Internet access (HSIA) becomes ubiquitous. But so far, phone calls represent 75% of mobile revenues[3]. Voice is leveling off and data traffic will rise as cell phones are used to access the Internet with a growing offer of more services and phones can also be used to connect a PC to the Internet. The success of smart phones, e-readers, and tablets challenge the telcos that must provide bandwidth for these devices and it is estimated that mobile devices will become the more important way to access the Internet[4]. People all around the world want to be connected and present on platforms where they can communicate, show and download pictures, videos, music, and other digital content.

2.1.4. The possibility of another Internet bubble

Economic history tells us that new and promising technology translates itself into a speculative bubble in the stock market like electricity in the 1890s. As the Internet was emerging, dot-com companies were able to raise a lot of money in the capital market and their value was rising in the stock market. Expectations of soaring profits and a great potential for takeover targets were driving the value of "Pure plays" in the capital market. But the lack of sound business models, unrealistic predictions about the growth of e-commerce, and a lack of experience when dealing with the real world economy such as delivering goods to customers led to the dot-com crash of the early 2000s. Only a few firms managed to survive.

---

[1] http://www.businessinsider.com/google-to-acquire-mobile-ad-network-admob-for-750-million-in-stock-2009-

[2] J. Hagel III and J.F.Rayport were among the first to talk about web infomediation services customized for web users in "The New Infomediaries", McKinsey Quarterly, no 4, 1997.

[3] Scott Woolley, "Going from Calls to Connections", MIT Technology Review, November-December 2010.

[4] Stephen Cass "Can Mobile Operator Afford to Keep Up with the Demand for Bandwidth", MIT Technology Review, November-December 2010.

Despite the crash, it seems that investors have remained remain very bullish about Internet companies. In 2002, eBay bought Paypal which was a key player in its ecosystem. It paid $1.5 Billion and later in 2007, it posted a write-off and valued PayPal at $1.2B indicating that it had overestimated the value of this company. In 2005 eBay Inc. bought Luxembourg-based Skype Technologies, a global Internet communications company, for approximately $2.6 billion[1]. Skype was supposed to generate substantial synergies and make eBay's core business much more profitable. Other new profitable activities were expected from the integration of Skype in the eBay ecosystem. In 2009, eBay sold Skype for $2.7 billion as the expected synergies did not materialize. eBay kept 35% of the shares in the firm[2]. Google bought YouTube in 2006 for $1.65 billion. Since YouTube was still an unprofitable startup, analysts at that time estimated that it was a very high price to pay. Google's CEO Eric Schmidt said later that Google estimated the value of YouTube at $600 million and thus paid a premium of $1 billion to make sure no other investor would make an offer[3]. It is indeed a very high premium to pay to buy a non profitable start-up.

The rise of social networks began among other start-ups with MySpace and later on with Facebook. The investment of Goldman Sachs in Facebook in the fall of 2010 was implicitly putting the value of Facebook at $50 billion. Facebook's value more than tripled in 2010 only and in the secondary market, the value was reaching $76 billion on January 23, 2011. Some argue that network effects are so huge that Facebook is very appealing for advertisers, but other analysts doubt that a firm with an unproven business model is worth more than media giants[4]. Three years ago, the leader in social networking was still MySpace, which was bought for $580 million by News Corp. in 2005. Recently, News Corp mentioned that the subsidiary is up for sale after cutting 50% of its employees.

Social networking is the new appealing sector and now you can find a 4 year old firm like Zynga, which is a social video game firm, with a value higher than Electronic Arts (EA), which is a 28 year old firm in video games. Zynga is now valued on the secondary market at $5.65 billion on SharesPost[5]. EA is worth $5.02 billion in public trading on the Nasdaq stock market[6][7]. Zynga got the momentum when in the middle of 2009 they launched FarmVille, which is still the No. 1 game on Facebook with 57.4 million monthly active users. With such popular games, Zynga can cross-promote its titles and advertise them as well, allowing it to turn lots of its games into huge hits. EA bought Playfish for $400 million in the fall of 2009, but is still behind Zynga in that area. However, EA's online game revenue at $750 million in the current fiscal year, or around 20 percent of their overall revenue, is significantly bigger than Zynga's online game revenue, the only source of revenue for Zynga. It appears that Zynga's market values is equal to EA in market shares, so the market is deeply discounting the rest of EA's nearly $3 billion or so in traditional video game console and PC game revenues. It seems that Zynga is truly overvalued.

---

[1] http://about.skype.com/2005/09/ebay_to_acquire_skype.html

[2] http://news.bbc.co.uk/2/hi/technology/8231072.stm

[3] http://www.huffingtonpost.com/2009/10/06/eric-schmidt-on-youtube-d_n_311390.html

[4] "Is Facebook really worth $50 billion?"' The Economist, January 8th 2011.

[5] As of January 23, 2011.

[6] As of January 23, 2011.

[7] The SharesPost listings are thinly traded compared to EA's stock, but it is perhaps the only real measure of the value of Zynga's stock at any given moment.

Another example of exuberance is Apple, which has 75% of the market capitalization of Exxon Mobil[1]. Apple has a P/E ratio of 18 and Exxon has a low 14. Apple is one of the best innovators in the world and has created a dependency for its customers toward its proprietary platforms, such as iTunes and the Apple Apps store. Apple is now more a telecom and a content firm than a hardware firm. The potential of its mobile advertising network is huge. The question is: can Apple create, on the long term, 75% of the profits of a firm such as Exxon Mobil?

Groupon is another example as the firm rejected a $6 billion offer from Google in December 2010. Now, it seems that they are planning a spring 2011 IPO of around $15 billion. Groupon is the fastest Internet firm to reach 2 billion in revenue in its 2 years of existence. Gross and net margins are very high. Groupon and number 2 LivingSocial have around 92% of web traffic in the group-buying discount market.

However, there are more than 500 imitators with the same business model targeting national, regional, and international markets. Most of them even have the same web site design as Groupon. It is very easy to copy the recipe of success of Groupon. The leader already has the first mover advantage and is an important player in many cities in creating critical mass and network effects. They have hundreds of salespeople to get new daily discounted group-buying deals. They also enjoy temporary winning conditions. For instance, when web advertising started to be implemented with high speed Internet, banner ads and video ads were very lucrative. Now, with the proliferation of Internet advertising, these online ads are not so lucrative per click. Google is planning to compete against Groupon with the launch of Google Offers. So the future of Groupon is uncertain and the value of the firm may be much less than initially expected.

In the short term, nothing currently seems to suggest a sudden drop in P/E ratios of these Internet firms. However in the medium and long term, there might be an important depreciation of overvalued Internet stocks. Competition is moving fast, technological changes are disturbing established markets, and business practices and easy to copy business models face severe threats. A considerable drop in the value of Internet companies would impact all communication companies. Internet firms can be partners or competitors for the telcos. So the latter must be prudent either in partnering or investing into Internet firms to keep or enlarge their client base or to grab higher value added operations.

2.2. Telcos in this New Brave World

Since IT and the Internet are ubiquitous everywhere in the economy, there is a huge and expanding demand for bandwidth, landline or mobile HSIA, software and hardware, data storage and processing, privacy and security systems, web apps, and other services related to all kinds of electronic exchanges. The growth of social networks is phenomenal at a moment where wireless technologies are substantially improving. This is at the origin of social changes where people will keep in touch with each other locally, nationally, and across borders[2]. The demand for wireless services will expand enormously[1].

[1] At January 23rd, 2011.

[2] In "Grown up Digital", op.cit, Don Tapscott describes how the Net generation is changing the world as we know it.

The extent to which telcos and their immediate competitors like the cable companies can offer all these services depends upon their business strategy, their profitability, and access to the capital market. Their strategy must take into account the competition coming from a large variety of companies like IT consultants, niche players, new comers like Google, and others. Regulation will also have an impact on telco's strategy. In particular, the debate about Net neutrality will also have a significant impact on the strategy retained by telecom service providers. Net neutrality will be discussed in the regulation section.

3. TELECOM PROVIDERS' STRATEGIES: THE CHOICE OF BUSINESS MODELS

The telecommunications landscape was substantially altered over the last decade. New technologies have increased competition and have commoditized many services offered by telcos. Several new devices as endpoints have made telecom services more complex. Several open platforms of communication are available and must be interoperable to attract users. As indicated before, a whole array of new technologies became available in the recent past. Just to illustrate the acceleration of technical progress and to show the challenge for telco strategies, let us sum up as follows:

The entry of Vonage followed by many others offering VoIP. This technology "has killed the phone business as we knew it"[2] as time and distance have no more value added.

In 2008, RIM has taken its Blackberry into 3G territory whereas Apple has launched its iPhone.

By developing a line of unified communication products that deliver complete solutions, IBM in 2008 wins the Frost and Sullivan Award designed to reward innovation. For telcos, information technology consultants are strong competitors for value added services.

In 2008-2009, Skype entered the wireless market as well as Google with Android. Once very lucrative, termination and roaming charges are now being commoditized by these innovations.

In 2010, Google launched its Nexus one phone and Apple its iPad Tablet. As a consequence, market power has shifted from telcos to service providers and equipment producers. On the demand side, the customers are in the driver seat as they select devices and services from many competing offerings. The growth in data communications has outstripped voice, and wireless services are to become the main technology to access the Internet worldwide.

It has become very difficult for any communications company to stand alone in a business environment made turbulent by incremental and disruptive innovations. Firms are then opening their frontiers and thus going from a push to the market approach to a pull one where partners including customers are contributing to define products and services adapted to market needs[3]. This strategy is the very opposite of the former telco incumbents' ways of

[1] Scott Woolley, op.cit.

[2] From the front page of The Economist, September 15th, 2005.

[3] John Seely Brown and John Hagel III, "From push to pull: the next frontier of innovation", Mckinsey Quarterly, August 2005.

doing business where they were deciding what was best for customers. Telcos have to migrate from proprietary protocols and customers control to open Internet platforms. They also have to redefine their relationship with partners and let new comers from equipment makers to software or Internet companies be part of their ecosystem, even if some of these partners can be competitors in other markets. This is co-opetition[1], namely letting competitors be your partners in your ecosystem for specific products or services.

As carriers were attempting to get scale and network economies to face the commoditization of bandwidth and basic IP services, a wave of US carrier's consolidation unfolded in the first decade of the 2000s. After the horizontal integration of RBOCS like Pacific Telesis and Ameritech, SBC Communications made a vertical integration by merging with AT&T. The new AT&T bought Bell South and rebranded Cingular Wireless as an AT&T product. In 2006, Verizon bought MCI to acquire its networks and rebranded MCI services as Verizon Business. Reconstructed new giants AT&T and Verizon are new strong competitors, particularly in the wireless market.

In Canada, consolidation occurred earlier as some new LD carriers were bought by incumbents. The year 2003 was the last step in the consolidation wave as AT&T Canada formed from the merger of MetroNet, AT&T Canada Long Distance Services, Netcom Canada, and ACC Tel Enterprises, which came back into the market by merging with Manitoba carrier MTS under the name of MTS Allstream.

We now review stylized versions of telecommunication company's current business models and look at what strategies telecom providers may have to put forward in order to stay competitive in this new environment described above.

3.1. One Stop Shop (OSS) Versus Broker Strategy

The OSS[1] strategy aimed at enabling a company to establish a major presence or even a dominant position in the communications market. This strategy is inspired by the monopolistic culture of traditional telcos, which are used to offer all services from landline to mobile communications. The former incumbents generally still have an important client base, financial resources, and a large assets portfolio enabling them to put forward a marketing strategy consisting of bundling many services for a flat price in order to maximize the number of subscribers. Currently, after the consolidation of carriers in the US, Verizon and AT&T have a comprehensive OSS offer including landline phone, HSIA, VoIP, TV, and wireless services for residential, businesses of all sizes, and governmental customers. In this OSS approach, the telco owns all the assets or many of them delineating its field of operation. In contrast, the broker, or syndication approach consists in offering customers a full range of services while not owning all the infrastructures, but only some key assets as indicated below.

[1] The co-opetition concept is not new as presented in G. Hamel, Y.L. Doz, and C.K. Prahalad, "Collaborate with Your Competitors--and Win". Harvard Business Review, 67(1), 1989. More recently, co-opetition is integrated into open innovation business model or ecosystem, see for example: H. W., Chesbrough, and M. M. Appleyard "Open Innovation and Strategy" California Management Review, 50(1), 2007.

As customers are solicited by competitors, a bureaucratic approach when dealing with them must be absolutely avoided. Large telcos proposing an OSS wicket to their customers must eliminate silos in the organization and become flexible by eliminating rigid business practices inherited from a monopolistic past and make all relevant information for all the services offered easy to access and available on any channel (web site, telephone, physical facilities) offered to customers. A well-designed interactive web site is a relatively inexpensive channel to interact with customers, allowing human resources to be deployed in more value added functions like marketing tasks to gain new customers. A consistently satisfactory experience must be available to all customers to sustain competition. In addition, telcos should unify all the information about a customer in one folder in order to do some infomediation like Internet firms. Making a personal offer for new products or services or new service bundling that may interest a customer helps in developing client loyalty.

In a highly competitive market and with prices for basic services driven down to commodity levels, telcos must climb a staircase of value addition to increase margins and grow[2]. Since most incumbent carriers like AT&T and BCE have divested their R&D arms (Lucent and Nortel in occurrence), they are now followers in terms of technology. The danger is to be reduced to supply dumb pipes at commodity prices. With innovative new comers in the telecom market, large telcos should welcome quite a few of them as part of their ecosystem and also extend their relationships with their current partners. Offering Apple's iPhone or Google's Nexus one and making the necessary investment in infrastructure to support these devices is one way for acceding innovative products[3]. Five years ago, AT&T through a partnership with Apple, was offering the iPhone. Verizon is now trying to catch up with its main competitor and currently partnering with Apple to offer the iPhone[4]. Large carriers opting for the OSS strategy will have more opportunities than ever to open their ecosystem to innovative partners.

In December 2010, Google with its partner Samsung launched its Nexus S and Android 2.3 Gingerbread. Since the device is offering wireless VoIP, again large carriers should see if offering that device would be the best strategy to employ. But Telcos should also be part of the innovative process. In an open innovation ecosystem, every member participates in the innovation process[5] by exchanging information. Thanks to their many contacts with a whole variety of customers widely spread through the economy, the Telcos can be a good source of

[1] The acronym OSS for "One Stop Shop" used in this chapter must not be confused with the technical acronym "operations support system" or OSS. We use OSS here as a business strategy for telecom companies.

[2] M. Baghai, S. Coley and D. White, "Staircases to Growth", The McKinsey Quarterly, Number 4, 1996.

[3] For example, AT&T was the first partner selected by Apple to provide the mobile network necessary to support the use of the iPad. But other carriers were eager to offer the wireless data services necessary to use the tablet. A comparable scenario happened when AT&T offered Google's Nexus one. It shows the competition that large carriers have to deal with. See: Olga Kharif, "Freeing the iPad from At&T", Business Week, March 30th 2010.

[4] Rosabeth M. Kanter, "Verizon, the iPhone and the Power of Second Chances", Innovation and Strategy, Harvard Business Review, Weekly Hotlist: Thank You for Doing Your Job, January 18th, 2011.

[5] H. W. Chesbrough, "The Case for Open Business Models", MIT Sloan Management Review, winter 2007.

information for their innovative partners. Joint ventures in delivering new services to customers can facilitate large carrier access to higher value added products.

Media companies have for a long time been Telco's partners. But the media industry is changing rapidly as underlined above. Web 2.0 web sites like You Tube, smartphones connected to the Internet, and bloggers and social networks are now present in the media landscape. These new comers are raising issues about Net neutrality discussed above in the regulation section. But putting this aspect aside, telcos would gain by partnering with those new medias that are bound to grow as the digital generations grow older. Offering space and web apps on different platforms to new media could allow telcos to get some advertising revenue and compete with Internet companies. Advertising revenue is one way to compensate for commodity prices in basic telecom services.

3.2. Convergence: Telcos and Content

The so-called convergence approach was put forward in the media by the merger of AOL and Time Warner in 2000. Previous acquisitions of Rogers with MacLean Hunter in 1994, Disney and ABC, Viacom were previously "convergence acquisitions", but not mediatized like AOL and Time Warner, which was the biggest merger in the history of infocom industries. AOL was to give an Internet access to the content of T-W whereas the cable assets of T-W were to allow AOL to offer HSIA. It was in fact nothing more than a traditional vertical integration strategy. Historically, this strategy has not always been successful[1]. One of the factors behind these findings is the fact that the success of a merger is not determined by costs or hypothetical economies of scale, scope, or network, but rather by the growth in revenues of the new entity. The problems created by merging different corporate cultures tends to complicate the situation even further, eventually leading to the disposal of assets and, at the very least, a partial dismantling of the conglomerate. Instead of relying on T-W content only, AOL would have probably fared better by making alliances with several content providers and by making partnerships with telecom or cable companies to get access to HSIA services. This first convergence merger in the Internet era was an unfortunate failure.

In Canada, during the telecom bubble era in the late 90s, BCE combined the OSS approach with the convergence strategy by making several acquisitions. Bell was offering local and LD landline services to companies, households, institutions, and government. It was also present in mobile communications (Bell Mobility) and satellite cable distribution (Bell ExpressVu). The acquisition of Teleglobe gave Bell important capacity in Internet transmission at the world level. To complete this OSS strategy, BCE invested in an e-commerce service company (Emergis) and in a systems integrator (CGI). Following the convergence approach and getting content to be distributed over its transmission capabilities, BCE invested in two TV stations (English CTV and French TQS), in a newspaper (Globe and

[1] See for example: M.M. Bekier, A.J. Ogardus and T. Oldham, "Why Mergers Fail', McKinsey Quarterly no. 4 (2001); S. A. Christofferson, R. S. McNish et D. L.Sias, "Where Mergers Go Wrong : Most buyers routinely overvalue the synergies to be had from acquisitions. They should learn from experience", The McKinsey Quarterly, May 2004 and T. Straub, "Reasons for Frequent Failure in Mergers and Acquisitions : A Comprehensive Analysis", Gabler Edition Wissenschaft, 2007.

Mail), and finally in a sports center in Montreal. “Connectivity, commerce, and content” was the advertising slogan used to describe the strategy behind the conglomerate built by BCE’s management. The strategy relied on the synergies among the different assets making up the conglomerate and enabling BCE to offer infomediation activities that were supposed to lead to a broad range of transactions by its customers. The results of this strategy were not very different from the historical ones referred to previously. The synergies achieved between the components of the conglomerate and the results of cross-marketing efforts focused on products offered by BCE were not up to expectations. Notably, the competition from Internet companies like e-Bay, Amazon, and Google was such that the expected network effects generated by BCE portals for businesses, households, and others clients did not materialize. Then, the management, after the burst of the telecom bubble, decided to put assets related to the convergence strategy on sale. Some were sold whereas for others like CTV, BCE did not find buyers at a satisfactory price.

Like in the AOL case, BCE could have followed a syndication approach[1] by forming partnerships with different content providers without investing in these companies, except perhaps in some firms whose services are estimated to be essential to the success of the strategy. For example, e-Bay bought PayPal but not all the other partners involved in its value chain. Before any agreements are made with partners, potential suppliers can be made to compete so as to generate the best quality/price ratios. The company can thus secure a wide range of content needed to satisfy its customer’s demands. There is always the possibility to change suppliers to adjust to new market trends. The syndication approach carries transactional and complexity costs for the leader, but it remains far more flexible than the vertical-integration approach. It is a strategy that is well known in the information industry. For example, a newspaper company must deal with many information sources. It may own some essential providers, but in general it will reach agreements with various suppliers.

But in the view of BCE management, the convergence approach is not dead. On the 10th of September in 2010, surprisingly enough, BCE announced it was taking complete control of CTV[2]. The strategy here is to broaden the media corporation's content assets as those assets will be used across Bell platforms (mobile, online, and television services). BCE management believes the adoption of mobile TV is set to accelerate and video is going to be an integral part of Bell's product offerings and a key growth driver in the near future.

However, BCE is not the only carrier in Canada to own content providers. Cable and wireless operator Rogers is also combining an OSS business model with an impressive content business[3]. Rogers is a profitable company offering cable telephony, cable TV, and HSIA to the residential and business sectors. It is also a leader in the wireless industry with infrastructure across Canada and offering wireless HSIA, Blackberry, and Apple’s iPhone as well as others brands of wireless devices. Furthermore, Quebecor Media (QMI) is another

---

[1] See for example: K. Werbach, “Syndication: The Emerging Business Model in he Internet Era,” Harvard Business Review, May-June 2000.

[2] http://www.ctv.ca/CTVNews/TopStories/20100910/ctv-bce-100910/

[3] Rogers media group is as follows. The Radio arm of Rogers operates 54 radio stations across Canada, while its Television properties include the five television station and The Shopping Channel. Rogers Sportsnet, a specialty sports television service licensed to provide regional sports programming across Canada and Media's Sports Entertainment assets include the Toronto Blue Jays Baseball Club and Rogers Centre, Canada's largest sports and entertainment facility. Media's Publishing group produces 70 well-known consumer magazines and trade and professional publications in Canada.

company integrating a cable company (Videotron), French Canadian television stations, and a large array of publications both in French and English across Canada. QMI has a strong Internet presence in both official languages and is offering cable telephony and HSIA to consumers and businesses. It has recently invested in a new wireless network in Quebec that is to be interconnected with other carriers in Canada. QMi is profitable and promotes its convergence model and is a strong competitor to BCE and Rogers mainly in the Quebec market.

Telus, the main telco competitor of Bell and Rogers in Canada, has no intention to follow the convergence strategy. Telus does not plan to own the content it puts on its TV, and its online and wireless services will not match competitors such as Bell and Rogers by buying up media assets. Telus is following the syndication approach as it plans to aggregate and collect the best content, whether it comes from capabilities in Canada, the U.S., or internationally[1]. In addition, Telus is poised to offer IPTV services with its cutting edge fiber optic network[2] and will challenge cable competitors by offering an OSS package to theirs residential customers.

In the US, on January the 18th of 2011, the FCC after long debates finally approved the purchase by Comcast of a majority stake in NBC Universal. The largest American cable company will be offering all its cable services and local TV stations, getting content from a film studio and will be competing with Disney as one of the world's biggest media companies[3]. This is another attempt at vertically integrating content with carrier assets to get as much viewership as possible.

3.3. Niche Players

The large carriers in the US control about 90% of wire-line service revenues[4]. There are several companies offering local, LD, Internet access, and international services. Many of these small firms are reselling carrying capacity belonging to larger carriers. On the wireless side, about 900 firms provide cellular, paging, and other mobile services[5]. With a regulation that has lowered barriers to entry, small and midsize players rent or purchase bandwidth for the local and LD markets with a relatively small amount of capital expenditure. They can offer calling cards as well as data cards and Internet access at prices that create a downward pressure on subscriber rates. Large carriers do not welcome these small companies that are targeting lucrative markets like large cities and have no interest to serve less populated areas. Incumbents underscore that this mechanism creates asymmetric competition with little value added for customers. Nonetheless, these small players, by pushing down prices, force large carriers to improve their productivity and change business practices to offer competitive prices to their customers.

---

[1] The Canadian Press, September 14th, 2010.

[2] Iain Marlow, "Telcos make assault on cable's stranglehold", Report on Business, The Globe and Mail, January 24th, 2011.

[3] "Big bad media, poor little internet", The Economist on-line, January 19th, 2011. The transaction still needs to be approved by the Department of Justice.

[4] http://researchwikis.com/Telecommunications_Marketing_Research

[5] Ibidem

But niche players also refer to high value added services. Since IT and the Internet are at the heart of communications, business models, and public services, there is a huge market for specialized services. Software producers, system integrators, and other IT consultant companies can offer a long list of specialized services, given the multitude of electronic exchanges at the local, national, or world level. Niche players can specialize at an industry level, for a certain layer of services, or certain electronic platforms and so on.

3.3.1. Cloud computing

Currently, cloud computing including secure data storage and several web based applications is a growing service. In fact the "cloud of clouds" [1] has three layers:

- Software as a service or SaaS: it offers web-based applications such as Google's different services and also Salesforce.com services helping companies to follow their customers. There are many providers in this layers offering specific services and according to Forrester Research, the value of services sold in this layer was $11.7B in 2010[2].
- The second and deeper layer is "platform as a service" or PaaS, which is an operating system embodied in the cloud offered by Google, Microsoft, and Salesforce.com. There is a limited number of firms involved in this layer and Forrester Research estimates revenues at $311M only.
- Finally, the third layer called "infrastructure as a service" or IaaS is better known, and studies on cloud computing are referring to it. IaaS offers computing services, data mining, and storage that customers combine to obtain adaptable computer systems. Currently, GoGrid, Rackspace, and well known Amazon Web Services are the main players in that layer.

There is a potential market for providers that could own some carrying capacity and offer new services in one of the Cloud Computing layers. The first and third ones are probably easier to penetrate and the main value added resides in the services offered. For corporations embracing the business model presented in Chart 1 above, a firm can offer a secured virtual private network (VPN) tailored to the needs of their ecosystems and offer computing services to the entire ecosystem.

Some players offer e-commerce platforms made for a certain industry. Or niche players can adopt a horizontal platform strategy offering specialized services in many industries. New Wireless ISPs are also another market niche that can be complementary to Mobile Network Operators but also disruptive for ILEC business[3].

A firm offering a value added service like security can be part of a few larger carrier ecosystems. In that case, there is always the possibility that a large carrier will buy a niche company if the provided services become an essential component of the carrier business model. A niche player can also team up with equipment producers. For example, Intel sees itself as a catalyst for change and wants to accelerate the arrival of open multi-vendors of

---

[1] "Information technology goes global: Tanks in the cloud", The Economist, January 1st, 2011.
[2] Ibidem
[3] See: "Towerstream's Manhattan Wi-Fi Network – Close To True 4G", At: Http://Www.Glgroup.Com/News/Towerstreams-Manhattan-Wi-Fi-Network---Close-To-True-4G-51928.Html, December 17TH, 2010.

cloud computing solutions[1]. In high-tech industries, like aerospace and health, there is a need for very sophisticated services that large carriers are not offering.

Google[2] and Apple are considered niche players in the telecom market, partnering with large carriers to offer their products. Both companies are exploiting their strong brand and image to create links with the final customer and redefining the way a niche player can be part of large carrier ecosystem.

3.3.2. Infocom industries realigning

Another way to look at niche players is to define all the communication industries as Content, Appliances, and Transmission or carrying over networks (see Figure 2). Then a firm like Google can be considered a horizontal integrator. Google is partnering with appliance producers by offering Nexus one[3] (made by Taiwan's HTC Corporation) or a new notebook Chrome OS to be launched in 2011. Then Google needs partners to support Internet traffic over wireless networks. Google will supply content for these appliances. For example, with network reach and reliability reaching a point where cloud-based solutions can be considered viable, and remote servers already being used to allow the mobile Internet and email, it is possible that Google will launch the first cloud OS service for mobile devices in 2011. Other companies can act as horizontal integrators for various services. Many analysts think that the year 2011 might be the year of the smartphone for Google, which will dominate the wireless landscape and force telcos to define a strategy to either make strategic alliances with the Internet firm or to compete against it with other partners.

As for horizontal integrators, different industries would benefit from an integration of appliances/content/transmission. The health industry where highly specialized services are required is likely a promising market for horizontal integrators.

From another perspective, the business model of Google can be qualified as multi-sided platforms[4]. The main strength of the model resides in extremely targeted ads around the world. With AdWords, advertisers can publish advertisements or sponsor links on Google's search page and affiliated content network (auction service for key words). Thus, they can tailor online campaigns to specific segments of the population. With a critical mass of online consumers using Gmail, Google Maps, Picasa (online album), and its powerful search engine, Google attracts many visitors to its network. For publishers, Google AdSense offers third parties the opportunity to earn advertising revenue adjusted to the content of their web sites. Thus, Google's business model makes money from one customer segment: advertisers, while it is also subsidizing free offers to two other kinds of segments: web surfers and content owners. Finally, Google's core competencies are tightly concentrated with highly complex

---

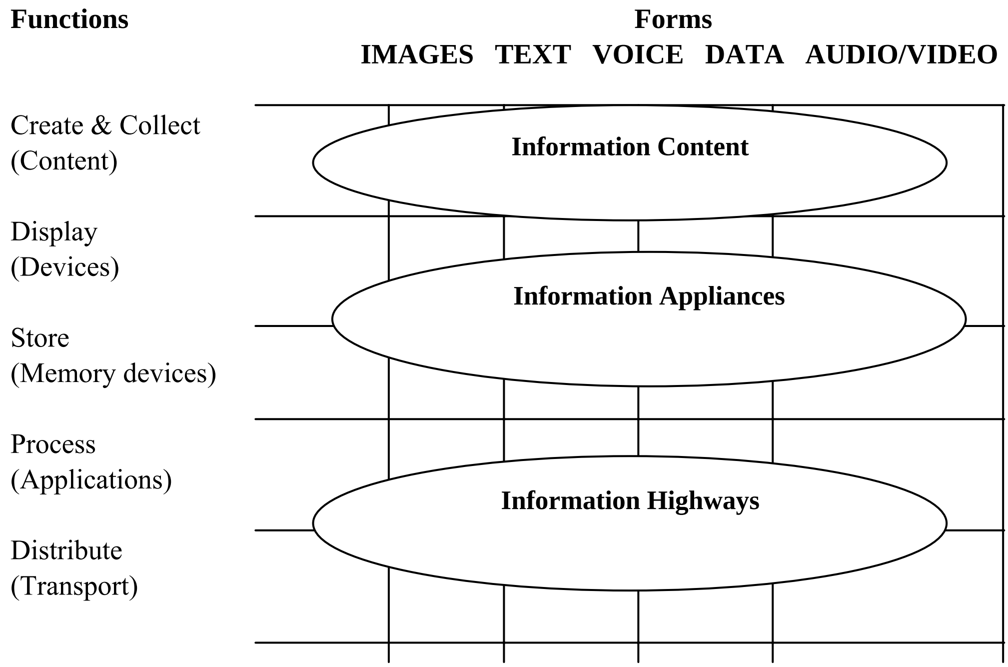
[1] http://www.intel.com/itcenter/cloud/index.htm?cid=cim:ggl|cloud_ca_gen_content|as5B99DC

[2] It is even possible that Google buy wireless capacity and compete head on with Verizon and AT&TMobile. See: Rob Pegoraro, "Will Google compete with Verizon, At&T in mobile market", Washington Post, December 30th, 2010.

[3] "Google's Android dethroned Nokia's Symbian as the most popular smartphone platform in the last quarter of 2010, ending a reign that began with the birth of the industry 10 years ago", in "Google topples Symbian from smartphones top spot", by Tarmo Virki, Reuters, January 31th, 2011

[4] Osterwalder, Alexander and Yves Pigneur, Business model Generation, 2010, John Wiley & Sons. New Jersey.

proprietary search and matchmaking algorithms and an engineer culture, which are supported by an extensive network of IT infrastructure[1].



Source: The New Information Industry, Web Proforum Tutorials, www.iec.org)

Figure 2. How the information industry is realigning: opportunities for vertical integrators.

3.3.3. Focus players

Focus players are another category of niche player. Skype is a good example of a niche player of that sort. Skype, like Adobe, is offering its Internet telephony services for free like other Internet firms . Customers pay premiums for more advanced options or additional capacity and other services. The Skype worldwide network effects are so large that it takes only a small percentage of customers willing to pay for more services to make it profitable. TeleGeography was putting Skype shares of all international calls at 8%[2] in 2008 and at 13% in 2009[3]. Skype traffic is growing constantly and Skype is now the largest provider of cross-border voice communications services in the world, according to Telegeography.

Skype uses wholesale carriers such as iBasis and Level 3 for handling its network traffic. As any VoIP new player, Skype is a threat to incumbents in all countries. But it can move to offer other services like local services such as Skype Find, extending its reach in markets

---

[1] Osterwalder, Alexander and Yves Pigneur, Business model Generation, 2010, John Wiley & Sons. New Jersey.

[2] Om Malik, "Skype, Now the Largest Long-distance Phone Company", GigaOm, march 24th , 2009.

[3] http://www.inquisitr.com/71802/skype-commands-13-percent-of-international-calls/, May 3th, 2010.

dominated by local carriers. Skype is also partnering with equipment makers like Nokia as its wireless phone allows Skype's customers to call directly over the Internet. In addition, Skype can also get revenues from advertising on its Web site[1]. On October 14, 2010, Skype 5.0 for Windows was released with a number of improvements and feature additions, including a Facebook tab to allow users to SMS, chat with, or call their Facebook friends via Skype from the News Feed[2]. On January 14th, 2011, Skype announced the acquisition of Qik which is a Video Sharing mobile platform[3].

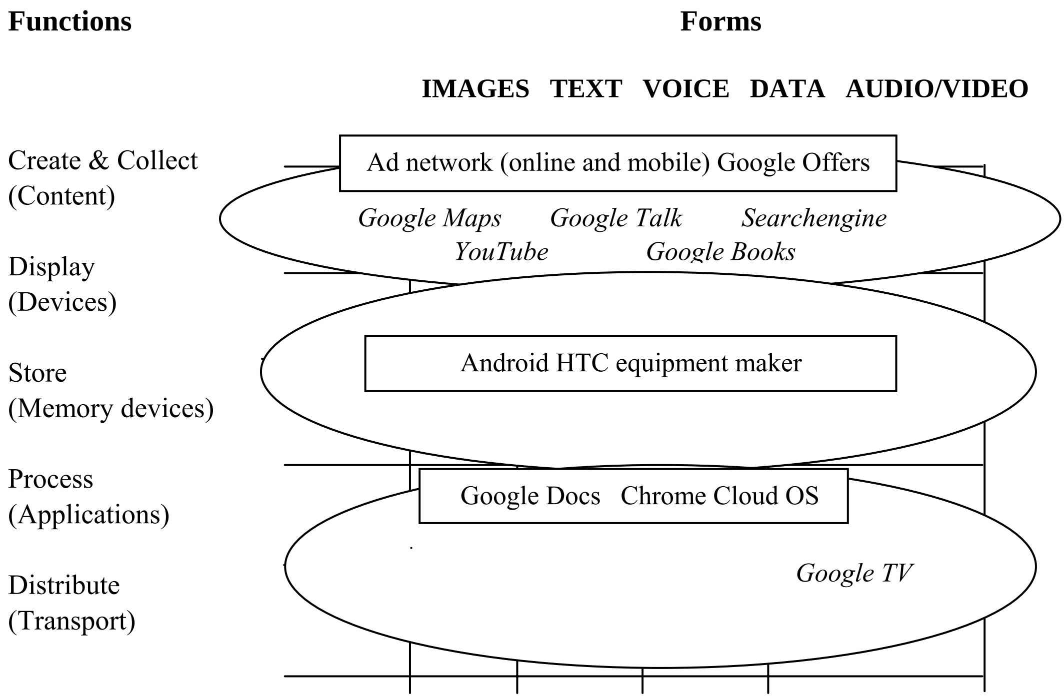


Figure 3. Google's business model: involved in almost all segments of the Information industry.

The Internet phone company is expected to go public in 2011[4]. Therefore, wire-line and wireless telecommunication operators have to take into consideration the growing and disturbing influence of Skype in their business strategy.

Another focus player is Facebook, which now has more than 600 million users around the world. Facebook earns almost all of its revenues from online advertising. The firm's expenses are rising (servers, storage, bandwidth) while it adds millions of new users. Over the past year, Facebook launched several versions of the site in foreign languages.

Facebook is buying its tech equipment from a variety of established manufacturers. It represents: "a simpler yet more expensive approach than that of Google, which saves money by building its own servers and storage systems and writing complex new software to tie

[1] Erick Schonfeld, "A New Business Model For Skype: Turning Phone Numbers On The Web Into Paid Ads", http://techcrunch.com/, April 17th, 2009.

[2] "Skype gets serious Facebook Integration upcoming IPO". Mashable.com. Oct 14 2010.

[3] http://digitalanalog.in/2011/01/15/skype-buys-out-qik-video-sharing-mobile-platform/

[4] http://www.cnbc.com/id/41276513/Skype_Delays_IPO_to_Second_Half_of_2011_Sources

everything together... They are being forced to build at scale with expensive technology[1].” Facebook’s annual revenue has more than doubled since last year and is estimated at 2 billion in 2010[2]. It appears that the company now competes with Google for online ad dollars from major brands as well as small businesses. Facebook supplied around 50 billion display ads per month and is on track to serve 1 trillion display ads for all of 2011[3].

A new trend for Facebook is relying heavily on virtual currency (revenue that comes, in a roundabout way, directly from Facebook users themselves). The popularity of online games from supplier Zynga (presented above) enables Facebook to earn money when users pay with Facebook Credits; for every dollar the user spends on Facebook Credits[4], the social network gets 30 cents. However, the majority of Facebook’s revenue still comes from advertising and marketing. Google and Facebook are both targeting location-based ad dollars, especially where small and medium-sized businesses are concerned. According to Mashable: “How Facebook performs in 2011 may have a lot to do with how it competes in offering a compelling and effective location-based marketing platform for businesses of all sizes.” Online advertising analysts say that while the majority of the growth in social networks is coming from overseas , the relative immaturity and small size of those ad markets makes it difficult to earn strong revenues from that growth[5]. There is a large discrepancy between the secondary market value of Facebook, (which reach 76 billion in January 2011 as indicated above) and the level of its revenues (about 2 billion in 2010). Given the size of Facebook’s customer base and the traffic generated worldwide by the site, telcos will have to determine in their strategy what kind of relation they envisage with this Internet firm. Will Telcos want to have Facebook in their ecosystem as a partner, and if so, under what conditions given the present market forces?

3.3.4. Consumer Internet business models

Consumer Internet business models are emerging as interesting new niche players. The following table provides some with key monetization drivers.

In the 13 consumer Internet business models, there are 3 or 4 key monetization drivers for each, and it gives a company the scale it needs to achieve to get to an annual revenue run-rate of $10 million. These are among the most important drivers to build a sustainable consumer Internet business[6]. Among the most interesting niche players are Groupon (presented above) and Twitter. Twitter is a social networking and microblogging service, enabling its users to send and read messages called tweets. It is sometimes described as the “SMS of the Internet.[7]”

---

[1] http://www.businessweek.com/technology/content/mar2009/tc20090326_604141_page_2.htm

[2] http://mashable.com/2010/12/16/facebook-2-billion-revenue/

[3] Jennifer Van Grove, “ Facebook Now On Pace to Serve 1 Trillion Display Ads Per Year”, Mashable.com, November 8th, 2010

[4] Mashable.com/tag/Facebook-Credits

[5] http://www.businessweek.com/technology/content/mar2009/tc20090326_604141_page_2.htm

[6] Source: http://tctechcrunch.files.wordpress.com/2010/10/chart-2-consumer-internet-startup-models-overview.jpg

[7] http://en.wikipedia.org/wiki/Twitter

Startup Type	Examples	Key Monetization Driver(s)	To Generate $10MM In Annual Revenue
1. Search	Powerset, Hunch	Monthly Uniques, Queries Per Month, % Clicks, Revenue Per Click	5MM, 10, 5%, $.50
2. Gaming	Zynga, Playdom	Monthly/Daily Average Users, Conversion Rate, Monthly Spend	10MM MAU, 1.5% Conversion, $5.50/Month
3. Social Network	MyYearBook	Unique Visitors, Ad Impressions, Sellthrough Rate, CPM	3MM, 450MM, 75%, $2.00
4. New Media	Twitter, Foursquare	Unique Users, Actions, % Monetizable, CPM, CPA	3MM, 10/Month, 1.5%, $1.50
5. Marketplace	Etsy, Prosper	Listings, Listing Fee, Sales, Commission	2MM, $.20, $12.5MM, 3.5%
6. Video	5Min	Unique Viewers, Ad Impressions, Sellthrough Rate, CPM	120MM, 108MM, 90%, $8.00
7. Commerce	Groupon, Guilt, Modcloth	Uniques, Conversion Rate, Average Spend, Gross Margin, Acquisition Cost	4MM, 10%, $50.00, 50%, 30%
8. Rental	Chegg	Uniques, Conversion, Average Rental, Customer Acquisition Cost	1.5MM, 10%, $140, 30%
9. Subscription	Box.net	Uniques, Conversion Rate, Customer Acquisition, Churn Rate, Life Time Value	1MM, 1%, 40%, 5%, 20 Months
10. Audio	Pandora	Uniques, Ad Impressions, CPM, Conversion Rate, Upsell Value	10MM, 320MM, $2.00, 1%, $2.50
11. Lead Generation	Mint	Unique Visitors, Offers Viewed, Conversion Rate, Affiliate Cost Per Action	3MM, 3MM, .6%, $50.00
12. Hardware	Kno	Units Sold, Gross Margin, Marketing	175K, $200/Unit, 40% Gross Margin, 30% Marketing
13. Payments	Zong, Venmo	Unique Users, Average Payment, Transaction Fee	1MM Users, $25 Average Payment, 3.5% Fee

(c) Steven A. Carpenter, 2010

Source: http://tctechcrunch.files.wordpress.com/2010/10/chart-2-consumer-internet-startup models-overview.jpg

According to the founder of Twitter, the company did not have a real business model until recently since they are now targeting advertising revenue. They concentrated on getting a critical mass of subscribers first which then gave rise to an exponential traffic growth. They now have more than 190 million users, which are tweeting 65 million times a day[1]. Twitter's market value has reached $4 billion in January 2011, just a month after it raised $200 million in Venture Capital funding[2]. Twitter's new business model will rely on charging for access to its full data stream. According to Giacom[3], it will feature promoted tweets in search results and promoted trends in its trending topics[4]. It will publicize sales and other deals. Twitter is even considering inviting users to pay to promote their Twitter accounts. Promoted tweets would represent a Twitter-specific version of Google AdWords. Thus, Twitter is still experimenting with new sources of revenue. As for Facebook, Telcos will have to determine if they want to have Twitter as a partner and at what conditions.

3.4. Telecom Broker and Convergence of IT Technology and Telecom

A McKinsey survey showed that large corporations CIO in order to reduce cost and organizational complexity, and would like to sign one contract to obtain end-to-end service from a provider offering a combination of IT and telecommunications services rather than two or more agreements with a few suppliers[5]. They also want "standardized services that can

---

[1] Erick Schonfeld, "Costolo: Twitter now has 190 million users, which are tweeting 65 million times a day", TechCrunch, June 8th, 2010.

[2] http://mashable.com/2011/01/25/twitter-now-worth-4-billion/

[3] http://gigaom.com/2010/04/12/the-twitter-ad-model-revealed-what-were-you-expecting-a-pony

[4] http://gigaom.com/2010/06/17/when-is-a-twitter-trend-not-a-trend-when-its-promoted/

[5] Andreas Franz, Claudia Funke and Katrin Suder, "A rising demand for integrated IT and telecom services", McKinsey Quarterly, May 2006.

be developed inexpensively and delivered wherever their employees conduct business."[1] The broker business model could be the answer to such "a not so easy to satisfy demand."

The broker who may own telecom assets will also use the services of other suppliers who own telecom or other types of assets like software, content, and so forth. In a rapidly changing environment, a telecom broker must have in-depth knowledge of all telecommunications aspects as well as IT development in order to meet the needs of customers and be able to manage an entire ecosystem of service providers. The telecom broker is assuming the complexity and transaction costs in assembling the best solutions for its clients. It will have to find suppliers in its ecosystem capable of delivering the value added services needed by its customers . If necessary, it will have to find and put under contract new providers for certain services.

For example, in the context of the fragmentation and specialization of the computer market, IBM sold many of its hardware assets and in the 1990s adopted the broker approach when it created its Global Services consulting business, in which specialists determine the best solutions for customers, including products from suppliers other than IBM. The company does not have to be a leader in all market niches, but instead focuses on its vast knowledge of computer technology. In doing so, it can innovate by creating tailored made solutions for customers. As indicated above, IBM was honored in 2008 for producing a line of unified communications products that deliver complete solutions for their customers.

The British carrier BT is among the early adopters of the IT telecom convergence business model. In 2005, it made a few strategic acquisitions to enlarge its presence at world level. With a network sprawled across 120 countries and with a large number of experts in its world organization, BT intended to be by 2010 a global supplier of IT and telecom services capable of offering an OSS service for large corporations and other international organizations. In 2005, BT was considered a leader and a visionary with a top level ability to execute by the telecom consulting company Gartner in its so-called Magic Quadrant for Global Network Service Providers. In 2005, it launched its BT Fusion services offering fixed and mobile network partnering with Vodaphone. BT offered a network to large corporations: the "21st Century Network", the largest IP network in the world. This resulted in BT owning a high density and cutting edge IP and MPLS world network.

BT contracts with Unilever and CEMEA (a branch of Visa) illustrate the value brought by a competent telecom broker. Instead of dealing with 200 suppliers, Unilever had with BT a single contract spanning over 7 years and estimated at 1B Euros[2]. In July 2010, Unilever extended its contract with BT's Global Services. The contract for the four year covers integrated communications services, supply chain management, and wireless technology in more than 100 nations where Unilever is present[3]. UK Top News describes the nature of the agreement as follows:[4]

"*The span of the service entails the drawing, management and operation of a safe, fully incorporated back to back IT networking infrastructure, providing data, voice, video and mobility services to approx 1,000 Unilever locations.*"

---

[1] Ibidem

[2] http://www.computerweekly.com/Articles/2006/05/16/215973/Unilever-signs-163270m-extension-to-BT-contract.htm

[3] http://topnews.co.uk/28941-unilever-widens-bt-contract

[4] Ibidem

The contract with CEMEA[1] implies the deployment of a solution network entirely secured and covering 256 sites. The solution reduces waiting time by 40% and is almost available all the time.

Since 2005, BT has continued to make acquisitions in order to enlarge its global network. The BT business model comprises some risk. There is, like in most business, a cyclical risk and BT was hit by the global recession. As a former monopolist, BT has to respect previous agreements with its employee unions. Pensions to be paid to retired employees are a heavy burden in recession time and BT made a loss in 2009 and was forced to cut jobs. In addition, by providing a large array of services, it has to deal with complex situations and some unexpected problems can crop up and be expensive to solve. It is crucial to satisfy its clients and confidence is at the heart of long term contract with large corporations.

In its 2010 Magic Quadrant, consultant Gartner saw the global network service provider market as very competitive with new comers and strong competition from regional carriers[2]. The competition ensures downward pressure on prices and the possibility for companies to get advantageous contracts from carriers. In 2010, AT&T and Verizon Business were in the leaders and visionaries quadrant and both had an ability to execute, estimated by Gartner, somewhat better than BT. The BT model could then be challenged by large contenders in the years ahead.

4. Regulation in Canada and the US

4.1. Opening Competition: Brief Reminder

Opening competition in the US LD market occurred in 80s when AT&T was broken up into local operators (RBOCs) and an LD provider who kept the AT&T Company name.

Major competitors of AT&T, such as MCI or Sprint, were able to achieve profitability in the LD market before the Internet boom. Opening the competition in the Canadian LD market came in 1992 and in 1994 the CRTC ended the partitioning of the communication industry by allowing cable companies to offer telephony and the telcos to offer cable and video services. Moreover, there was no break up of the incumbents which remained present in both local and long-distance markets. The CRTC has adopted a number of regulations designed to prevent incumbents from shutting out rivals through predatory pricing policies. As a result, a certain number of companies were able to operate in the LD market in Canada.

The FCC opened competition in the local loop in 1996 and the CRTC followed in 1997. When an innovation like the Internet is changing telecom technology, rate-setting to allow new entrants to rent bandwidth from incumbents is a complicated task.

The new technology based on IP data transmission is substantially less costly than the technology associated with the PSTN networks inherited from the past. In theory, prices should reflect network access costs obtained in a competitive environment. But if the regulator allows incumbents to claim its historical costs, new entrants do not pay only the

---

[1] http://www.docstoc.com/docs/68195491/Integration-Planning

[2] Neil Rickard and Robert F. Mason, “Magic Quadrant for Global Network Service Providers”, Gartner RAS Core Research Note G00174070, March 8th, 2010.

marginal cost of the service received, but must also assume a part of the depreciation cost of past investments. Depreciation is irrelevant since prices are determined by current supply and demand conditions. The answer given to this question by the FCC in 1996 was to use the concept of "forward looking cost," an approach based on the costs of the most recent technology to set access rates. The RBOCs vigorously contested this approach and won in part on constitutional grounds. In Canada, the CRTC attempted to follow an intermediate course by allowing incumbents to recoup a substantial portion of their historical costs by using the concept of long-term incremental costs as a basis to set access rates. In fact, this a complicated issue where shareholders must be taken into account.

Once the rules were in place, the CRTC followed the "forbearance principle" by which the regulator let market forces determine prices and other characteristics of the service. The LD market for example is free from any regulator's rules and market forces determine prices. In the US, the FCC referred also to market forces in the telecom market as the foundation of its decision. However, in 2005, the FCC removed requirements that incumbent local carriers must provide DSL broadband access to third-party ISPs under the same terms and conditions as their own ISP. The FCC referred to a Supreme Court decision, saying that the Internet is classified as an "information service," and not a "telecommunications service," as the basis of its decision[1]. This decision was likely to be an important factor in a certain lack of competition in broadband service in the US[2].

4.2. Technology as a Source of Competition

The result of these rather complex rules was a first step toward a competitive telecom market. The impact of competition opening by regulators spanned over a decade and featured a period of chaotic pricing for bandwidth. Then, a good deal of telecom competition came from technology. VoIP, cable telephony, and high speed wireless Internet access (HSWIA) are innovations that have brought serious competition in the telecom market and have contributed to commoditize the price of bandwidth.

Cable companies initially offering HSIA were somewhat slow to enter the telephony market. The VoIP technology gave the cable industry the opportunity to offer IP telephony and bundle television, HSIA, and telephony to their subscribers. The entry of companies like Vonage offering VoIP as long as the customer was subscribing to a HSIA raised other regulation questions. One is the possibility by incumbents to offer VoIP at prices low enough to crowd out all new comers. To avoid that sort of predatory pricing, in 2005 the CRTC in Canada set some price setting rules for incumbents. But incumbents were arguing that the market is already competitive enough to let its forces determine prices for VoIP services. The

---

[1] B. Washburn, "The FCC Eliminates DSL Internet Sharing, Handing RBOCs Freedom over Broadband Resale", Current Analysis, Advisory Report, August 11, 2005, http://www.currentanalysis.com/r/2005/reports/files/CIR_14295.pdf

[2] According to research firm Strategy Analytics, The United States ranks 23rd out of 57 countries in terms of "broadband development". At: http://www.strategyanalytics.com/default.aspx?mod=PressReleaseViewer&a0=4930. See also: "The Web New Walls, The Economist, September 2th, 2010.

incumbents asked the federal government to reverse the CRTC decision and in a rather surprising move, the minister's cabinet did reverse in 2006 the CRTC ruling and thus, let the market determine prices.

In the US, the VoIP debate raised different issues. When offerings of VoIP were hitting the telecom market in 2004, SBC Communications was planning to place higher fees on rival Internet phone companies. The FCC at that time found that Vonage's VOIP phone service is an "interstate" service that is not subject to state jurisdiction, but did not deliver a comprehensive framework setting the rules for VoIP providers and local incumbents. But federal regulators in 2004 stated that that voice communications flowing entirely over the Internet are not subject to traditional government regulations. Then one VoIP provider, Pulver.com, was declared to be immune from government rules, taxes, and other requirements that were applying to traditional telephony[1].

In the US, the FCC claimed there should be no restraint from incumbents for those players offering VoIP services[2]. However, most of the calls through a VoIP provider will be terminated through public-switched telephone networks (PSTN). It thus becomes more complicated when one takes into account the termination of the call, but the FCC did not provide a ruling on that aspect when it published its ruling about VoIP providers.

4.3. Net Neutrality

The Net neutrality issue has been discussed for a few years. There is no definitive version of Net neutrality[3]. Technically, it means that the Internet network should be delivering packets of information from one place to another, whatever the content, origin, and destination might be. The democratic principle behind neutrality is freedom of speech and the universal right to access all the information available in the world.

But it raises quite a few economic issues. The price mechanism remains the more efficient way to equilibrate supply and demand of bandwidth. Without it, some users might hog all the bandwidth and make it difficult for other users to exchange information. Carriers will lack the capacity to satisfy demand. Or some users might be ready to pay more than others to obtain more services like faster transmission of their data on the Internet. Then, if some users are ready to pay more, carriers will invest more to meet that demand by extending capacity or by being more innovative. So it is not easy for the regulators to make rules that will preserve to some extent the basic principles of Net neutrality and at the same time not harm the economy.

In addition, assuring equilibrium between supply and demand is more difficult with wireless technology. When wired networks become congested, phone or cable operators can

---

[1] Declan McCullagh and Ben Charny, "Pure" VoIP not a phone service",
http://news.cnet.com/2100-7352_3-5158105.html, February 12, 2004.

[2] "New FCC Rules Should Benefit VoIP Providers and Users",
http://voxilla.com/2009/09/21/new-fcc-rules-should-benefit-voip-providers-and-users-2337, September 21st, 2009.

[3] "Ask five geeks and you may well be given six definitions of it", citation from "A Tangled Web", The Economist, December 29th 2010.

add more connections which cannot be done by wireless operators since they are licensed to use fixed portions of the radio spectrum[1].

This is why Google, historically a strong supporter of Net neutrality, changed its position in 2010. Google recognised that it would not be optimal to apply the neutrality principle to all wireless networks since when "capacity is tight, there a risk of huge traffic jam that benefits no user.[2]" Google joined Verizon to submit a proposal that is presented as a compromise. Google argues that broadband providers should not be able to discriminate against or prioritize lawful Internet content, applications, or services in a way that causes harm to users or competition[3]. However, wireless broadband providers would be free to manage their network since the mobile marketplace is more competitive and changing rapidly and subject to the technical constraint mentioned above[4].

Given that debate, the FCC came up with middle of the road rules maintaining that regulation is needed to prevent network operators discriminating in favor of their own services. The new Net neutrality rules adopted by the FCC create two classes of service subject to different rules: one that applies to fixed broadband networks and one for wireless networks. The FCC says this is necessary because wireless networks are technologically different from fixed broadband networks.

One rule[5] prohibits the blocking of lawful traffic on the Internet and allows reasonable network management. Another rule requires both wireless and wireline providers to be transparent in how they manage and operate their networks. This way, consumers and corporations can see what content might be blocked and the reasons put forward to do so. Operators can manage spam or take actions against attacks on the network. A third rule is prohibiting "unreasonable discrimination," a statement that puts a lot of discretion in the hands of the FCC. In other words, the FCC has not banned fast-tracking, but does not indicate to what extent it will be tolerated. An interpretation of this rule would be that some packets of information cannot run "unreasonably faster" than others on the network. In addition, the FCC states that if customers pay for having some priority, it would raise significant cause for concern. This rule will be difficult to interpret and apply and may lead to sterile squabbles where resources will be allocated to legal battles instead of being directed to improve technology. And it goes against existing practices that do not hamper the free circulation of information on the Net. Consumers can subscribe to different sorts of Internet access. A consumer can pay more for a superfast Internet connection without causing harm to other customers who subscribe for the normal HSIA from carriers. Large corporations operating worldwide pay for Internet business services that may include special data delivery networks faster than the average traffic speed on the network.

---

[1] Stephen Cass "Can Mobile Operator Afford to Keep Up with the Demand for Bandwidth", MIT Technology Review, November-December 2010.

[2] Paul Boutin, "Google Misses You", MIT Technology Review, November-December 2010, p. 76.

[3] A classic example of discrimination would be a cable company offering broadband internet access and television services might try to block internet-based video to avoid competition with its television content.

[4] Omar El Akkad and Iain Marlow, "Google, Verizon target pricing for wireless usage", The Globe and Mail, august 9th, 2010.

[5] A summary of the rulings done by CNetNews.com is reproduced in annex 1.

As for the wireless technology, a field currently exhibiting strong growth since the launch of smart phones will remain much less regulated. The FCC was therefore open to the remarks and proposals of operators and new comers like Google. Then, for the wireless networks, operators will be permitted to block some traffic, but there are no economic measures proposed to facilitate the network's management. A usage-based pricing would be an efficient way to do the job. A monthly flat rate offered for a certain data limit beyond which the customer pays a certain rate per megabyte is a classic way to avoid excess traffic. But the FCC does not indicate if this measure is acceptable.

For wireline, there is also a legal problem with the ruling since cable giant Comcast won a federal court case against the FCC. The court ruled Comcast could essentially regulate traffic on its system, slowing down certain types of data. Hence, the court ruling questions the FCC's ability to enforce network neutrality principles[1]. This judgment will likely be an incentive to carriers opposed to the rulings to go to court to neutralize the FCC's rulings.

To some observers, the debate about Net neutrality underscores the lack of competition in the supply of Internet access[2]. In other industrialized countries, consumers have a larger choice of ISPs and therefore they can change their provider if they are not satisfied with its traffic management. The key here is the regulation compelling incumbent carriers to open their network on a wholesale basis to ISPs. This ensures competition between several providers offering competitive prices and service. Getting cable operators and incumbent carriers in the US to open their networks on a wholesale basis could have been more efficient in obtaining an open Internet rather than neutrality regulation. But the FCC as we have seen has gone in the other direction by eliminating the obligation for incumbents to give access to competitor ISPs to their DSL network.

In Canada, the CRTC has already adopted a similar regulatory framework. ISPs and carriers are not allowed to alter Internet traffic, but they can take measures to manage their networks to make things run smoothly. But the practice must be transparent to users and take place only when necessary[3]. The CRTC also allow ISPs to charge more for bandwidth, or giving discounts in off-peak hours, which are standard practices when running networks with limited capacities . Furthermore, opening incumbent networks on a wholesale basis exists in Canada and the large presence of cable operators in the market secures a relatively high level of competition for the benefit of residential and to some extent business customers. Since Canada is a huge country with a small population, HSIA availability remains a problem in remote areas. Wireless Internet access is however helping to address this problem.

Finally in Europe, countries of the E.U. will not adopt Net neutrality rules because regulators believe competition is strong enough among ISPs to prevent carriers from blocking or prioritizing certain content[4].

4.3.1. Planning with neutrality rules

The FCC rulings seem to open up the possibility for High-speed Internet providers like the cable operators and large carriers to "reasonably" manage their networks and perhaps

[1] http://www.reuters.com/article/idUSN0622428720100406 April 6th 2010.
[2] "A Tangled Web", The Economist, December 29th 2010.
[3] Richard Blackwell, "CRTC sets Web 'throttling' rules", The Globe and Mail, January 8th, 2010.
[4] Cecilia Kang "E.U. won't adopt net neutrality law", Washington Post, November 11th , 2010.

charge consumers rates based on levels of Internet usage. In that case, the impact on day to day carrier operations would be minimal.

However, the remaining uncertainty might affect investment and innovations. Telcos and cable operators might delay their investment if doubt remains as to how they will be able to run their network. For example, some carriers are investing in fibber-optic technology to upgrade, at high cost, the "last mile" of their network. Then the strategy to get a return on the investment would be to offer the bundle "HSIA, IP telephony, and IP-TV" covering virtually all the needs of their customers for wired services. But in peak hours, the carrier will have to give priority to their customers for IP-TV which requires a lot of bandwidth[1]. Under the proposed rulings, that could be forbidden and make uncertain the investment profitability.

Under these conditions, before engaging corporate funds into major investments or partnering with equipment producers to finance R&D to support new technological advances, carriers will go to courts to clarify the FCC rulings[2]. Since legal battles can be a long and expensive process, investments and innovations might be delayed and funds allocated to unproductive activities.

The FCC proposal gives a positive signal to wireless network providers and equipment companies and new comers like Google and Apple. So, as demand for wireless devices and services will continue to grow, network providers will have to invest in new infrastructures and also bolster existing capacity. But carriers will have to make a careful planning of their capital expenditures. Expanding wireless networks will have to be supported by backhaul links and avoid bottleneck problems, as investment will be needed as an expansive fibber optical backbone. In addition, wireless capacity may pose a problem as people will be more and more connected to the Internet with mobile connections[3]. Securing the balance between demand and supply in a wireless capacity will be a challenge for all carriers in the coming years.

5. COMING TRENDS

("Forecasting is something very difficult, particularly when it deals with the future." Yogi Berra)

We now look at main trends coming in the near future with short comments. Like wireline operators in 1998-2001, wireless operators are seeing increases in traffic, which is not associated with increases in revenue. However, more devices are becoming available on their networks, and not just smart phones: laptops, embedded devices, machine-to-machine modules, iPads, and electronic readers, each with their own pricing models, subsidies (or not),

[1] "Resource Management for IPTV distribution", ERCIM News, April 22th, 2009.

[2] Verizon has already announced that it is challenging FCC's rulings. See: Cecilia Kang, "What next for net neutrality? Lawsuits galore", Washington Post, January 24th, 2011.

3 Tom Simonite, "Feeding the Brandwidth Beast", MIT Technology Review, November-December 2010.

and revenue opportunities. Using the network intelligence, telcos can find new business models and revenue opportunities[1].

Every player in the communications industries wants to get closer with the end user, including new entrants such as digital media companies. There is also increased monitoring by government at the federal, state, and local levels; more competition for the customer from device manufacturers, operating system providers, and even the channel itself[2]. It appears that with penetration rates soon reaching saturation, North American operators are looking for new devices and new verticals to gain customers. Those verticals are healthcare, smart-grid solutions, and the automotive industry. However, each one of those segments likely has its own business model and strategy. Tiered pricing plans are interesting options. Customers will pay a premium for some mobile video, much like people pay modest premiums to get high-definition video on demand from their cable providers. A probable scenario could be that the mobile operator will deliver the video to the end user behind the scenes, and the end user will deal directly with a third party content provider. New business models open up with those types of agreements. Major opportunities reside with video content, including peer-to-peer sharing[3].

5.1. Six Major Trends Influencing Communications Players' Business Models:

- Demand for content will rise because of the proliferation of apps, mobile platforms, widely adoption of high-speed Internet, cheaper smartphones, new powerful tablets, and cheaper eBook readers:
    - Demand for video content will rise
    - Demand for quality specialized texts will rise
- With fixed mobile data plans with unlimited bandwidth, demand for mobile network capacity will rise:
    - Suppliers become more strategic

Global wireless traffic increased by 160% to 90 petabytes per month in 2010, according to Cisco Systems. It has been referred to as the "Mobile data tsunami,". Operators have to manage network congestions. AT&T Mobility plans to invest $7 billion in 2011 for an upgrade to its wireless network, because data traffic increased 200% in 2009[4].

With new smartphones and tablets, Cisco forecasts mobile traffic will increase by 26X by 2015 (from 2010). It would reach 6.3 exabytes per month, from 237 petabytes in 2010.

---

1 Ford, Tracy, RCR Wireless, July 2010, "The forces of change, Wireless operators ride the wave of content, connections"

2 Ford, Tracy, RCR Wireless, July 2010, "The forces of change, Wireless operators ride the wave of content, connections".

3 Ford, Tracy, RCR Wireless, July 2010, "The forces of change, Wireless operators ride the wave of content, connections".

4 Pearson, Chris, RCR Wireless, January 2011, "The mobile data tsunami meets the spectrum shortage".

Cisco anticipates that 5.6 billion devices will connect to mobile networks in 2015 with video streaming comprising 66% of global traffic (50% in 2010)[1].

- With the proliferation of free or cheap unlimited video content offerings (ie. YouTube, Hulu, Netflix), and fixed Internet data plans, demand for Internet capacity bandwidth will rise:
  - Suppliers become more strategic.

It is projected that by 2012, the sum of all forms of online video, including TV, video on demand, Internet, and peer-to-peer (P2P) file transfer, will account for nearly 90 percent of all consumer Internet traffic, a large portion of which will flow through OTT[2] applications[3].

- The intersection between social networks, mobile commerce, and location-based services will offer several new growth avenues:
  - Enterprise spending on Web 2.0 collaboration technologies is forecast to grow to $4.6 billion globally by 2013, with social networking as the top spending category[4].

Unique visitors to social networking sites in June 2008 represents approximately two-thirds of the world's Internet audience[5]. In June 2010, Nielsen reported that social networking is used more than twice as much (online time) as any other activity. Sites like Facebook and Twitter now account for 22.7% of time spent on the web; the next closest activity is online games, which make up 10.2%[6].
Forecasters estimate that by 2012, mobile social networking will represent a market opportunity of between $22.5 and $52 billion[7].

- Firms offering strategic tools for making tailored advertising (Internet online and mobile) with mass personalization and strategic tools for business intelligence of consumers' evolution and forecasting will prosper:

---

1 Lalonde, Denis, February 2th, 2011, "Le trafic mobile mondial multiplié par 26 d'ici 2015", Direction informatique.

2 OTT: "over the top" providers, such as social networking sites, are becoming ever-more viable platforms for communication services – and consumers are responding eagerly. van den Dam, Rob, Ekow Nelson and Zygmunt Lozinski, "The changing face of communication, Social networking's growing influence on telecom providers", IBM Institute for Business Value, 2010

3 "Cisco Visual Networking Index: Forecast and Methodology, 2007-2012." Cisco. June 16, 2008. http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360_ns827_Networking_Solutions_White_Paper.html.

4 "Enterprise Web 2.0 worth $4.6 billion in 2013." searchviews. April 21, 2008. http://www.searchviews.com/index.php/archives/2008/04/enterprise-web-20-worth-46-billionin-2013.php.

5 "Social Networking Explodes Worldwide as Sites Increase their Focus on Cultural Relevance." comScore. August 12, 2008. http://www.comscore.com/press/release.asp?press=2396

6 http://mashable.com/2010/08/02/stats-time-spent-online/

7 "Mobile Social Networking Revenues Could Reach US$52 Billion by 2012." Cellular-news. February 11, 2008; "Mobile Web 2.0 revenues to reach $22.4bn by 2013, driven by User Generated Content and Social Networking." Juniper Research press release. May 14, 2008.
http://juniperresearch.com/shop/viewpressrelease.php?pr=91

  - o The latest generation of Software as a Service (SaaS) technology is transforming the landscape, bringing personalisation to the mass market.

There should be huge opportunities for companies willing and able to leverage low cost SaaS solutions to move towards true personalisation, enabling organisations to achieve online what is already being done offline with “propensity modelling and other business analytics.” According to *Mark Simpson of Maxymiser,* it represents a significant development beyond the segmentation, recommendation, and retargeting, that can already be achieved online, providing organizations with an opportunity to truly get close to the online customer for the first time[1].

- Useful cloud computing initiatives will provide major cost savings for customers:
  - o Verizon Communications bought Terremark Worldwide for $1.4 billion. It follows a trend toward more operator involvement and ownership of enterprise services delivered in the cloud. Communication service providers are transforming themselves into technology providers.

“Cloud computing (as indicated previously) continues to fundamentally alter the way enterprises procure, deploy and manage IT resources, and this combination helps create a tipping point for everything as a service,” said Lowell McAdam, president and COO at Verizon. Telecom operators provide 9%, or $1 billion, of global cloud services revenue to enterprises, but analysis projects telcos’ share in the market to rise to 23% by 2015 to around $8.2 billion[2].

5.2. Skype Influence

Skype has 850 employees, mostly engineers in Estonia, and is also headquartered in Luxembourg; their marketing department is in London, and the audio-visual works department is in Stockholm, Sweden. New CEO Bates plans to recruit up to 400 employees, of which 80% will come from Silicon Valley: “Business is changing rapidly...What is important is not the communication inside the firewall, or contact with one or two partners, but the coverage of communications services.” Thus, Skype wants to meet the requirements of most companies when considering the development of a service level agreement (SLA) to ensure service quality, and to provide new services. Recently, Skype experienced a major shutdown for a day. CEO Bates said the incident does not indicate that Skype’s peer network’s “infrastructure” is flawed, but that some parts of it have problems related with the carrying capacity[3].

According to Kijl et al., **“**From a technological capabilities view, technological change can be seen as competence-enhancing or incremental from an organizational perspective if the capabilities (e.g. skills, knowledge, assets, and resources) required to exploit new technology are built on existing firm capabilities. New entrants do not suffer from these problems, and

1 http://www.utalkmarketing.com/Pages/Article.aspx?ArticleID=19477

2 http://www.rcrwireless.com/article/20110128/ENABLERAPPS/110129948

3 http://sillyfox.wordpress.com/2011/01/17/skype-plans-to-increase-the-number-of-employees-and-50/

therefore are more likely to develop products and services based on new technology (i.e. Skype)."[1]

Afuah and Tucci (2003) propose the Internet technology life cycle model to offer a global perspective on the phasing process in which companies in an emerging industry are developing their services and business models[2]. They distinguish the following three phases:

- The emerging or fluid phase: in this phase, (a lot of) new entrants as well as incumbent players choose their profit sites and value network positions. There is competition between new and old technologies and different designs using new technology. Product quality is low, costs and prices are high, market penetration is low with mostly lead users and high-income users as customers. Since product/service and market requirements are still ambiguous, there are few failures in this phase.
  I.e. Mobile apps, Cloud computing
- The growth or transitional phase: in this phase, a standard or dominant design defines a critical point in the life cycle of the innovation. The customer base moves to mass market. Competition and disappearing requirement ambiguity forces many firms to exit or make important changes in their business models. Firms with the best adapted business models will survive.
  I.e. VoIP
- The mature or stable phase: in this phase, companies focus on keeping and improving their competitive advantages. In markets where imitation is easy (e.g. most (mobile) Internet services), companies continuously make (incremental) innovations to their business models.
  I.e. Voice 3G, roaming revenues

The influence of Skype over wireless operators is becoming critical. Evalueserve predicted a long time ago in 2005 that wireless operators will experience reduced mobile revenues due to:

- Reduction in revenues from roaming fees due to Skype/WiFi combinations
- In the case of Skype introducing a solution that can be used over 3G/GPRS, the mobile voice traffic will decline further[3].

Skype just recently launched new services enabling some customers from Verizon Wireless to make free WiFi Skype-to Skype phone calls and for a low fee, with Skypeout (Skype-to other customers). They are starting with Verizon customers and plan to expand globally.

---

1 Kijl, Björn et al., "Developing a dynamic business model framework for emerging mobile services", Working paper, userpage.fu-berlin.de/~jmueller/.../Kijl_Bouwman_Haaker_Faber.doc

2 A. Afuah, C.L. Tucci, Internet Business Models and Strategies, Text and Cases, second edition, McGrawHill, New York, 2003.

3 Evalueserve, "Impact of Skype on Telecom service providers", January 2005.

5.3. Social networks influence

Four categories of players are involved in the social networks dynamic: see Figure 4[1].

1-Traditional telcos: firms which have historically dominated two-way interpersonal conversations, are increasingly facing new market entrants that use open platforms to meet diverse and rapidly changing user wants and needs[2]. Thus, there is a shift in communication patterns from point-to-point, two-way conversations to many-to-many, collaborative communications. Moreover, control of the communication environment is transitioning from telcos to open Internet platform providers, pushed by better and cheaper technology, open standards, greater penetration of broadband services, and wireless communication networks.

2-Open and Free: refer to companies that offer one-to-one communication services, but through an open Internet platform and at no, or for a low fee. These services potentially threaten profitable traditional services, such as long-distance calling and mobile roaming. i.e. Skype, Google Talk, and MSN Messenger[3].

3-Gated Communities*:* *c*ompanies that focus on many-to-many communications, rather than point-to-point, within telecom-controlled environments. They are, essentially, a “walled-garden” for operator-led collaboration services and are likely to appeal to users and enterprises that desire secure and reliable communication environments[4].

4-Shared Social Spaces*:* firms that facilitate collaboration on the open Internet. i.e. MySpace, Twitter, and Facebook. These providers have the potential to become “de facto” integrated communication platforms, bringing together social networking, voice communication, e-mail, instant and text messaging, as well as content. They are getting attention away from traditional Telcos and contributing to the fragmentation of the market. While gaining an audience share, these services pose a serious threat to telcos as they “piggyback” on the existing communications infrastructure, imposing network capacity issues and increased costs for the network providers[5].

It appears that in the short term, as the industry experiences more open and collaborative communication models, the traditional model will likely remain dominant. However, in the long term, the industry can expect a paradigm shift towards models that facilitate collaboration and sharing, with “Shared Social Spaces” attracting a more significant portion of people’s communication time.

Telcos should embrace a more open and collaborative future. Opportunities in partnerships with, or acquiring, existing players to proactively develop the infocom capabilities required for success, or enabling other participants in the value chain. They can reduce the cost of delivering high-bandwidth content, and potentially capture value from it, by using network and computing infrastructure optimization techniques. Telcos are well

---

1 van den Dam, Rob, Ekow Nelson and Zygmunt Lozinski, “The changing face of communication, Social networking’s growing influence on telecom providers”, IBM Institute for Business Value, 2010.

2 van den Dam, Rob, Ekow Nelson and Zygmunt Lozinski, “The changing face of communication, Social networking’s growing influence on telecom providers”, IBM Institute for Business Value, 2010.

3 van den Dam, Rob, Ekow Nelson and Zygmunt Lozinski, “The changing face of communication, Social networking’s growing influence on telecom providers”, IBM Institute for Business Value, 2010.

4 van den Dam, Rob, Ekow Nelson and Zygmunt Lozinski, “The changing face of communication, Social networking’s growing influence on telecom providers”, IBM Institute for Business Value, 2010.

5 van den Dam, Rob, Ekow Nelson and Zygmunt Lozinski, “The changing face of communication, Social networking’s growing influence on telecom providers”, IBM Institute for Business Value, 2010.

positioned to enable quadruple plays: cross-platform, fully integrated experiences across mobile, fixed, and Internet Protocol television (IPTV) services. Thus, the status quo option is not a luxury many providers can afford.

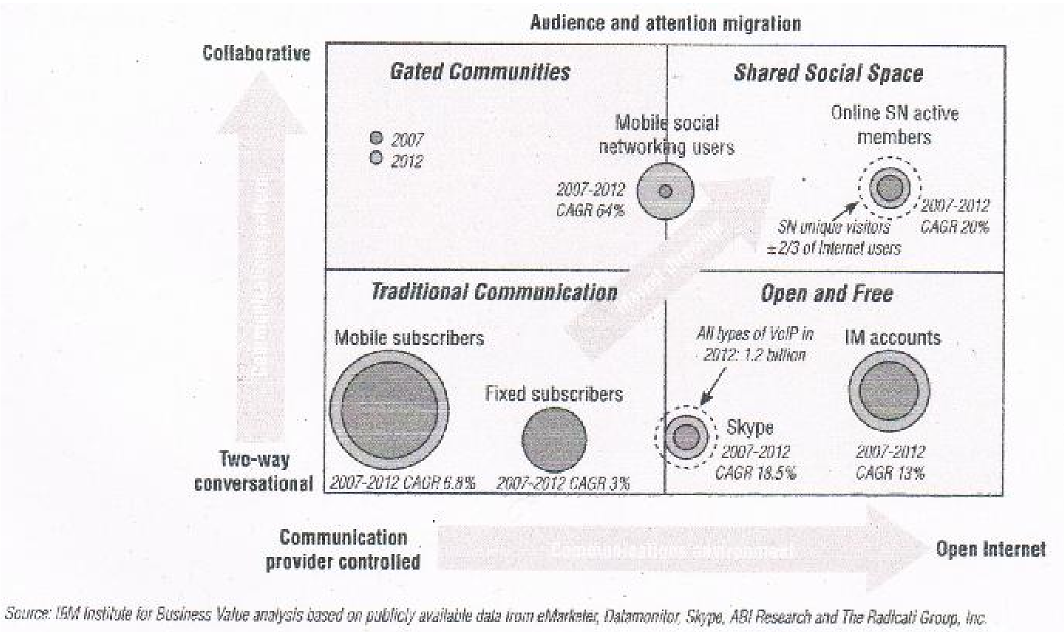


Figure 4. Shifts in communication control and patterns give rise to new business models.

5.3.1. An example of a telco initiative embracing social networking: Vodafone

Vodafone 360 brings together a customer's contacts, status updates, and messaging services from the mobile phone, social networks, and other Internet accounts, enhancing the customer's experience and use of social media. Customers have integrated contacts, music, photos, and mapping services and can share their favorite music choices – and even their physical location – how and when they choose with chosen groups of friends. The service is automatically backed up and synchronized, regularly and wirelessly, between a user's mobile device and computer[1].

Contrary to traditional communication models, emerging models are based on open platforms that support many-to many and/or collaborative communication patterns (see Figure 4). Traditional telcos are actually the largest segment in terms of revenue and subscribers, but they are showing signs of slow growth as other models take hold. Wireline revenue is declining and, although, according to Gartner Inc., global mobile services revenue is forecasted to grow 7.6 percent from 2007-2012, the mobile subscriber base has reached saturation in key developed markets[2]. It is essentially a "walled-garden" approach in which operators facilitate collaboration services that will appeal to users and enterprises with a

1 "Vodafone announces Vodafone 360." Vodafone press release. September 24, 2009.

2 Hahn, William L. and Nhat Pham. "Dataquest Insight: Global Telecommunications Market Take." Gartner, Inc. July 22, 2008.

preference for the more secure and reliable communications environment traditionally provided by telecoms[1].

On the other hand, "Gated Communities" include both fixed (online) and mobile social networking. The IBM Global Business Services report suggests that walled-garden social networking sites are unlikely to be successful as the majority of consumers have demonstrated a preference for open social networking sites. However, telcos possess a window of opportunity in this area, by enabling mobile social networking on existing mobile network architectures, which, for the most part, remain closed platforms. Proliferation of new powerful and cheaper devices will gradually blur the advantage of mobile exclusivity.

5.3.2. Facebook partnership

Several traditional telcos providers have formed partnerships with open social networking providers to incorporate such services into their telecom environments. For example, in October 2008, Telefonica signed a global agreement with Facebook allowing it to integrate access to Facebook's mobile service and applications from all of Telefonica's mobile portals[2].

It appears that towards the long term, the majority of incremental communication time will be supported by new providers and models. Traditional telcos face many challenges and they can become more competitive in the face of changing user sentiments and demands if they take strategic measures to adapt to this evolving marketplace[3]:

1- Exploit opportunities in the Gated Communities by enabling mobile social networking in current "closed" networks to capture a share of potential revenues, which are forecast to reach between $22 billion and $52 billion by 2012[4].

2- Partner with or acquire existing social networking players to proactively develop the capabilities required for success, including experimenting with new revenue-generating services.

    i.e. Vodafone's acquisition of Zyb: The social networking and online management tool for backing up and sharing contact and calendar information – demonstrates how a mobile operator can extend social networking from PCs to the mobile device with a view toward increasing data services revenue while delivering a much richer and unified communication experience[5].

3- Enable other participants in the value chain, such as advertisers, virtual operators, and application developers, to benefit from telecom capabilities such as location,

1 van den Dam, Rob, Ekow Nelson and Zygmunt Lozinski, "The changing face of communication, Social networking's growing influence on telecom providers", IBM Institute for Business Value, 2010

2 "Telefonica, Facebook strike global partnership deal." Telecom Paper. October 6, 2008. http://www.telecompaper.com/news/article.aspx?cid=639352

3 van den Dam, Rob, Ekow Nelson and Zygmunt Lozinski, "The changing face of communication, Social networking's growing influence on telecom providers", IBM Institute for Business Value, 2010

4 "Mobile Social Networking Revenues Could Reach US$52Billion by 2012." Cellular-news. February 11, 2008; "Mobile Web 2.0 revenues to reach $22.4bn by 2013, driven by User Generated Content and Social Networking." Juniper Research press release. May 14, 2008. http://juniperresearch.com/shop/viewpressrelease.php?pr=91

5 "Vodafone: ZYB acquisition facilitates more mobile social networking." Datamonitor. May 2008. http://www.datamonitor.com/industries/news/article/?pid=6A58957F-4BB1-42CD-A1EE-62D53B810155&type=CommentWire

presence, text/multimedia messaging services, and conference calling, while providing access to customer analytics to help enhance services and offerings.

4- Reinforce the capability to deliver fully integrated enterprise communication services that combine voice, Internet-based communication, and collaborative communication models.
   Partnerships with enterprise communication providers, software vendors, and/or system integrators can provide an efficient path to delivering these services.

5- Work more closely with CDN[1] and/or OTT providers to reduce the cost of delivering high-bandwidth content (e.g., video, music) in response to increasing demands for such services. This can be achieved through network and computing infrastructure optimization techniques such as traffic shaping, use of faster-than-real time progressive downloads in place of real time streaming, and caching of content close to the edge of the network using CDNs[2].

Such measures can lead to new business models that capture more value from this increasing social network traffic, which can result in an enhanced user experience through improved quality of service and innovative service partnerships, based on this optimized environment.

5.4. Open Innovation

Given this rapidly moving environment, if not completely done, incumbent telcos must move from a push approach where telcos decided what is good for their clients to a pull approach where a company is reaching out to take advantage of partners and clients in shaping its strategy to increase client base and value added from its offerings. In addition, as indicated before, since most telcos have spun off their R&D unit, they must make sure they have innovative partners in their ecosystem. An open innovation ecosystem will be key in the future and will evolve and improve over the years[3]. Telcos should part of it as innovation will continue to flourish among communications and Internet firms.

CONCLUSIONS

Since the mid-90's the telecom industry went into a period of sustained disruptive innovations which, combined with deregulation, led to a lot of turbulence and a sometimes difficult redefinition of business models. Some key facts emerged: mobile app downloads grew 16 times from 2009 to 2010; since 2000, the prices for fiber-optic communications have

[1] CDN : Content Delivery Network: i.e. intermediaries between telcos and OTT providers. Akamai, AT&T limelight network.

[2] van den Dam, Rob, Ekow Nelson and Zygmunt Lozinski, "The changing face of communication, Social networking's growing influence on telecom providers", IBM Institute for Business Value, 2010.

[3] About open innovation, John Hagel III and John Seely Brown are saying: "the initial successes, encouraging as they are, represent only the beginning", in Open Innovation's Next Challenge: Itself, Harvard Business Review Blogs, Thursday February 4th, 2010.

dropped considerably; there will be more new Mobile Internet users than online internet users in the world soon; outsourcers providing Cloud computing services are becoming critical.

We provide a strategic framework to help management deal with these issues. Here are the highlights of our chapter.

- The "fully integrated value chain" based on electronic exchanges (FIVC) became a strategy to re-assemble an unbundled value chain to deliver the best products at the most competitive price to customers.
- Pioneer corporations like Dell, Ford, and Jet Blue have incorporated social media networks into their business model.
- Google is now in a position to offer what telecom companies would like to offer as a value added service, namely infomediation service[1], with its developing technology combined with its formidable data bank on web users.
- New technologies have increased competition and have commoditized many services offered by telcos. i.e. VoIP has killed the phone business as we knew it. It is reducing the roaming and voice revenues of wireless operators and making an entrance in the local market with Skype Find.
- With new smart phones and tablets, market power has shifted from telcos to service providers and equipment producers. On the demand side, the customers are in the driver seat as they select devices and services from many competing offerings. The growth in data communications has outstripped voice communication, and wireless services are to become the main technology to access the Internet worldwide.
- Telcos have to migrate from proprietary protocols and customer control to open Internet platforms. They also have to redefine their relationship with partners and let new comers from equipment makers to software or Internet companies be part of their ecosystem, even if some of these partners can be competitors in other markets. This is co-opetition, namely letting competitors be your partners in your ecosystem for specific products or services.
- The OSS[2] strategy: the former incumbents generally still have an important client base, financial resources, and a large asset portfolio enabling them to put forward the marketing strategy consisting of bundling many services for a flat price in order to maximize the number of subscribers. OSS offers include landline phone, HSIA, VoIP, TV, and wireless services for residential, businesses of all sizes, and governmental customers. Large carriers opting for the OSS strategy will have more opportunity than ever to open their ecosystem to innovative partner.
- In contrast, the broker, or syndication approach consists in offering customers a full range of services while not owning all the infrastructures but only some key assets. The syndication approach carries transactional and complexity costs for the leader, but it remains far more flexible than the vertical-integration approach.

---

[1] J. Hagel III and J.F.Rayport were among the first to talk about web infomediation services customized for web users in "The New Infomediaries", McKinsey Quarterly, no 4, 1997.

[2] The acronym OSS for "One Stop Shop" used in this chapter must not be confused with the technical acronym "operations support system" or OSS. We use OSS here as a business strategy for telecom companies.

- The broker business model could offer "standardized services that can be developed inexpensively and delivered wherever their employees conduct business."
- With a network sprawled across 120 countries and with a large number of experts in its world organization, BT combines the broker model and the convergent IT-Telco model in order to be a global supplier of IT and telecom services capable of offering an OSS service for large corporations and other international organizations. Considered a leader and a visionary with a top level ability to execute by the telecom consulting company Gartner, BT offers a network to large corporations: the "21st Century Network," the largest IP network in the world. BT owns a high density and cutting edge IP and MPLS world network as a result. But in 2010, according to Gartner, Verizon and AT&T can now compete with BT and become worldwide players.
- Since most incumbent carriers like AT&T and BCE have divested their R&D arms (Lucent and Nortel in occurrence), they are now followers in terms of technology. The danger is to be reduced to supply dumb pipes at commodity prices.
- Offering space and web apps on different platforms to new media could allow telcos to get some advertising revenues and compete with Internet companies. Advertising revenues is one way to compensate for commodity prices in basic telecom services.
- Niche players can specialize at an industry level, for a certain layer of services, or certain electronic platforms and so on.
- Google can be considered a horizontal integrator since it made inroads in the three horizontal information sectors: Content, Appliances, and Highways[1].
- Skype is now the largest provider of cross-border voice communication services in the world, according to Telegeography. As any VoIP new players, Skype is a threat to incumbents in all countries. Therefore, wire-line and wireless telecommunication operators have to take into consideration the growing and disturbing influence of Skype in their business strategy.
- More than what was achieved by sometimes complex regulations, technology was a major source of competition. VoIP, cable telephony, and high speed wireless Internet access (HSWIA) are innovations that have brought serious competition in the telecom market and have contributed to commoditize the price of bandwidth.
- The FCC came with middle of the road Net neutrality rules maintaining that regulation is needed to prevent network operators from discriminating in favor of their own services. The new Net neutrality rules adopted by the FCC create two classes of service subject to different rules: one that applies to fixed broadband networks and one for wireless networks. The FCC says this is necessary because wireless networks are technologically different from fixed broadband networks.
- Given the FCC's proposed rulings, Telcos and cable operators might delay their investment if doubt remains as to how they will be able to run their network. Before engaging corporate funds into major investments or partnering with equipment producers to finance R&D to support new technological advances, carriers will go to

[1] See Figure 2.

courts to clarify the FCC rulings. Since legal battles can be a long and expensive process, investments and innovations might be delayed and funds allocated to unproductive activities.

- On the other hand, the FCC proposal gives a positive signal to wireless network providers and equipment companies and new comers like Google and Apple. So, as demand for wireless devices and services will continue to grow, network providers will have to invest in new infrastructures and also bolster existing capacity.

In the future important trends, telecommunication operators must take into account:

- The high demand for quality content (text, images, videos).
- That suppliers are becoming more strategic.
- There is a convergence of infocom sub-sectors enabling important growth opportunities.
- The influence of Skype is becoming critical, since the firm is investing massively in the improved quality of its services and new product development.
- The influence of social networks is decisive in their future profitability. Figure 4 illustrated the shifts in communication control and patterns that give rise to new business models.

Thus, telecommunication operators must cope with these trends to fine tune their business models in this period of high turbulence.

APPENDIX

CNetNews.com summarizes the rules as follows:

"The first rule requires both wireless and wireline providers to be transparent in how they manage and operate their networks.

The second Net neutrality rule prohibits the blocking of traffic on the Internet. The rule applies to both fixed wireline broadband network operators as well as to wireless providers. But the stipulations for each type of network are slightly different.

For fixed broadband networks, operators cannot block any lawful content, services, applications, or devices on their network. Wireless providers area also prohibited from blocking Web sites, but the rule is slightly more lenient when it comes to blocking applications and services. The rule only prohibits these companies from blocking access to applications that specifically compete with a carrier's telephony voice or video services. In each case, the blocking rule also allows fixed and wireless broadband providers to reasonably manage their networks.

And finally, the last rule applies only to fixed broadband providers. It prohibits fixed wireline broadband providers from unreasonably discriminating against traffic on their network."[1]

[1] http://news.cnet.com/8301-30686_3-20026283-266.html, December 21, 2010.

In: Advances in Communications,Volume 8
Editor: Anthony V. Stavros ISBN: 978-1-61324-794-5

Chapter 2

MOBILE PHONES, MULTIMEDIA AND COMMUNICABILITY: DESIGN, TECHNOLOGY EVOLUTION, NETWORKS AND USER ISSUES

Francisco V. Cipolla-Ficarra*
HCI Lab – F&F Multimedia Communic@tions Corp
ALAIPO – Asociación Latina de Interacción Persona Ordenador (www.alaipo.com)
AINCI – Asociación Internacional de la Comunicación Interactiva (www.ainci.com)
Via Pascoli, Italy

Abstract

In our research work we focus on the users who have a greater difficulty in understanding the functionality of mobile phones with the triad of communicability, usability and usefulness. It is a detailed study which starts with the models of the 90s and the first decade of the new millennium, and tries to underline the area where the formal and factual sciences interact. To this end an unprecedented evaluation methodology is established where there is a technological and communicative intersection. The participation of real users allows us to put into use a series of heuristic research techniques to face these problems from several points of view of interactive and communicational design. In this work we make a special analysis of the subject of the choice of the random samples inside descriptive statistics, since we are working with it to cut down costs in the methods and evaluation of the quality in the interactive systems, where the mobile multimedia phones are included, for example. Additionally, the obtained results allow not only to know the weak and strong points of the current multimedia mobile phones but we set up a little vademecum where the sociological and technological aspects of the users in the coming years are summed up.

* E-mail address: ficarra@ainci.com

INTRODUCTION

Communication is an essential element in the human being inside a community since one tends to go from an individual communication to a social communication. The social communication media have tried to foster this phenomenon of socialization of people through the press, the radio, the television, the cinema, etc. in the last decades of the twentieth century. In this socialization process the phone has also been present, first landline and currently mobile and multimedia [1]-[4]. However, the new technologies show a constant cycle of elements which act as centripetal or centrifugal forces according to the technological novelties which are included in the personal communication devices [5]-[11]. Many of these components are supposedly aimed at solving communication problems and quicken socialization among their potential users. However, in many cases these problems do not only persist, but they grow worse. It is an exponential worsening deriving from the passing of time and the potential users of the new technologies [12]-[14]. Currently we are witnessing a geometrical increase of the difficulties detected with the triad of mobile phones, multimedia and communicability deriving from factors such as ergonomic design, technology, the networks and the users [15]-[18]. In other words, problems regarding communicability, usability and usefulness. In both cases the triadic components keep a bi-directional relationship among themselves. The magnitude of these problems is such that in some cases it can lead to non-communication among people, even with all the technological aspects working at a 100%.

In the mobile multimedia computing related to phones it is made clear how the user does not know the full 100% potential of the hardware and the software of the device in his hands [19]. This situation was common in the 90s in the context of the 2D and/or 3D designers. They used and knew part of the options of the commercial software and hardware aimed at entertainment and leisure time, for instance in the creations of computer animations for cinema and videogames [20]. The speed of novelties in hardware, new versions of software and the work to be developed prevented them from finding out the details of the tools with which they interacted daily. The same has happened with the evolution of mobile phones in Southern Europe. At first it meant the liberty to communicate from any place in the home or the office [21], [22], whether it was with spoken word or text messages: SMS (Short Message Service). However, little by little mobile phones have become multimedia instruments with which many senior users can't easily communicate, remaining totally sidelined upon the arrival of the new phone models. Those who know some of their basic functions have again the same problems after a short time because they are forced to change the phone because the batteries over time loose capacity, and they are no longer made. Besides, in Europe there are no training centres to teach how to use multimedia phones for users with difficulties.

Consequently the apogee of the mobile phone which has led to many countries in Southern Europe to have the highest fees in landline communication with the alleged purpose of boosting the phone communication network, has left part of their population cut away from these breakthroughs, and sidelined technologically speaking. For instance, Telefónica from Spain had the most expensive fee between Lisbon and Moscow in the first half of the 90s. Obviously the cost factor and low purchasing power has also prevented some users of mobile phones from updating and getting acquainted more closely with new models to not stay in the technological rearguard. To this contextual factor we may add other realities from France or

Italy where many people usually have more than one mobile phone in order to split work, family and leisure time, for instance. Some of them use the same model of mobile phone, others, in contrast, prefer different models in order to always be acquainted with the latest technological novelties which fashion thrusts upon them.

Using more or less mobile phones simultaneously is not synonymous with greater social communication as it happened with the fruition of the traditional media: movie houses, family gatherings to listen to or watch a soap opera on radio or television in the last century. The multimedia mobile phone user tends to participate unwittingly in the autism phenomenon in the current era of global communications. Oddly enough, the advantages of freedom and socialization of the individual entailed by the communication media disappear, because of the complexity of the access to the contents and the degree of complexities or functions of it among given potential users [23]-[25].

The current work is structured in the following way: a constant diachronic analysis of the new technologies focused on computer science, telecommunications and the communicability of users with mobile phone systems and other interactive communication systems, in the daily life of their activities, description of the techniques, methods and strategies used deriving from the intersection of the formal and factual sciences, with special emphasis on the correct realization of the samples of study, bearing in mind the details of descriptive statistics: study of the main software and hardware characteristics of the mobile phones, in the past, present and with a particular look at the future; presentation of the obtained results and future lines of research.

TELEPHONE EVOLUTION, MASS MEDIA, COMPUTER SCIENCE AND GLOBAL VILLAGE

Since the dawn of the telegraph, the radio and other technological breakthroughs of telecommunications throughout the 20th century, there has always been an attempt to unite the territory of a nation and its inhabitants. That is, it was an implicit and/or explicit goal of a long series of discoveries and inventions in telecommunications until reaching the current global village, where the mobile phones and the Internet have an overruling role in the social evolution and the economic development of societies [26]. First, it was the audio in the radio, then the radio was joined to the image with television, until the appearance of interactive communication with the Internet in the mid 90s and then the multimedia mobile phones. In an early stage of social communication it was usual to go from one situation of predomination of the emitter over the receptors, since there was no possibility of interaction with the media, as it was the case with analogical radio and television.

Later on, the democratization of the use of the PC connected to the Internet broke with that verticality in the communication process towards a horizontality and the promotion of the free access of contents for everyone in the early 90s, first with the computers and now with mobile phones. In the new millennium, one can see situations of a return to a position of dominance by some members in the social groups through persuasion. It is the case of the dynamic persuader and interactive persuaded person [27] in Web 2.0, for instance. An individual who is capable of breaking up the democratic horizontality of the web for his/her own benefit or for the benefit of an elite to which he/she belongs. This is achieved through the constant manipulation and persuasion made among the members of his social networks. That is, he generates small spaces of vertical communication and with unethical and dictatorial behaviours. Mc Luhan's global village based on the democratization of information is being attacked by the behaviour of the dynamic persuader, who moves freely inside the social networks [27]. This phenomenon is much more serious than that foreseen by those who claim that the Internet was invented for purely military reasons.

Many of these behaviours are related by some to the origins of the research projects in the new technologies, that is, the military context, such as have been the first text messages sent via the net [28]. However, in the ARPA-Net project, which was born with the purpose of joining the university researchers it has nothing to do with the use of the net for purposes of national security which was later applied to it. In the context of the mobile phones there are those who see the radios used in the Second World War (some models had a phone pipe) as the origin of its diffusion. However, it is necessary to remember that the radio and the phone are two very different entities. The phone is a telecommunication device designed to transmit acoustic signals through a long distance. It is very similar to the teletrophone. From the technical and historical point of view, the teletrophone is the name that was given by the Italian inventor Antonio Meucci to his version of the phone in 1870, who had built the first electromagnetic phone in 1856 [29]. Alexander Graham Bell was the first in patenting it, in 1876, in contrast to Meucci, who didn't have the financial resources to do it in the U.S. Patents Office and who could only hand in a short description but couldn't formalize the patent because of his lack of financial resources [30]. This dispute between both inventors of the phone was solved in 2002 when the U.S. congress passed resolution 269, which recognized that the inventor of the phone had been Meucci and not Bell.

In our case, the mobile phones, also called cell phones, are basically made up by two main parts: a communication network (or mobile phone network), and the terminals (mobile phones) which allow access to said network. Here it is interesting to see how the word "cell" or "mobile" divides from the linguistic point of view several countries in the Mediterranean basin, and the countries in the American continent. Just to mention a few examples, they are called mobile in Spain, Portugal, etc. and cellular in Italy, Argentina, Chile and Uruguay. Now the mobile phone is a wireless electronic device which allows to have access to the net of cellular or mobile phones. It is known as cellular due to the masts that make up the net, each one of which is a cell, although there are satellite mobile phone nets. Their main feature is their portability, which allows one to communicate from almost any place. Although its main function is voice communication, like the conventional phone, its fast development has incorporated other functions, such as a photo camera, notebook, access to the Internet, video reproduction and even GPS (Global Positioning System) and MP3 player. Obviously, another of the factors to be considered at the moment of speaking of the evolution of the mobile phones is the weight of the first non-commercial devices which surpassed the kilos as compared to the current ones which have the weight of just a few grams, like those which have the shape of a wrist-watch.

Multimedia communication has been boosted through the use of mobile phones [31]. Obviously, this has been a real revolution in the interactive communication of the users. The first forerunner of the mobile phone was from the Motorola company, with its Dyna TAC 8000X model. The model was designed by Rudy Krolopp of the Motorola company in 1983. The model weighed little less than a kilo, and had a market value of almost 4,000 US dolars. Krolopp would later join the research and development Motorola team lead by Martin

Cooper. Both Cooper and Krolopp appeared as owners of the original patent. From the DynaTAC 8000X onwards, Motorola developed new models such as the Motorola Micro TAC, launched in 1989, and the Motorola StarTAC, launched in 1996 to the market.

The design of the first mobiles mainly emulated the dimensions or functions of the traditional or landline phones (Figure 1). However, little by little, as has also happened in the technical evolution of computing, a diminution of the size of the telecommunication device has taken place. Obviously, this design factor has also joined the usability and communicability factor of the first mobile phones. In these the functions were limited to the utmost, there were no color screens, nor multimedia functions, for instance. In some way there was a bid for the bigger public to get acquainted with the possibility of being able to call by phone without using a public booth when one was outside the home or the work office. Until this point such devices existed in the collective imagination of a society through the American television series. Where the mobile phone was joined to a high priced car of in a spy's shoes, like Maxwell Smart in the 1965 'Get Smart' TV series [32]. That was already a reality in the hands of millions of users. Later on, the costs of the phones started to decrease, the same as happened with the professional computers, which went on to become personal computers until in less than a decade in some economically developed countries almost every home has it available as if it were a Digital Terrestrial Television (DTT or DTTV) or a radio.

Figure 1. First commercial mobile phones in Spain and Italia –Motorola, Alcatel and Sony Ericsson (1991, 1998 and 2005).

Both in mobile phones as in the personal computers there is a tendency to diminish the size of the hardware (CPU, peripherals, etc.), and of the programs for their functioning. This tendency leads to the currently taking place parallelism between the personal computers and the multimedia phones. Evidently, if we consider the breakthroughs in the new technologies as a pyramid we have in its upper part the R+D area and the sectors of the population with a greater purchasing power. Whereas in the lower parts of the pyramid are the users who push for free access to information and the last technological advances with the least possible cost. Currently it is easy to detect how in the different communities the university and/or industrial labs are orienting their works only at the summit of that pyramid., especially in the great cities of the European Mediterranean basin, such as can be Rome or Barcelona. It is easy to find in those places because of fashion motives the use of mobile phones in a watch. Some models are GSM (Groupe Special Mobile) Triband, with double SIM (Subscriber Identity Module), photo camera and video (Figure 2).

Figure 2. Mobile phones –Motorola, only SMS (1991) and Samsung SGH-L170–Multimedia phone (2008).

Analyzing some of its main characteristics, we can see how the microphones and the microcomputers converge on the wrist of the arm of a user under the function of a watch: a color screen 1.3”, TFT (Thin Film Transistor) 360K, touch screen 128*160, 1,3 MPixel digital camera, MP3/MP4/3GP play, storing of JPG/GIF photos, Dual Stereo Bluetooth, spot for T-Flash card with memory until 2 GB, GSM, /GPRS/ WAP, videogames, alarm function, email POP3/SMTP, battery duration for conversations between 3 or 4 hours, 64 ring tones, etc. All of this with a 61x45x16 mm and a weight of 75 grams. To this device we can connect a Bluetooth battery load or a network adapter, for instance. These are the peripherals which are pacing the rhythm between the purchase of computers or multimedia mobile phones, since both started decades ago the miniaturization process of the CPU and the peripherals. From the point of view of communicability and usability we find two main elements which may lead the adult users to purchase mobile phones instead of small computers (Pocket PC, Tablet PC, etc.), the VLK (Virtual Laser Keyboard) and the new generation of flexible screens (Figure 3).

- The VLK allows to use a smartphone or PDA as a normal computer, with an enlarged keyboard, even in the dark. This device projects a QWERTY keyboard on a flat surface and allows to digit the keys as in a normal computer. Its main characteristics, which may vary among the different models, can be summed up in the following way: laser source, over 60 keys keyboard in full format QWERTY. Size of the virtual keyboard 295x 95 mm. Possibility of recognizing up to 400 characters per minute via Bluetooth. It usually works with a lithium-icon battery. The system is compatible with PalmOS 5, Pocket PC 2003, Windows Smartphone, Sybian OS, Windows 2000/XP.

- The new generation of flexible screens are made from an organic semiconductor material with transparent electrodes. Which allows them to be bent and to be very thin (almost of the thick of a hair). The same as with the early CDs, it is a technology patented by Sony, who has not announced yet to what products it will be applied. The Japanese brand Kyocera is also developing EOS, a flexible mobile phone which looks like a wallet. Nokia has also developed a prototype, the 888 model, a spectacular mobile phone with the shape of a wristband. The lightness of these screens will make it possible to place them on almost any object, which makes the only limit to be creativity or imagination, especially in the multimedia sector: newspapers with flexible OLED (Organic Light-Emitting Diode) sheets, soda packages wrapped in this technology that shows ads, or television sets which will be no thicker than a poster hanging on the wall.

Figure 3. Mobile phones and micro information and communication technologies: iPhone (1), VLK (2), BlackBerry (3), Watch & Chronophone Mobile (4), and OLED sheet (5)

Once again the peripherals set trends among the different kinds of users and the different reasons for which one interacts with the interactive systems. Evidently, the goal is that the multimedia phones go down to the basis of the pyramid as soon as possible to boost the expansion era we are going through. Now in regard to the communicability it is necessary to differentiate the aspect of the design of use of the interactive systems which belong to the classical environment of computers usability and in this case it would be microcomputing applied to the telecommunications systems. In our work methodology, which has been developed for years in several sectors such as education, tourism, e-commerce, etc., one of the design categories of the interactive systems is connectability. This notion allows to differentiate what is traditionally associated to the hardware and software such as is the practical acceptability stated by Nielsen from system acceptability [33]. That is, he contends that the acceptability of a system is made of the social acceptability (without developing its content), and the practical acceptability, in which we have cost, compatibility, reliability, usefulness (in which usability is inserted and its five classical principles efficient to use, ease to learn, few errors, easy to remember and subjectively pleasing [33]. In contrast, compatibility is another component to be considered at the moment of designing an interactive system, especially in the current competition between computers and mobile phones. By correctly dividing the design in the interactive systems in several categories it is feasible to reduce their production costs, since it is possible to quickly detect where are the failings in the realization stage of the prototypes of the mobile phones, for instance.

COST FACTOR, TELECOMMUNICATIONS DIFFUSION AND DESIGN ASSESSMENT

The cost factor in the purchase of the new technologies by the basis of the pyramid: the population is essential for the expansion of communicability, where the dynamic and static means of interactive communication play an essential role. In the case of phones, we can see how the devices have been adapted first for the text messages and then little by little for the multimedia. At the start the mobile phone was used for conversations and avoiding the possibility of sending SMS messages [34]. Today, some mistakenly claim that the emails and the SMS tend to disappear as a result of the apogee of the web. However, the text in the communications is the basis of the hypertextual communications and will remain because when confronted with doubt in the veracity of the dynamic and static images the human being tends to believe in the written word.

There were two reasons: the costs of the phone reloads and the lack of experience of writing on such a small device for millions of young and adult people. In regard to the cost of the service it was very high in the 90s in the countries of the Mediterranean basin. This was due to the monopoly of the state enterprises in the telecommunication field which forced the users at the moment of the mobile phones reloads to pay additional sums for the simple fact of making a reload ("ricarica"), as it can be seen in the following image of the Italian mobile phone cards (Figure 4). That arbitrary plus of the phone companies was in use until 2008, where the private companies appeared in Italy. In the Spanish case, the high costs were due to the VAT percentage applied to the reloads ("recarga"). In others words, since the times of the peseta in Spain additional services were never paid for reloading the mobile phone (Telefonica –MoviStar). In contrast, in Italy, at the time of the lira, the directors of the public and private companies made people pay for this service "ricarica", even with the arrival of the euro (Telecom –TIM).

Even these enterprises in the prevailing position of the mobile phones achieved financial records such as those reached by the Telefonica phone company. They got an extra 40% profit in the 2004-2005 fiscal year, (over 5 billion euros). However, with the passing of time and the democratization in the spread of these devices among the populations of the economically developed EU countries, these abusive costs of the mobile phones were accepted by almost everyone as a way of paying for the additional advantages that were inherent to the possibility of freely communicating from any place in the planet. That is, a kind of apocalyptic-biblical principle applied to technological innovation: before enjoying the technological breakthroughs it is necessary to suffer.

Now the cost problems and the way of managing the origin of the diffusion of mobile phones expanded through several countries in the South of the American continent. However, many of these communities are used as labs of the new technologies to be later on applied to the developed countries. Even in trivial issues such as the marketing of mobile phones. In the first years of the new millennium more mobile phones were sold than could actually work because of the infrastructure that existed in those places. A structure of telecommunications that were developed and managed by the very same firms of mobile phones, with their headquarters in Spain, France and Italy. The real users of mobile phones were submitted to real tolerance tests concerning the malfunction of mobile phone services and which later on were applied to the diffusion and marketing of the Internet. All these studies of public reaction or of users have served to automate mobile phone services, landline, Internet, etc., in Europe.

Figure 4. Evolution of the prices in the phone cards for mobiles between Italy (TIM) and Spain (MoviStar), including the swap of national currency (peseta and lira) towards the euro.

For instance, with the programming of the virtual assistants to solve the technical problems of Internet connection using mobile phones. In some regions known as traditionally as the engines of Europe such as Baden-Württemberg, Catalonia, Rhône-Alps and Lombardy it is possible to see how mobile phones are used to cheat the final user and increase the costs in the internet services. For instance, in Lombardy it may happen that the subscribers to the phone lines surpass the available lines for the high speed internet services, consequently there are subscribers who do not have the service continuously available, as it happens in many developing countries. That is, cyclically some subscribers are continuously disconnected from the internet. To solve this problem, the subscriber activates a second or third line of high speed internet of high speed to have the internet services continuously available. Automatically he/she will keep on receiving fake SMS on his/her mobile phone which say that his/her service has been activated (Figure 5), when in fact, 10, 20, 30 days may pass without Internet service, and having to pay for a non-existing service. Additionally, the virtual assistant will communicate that the failing is due to problems in the central phone service. Consequently, there are cities inside the so-called engines of European economy where the user is deceived with SMS and has to multiply by two or three the costs of the phone and internet services. This reality slows down the diffusion of mobile phones in a wide sector of the European population, for instance

Figure 5. Italian fake SMS about Internet service (ADSL) on mobile phone.

Aside from these negative realities for the progress of the new technologies in Europe, most users started to invest in the new technologies inside telecommunications in the 90s in a cautious way because of the novelty, plumping for the basic functions, that is, those that were cloned in the new devices and which offered that additional freedom of movement for city telecommunications, inter-city and international (late 90s). Then the public started to use the phones for sending SMS during almost the whole first decade of the new millennium, until the multimedia mobile phones allowed them to participate in the social networks from his/her mobile [35]. This is an interesting area to be analyzed in the future, whether it is from the sociological and/or social psychology point of view, seeing the negative impact among the Human-Computer Interaction, the communicability and the users with the traditional and multimedia mobile phones. In the diffusion of mobile phones not everything is focused on communicability and usability of the new devices, but in the human and economical factors which are included in the new technologies. Consequently, it is important to approach these problems from the convergence of the formal and factual sciences, for instance.

HEURISTIC EVALUATION: TOWARDS A RELIABILITY OF THE RESULTS

At the beginning of the mobile phone era, the quality demanded by the users focused on the usability of the device and the biggest possible coverage of the mobile in the phone network. With the passing of time and the incorporation of the dynamic and static means in the mobile phones, usability was replaced by communicability for the user himself/herself. Later on, communicability requisites have been increasing as the multimedia mobile phones could be wired to the net and use the last novelties deriving from the Web 2.0. Along the time a series of heuristic evaluation techniques have been developed. However, there wasn't in the first years of the mobile phones an evaluation methodology of interactive design. With the incorporation of the dynamic and static means, the mobile phone became a multimedia device entailing an endless variety of design of the interactive systems of the computers. In this regard, a set of techniques and methods of heuristic assessment for the multimedia system, such as: MEHEM (MEthodology for Heuristic Evaluation in Multimedia) and MECEM (MEtrics for the Communications Evaluation in Multimedia) [36]-[38]. This set was based on a set of quality attributes deriving from HCI, software engineering, interfaces, primitives of design of hypertextual systems, multimedia and hypermedia, and the semiotics or semiology. These quality attributes have been increasing along the years in regard to the different applications and/or heuristic assessments which were carried out for education, tourism, e-commerce, etc. [39]. Obviously, these are quality attributes that can be broken down into metrics to assess the quality of the interactive systems. An assessment of the whole system or part of it, with the results always 100% reliable, carried out by evaluation experts, and verified in some cases in usability laboratories. The goal was to establish quality attributes validated by other heuristic evaluation techniques, such as videotaped sessions, guided interaction, beta-testing, user feedback, user surveys, focus groups, etc.

After the establishment of the set of quality attributes for the design of the multimedia/hypermedia systems there is a set of them which are related with a greater interdependence intensity or "subordination" among themselves. This interdependence makes it possible to elaborate a sketch of those attributes which have a higher degree or level of interdependence. For instance, the factual function is inside or derives from accessibility, or the motivation depends on the wealth or what is now called "augmented reality". A grouping of such attributes is to be found in the Figure 6.

The use of heuristics for the generation is due to the fact that they respond to the following factors enunciated by Nielsen [33]: Speed, ease of implementation, and low cost. Although as a denomination some of them exist inside software literature, the here presented attributes are aimed at the multimedia systems in mobile phones. The heuristic attributes which have been presented are not orthogonal among themselves, but they are based on a communicability model and in a procedure which has been in a continuous definition, correction and verification. The use of a terminology stemming from the design models in the on-line and off-line interaction systems, semiotics or semiological, the primitive actions of interactive communication and in a special way the evolution of the hypertext, multimedia and hypermedia, allows us to hone even more the definitions although sometimes we resort to a broad terminology. For instance, the notion of the richness in which it is necessary to set down the limits of its scope.

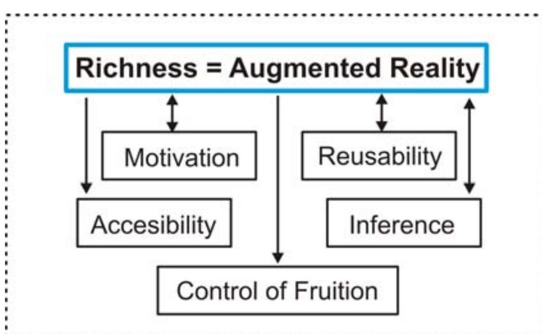


Figure 6. Richness and the quality attributes relations.

The establishment of the accurate limit of the significations is not easy, especially when one works with techniques which stem from the social sciences, such as is the case of heuristic assessment. In regard to the richness attribute, which was presented for the first time in 1999 [40], together with other quality attributes, it has been derived to the current notion of augmented reality [41], [42] especially in the interfaces at the moment of navigation, when the user can decide to incorporate or eliminate resources, or change the modalities of visualization of the information, for instance, in Google maps with its traditional options to locate a street or a city on the different maps, that is, Earth, satellite and street guide. However, there is no reference to the criteria of quality of richness or the accessibility (Figure 5), another criteria of quality presented before finishing the past century [40], [43]. In the context of interactive design pioneers are often unfairly forgotten by resorting to the use of synonyms to define the concepts. The reason in these cases is the personal competence by certain clans inside the university sector in the Southern Mediterranean and for economical purposes, since the authors can present their works in publications of lesser international circulation, such as may have other publications of international associations, related to computing or electronics, for instance.

However, with such attributes one intends to set down the basis of a previously unseen methodology in the evaluation of the interface of a multimedia/hypermedia system. Their importance is even greater because there is still no “guidebook” to get 100% success on designing an application. The solutions that are suggested still require high investment, such as the creation of labs and counting on plenty of human resources [33], [39], [44], [45]. Once again the cost factor demands thinking of solutions which take into account the high quality equation of the results plus the least possible time and with the least cost. Under these conditions, it is necessary to find the best possible solutions derived from the social sciences and descriptive statistics, such as can be the straight observation techniques and the use of samples with the lowest possible error in the making of the universe of study.

HEURISTIC EVALUATION: ACHILLES HEEL

The research in the field of the factual sciences and where many of the heuristic sciences of evaluation of usability engineering confirm a kind of subset of the former. Social research, as Mario Bunge [46] and Ezequiel Ander-Egg [47] say, is a reflexive, systematic and critical procedure. This entails the need of organizing beforehand the whole research process and planning its execution in several stages. The organization and programming of research consist in projecting work in keeping with a logical structure of decisions and with a strategy that gives an aim to the way of obtaining the adequate data to the planned objectives. Neither the planning nor the stages to be followed by the evaluators are not described in those works related to the education and the heuristic evaluation of the interaction of the young users of new technologies in Catalonia or in Lombardy. Perhaps the word "planning" has implications that are not liked by the mentors of these works. In contrast to this modus operandi, a correct planning is indispensable to cheapen the research costs. In the current work the notion of organization includes the scientific and technical aspects of the different stages and phases which must be implemented in order to carry out a research and which in our case can be summed up in the following way:

1. Determination of the main and secondary goals to be reached.
2. Exploratory phase. Consult and documental compilation.
3. Structuring of the research design: elaboration of a theoretical framework, shaping of a team of evaluators and users, task coordination, evaluation techniques and methods, organization of the consult and research material, determination and choice of the sample, management of the financial resources to cover the expenses of the lab (use of computers, payment to the users, etc.)
4. Preparation of the work group.
5. Post-evaluation work: classification of the obtained results, analysis and interpretation of the data and/or information, writing of the final reports.

In the organization and planning of the research work in the social context one can start with the classical Platonic rhetoric questions and then add others. That is, the questions that we ask ourselves at the beginning of the work and the goals we must set. In this regard the following parallelism can be established: what (issue or subject), what for (goal), why (problem), where (area of the factual and/or formal sciences), when (activities chronogram) how much (intensity degree of the study, for instance, the whole universe or a part of it, that is, a sample), how (methods and techniques), who (human team), with what (technological resources and implicitly economical). In few words, it is about making operational the scientific method applied to a given field of research.

Currently there is no guide within the usability and usefulness concept which can serve us as a kind of compass in the face of the problems to be solved and all the decisions that must be made in order to carry out such research. In view of such a situation we will also try to include communicability, usability and usefulness of the mobile phones in this kind of guide where are indicated the stages or logical steps of the research process. It is necessary to make clear that these stages or steps which are taken in the context of the social sciences and in our case, by working with real users, some of the here indicated stages may overlap with other analogous situations where communicability is not considered. Now it is important to point out that the organization of the research work in the heuristic research environment including its methods and techniques in the formal and factual sciences consists in making available all the necessary operations for its realization. Carrying out a research task without an adequate preparation as may be the case of a usability expert but who is inexpert in communicability may require a longer time than what is actually needed. In well-prepared research, there is neither hurry nor unnecessary waste of time in the preliminary tasks, which in some cases entail high costs in relation to the profits or obtained results. In these preliminary tasks we can insert those communicability elements that the team of experts, jointly with the users have to locate regarding the interaction of the mobile phones system and which need an explanation or previous example.

Formulation of the aspects to be researched. In this context it is important to determine what is the goal of the research, the purpose and the delimitation of the field to be researched. Everything has its beginning in a good formulation of the problem that delimits the research and is its target. Once the problem has been formulated it is necessary to subdivide into several subsets or sub problems, as many as possible. This operation includes the task of explaining the aspects, factors or relevant communicability components, related to one or several categories of the design that is the aim of the research, for instance. Once the main goal or problem of the aspects to be researched has been formulated in an accurate and operational way, the evaluator will start to solve each one of these sub problems until encompassing them in their totality. It is important to point out that each one of these subproblems inside the scientific methodology are denominated dimensions and variables of the fact to be researched. For instance, when the problem of evaluating the design of the interactive systems was posed, it was divided into several subcategories to facilitate its resolution in the creation of methodologies, techniques, instruments, etc., and bearing in mind several factors related to the potential users and their cultural aspects, among others. Once again the rhetoric questions may help the evaluator as it is our case. Starting from the fact that the ability to formulate problems is something that is related to the scientific talent, some rhetoric questions that may help develop this talent are:

- What is the problem? What are the data of the problem? What are the main aspects or elements of the problem? (identification of the problem in a clear and accurate way, thus decomposing it into its components, establishing categories and variables or dimensions).
- Is it an unprecedented problem or not? (to research the existing bibliography in the issue that will be approached)
- Which are the relationships among the different components of the problem? Which are the issues related to the problem? (translation of the approaches into variables that can be measured).
- Has lexical ambiguity been eliminated? (a clear and accurate conceptual framework must be generated so that then the different agents that will intervene in the later stages do not generate confusion and waste time. Wasted time is equal to economical resources).
- Has the finality been defined? (each one of the operations that will be made in the evaluation must lead to solving the main problem).

Consequently, a correct formulation of the field to be analyzed or examined must respond in a clear and accurate way to the what and why of the evaluation. Additionally, the evaluation field must be limited, for instance to a whole universe of study or a sample. In this latter case, the mechanisms must be established through which the elements are randomly chosen that will be researched to guarantee the greatest possible objectivity in the results which are later obtained. In our case, we have had to define the groups of adult users and the models of mobile phones which were used in the experiments made. That is, several types of users divided according to age, studies, previous computer knowledge, etc. and several models of phones classified by decades, functionalities of said phones, presence or absence of the dynamic and static means, etc. (our users were to be found in the cities of Barcelona – Spain, and Bergamo – Italy, and the remarks made refer to the models of mobile phones which go from 1998 until 2010). Once this phase has been left behind, we have to analyze the existing references in relation to our problem, bearing in mind taking into account the originality of the research and problems resolution process. In this exploratory stage, the review of the existing literature is of great importance. In the social sciences, some authors [47], [48] contend that the successive stage is called research design or tasks to be carried out, that is, a kind of outline in which is indicated the set of activities to be carried out to steer the course of a heuristic evaluation of an interface in a multimedia interactive system, for instance. Therefore, we are going full steam into a set of tasks which can be grouped in the following way: elaboration of a theoretical framework, confirmation of a work team (especially in the usability labs). In our case, the communicability expert does not require that work team, however, it can be made up to verify the obtained results through some heuristic evaluation techniques where users intervene, for instance, interviews, video sections where the interaction with the mobile phones is recorded, etc. Continuing with the eventual need of working with a researchers and/or users team, it is necessary to coordinate the tasks of the researchers or project advisers, the surveys and stats team (if there were), computing and telecommunication experts (technical issues in the interaction process with the mobile phone models), auxiliaries for the data compilation and writing of partial reports, etc. Then there are the selection stages of the methodological instruments, for instance evaluation tables, guides, discs, etc. the organization of the consultation material for the members of the evaluating process, choice of the field of study (total universe or sample).

SAMPLE FOR HEURISTIC EVALUATION

The meaning of the sample designates the part or representative element of a set. If we use the term in an exact sense, as it is used in statistics, the sample is part of a population which is the target of a research. For instance, the users in our case stemmed from professional training courses and they were taking computing courses for resinsertion in the labor market. Others, in contrast, attended language courses and were handpicked following each one of the steps which will be analyzed later on, in order to obtain a random sample in the strict sense of the notion. This sample is chosen by certain procedures, and its study leads to conclusions which are extensive to the whole of the members of the set with a significant economy of cost and with a greater speed of execution. Now, when one chooses to use samples instead of the whole of the universe of study, said samples have to be planned and carried out in a careful way since the value of the conclusions depends on the representativeness of the sample. From the conceptual point of view, in the sampling process are used a series of basic concepts which are next described to eradicate ambiguity [47]:

- The sampling technique is the set of operations which are made to choose a sample.
- The unity of the sample is made up by one or several elements of the integrants in which is subdivided the basis of the sample and which inside it are defined in a very accurate way.
- The basis of the sample is the set of individualized units that make up a universe of study.
- The universe of study is made up by the whole of a set of elements that are being researched and from which a fraction or sample will be studied which is expected to gather the same features and in the same proportion.

Some authors distinguish between universe and population [48], [49], that is, with the first term a set of elements is denominated, beings or objects, whereas with the second a set of obtained numbers by measuring and counting certain features of said objects. In the current work and with regard to communicability we prefer to use the first notion, although we regard both as equivalent or similar [39]. Although it is true that we are examining real users at the moment of the interaction with the mobile phones, that is, population belonging to one or more communities, the notion of universe of study is more correct in the context of communicability.

Now the sampling method is based on a set of laws that provide it with scientific foundations, that is, the law of large numbers, and calculation of probabilities, excepting the case of empirical or non random samples. The law of large numbers, formulated by Jacques Bernouilli, is expressed in the following way [49]: If in a test the probability of a happening or event is ‘p’, and if this is repeated a great number of times, the relationship between the times in which the event takes place and the total amount of times, that is, the ‘f’ frequency of the event, tends to get closer every time more to the probability ‘p’. More exactly, if the number of tests is sufficiently big, it is totally unlikely that the difference between ‘f’ and ‘p’ surpasses any predetermined value however small it can be. The probability or a fact or event is the relationship between the number of favourable cases ‘p’ to this fact with the amount of possible cases, assuming that all the cases are equally possible. The way of establishing the probability is what is called calculation of probabilities. However, it is from these two main laws of statistics that are inferred those which are the cornerstone of the sampling method, such as:

- The law of statistic regularity, according to which a set of units, randomly taken from a N set are almost certain to possess the features of the bigger group.
- The inertia law of the large numbers, which is a corollary of the former, and it refers to the fact that in most phenomena, when a part varies in one direction, it is likely that an equal part of the same group varies in the opposite direction.
- The law of permanence of the small numbers, that is, if a sufficiently large sample is representative of the universe of study, a second sample of the same size must be similar to the first. If in the first sample few individuals with odd features are found, the same proportion is to be expected in the second sample.

Consequently, the qualities of a good sample for it to have technical-statistic validity can be summed-up in the following way:

- Being representative or a general reflection of the studied set or universe, reproducing the most exactly the features of said universe.
- That its size is statistically proportionate to the magnitude of the universe.
- That the sampling error is kept inside the limits regarded as admissible.

In the current work with users and mobile phones we have established a reduced universe of study, due to the fact that our methodology does not require analysis labs, because essentially we rely on the presence of a new professional known as communicability expert. Consequently, the study made with real users has been useful to verify some of the results obtained beforehand and in an autonomous way.

The task of determining a sample is included inside the different stages of research, consequently the sampling task (design and collection of the data that make up the sample of previous and latter tasks of what is known stages for the selection of the sample. These stages can be summed up in the following way:

- To keep always in mind the main and secondary goals of the research.
- To have the most possible available information about the set from which the sample will be taken.

- Allowed error of estimation.
- To count the available human resources and instruments.
- To examine the methods and techniques to be used in the research.
- To compile and analyze the data.
- To measure the representativeness of the sample.

Now the design of the sample and the sampling plan is an operation that demands special attention and training. Therefore, in many occasions the researchers need to count on specialists in sampling. The latter and former tasks are part of the general stages of the research but they must also be known by the statistician responsible for the selection of the sample. For instance, currently there are many works in some provinces of Northern Italy, such as can be Bergamo, where alleged works of educational research of the universe of secondary studies of the province sound the false alarm of the damage that the Web 2.0 does among teenagers. However, these studies lack scientific value because they do not follow the technical and theoretical patterns of the samples, like the stages for their selection, for instance. The falseness of the result is due to the fact that the sample does not represent the principles and main laws of descriptive statistics. Inside this context of stages in the selection of a sample, it is important to differentiate the types of it, the selection procedures and the representativeness. If one considers the structure and the selection procedures, one can distinguish two types of samples and inside each one of them there are several kinds [50]:

1. Random sampling, also known as probabilistic or randomly, is based on certain laws, and it is rigorously scientific. In our work method we are always based on this kind of sampling.
2. Non-random sampling, also known as empirical or erratic, which in contrast to the former does not have a statistic-mathematic basis, and may take two forms: intentional sample or circumstantial or without rule.

The random samples are obtained through procedures based on the law of large numbers and the calculation of probabilities, thus eliminating possible arbitrariness with a random determination. It takes for granted that the universe of study can be subdivided into units known as sample units. These sampling units may be natural, for instance, the users of multimedia mobile phones of a city; natural numbers, for instance; university students; or artificial units, such as can be the surface in meters or square kilometers that a university occupies. By claiming that a sample is randomly determined, one is stating that any of the units or elements that make up the set has the same possibilities of being included in the sample. It appears obvious that this procedure has higher chances of validity the more homogeneous is the set. Goode and Hatt [51] contend that a random sample is that which is obtained in such a way that from all the pertinent points of view the researcher does not have any reason to believe that it may cause any propensity or tendency. That is, the units of the universe have to be arraigned in such a way that the selection process yields an equal probability of selection to each and every one of the units that make up the universe of study. Although there are several procedures for the random sampling which give rise to different kinds of samples, in the current work we will focus on those which we frequently use in the set of techniques and/or methods of the communicability and the quality of the design of the interactive systems, including mobile phones.

- **Simple random sampling**: It is the foundation of any probabilistic sampling and consists in that all the elements have the same probability of being directly chosen as part of the sample. Two cases may turn up:
  1. The sampling without replacement, used if the population is finite, in which all the samples of ‘n’ elements are equally probable.
  2. The sampling with replacement, in which every selected element returns to the set or universe. These are then finite populations or universes.
- **Stratified sampling:** With the purpose of improving the representativeness of the sample: when certain characteristics of the set or universe of study are known one groups them in strata or categories, the sampling units which are homogeneous among themselves. Inside each one of the strata one proceeds to a random selection, that is, random simple. The distribution of the sample in every strata (technically denominated affixation) can be carried out through a random selection and in such a way that:
  1. Each strata has a sample of the same size (uniform affixation).
  2. Every strata is proportional to the number of elements of every strata (proportional affixation).
  3. The sample of every strata is proportional to the number of elements and the standard deviation. In that case the optimal size of the sample for every strata has to be determined.

Finally, there are the non-random or empirical samples which are not based on a mathematic-statistical theory but which depend on the researcher’s judgment. Very used in the regions of Northern Italy in the university educational context and the new technologies, especially when one tries to study wide territories of population, such as can be a whole province. In relation to the random samples, this method possesses advantages in regard to the cost-time equation, but it is more difficult to control the validity of the results. There are two modalities of non-random samples: the intentional sampling and the erratic, circumstantial and without rule sampling. The first one offers certain minimal guarantees in regard to the results because there is a person who chooses the sample intending it to be representative, doing it in agreement with his intention or opinion (it is the classical modus operandi of a star enunciator devoted to the research in the university context, for instance). In contrast, the second is the common denominator detected in the centers of university education, especially in the area of the Mediterranean, where numerous works of the usability evaluation are carried out by using the cases that are at hand or which have to be selected arbitrarily.

The sampling error is present even at the moment of using the best procedures since no sample can be an absolute guarantee of being an exact replica of the universe of study it represents [39], [47], [52]. That is, errors are unavoidable in any sample. This difference between the real universe and the sample of study is called sampling error. The important thing is to be able to determine the order or margin of the errors and their frequency inside the set since the latter allows us to establish the confidence interval inside which we are moving. Usually two kinds of mistake are distinguished: systematic or accidental. Systematic mistakes stem from several sources which, although they are alien to the sample, generate distortions. These distortions make the obtained results veer towards a direction in particular. Some examples of these systemic mistakes are the omission errors due to observation errors of the researcher, which is why they are usually called human factors; the insufficiency in data compiling, especially in the questionnaire technique (lack of responses), the inadequate replacements of the elements that make up the sample, that is, when it is not possible to have access to the correct elements and one resorts to those within reach, for instance. The second kind of error, called accidental or random mistake, is that in which the average values obtained from the sample will differ from the real average values of the universe. These differences are generated in the randomisation or in the measurement instruments. The sampling error depends on two factors:

1. The size of the sample: the greater the size of the sample, the lesser will be the error.
2. The dispersion or typical deviation of the sample: the greater the dispersion, the greater the error.

Consequently, the communicability evaluators must avoid to the best of the inability the systemic or random errors, starting with the personal human factors and the work team (evaluators and users of new technologies), if there was at the moment of carrying out the tests or remarks, for instance. The human factors in some places can severely damage the validity of a heuristic research, because those responsible are guided by the principles of what Saussure calls "parochialism" [39], [53]. That is, being against the interrelations among human beings beyond the local gridlock geographically speaking and even outright rejection of novel things, deriving from scientific progress. The notions and solutions deriving from computing, software engineering, usability engineering, human-computer interaction, ergonomics, etc. to eradicate that source of error and increase in the production costs of the interactive systems in the last two decades.

SOFTWARE FND HARDWARE EXPANSION FOR MULTIMEDIA MOBILE PHONE: NEW TECHNOLOGIES, USERS FND MARKETING

The mobile phones have developed so quickly in the last five years that the great majority of them currently possess a set of functions, similar or superior even to a Pocket PC of the early millennium. Conversely to what has happened in the evolution of the computers among the operative issues for Macintosh or Microsoft compatible with IBM PC, in the context of the mobile phone each hardware manufacturer presents his alternative in software for the functioning of their products. Obviously this is a great advantage for the advance of the potentialities of the device in the telecommunication issue, but it is also true that the adult public can't interact quickly with each one of these models without a previous experience. That is, there is no standard universal in those operative systems, for instance. Here is the reason why one of the design categories of the interactive system such as compatibility has a very important role.

One of the main components of compatibility is connectivity [39]. Great part of the economical and industrial activity regard the mobile phone as a device to be connected. In this sense connectivity has been boosted. Technologies such as Bluebooth or the infrared have lost importance with regard to the WiFi and 3D connectivity. Even the possibility of having access to the Internet from the mobile phone is becoming some necessary for a growing number of users. Consequently, in the middle of the current decade almost all mobile phones will need to be wired to the Internet. The appearance of the 2.0 Web is a determining factor in this phenomenon. Obviously, this factor can be either slowed down or boosted if free accessibility to information is maintained. However, from the sector of on-line information, such as the digital press, they have started to close down the free access to on-line contents. The user can only have access to the first page of the digital paper, for instance. However, the size of the screens of the mobile phones have not been convinced to supplant an E-book, for instance. Now there is no operative system that doesn't count with several alternatives to be able to connect with Facebook, Twitter, etc. We have gone from the massive communication era to the personal communicability era or in virtual communities, where the software and the hardware of the mobile phones is a determining factor in daily life [54]-[56].

Inside this new expansion age of communicability a variegated spectrum of applications for the mobile phone has emerged. In the same way as happened with the databases in 1980 and which has been maintained along the decades, the important thing is transportability and/or compatibility of these data or files as the software of the applications evolve [57]. As happened at that time to the databases, the present is a period of great activity in the design environment and the programming of applications in the mobile phones for millions of users across the planet. The origin of the current programming phenomenon is to be found in Apple's App Store. Since its conception, all the multimedia mobile phone makers have had their own services available. Some of them are works in progress, such as is the case of App Catalog of WebOS, others, in contrast, have a trajectory in the sector such as Android Market. Obviously, these examples like others, do not reach the market figures as for services of the Apple iPhones. The great advantage of this phenomenon is that so far the operating systems of the commercial mobile phones were regarded as a fenced-in system, that is, without enlargement possibilities. In contrast, now we can widen its functionality to carry out any kind of tasks, since the terminal itself. The updating or widening of functions will be made in an immediate way. For instance, we can get the street guide of a bordering country if we want to circulate through its highways and/or roads network. We can also update the videogames in the case that we have missed a train or a plane and we have to wait for the next one to leave towards our destination, etc. However, in the view of this great possibility of updating, it is always necessary to make a diachronic analysis of the applications as it is made in the multimedia interactive systems and the mobile phones are not an exception. The reason is that many applications for which the provider cashes in from the potential user are just adaptations to the already existing software. The problem lies in the interfaces of said applications because in some cases they have not been adapted to the new screens of the multimedia mobile phones. For instance, Windows Mobile needs to get updated, otherwise it may become obsolete from the communicability and usability point of view. Currently WebOS has been devised to be always wired to the net, to the extent that it has been conceived since its design stage to be always informed about the latest news and keeping the possibility that our data are accessible from any place in the world. That is, one is boosting the artificial life but reality on-line of the information stored in our mobile phone devices. We say artificial life because in some way the mobile phones will exchange data with the network in an automatic way, including the updating and real because in contrast to artificial life with artistic purposes, such as the fractal algorithms to generate tridimensional images, the data that we handle are vital to the telecommunications with commercial, industrial, educational purposes, etc. That is, the daily life of the human being in the economically developed societies.

HUMAN-COMPUTER INTERACTION AND INTERFACE DESIGN

In this daily life there are automatic cashpoints with tactile screens where some Southern Europe countries have the greatest network of international banking information such as in Spain. That is, information and banking operations have been within the reach of everyone since the 90s [14], [58], [59]. This technology already existed in some screens of the Hewlett Packard computers of the 80s, always with a commercial and/or financial purpose. Simultaneously, the first stands of tourist information were also created in the buses, trains, airports stations, etc, with the use of tactile screens. Later on, this way of interacting with the interactive systems has been reduced in size such as mobile computing screens like Palm, Tablet PC, E-book, etc. [60]-[64]. In these small size devices the tip of a pencil is used to activate the several functions of the programmes instead of the mouse, for instance. In the first models of the mobile phones keys were always used like in the keyboard of a calculator. However, from 2007 onwards we have entered the tactile revolution of the mobile phones like the iPhone, multimedia phone with Internet connection, tactile screen (multitactile technology) and a minimalist interface. It was from then on that almost all the mobile phone makers had to adapt their products to the new requirements of the users to have a tactile screen available.

Obviously, this also entailed the adaptation of the operating system to the new demands. In this speed of adaptation the communicability of many interfaces has been left aside to face the international demand of the new phone models which still maintained in their tactile screens the same old functionalities, shapes, topology of the interface components, etc. as in the classical phones with keyboards. In short, a kind of opportunism took place in the design of the interfaces, that is, to change everything so that nothing changed (gattopardism). To such an extent that the fast remodelling of the software to adapt it to the demands of the international market has prompted that we are in the face of interactive systems for mobile phones which although are handled with the touch, they still keep the same design as before. In an endless number of cases, these are hard to use and require expert users in computing. Consequently, the communicability and usability of these telephonic devices gets lost.

In the constant evolution of the new technologies and in a particular mobile microcomputing joined with telecommunications, the changes that take place daily, can make totally obsolete as a novelty a product that has been promoted for a week on digital television, for instance. The reason for this has to be looked in the hardware revolution that is always ahead of software. In the context of mobile phones, the revolutionary changes linked to software and hardware can derive both from the design of the interactive software or the ergonomic point of view, the appearance of new components to cut down even more the size of the phone devices, the incorporation of new materials to lengthen the load of the batteries or make more agreeable the vision of the information on the mobile phone screen, the incorporation of new devices or the adaptation of others which exist for constant connection, such as can be Bluetooth , etc. (see Figures 7 and 8). In all these cases, you always have to differentiate the inventions from the technological discoveries. That is, that which didn't exist before in the context of the new technologies, for instance, or the new use of something which already existed. In this sense, mobile phones is an interesting field of intersection of inventions and discoveries.

In the history of the new technologies, as a rule, a consensus was usually generated among the different manufacturers of the new products to set up a series of quality standards, not only from the point of view of the product and/or services, but also of the potential users [65], [66]. Nevertheless, this modus operandi has all but disappeared from the mobile phone scene, where each manufacturer, Asian, European and North American has set up his/her own rules, trying to impose them to the rest of the potential users. Evidently, there are elements which are some kind of common denominator: keyboard, screen, batteries, servicing costs, coverage, etc. In principle, this common denominator has been easy to manage, both from the hardware and software point of view, for instance. In contrast, the cost factor of the service has served to slow down the democratic spread among the population. This factor deserves a detailed analysis, as it will be seen in the next sections, because it prevents in many cases the evolution of the population's quality of life through the new technologies.

Figure 7. "Motorola" and old mobile phone: heavy and long –16 centimetres (no multimedia) and a new mobile phone "Samsung" E1150 (light and short but closed –8 centimetres).

Figure 8. Once the Samsung model is open (16 centimetres), it has the same size as the Motorola. Because of ergonomic reasons, the users whose age surpasses 60 years prefer that or another similar model.

The operative system in the mobile phones has served to gain the potential users of the hardware. In this regard as usual the lines followed by Microsoft and its operative system Windows on one side, and Apple on the other side are worth mentioning. From the point of view of the design of the interface component it is easy to detect in the first operative systems how there is a transfer of an endless number of features (dropdown menus, scrolling bars, activation and applications shutting, etc.), similar to the Windows operative systems or those used in Apple. Consequently the features of the design of the interface, the communicability, the software and its updating (applications located in on-line virtual shops), the hardware and its potentialities in the user's mobility [5], among other main features lead to the trends of the multimedia mobile phones being concentrated on a few brands, mainly distributed across Asia, America and Europe.

ROCOCO OR MINIMALIST DESIGN FOR MOBILE INTERFACES

On the screen of the different models it can be seen how there is a prevailing style in regard to the amount of functionality of the icons present in the interface [67]. In the interface of multimedia mobiles and for space reasons we can find either few or many icons. That is, a minimalist style, in the first case, which sometimes implies speed in the performance once the user knows how to navigate through the icons and even the same icon can have several associated functions, unfolding a menu, for instance. In the opposite sense, we can come across screens where each icon has a given function and consequently the required number of icons is bigger and it may be that they do not unfold associated windows [16]. If there is communicability in the design, the main advantage in this case is the ease of use, especially for those adult users who do not have a usual or constant familiarity with mini screens of the mobile computer, for instance.

In the wide range of the operating systems that work in the mobile phones it can be seen how the dynamic and static means are present in computer devices aimed at the small-sized telecommunication. These operating systems allow us to carry out tasks such as recording high quality video, obtain high quality pictures to generate a high quality personal agenda, etc. In the current work we will focus on some of the following mobile phones with tactile screens that are next listed. Evidently many of these features will be outdated by the brands that are mentioned at the moment of reading of the current chapter, however their main features have established different differentiation lines which throughout time will have either joined or excluded each other. The models and their makers are (in brackets are the initials of the different models):

- Android 1.6 (HTC Tatoo) is an advanced model inside the Android range. Additionally, the low resources consumption of the operative system allows for a remarkable speed in the different applications which can be stored in it. The only drawback is the size of the screen, which is smaller in regard to the HTC-HD2 model.
- Blackberry OS 5 (Blackberry Store 2). The Canadian firm incorporates a tactile screen with a haptic technology which acts like a great button and which upon writing triggers a pleasant feeling to the user as if he/she were writing on a keyboard.

- iPhone OS 3.1 (Phone 3GS 16Gb). This embodies the evolution of Apple's hardware. It has a high RAM memory and a very fast processor.
- Symbian S60 V5 (Nokia 5800). It is a mid-range model with a high quantity of multimedia functions.
- WebOS (Palm Pre). Through the current model the firm Palm has released to the general public its new operational system. On it is a slide design keyboard is incorporated and the touchstone technology allows to reload its battery in a wireless way.
- Windows Mobile 6.5 (HTC HD 2), the HTC initials refer to the manufacturer in Taiwan, that is, High Tech Computer Corporation. It is a model with a high number of functionalities and with a 4,3 inches screen. It has a IGHz Snapdragon processor which allows to make work quickly the operative system it carries.

It has been seen once again how in the evolution of the software and the hardware applied to the new technologies in telecommunications, the designers, manufacturers and marketing work in an interdisciplinary way to obtain the highest profit in the least time. However, in the initial process of the design and conception of the new products and services in mobile phones it is important to consider communicability between the new devices and the potential users.

In the context of the early mobile phones and those of the first of the new millennium the following phone models were analyzed (users in the HCI lab): Siemens C45, Siemens A65, Motorola RAZR V3x, LG U81200, Samsung E1150, Samsung SGH-L170. Besides, we worked with Blackberry Storm, HTC HD2, HTC Tatoo, iPhone 3GS, Nokia 5800, and Palm Pre. In all these models there are several common denominators, such as the phone card, the connexions to the different peripherals (loudspeakers, keyboards, earphones, etc.), the batteries [8], etc. The energy consumption of this kind of phone is the Achilles heel of several models since they require to be connected several times every day to the electric network. The ideal solution today is that presented by WebOS, with the Touchstone technology. The multimedia functions require a high activities load for the processor, which entails a high energy consumption. It is the case of Internet navigation through Wifi or 3G which quickly empties the battery as also through the use of the GPS location systems. The mobile phones are replacing little by little the devices for such ends or the Palm PCs. This Achilles' heel has not evolved at the same speed as the rest of the hardware components. Equally, from the point of view of software a series of strategies have been activated to lengthen the duration of the battery loads as will be seen further on. Next the last generation multimedia mobile phones will be grouped. However, many of them do not emulate anymore the early models of the mobile phones. Consequently, for the adult population, some manufacturers have gone back to the early models, reducing considerably their size, thanks to the new materials for the hardware, and with accessible costs for most of the population, such as is the Samsung E1150, with a 65.000 colors screen.

SETS AND SUBSETS OF THE MOBILE PHONES IN THE COMMUNICABILITY ERA

A way to understand the state of the art of communicability inside the current technological context consists in presenting and breaking down the main characteristics of the mobile phones such as their main positive and negative aspects. These comparisons must take into account both the diachronic and synchronic of the technological evolution. Since the first monochromatic models of the 90s we have seen how the size of the original tube of the classical phone has been cut down until occupying the palm of a hand [61]. Then the colour screens have appeared the picture and video cameras have been incorporated, multimedia reproduction, ubiquitous data collection, etc. [15], [68]-[70].

Since 2010 the convergence of the tactile screens and the connection to the Internet we are in the face of a software generation which is trying to turn them into miniature computers. We can draw a dividing line in the set of the mobile phones generating two subsets. In the first are the new generation operative systems such as is the case of Android (Phone OS and WebOs). Whereas in the second is included that software that has been evolving along time, such as are Windows Mobile and Symbian.

In the first subset it can be seen how the design of the interface has a more modern appearance, with an intuitive navigation and the architecture to be wired to the net. At the same time we can see in the second group a whole operation of interface make-up, that is , the menus have been redesigned making more visible some of their components bearing in mind the greater size of their screens, for instance. Additionally, they have incorporated new functionalities. However, there are still menus of a classical or ancient style which still damage the communicability with the users. All of this happens inside the Windows Mobile. In new version 6.5 we can see how the marketing has bet on a make-up operation by calling it “Phone”. Obviously, the possibility of moving across its screen without using a pencil is something that the young generations of users appreciate. However, the scrollbars and basic menus of the operative system still remain dodgy to our fingers and hard to press. Now being a veteran in the software sector has some advantages such as counting on a high number of software developers who constantly publish its applications. Many of these applications are free.

Here it is necessary to keep in mind that a free software does not entail a lower quality, but a way which some developers have to get known in the national or international market. It is not for nothing that we can find in the Internet free quality tools to carry out any kind of task with Windows Phone. For many of these tools it would be necessary to install the Net Computer Framework so that the programmers could develop efficient applications and without so many complications. The problem that this operating system currently has, but which derives from the manufacturer’s commercial and marketing environment, that is, Microsoft, is the lack of a virtual shop to concentrate the sales of these applications, in the way that the Apple’s App Store does (http://store.apple.com). These virtual shops allow millions of users to navigate through its catalogue and purchase those applications they are interested in having in their mobile. In contrast, Nokia is trying to generate a 100% renovated system aimed at Linux [17]. Between both subsets we can insert a third, which is the case of the Canadian Blackberrys, with a sober-looking interface but which reaches very well the communicability goals among a wide range of potential users of these devices in the adult population, for instance.

INTERNET NAVIGATION, APPLICATIONS AND E-COMMERCE MISCELLANEOUS

In regard to Internet navigation, all these models have the necessary components incorporated to implement this goal [71]. However, the current hurdle is the cost of the navigation through the mobile phones. As it happened in the dawn of the mobile phones age, it is necessary that the services offer diversifies in order to cheapen costs. Consequently, the use of the portable computer is a priority for many professionals who require hours of internet connexion. Inside the said context, Google, with the promotion of the Open Handset Alliance and the conception of Android as a free operative system has started to seek the free phone. So far both the hardware manufacturers and the mobile phone operators seemed to have agreed on making it difficult for customers to have free use of the purchased hardware. It suffices to remember the existing barriers inside the EU that forced to buy mobile phones in each country of the zone. Now, all the platforms have opened their gates to the developers so that they experiment with them through the different development kits, allowing a higher or lower freedom degree according to the case. The manufacturers know that the success or failure of their on-line trade is linked to the future of their operative systems. Consequently, the contribution of third firms or independent developers is a cornerstone of this new way of approaching the perspectives of the mobile technology in the coming years.

Google intends to develop a new operating system for mobile phones with an open code based on Linux. An aspect on which Apple and Google diverge are the software restrictions. Apple intends to keep a ring-fenced environment in the control over all the applications that work in the iPhone [1]. Additionally, it is necessary to pay certain costs for the SDK (Software Development Kit). Whereas the Android source code is freely available in the net within anyone's reach, and the development environment can be obtained for free. In this way all the programmers who wish to can contribute to the development of the operative system. Android 1.6 is a fast, stable and 100% wired to the net system. In contrast to what happens with Microsoft phone, there is on-line the Android Market which allows to have a priori a global view of the applications that can be enclosed in the mobile phone. Evidently it does not reach the level of Apple's App Store. One of the problems of Android is the safety for it is the user who determines the permits for each application that he/she encloses in the mobile, for instance, Internet communication, sharing the stored personal information, the phonebook , etc. Although it is the user who makes the decision, it is possible that there are failings from the point of view of the safety of the information.

Since the 90s the name Palm has been linked to the pocket computer of millions of users in the whole world, especially for the handling of the personal agenda and the reading of e-mails, once downloaded from the computer [72]. At the start of 2000 it was also linked to the smartphones thanks to Palm OS. However, in the international commercial context Windows Mobile and iPhone have shifted it from the first spots in the preferences of the users in the mobile phones market. Yet it has been seen in the last few years how Palm is gaining ground inside the mobile phone sector with WebOS. On-line integration and the multitask applications management are the strengths of this system based on Linux. Upon navigating through the interface of the operative system we can see how the Cards interface (a windows in which every application will run on a different card and which allows to switch from the one to the other without wasting the work that has been made). On the desk you can switch from one card to the other and see them work, although they have a small size. The multitask system is very good, since one can see the changes in real time if in the middle of a process. Additionally, the implemented method is qualitatively high for the notifications of system events and of applications that are functioning on a background. In the case of the e-mail, they are located in the lower part of the screen, occupying a small space.

The devices presented so far that have WebOS incorporated are Palm Pre and the Pixi (gestures systems). Both present a zone outside the tactile recognition screen known as "gestures area" and in our days they make up an important part of the communicability and the usability of the operative system. The gestures are movements that the user makes with the finger in said area and on the screen to carry out certain operations, such as going backwards (it means to go back in the menus or getting out of the main menu from an application) through the shifting of the finger from left to right. Obviously here the cultural factor of reading is taken into account, since it differs with Eastern users. Not bearing in mind this aspect may be practical from the system's usability point of view, but it is little efficient from the point of view of communicability for a system conceived in Western culture and used by millions of people in China or Japan, just to mention two examples. Once the commands have been learned by the users, these are carried out in an almost natural way, and they remarkably speed up the interaction with the mobile phone, because you can do without the keys and the menus for some functions.

This operating system has been designed to be wired to the Internet 24 hours a day [4] to such an extent that several of its components do not work if we are not wired to the Internet (connecivity around the clock). That is, a great part of its applications have been thought to keep a constant data flow, carrying out periodical synchronization operations with Gmail or Google Calendar. Evidently, the cost factor of the constant phone connection has to be previously analyzed at the moment of plumping for such an operative system. Also the energy consumption that makes the batteries to have a short duration and force the user to constantly connect to an electrical power source to reload them [8].

It is easy to see in the Internet how Palm hasn't put yet at the disposal of the developers of the applications to foster their offer a portal that groups them up as in the case of App Catalog. Obviously in this portal it is possible to detect those which are not yet 100% developed, and they have the Beta heading. Others in contrast present an interesting chance of incorporating additional functionalities to the phone, such as a language translator –a dictionary in nine languages), Pocket Mirror (synchronization with calendars, notes and contacts from Microsoft Outlook), Missing Sync (automatic synchronizer through the wireless net, the common programs data, for instance: Outlook, Windows Media Placer, iPhoto, etc.) [9]. In short, the interface is quick to answer and quite attractive visually. However, its applications are not sufficiently well categorized in the portal related to it.

iPhone like iBook or the iPad is always linked to software to set historic landmarks in the evolution of computing [1], [2]. The iPhone OS 3.1 is not an exception to this role due to the innovation of the software. Both the hardware and the software have set a pattern to be followed by the great majority of mobile phone makers. Although that star product was located because of its cost at the top of the pyramid of the democratic distribution of the new technologies when it comes to phones, now with the speed of the new operative system without the need of having available a powerful hardware, has moved it to the basis of the pyramid. The feeling of ease that you can enjoy in the menus is due to two main factors: the good design that Apple has traditionally developed along the decades and to the capacitating screen that these devices have to be handled through the fingers. However, operations like copying and sticking, something so usual in the alleged Latin creativity hasn't been possible until the third version of the operative system. What is really funny is that in the first version of this phone no multimedia messages, that is, MMS, could be sent.

Appe Store is a portal imitated by all the competitors of Apple. This virtual shop, since its opening until 2010 has surpassed two billions applications downloads. These are applications aimed at the basis of the technological pyramid where the secret lies in their cost, which is below 10 US dollars. Evidently, if we add to this the possibility of developing applications through the development kit, or iPhone SDK, the interest grows inside this pyramid of users which can be placed at its half, such as are programmers. The problem that has arisen is that many of these applications that are very interesting from the point of view of software engineering and especially in the context of the open software is that they harm the multinational's profit purposes. In this regard we can mention the possibility of carrying out free VoIP calls or the use of a terminal as a 3G modem for the desk computer. Without any doubt it is the portal where the user can find an endless and varied number of applications. Some of them are: the Evernote (it allows the user to keep in the form of notes a collection of textual or multimedia information that we want to remember, obtained from the Internet, for instance, and kept inside a cloud), Twitterrific (to send 140 characters in the communication), Quick Office (file editor in Word or Excel, visualize the contents of the files in Powerpoint, Pdf, Html, etc.), and Stanza (manager of the e-book library integrated to tuning).

The Subian S60 v5 can be found installed in other phones that do not belong to Nokia. This is another operating system that does not require a high hardware range for its correct functioning or obtaining the highest number of functions from it. In the new version there has been an attempt at leaving aside the use of the pencil as much as possible in order to adapt to the interaction through the fingers. Here is the reason why on the main screen we find big size icons for the main menus, but unfortunately the vertical scrolling bars are still used which prompts the use of a pointer. Another of the negative aspects is the inability to properly navigate through several applications which have been developed in Java in the previous versions, whose navigability was solved with virtual buttons on the screen. Although compatibility is insured in the installation of the applications developed in Java, it doesn't happen the same with the rest of the applications typical of S60, which can be installed but you can't navigate through them.

The quality of Nokia's virtual shop is lower than that of Apple's. Although there is the possibility that the user gets additional information to navigate through categories, ordination by popularity, description of applications, various comments, scores, etc. The information in this regard is not 100% reliable as in other cases of virtual shops, since it is made by the users of the virtual shop and the rest of members who make up the international community of mobile phone users, for instance. Nevertheless, it possesses an interesting community of application developers for this kind of hardware. Some of these applications which can be found in the virtual shop are: Fring (which focuses several communication services such as Skype, Facebook, Twitter, etc.), Battery Extender (it extends almost 30% the shelf life of a battery, disconnecting those features of the mobile phone that penalize it with a higher consumption, that is, GPS, Wifi, etc. in the moment in which they are not used). Easy Busy (it allows to avoid those incoming calls we do not want to answer), MobiSystems Office Suite 5 (an office automation solution that allows to create, see and edit Word files, Excel and Powerpoint, for instance). Qik (allows to broadcast that which is recorded in streaming, alive and direct), Handy W1 (to keep interesting access points when they are discovered and visualized through Google Maps).

In the current decade the Blackberry is a device which can be regarded as today's symbol of the high range of professionals inside the mobile phone context. The reason is due to the high reliability of the operative system and the safety of the data in the telecommunication process [17]. This latter issue is what makes it an ideal device for computer expert users. Now this doesn't mean that for its use you have to be an expert in computing, because precisely one of the main features is the ease of use. The menus are clear and simple with a minimalist interface and a truly practical conception design. With the incorporation of animations and 3D emulations of the keyboards in the interface the communicability of the new version is optimal. The possibility of having the QWERTY keyboard available. With the passing of time and the arrival of the new versions it was decided to eliminate this possibility by generating expectation among the users accustomed to it. The new strategy went through the Surepress system. That is, a capacitating screen which can be pressed and acts like a great button. Upon touching the screen this does not make a pulsation, but it would rather be the equivalent to move the pointer through it and it highlights the different objects of the interface. To activate it you have to press it by providing the user a feedback feeling.

In the Storm model they have incorporated 4 pressing points and it remarkably improves the user's experience. As the new versions of these models have appeared, it has been seen how the multimedia functions have been gradually inserted, for instance, in the video context, musical reproduction, etc. to the detriment of office or business automation. The aspects related to the Internet have also been improved. There are not many applications if compared to the Windows Mobile operative system, for instance. Currently there is the possibility of furnishing our mobile phone with plenty of videogames which use to the utmost the hardware potential. To such an extent that we have a variegated range of applications such as: Nice Office CRM (it is a very complete virtual office since from it you can manage the information in a very efficient way), Personal Assistant (it monitors the bank operations with the possibility of deactivating said application in the case of loss of the mobile phone), DriveSafely (it allows to visualize the messages received while driving and without the need of interacting with the phone), Call Blocker Professional (allows to hang up, send off the calls to the voicemail, keep them waiting, etc.), QuickLaunch (it admits the possibility of visualizing the interface and the menus or submenus of the most usual tasks, for instance), the Weatherbug (it anticipates weather forecast in the case of travelling, knowing beforehand the weather the user will find when reaching destination). To have access to the virtual shop to set an application known as Blackberry App World which sometimes doesn't work very well in some Blackberry models.

ASSESSMENT AND RESULTS

It is important to remember that in its origins usability engineering was not considered in a positive way, and even unlikely because of a series of methodology questions the possibility of defining quality attributes and starting from them to break them down in relation to the components to generate the metrics. Our metrics were aimed from the start at interactive communication and especially at quality, that is, communicability. These metrics have been examined and used in endless on-line and off-line interactive systems oriented at tourism, e-commerce, e-learning, etc. [20], [37]-[39], [43]. Besides, their constant use and positive verification with the tests carried out with other methodologies of usability engineering have allowed the establishment of a new professional called "analyst in communicability" [39]. For instance, in the evaluations made in Barcelona the evaluators of usability and communicability engineering previously to the evaluations made with the users, have carried out a series of planned tasks to evaluate the design of the multimedia mobile phones and be able to verify the tasks that had to be developed by the real users. All of them have come to trace the same errors in the interactive design of the multimedia mobile phones, for instance. The advantage of the proposed methodology is that we need no labs or special equipment, only a communicability analyst.

The universe of study is made up by users of mobile phones (classical, multimedia and of last generation), hailing from two towns in Southern Europe: Barcelona, Spain (Figure 9) and Bergamo, Italy (Figure 10). In all the real universe amounts to 200 users, which have been invited to participate in the experiments as volunteers (all the food and transportation expenses were covered during the experiments). Those who have volunteered, were later classified in relation to age (children –4/11 years, Junior –12/17 years, Adult 18/64 years, Senior –65 or more years), sex, working experience, knowledge in computer science and English, the phone models they had available at the time, through a draw system, using a lottery approach. In this way were established the groups of users with whom work would be started. At the same time, the experts in heuristic evaluation in usability for multimedia and communicability systems started interact with the different models of mobile phones and set down a series of activities that had to be carried out by the users (in annex # 1 can be seen the table recording them). These activities were timed in minutes.

The users have been split into groups of 20, which were subdivided into four groups of five. Each one of them has been given the tasks to be carried out, for instance, sending a SMS, using the different models (Figures 11 and 12). These operations had a maximum time to be carried out, and the result to be obtained was of the yes or no type. Some of the operations or tasks (in brackets the T indicates "task" and they are classified by the numbers in the Figures 9 and 10 of the final results) to be carried out at the moment of the evaluation were: Having access to the phone with the secret number or PIN –Personal Identification Number (T1); Identifying the lost or made calls (T2); Making an enlarge –zoom in picture of a painting in the room (T3); Using the phone as a videocamera (T4); Reading or answering the received messages (T5); Writing down a phone number in the agenda (T6); Making arithmetic operations with the calculator (T7); Changing the volume of the sound of the calls (T8); Making international calls (T9); Tracing some streets in Google Maps (T10); Reading the headlines of the news in the front pages (T11); Recording messages (T12); Activating and deactivating call diversion (T13); Listening to music (T14); Playing with classical videogames (checkers, chess, cards, etc.) or in the net (T15). The evaluation techniques used were: observation, questionaires, videotaped sessions, and interview. However, direct observation has prevailed in the whole evaluating process.

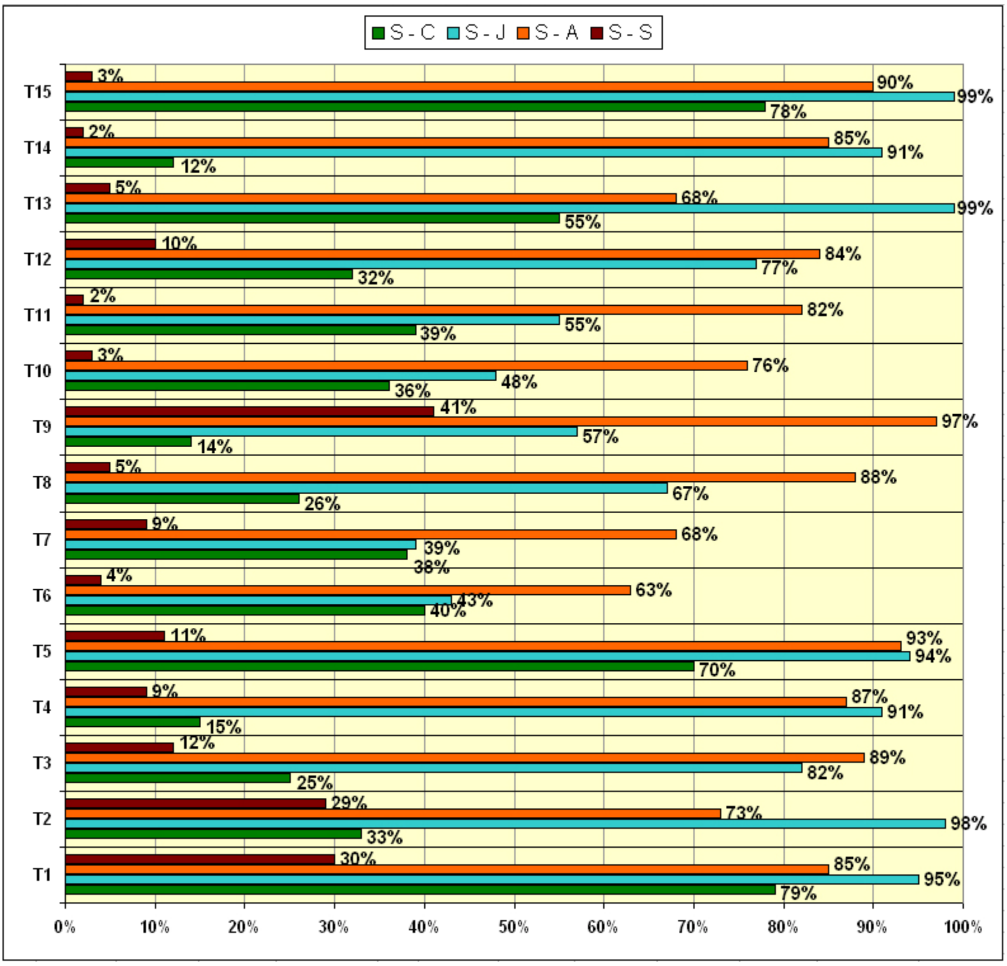


Figure 9. Results with Spanish (S) users (C = Children, J = Juniors, A = Adults, and S = Seniors).

The results obtained in the evaluation of the mobile phones with users inside the lab have made it possible to draw up a list of actions (vademecum) and conclusions which are next detailed:

- Usability must be split into several levels. One must start from very general problems which are common to all interactive systems and then aim at the specific problems of some given systems. That is, at a first level there are the mobile phones systems in general, at a second level the multimedia systems categories and at a third level the kinds of systems, for instance those aimed at communication or those aimed at work and communication such as a "Tablet PC" can be. At each level there must be a preparatory stage, a list of abstract tasks and a set of evaluation criteria.
- Upon evaluating a multimedia mobile phone system it is necessary to take into account first an inspection with an expert in heuristic evaluation of multimedia

systems and in communicability. The results of the inspection are some application outlines, a list of potential problems and a set of abstract tasks. The tasks must be specific activities for users in the empiric evaluation.

- Empirical evaluation is more efficient in quality results and costs when there is a list of abstract tasks.
- The expert's inspection must be efficiently combined with empirical evaluation.
- Finally, there is the need to contact with the production of the evaluated system to advice the designers on the upcoming versions of the mobile phones, for instance.

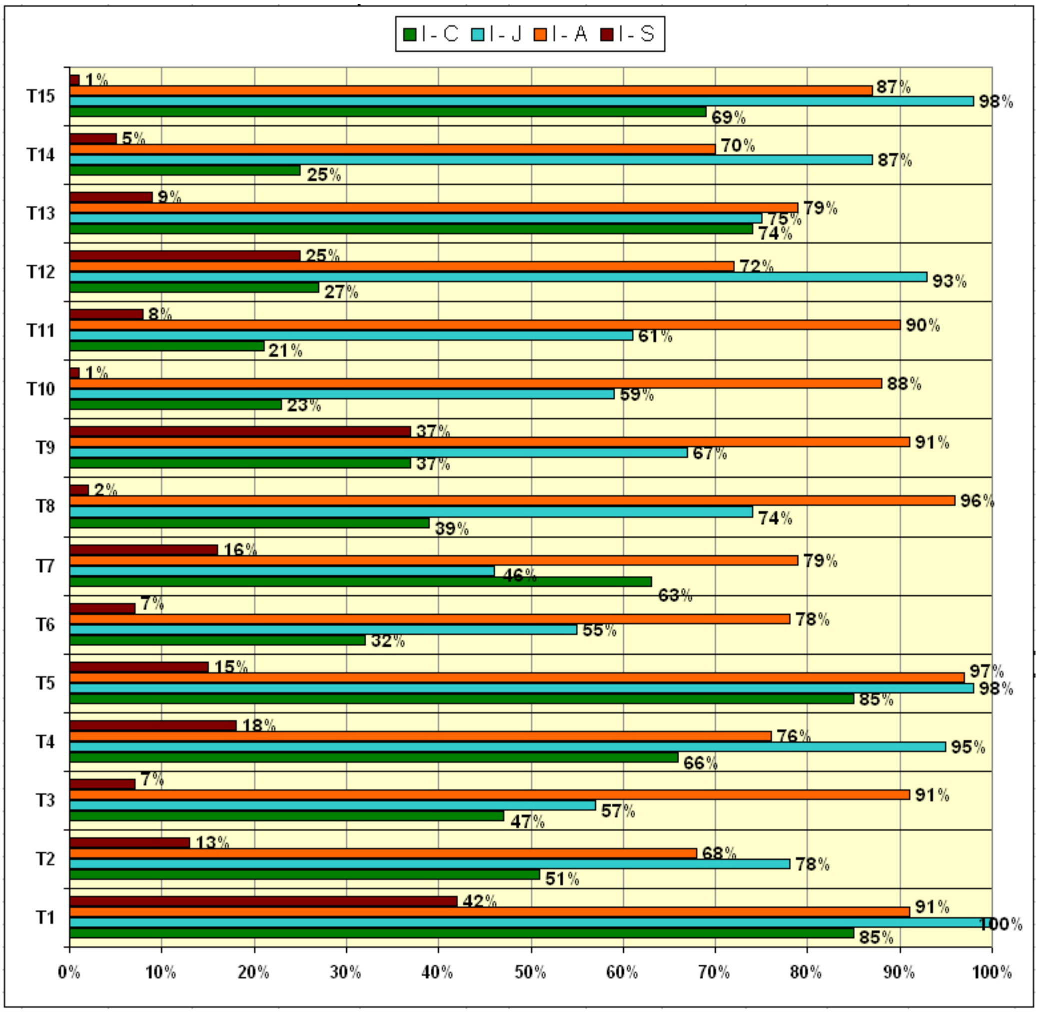


Figure 10. Results with Italian (I) users (C = Children, J = Juniors, A = Adults, and S = Seniors).

Figure 11. Senior user testing different models.

Figure 12. User analyzing the ergonomic aspect.

It has been seen in the current work how the adult users whose age ranges about 60 years prefer first generation mobile phones, that is, phones that serve to communicate through the voice. Only 15% of them use the phones to send SMS. Consequently, the Motorola RAZR V3x and Samsung E1150 models are a real commercial finding aimed at this kind of users. The Italian and Spanish users with a greater purchasing power and whose age ranges between 16 and 40 years prefer the latest technological breakthroughs from the point of view of mobile phones. However, it is important to stress that only 8% among them know or at least have tried to use all the possible functions from the multimedia mobile phone. Among the young knowing such functions means to establish a virtual community. Adult people used to resort to the Internet or the salesmen or the assistance service about the functions of the phone. Currently, among children whose age ranges between 7 and 10, they start to use the mobile phone as a means of communication for the location of adult people (relatives, friends and schoolmates of their age). Among teenagers the mobile phone is used in 93% of cases to send SMS among friends and/or classmates in high school with occasional calls between parents, grandparents, siblings, etc., videogames, and currently there is the trend to trying to navigate from the multimedia mobile phone to have access to on-line videogames, for instance.

Figure 13. Senior user talking by the phone after having chosen the model that best adjusts to her daily needs.

CONCLUSION AND FUTURE WORKS

In the communicability, usability and usefulness triad a spate of difficulties has been seen which persists among the mobile phone users in the European countries of the Mediterranean basin. The presented methodology has pointed out quickly which are the errors of this triad. Said errors may damage the current era of communicability expansion we are living in. The quality attributes and their corresponding metrics must be constantly updated to adapt to the daily evolution of this sector of the new technologies, where the hardware goes ahead of the software. Consequently, it is important that the operating systems and the applications which allow the functioning of these devices take into account the different kinds of potential users, distributed all across the planet. The presented little vademecum is a first guide in this sense, since in Southern Europe the multimedia mobile phones may replace by the end of the current decade the portable computers, especially with the mobile computer, among the young users who constantly interact with computers and videogames. Secondly, for this group of users there is the TDTV and the cinema. In contrast, the adult public will keep on interacting with contents on the bigger-sized screens, that is, the PC, the portable computer, the TDTV and to a lesser extent the Tablet PC, the e-book and the PDA, for instance.

In the future we will aim our research at the mobile computer which uses the human body for its interaction operation with the interfaces that appear in the arms and hands of the users, for instance. We think that this will be the field towards which will derive the current multimedia mobile phones in the watch format. These are multifunctional watches whose size and shape can take up several of the current functions of the sensors of the mobile computer muscles, which currently are adjusted in the upper arm, under the armpits. It is based on EMG (Electromyography) senses this muscle activity by measuring the electrical potential between a ground electrode and a sensor electrode. This is a natural environment of the current era which is known as the communicability expansion era, where the users, through micro computing and telecommunications mobile devices interact with the interfaces which are projected from those EMG devices, to which can be linked the dynamic and static means of the multimedia interactive systems.

ACKNOWLEDGMENTS

The author would like to thank Emma Nicol (University of Strathclyde), Maria, Carlos and Miguel Cipolla-Ficarra (Ainci & Alaipo) for their helps.

REFERENCES

[1] Bondo, J. Barnard, D., Burcaw, D. & Novikoff, T. (2009). *iPhone User Interface Design Projects*. Apress: New York.

[2] Macedoni, M. (2007). iPhones Target the Tech Elite. *IEEE Computer*, 40 (8), 94-95.

[3] Tsai, F. et al. (2010). Introduction to Mobile Information Retrieval. *IEEE Intelligent Systems*, 25 (1), 11-15.

[4] Smith, I. (2005). Social-Mobile Applications. *IEEE Computer*, 38 (4) 84-85.

[5] Faruque, S. (1996). *Cellular Mobile Systems Engineering*. Artech House Publishers: Boston.

[6] Ramacher, U. (2007). Software-Defined Radio Prospects for Multistandard Mobile Phone. *IEEE Computer*, 40 (10), 62-69.

[7] Seyff, N., Grunbacher, P. & Maiden, N. (2006). Take Your Mobile Device Out from behind the Requirements Desk. *IEEE Software*, 23 (4), 16-18.

[8] Jacoby, G., Marchany, R. & Davis, N. (2006). How Mobile Host Batteries Can Improve Network Security. *IEEE Security & Privacy*, 4 (5) 40-49.

[9] Reponen, E., Huuskonen, P. & Mihalic, K. (2006). Mobile Video Recording in Context. *Interactions*, 13 (4), 28-30.

[10] Rao, B. & Minakakis, L. (2003). Evolution of Mobile Location-based Services. *Communications of the ACM*, 46 (12), 61-65.

[11] Wang, A. (2006). The shazam music recognition service. *Communications of the ACM*, 49 (8), 44-48.

[12] Dimitriadis, C. (2007). Improving Mobile Core Network Security with Honeynets. *IEEE Security & Privacy*, 5 (4), 40-47.

[13] Balan, R. & Ramasubbu, N. (2009). The Digital Wallet: Opportunities and Prototypes. *IEEE Computer*, 42 (4), 110-102.

[14] Herzberg, A. (2003). Payments and Banking with Mobile Personal Devices. Communications of the ACM, 46 (5), 53-58

[15] Giguette, R. (2006). Building Objects Out of Plato: Applying Philosophy, Symbolism, and Analogy to Software Design. *Communications of the ACM*, 49 (10), 66-71.

[16] Ozkaya, I. et al. (2008). Making Practical Use of Quality Attribute Information. *IEEE Software*, 25 (2), 25-33.

[17] Balasubramaniam, S. et al. (2008). Situated Software: Concepts, Motivation, Technology, and the Future. *IEEE Software*, 25 (6), 50-55.

[18] Toyama, K. & Dias, M. (2008). Information and Communication Technologies for Development. *IEEE Computer*, 41 (8), 22-25

[19] Mashey, J. (2009). The Long Road to 64 Bits. *Communications of the ACM*, 52 (1), 45-53.

[20] Edvardsen, F. & Kulle, H. (2010). *Educational Games: Design, Learning and Applications*. Nova Science Publishers: New York.

[21] Isaacs, E., Walendowski, A. & Ranganathan, D. (2002). Mobile Instant Messaging trhough Hubbub. *Communications of the ACM*, 45 (9), 68-72.

[22] Lindholm, C. & Keinonen, T. (2003) *Mobile Usability: How Nokia Changed the Face of the Mobile Phone*. McGraw Hill: New York.

[23] Resnick, M. et al. (2009). Scratch: Programming for All. *Communications of the ACM*, 52 (11), 60-67.

[24] Biehl, M. (2007). Success Factors for Implementing Global Information Systems. *Communications of the ACM*, 50 (1) 53-58.

[25] Ladner, S. (2010). The Essence of Interaction Design Research: A Call for Consistency. *Interactions*, 17 (2), 48-51.

[26] Wilson, J. (2000). Toward Things That Think for the Next Millennium. *IEEE Computer*, 33 (1), 72-76.

[27] Cipolla-Ficarra, F. (2010). *Persuasion On-Line and Communicability: The Destruction of Credibility in the Virtual Community and Cognitive Models*. Nova Science Publishers: New York.

[28] Berners-Lee, T. (1996). WWW: Past, Present, and Future. *IEEE Computer*, 29 (10), 69-77.

[29] Trowbridge, C. (2010). *Marconi: Father of Wireless, Grandfather of Radio, Great-Grandfather of the Cell Phone, The Story of the Race to Control Long-Distance Wireless*. BookSurge: New York.

[30] Miller, F., Vandome, A. & McBrewster, J. (2009). *Guglielmo Marconi*. Alphascript Publishing: Beau Bassin.

[31] Ling, R. (2004). *The Mobile Connection: The Cell Phone's Impact on Society (Interactive Technologies)*. Morgan Kauffman: San Francisco.

[32] Green, J. (1993). *The Get Smart Handbook*. Macmillan Publishing: New York.

[33] Nielsen, J. (1993). *Usability Engineering*. Academic Press: London.

[34] Brown, J., Shipman, B. and Vetter, R. (2007). SMS: The Short Message Service. *IEEE Computer*, 40 (12), 106-110.

[35] Homquist, L. (2007). Mobile 2.0. *Interactions*, 14 (2), 46-47.

[36] Cipolla-Ficarra, F. *MEHEM: A Methodology for. Heuristic Evaluation in Multimedia*. In Proceedings of the DMS'99. KSI: Aizu, 89-96.

[37] Cipolla-Ficarra, F. *Communication Evaluation in Multimedia: Metrics and Methodology*. Universal Access in HCI. In Stephanidis, C. (Ed.). LEA: London, 567-571.

[38] Cipolla-Ficarra, F. (1997). *Method and Techniques for the Evaluation of Multimedia Applications*. HCI International ’97. Elsevier: San Francisco, 635-638.

[39] Cipolla-Ficarra, F. (2010). *Quality and Communicability for Interactive Hypermedia Systems: Concepts and Practices for Design*. IGI Global Editions: Hershey

[40] Cipolla-Ficarra, F. (1999). *Evaluation Heuristic of the Richness*. International Conference on Information Systems Analysis and Synthesis. ISAS: Orlando, 23-30.

[41] Miller, F., McBrewster, J. & Vandome, A. (2009). *Augmented Reality*. Alphascript Publishing: Beau Bassin.

[42] Antoniac, P. (2008). *New User Interfaces for Mobile Devices Using Augmented Reality*. VDM Verlag: Saarbrücken.

[43] Cipolla-Ficarra, F. (1997). *Evaluation of Multimedia Components*. In Wemer, A. (Ed.), International Conference on Multimedia Computing and Systems. Ottawa, Canada: IEEE Computer, 557-564.

[44] Muller, M., Matheson, L., Page, C., & Gallup, R. (1998). Participatory Heuristic Evaluation. *Interactions*, 5 (5), 13-18.

[45] Kuurniawan, S. (2007). Mobile Phone Design for Older Persons. *Interactions*, 14 (4), 24-25.

[46] Bunge, M. (1981). *The science: your method and your philosophy*. Siglo XXI: Buenos Aires.

[47] Ander-egg, E. (1986). *Techniques of Social Investigation*. Hvmanitas: Buenos Aires.

[48] Berg, B. (2001). *Qualitative Research Methods for the Social Sciences*. Allyn and Bacon: Boston.

[49] Yule, Y. & Kendall, M. (1964). *Introducción a la estadística Matemática*. Aguilar: Madrid.

[50] Canavos, G. (1988). *Applied Probablity and Statistical Methods*. McGraw Hill: Cambridge.

[51] Goode, W. & Hatt, P. (1972). *Métodos de investigación social*. Trillas: México.

[52] Hand, D. (2008). *Statistics: A Very Short Introduction*. Oxford University Press: Oxford.

[53] Saussure, F. (1990). *Course in General Linguistics*. McGraw Hill: New York.

[54] Smith, I. (2005). Social-Mobile Applications. *IEEE Computer*, 38 (4), 84-85.

[55] Halvey, M., Keane, M. & Smyth, B. (2006). Mobile Web Surfing is the SAME as Web Surfing. *Communications of ACM*, 49 (3), 76-81.

[56] Macedonia, M. (2005). Power from the Edge. *IEEE Computer*, 38 (12), 123-127.

[57] Ambler, S. (2007). Test-Driven Development of Relational Databases. *IEEE Software*, 24 (3), 37-43.

[58] Christou, I., Ponis, S. & Palaiologou, E. (2010). Using the Agile Unified Process in Banking. *IEEE Software*, 27 (3), 72-79.

[59] Holland, C. & Westwood, J. (2001). Product-Market and Technology: Strategies in Banking. *Communications of ACM*, 44 (6), 53-57.

[60] Yoshioka, A. et al. (2006). Point, Push, Pull: The FAU Interface. *Interactions*, 13 (4), 40-41.

[61] Maiden, N. et al. (2007). Determining Stakeholder Needs in the Workplace: How Mobile Technologies Can Help. *IEEE Software*, 24 (2), 46-52.

[62] Prey, J. & Weaver, A. (2007). Tablet PC Technology the Next Generation. *IEEE Computer*, 40 (9), 32-33.

[63] Schneiderman, B. (2002). Understanding Human Activities and Relationships: An Excerpt from Leonardo's Laptop. *Interactions*, 9 (5), 40-53.

[64] Kimel, J. & Lundell, J. (2007). Exploring the Nuances of Murphy's Law Long-term Deployments of Pervasive Technology into the Homes of Older Adults. *Interactions*, 14 (4), 38-41.

[65] Hinman, R. (2009). 90 Mobiles in 90 Days: A Celebration of Ideas for Mobile User Experience. *Interactions*, 16 (1)10-13.

[66] Kangas, E. & Kinnunen, T. (2005). Applying User-Centered Design to Mobile Application Development. *Communications of the ACM*, 48 (7), 55-59.

[67] Marsden, G. (2008). Toward Empowered Design. *IEEE Computer*, 41 (6) 42-46.

[68] Yi, J. (2010). User-Research-Driven Mobile User Interface Innovation: A Success Story from Seoul. *Interactions*, 17 (1), 48-51.

[69] Shilton, K. (2009). Four Billon Little Brothers? Privacy, mobile phones, and ubiquitous data collection. *Communications of the ACM*, 52 (11), 48-53.

[70] Yap, K. et al. (2010). A Comparative Study of Mobile-Based Landmark Recognition Techniques. *IEEE Intelligent Systems*, 25 (1), 48-57.

[71] Cao, J. et al. (2007). Data Consistency for Cooperative Caching in Mobile Environments. *IEEE Computer*, 40 (4), 60-66.

[72] Anokwa, Y. et al. (2009). Open Source Data Collection in the Developing World. *IEEE Computer*, 42 (10), 97-99.

Annex #1. Table of Heuristic Evaluation for Mobile Phone.

Age	Sex	Working Experience	Knowledge in Computer Science	English Knowledge

Task	Activities	Time (minutes)	Results		Categoria Design
			Yes	No	
1	Having access to the phone with the secret number or PIN –Personal Identification Number				
2	Identifying the lost or made calls				
3	Making an enlarge –zoom in picture of a painting in the room				
4	Using the phone as a videocamera				
5	Reading or answering the received messages				
6	Writing down a phone number in the agenda				
7	Making arithmetic operations with the calculator				
8	Changing the volume of the sound of the calls				
9	Making international calls				
10	Tracing some streets in Google Maps				
11	Reading the headlines of the news in the front pages				
12	Recording messages				
13	Activating and deactivating call diversion				
14	Listening to music				
15	Playing with classical videogames (checkers, chess, cards, etc.) or in the net				

Design Categories: Layout (L), Content (C), Navigation (N), Panchronic (P), Structure (S), and Conection (Co).

In: Advances in Communications,Volume 8
Editor: Anthony V. Stavros
ISBN: 978-1-61324-794-5

Chapter 3

THE PROCESS OF LIBERALISATION OF THE TELECOMMUNICATIONS MARKET IN THE REPUBLIC OF SERBIA AND THE RESULTS ACHIEVED

***Vladica Tintor*[1], *Vlade Milićević*[2], *Milan Janković*[13] *and Jovan Radunović*[3]**

[1]Republic Agency for Electronic Communications, Republic of Serbia
[2]Faculty of Economics, Belgrade University, Republic of Serbia
[3]Faculty of Electrical Engineering, Belgrade University
Republic of Serbia

Abstract

This paper aims at illustrating the process of liberalisation within the telecom market in the Republic of Serbia from 2005 to 2010. During that period, the telecom market was primarily a state monopoly market. As such, this market was unregulated, characterised by the abuse of the dominant position, the operation of broadcasters without a broadcasting license, low investment rates, high service prices, relatively low service quality and, by and large, a low intensity of telecom service use. Relying on the Telecommunication Law, (adopted in 2005 and harmonised with the 1998 EU Directives), closely monitoring the development trends of new technologies and taking into account the experience and the best practices of regulatory authorities from the developed countries of the EU, the national regulatory authority managed, in a relatively short period of time, to secure the legal and economic framework for the liberalisation and opening up of the telecom market – firstly within the market segment of broadcasting, then mobile telephony, Internet services, cable distribution systems, and finally, broadband access. The shared efforts of telecom operators, the regulator and the state have resulted in the telecom market now being regulated and characterised by the predictability of business environment, high investment rates, high quality services and lower prices, market entry of foreign operators, higher degree of availability and use of telecom services, and lastly, higher revenues for operators and higher tax rates for the state. All of the aforementioned would have been virtually impossible had it not been for

the appropriate regulatory framework, which had been efficiently implemented within the telecom sector.

Furthermore, this paper will focus on each telecom market segment separately. Within the domain of fixed networks, the socialist heritage is evident in the low prices of fixed telephony, which are still below the price threshold. Consequently, as these prices considerably lag behind those of the neighbouring countries, there has not been much interest in the initiation of the liberalisation and privatisation process within this telecom market segment. Given the circumstances, one license has been granted to an operator for fixed broadband network services with national coverage and one license for the provision of services via CDMA technology.

The liberalisation of the mobile telephony market segment was completed in 2006. This particular year was especially favourable as UTMS, a new technology in this field, had created the conditions for the introduction of new services and significantly increased the number of users. Moreover, these circumstances led to the issuance of new licenses and the privatisation of one part of the state property within this market segment. Three mobile telephony operators have introduced new, up-to-date and user-driven services and established absolute market competition where national traffic service prices rated among the lowest in Europe.

Prior to the implementation of the new regulation, the Internet market segment had been undeveloped and had a relatively small number of users. The new regulation ensured the commercial provision of lower bit-rate Internet service in the 2.4 GHz band, whereas the issuance of authorisations for international interconnection regulated the issue of wholesale Internet. A considerable decrease in prices of international traffic services became evident following the issuance of authorisations for VoIP services. Moreover, the technical conditions for the roll-out of next generation broadband networks were fulfilled as well.

Taking into account the regulatory, technological, as well as economic aspects, the development of broadcasting and cable distribution systems will be analysed under separate heading within this paper.

Additionally, the economic effects of market liberalisation, reflected primarily in the transparency of the market, its openness and attractiveness for new investors, along with the scope of realized investments which constitute some of the largest investments in the Serbian economy, the increase in the number of users, the constant enhancement of the quality of services, as well as the price-cuts are likely to secure the conditions for a competitive business environment. All these trends which appear in the Serbian telecoms market are monitored in the comparative overviews of the key telecom market indicators. The concluding remarks of the paper offer an insightful illustration of the further development of the telecom market, drawing special attention to the challenges that may lie ahead, especially within the context of the new Electronic Communications Law, which has been harmonised with the 2003 and 2007 EU Directives.

INTRODUCTION

The development of new technologies and the telecom market liberalization led to the growth in mobile telephony, followed by the mounting usages of Internet and hence the increasing use of data transfer telecommunications networks as opposed to voice transfer. The investments in the telecom infrastructure have become the prime mover of economic progress. Therefore, modern communications have become the foundation for building a modern society. The principal features of the modern society are the globalization of the world economy, rapid technological development and ICT convergence. Therefore, the development of telecommunications has become a basic need of the modern society and the measure of its economic, social and cultural development. Modern communications are characterized by a variety of services (telephone, fax, Internet, radio, TV, multimedia, e-services, etc.) available to users. The sudden technological development has enabled the convergence of these services making them available through a single data transfer system. Also, the globalization of telecommunications has made available all services all the time, to all. Telecom services have thus become a part of everyday life in a developed information society.

The globalization and convergence of telecommunications have brought about increased user demands concerning new services, higher quality and, indeed, lower prices. In addition to new technological development of ICTs, it also involves specific market relations, first and foremost the enforcement of demonopolization, liberalization, openness and level playing field. Under the umbrella of the World Trade Organization, it has been agreed, and made mandatory for the EU countries, that the member countries support and implement the necessary procedures in order to establish an open level-playing market regardless of ownership issues or technology deployed. The process entailed the adoption of modern telecom regulations and involved several phases [2]. The principal intentions of the regulations are to enable further development of telecommunications by creating a liberal, open and level-playing market. They stipulate that the same rules be applied to all market players. Furthermore, the operator with significant market power needs to be identified and the requirements to be fulfilled by the SMP need to be defined in order to establish a competitive environment as soon as possible.

The implementation of these regulations have led in most of the EU member countries to a regulated, liberal, open and level-playing market during the 1990s. The major change brought about by the regulations concerns the government authorities in charge of telecommunications. In addition to legislative, juridical and executive power, a new type of power has been introduced in the form of the regulatory authority, which is a national, independent and autonomous institution. This regulation procedure, which had been carried out and coordinated by the European Commission in the EU, has transformed the telecommunications sector from a monopolistic into an open and regulated sector where the prices are market oriented for the first time.

The term independent primarily refers to independency in decision-making such as to prevent any form of direct or indirect political pressure or pressure by any telecom market player. An independent regulatory authority contributes to the investors' trust in the impartiality of the regulator and legal certainty for their business operations. This encourages additional investments by market players and hence promotes further market development. When the majority share-holder of the incumbent operator is the state, the lack of genuine independency of the regulator raises concerns with the investors, new entrants in particular, that the government may influence the regulator and the incumbent operator, discriminating in favour of the state-owned operator. Such discrimination may be reflected in the interconnection prices that are unfavourable to the new entrants, in the frequency and number allocation that favours the incumbent, or in the delay of tariff rebalance while keeping the local calls charges artificially low [7]. This is why many countries needed to establish an independent regulatory body.

Unlike the USA model, where the Federal Communications Commission is in charge of the telecommunications service and network regulation throughout the territory of the country, the European Union does not have a supranational regulatory institution. The regulation in the EU is carried out through a complex relationship between a number of institutions. It is based on the principle of conferred competences, whereby the telecom sector is regulated by each Member State so as to address the national peculiarities. Taking into account the national peculiarities, The European Commission harmonizes the regulations through agreements, decisions, recommendations and other legal documents. They are subsequently implemented in the national legislations, whereas the market regulation is carried out by the national regulatory authorities. The twofold responsibility calls for a close cooperation of all players in the single European market.

The National Parliament of the Republic of Serbia adopted the Law on Telecommunications [11] harmonized with the 1998 EU recommendations, as part of the strategic goal of joining the European Union. The Republic Telecommunications Agency (RATEL) was established in 2005 as a national regulatory authority and an independent legal entity with a task to ensure an efficient implementation and enhancement of the formulated telecommunications policy in the Republic of Serbia, in order to foster further telecommunications development and create conditions for the implementation of the information society. By establishing RATEL, a new regulatory framework was put in place serving as the basis for the reform in the telecommunications sector. Since the very beginning of its work in 2005, RATEL has aimed its activities primarily at applying the principles laid down by the Law on Telecommunications, with the task of regulating the relations according to the principles of free and open market, prevention of monopolistic behaviour, equality and non-discriminatory position of all market players, the goal being the introduction of new services of higher quality and user protection. The Law on Electronic Communications [12] passed in 2010, has provided RATEL, now operating under the name of the Republic Agency for Electronic Communications, with new mechanisms facilitating the protection of market competition and a more efficient control over the operators' work, leading to end-user protection.

2. TELECOMMUNICATIONS MARKET REGULATION

Telecom market regulation is a rather significant segment for further development of telecommunications sector. Since the market growth depends on the number of new services and the number of users, it is evident that there are a number of necessary regulatory requirements that can ensure the market growth. Firstly, the issues of monopoly and market liberalization need to be tackled. Dependence on only one operator, regardless of the size or dominance, can prove to be rather risky in the globalized world. In this regard, the market needs to be regulated in order to prevent its collapsing due to bad and erroneous policy of the monopolist. If there is only one operator in the market, the development strategy of that operator becomes also the development strategy for the whole market, therefore if the strategy is bad it will affect the country's economy [3]. This is particularly important in a fast changing market such as the telecommunications one, where the technology is changing much faster than in other economic sectors.

Another reason for regulation implementation is the introduction of efficient competition. The problem of competition in the telecom market is rather specific because of natural monopoly especially in the segment of cable infrastructure, which cannot be duplicated. In the broader sense, the telecom market can be divided into telecommunications equipment, networks and services markets. The equipment market was the first to be liberalized due to increased number of manufacturers of terminal and network equipment. The service market was the next to be liberalized as the number of providers was growing. The most difficult process is the liberalization of the infrastructure market. It is not enough to attract new operators to the market. The operators need to be truly competitive and to compete against each other, by applying legitimate practices, in order to keep the old and attract new users [5]. The case where there are several small and one big operator deserves particular attention. The regulator needs to identify the operator with significant market power (SMP) and enforce necessary regulatory measures on the SMP operator. The regulator has an important role even if real competition exists, since it controls whether the methods applied to attract new users and keep the old ones are in line with the regulatory rules.

Additional reason for effective regulation is the protection of end-user interests. It is up to the regulatory authority to ensure transparent conditions applied to telecom service provision, and to provide protection in case the set conditions are not respected. Additionally, the competition itself leads to price cuts and better QoS, however in practice it is often subject to manipulation to the disadvantage of end-users. Therefore, it is necessary to administer constant control over the operators and monitor the conditions under which the services are provided to end-users.

The last but not the least important reason is the network and service availability. It is one of the major prerequisites for telecom market development. Since the market is characterized by natural monopoly, once the infrastructure has been laid out, for instance on the local loop, it is unreasonable and extremely expensive to lay out another set of cables. Therefore, mechanisms enabling all interested operators to access the infrastructure under non-discriminatory conditions need to be in place. In this way, not only is the market competition achieved, but also the infrastructure capacity utilization rate is increased thus increasing the return on investments made.

All the above given reasons for an efficient telecommunications regulations have led to establishing regulatory bodies all around the world and their activities have a direct impact on the further development of this sector.

The beginning of the liberalization process and the elimination of the monopoly in the telecom sector in the Republic of Serbia, as well as the beginning of the harmonization with the sector-specific EU regulatory framework, was marked by the adoption of the Law on Telecommunications in 2003 based on the 1998 EU Regulatory Framework. The principles underlying the regulation of relations in the telecom sector are, among other, the promotion of competition, cost-effective and efficient business performance, protection of user interests, the high quality of telecom services, rational and efficient RF spectrum utilization, interconnection under equal terms and observance of the international norms and standards.

The Law on Telecommunications defines the division of competence between the Government of the Republic of Serbia, the Ministry of Telecommunications and Information Society and RATEL, thus providing the separation between the political, operational and regulatory functions. The necessary condition for the implementation of the Law was the establishing of RATEL as an independent regulatory authority, which began with its work in late 2005.

RATEL's principal task is to create the conditions for an open market and equal treatment of all market players in the telecoms sector. The results of its work so far can be seen in the area of mobile and fixed networks development and higher utilization rate of the Internet and cable distribution systems, greater competition in the telecom services market, as well as in the segment of the switchover to digital television.

In the area of mobile telecommunications, the adoption and implementation of regulations have led to market opening. The tree mobile operators have provided new, modern services to the users. All three operators are providing data transmission services, i.e. Internet service, through a newly built UMTS network (3G). The competition has led to better quality of services and a price-cut. In order to further stimulate the game in this field, RATEL adopted the Rules on Number Portability, which should enable the mobile telephony users to change the operator while maintaining the same number.

With the adoption of bylaws pertinent to particular segments of fixed telecom services, the conditions for technological neutrality and the entry of new operators in the fixed telecom marked have been provided. In mid 2009, public bidding was carried out and two public fixed wireless access (FWA) licences were issued for voice service and lower bit-rate Internet service provision using the CDMA technology in the 420 MHz frequency band. It involves national coverage and should enable service provision in less developed areas and solve the problem of party-lines.

In late 2008 and early 2009, RATEL adopted a number of rules regulating the requirements for the issuance of authorization for Internet and other data transmission services provision. Commercial provision of lower bit-rate Internet service in free band was enabled. The fact that the problem concerning the Internet wholesale was solved at the same time is of particular importance. The conditions and the procedure for the issuance of international interconnection authorization and network authorization were laid down. Also, the issuance of authorization for voice over IP service provision, whereby international calls at significantly lower prices were enabled.

In order to solve the problem of monopoly present in the cable distribution systems, the newly adopted regulations have provided conditions that enable the application of new technologies, which provide direct-to-home (DTH) services via satellite transmission. Also, a set of technical regulations that should bring about technical improvements of these networks has been adopted.

Another rather important result of RATEL's work in 2009 is the adoption of bylaws enabling the licence issuance procedure for second operator to provide fixed broadband network with national coverage. In this way the conditions were established for further telecom market liberalization. Two operators will be able to offer bundle packages, as laid down under the Strategy for the Broadband Development, with minimum bit-rate of 4 Mb/s for the fixed, and 512 kb/s for mobile access. It is expected that the service packages offered by these operators will result in lower prices for end users in 2011.

In the broadcasting area, due to its activities related to international coordination, RATEL has created conditions for the beginning of the digital switchover with the possibility of broadcasting 48 different television programmes with national coverage. For further reference see [8]-[10], [16].

3. The Results of Telecommunications Market Regulation According to Services

The activities in the field of telecom market regulation have quickly resulted in specific financial performances. We shall briefly look back at the most important ones. According to the comparative data gathered by RATEL through a continuous telecom market analysis, the revenue from the telecom services in 2009 amounted to over 1.5 billion euros. Also, since the beginning of institutional telecom market regulation, the average annual revenue growth rate in the telecoms sector has been 13%. This data further gains on significance if we consider that the telecommunications have soon become the fastest growing sector in the Serbian economy. At this point, it is important to observe that the share of the revenues from the telecommunications in the gross domestic product has been almost 5% for the past three years. In the period covered by this text the telecoms sector has seen fixed investments of over 2 billion euros, while the Serbian economy is still to sell of the biggest domestic telecommunications operator in early 2011, which should significantly increase the investments.

More detailed information concerning the revenues and investments made in the telecoms sector during the past 5 years is available in the following two tables.

Table 1. Revenues from the telecommunications services in the Republic of Serbia in millions of euros [4]

Year	2005	2006	2007	2008	2009
Revenue	927	1300	1470	1611	1503
Share in GDP	4.6%	5.6%	4.7%	4.87%	4.76%

Table 2. Investments in the telecommunications sector in the Republic of Serbia in millions of euros [4]

Year	2005	2006	2007	2008	2009
Investments	350	190	899	362	284

During 2009, the revenues from mobile telephony had the biggest share in the total telecommunications services, amounting to around 55%, whereas VoIP services had the smallest share of only 0.04%. Also, it can be concluded that the structure of fixed investments in that year followed a similar pattern. Namely, the investments made in the mobile telephony reached as much as 52% of the total investments made, whereas the investments in VoIP accounted for barely 1% of investments. The structures of revenues and investments in the telecom sector are given below (see Figure 1 and Figure 2).

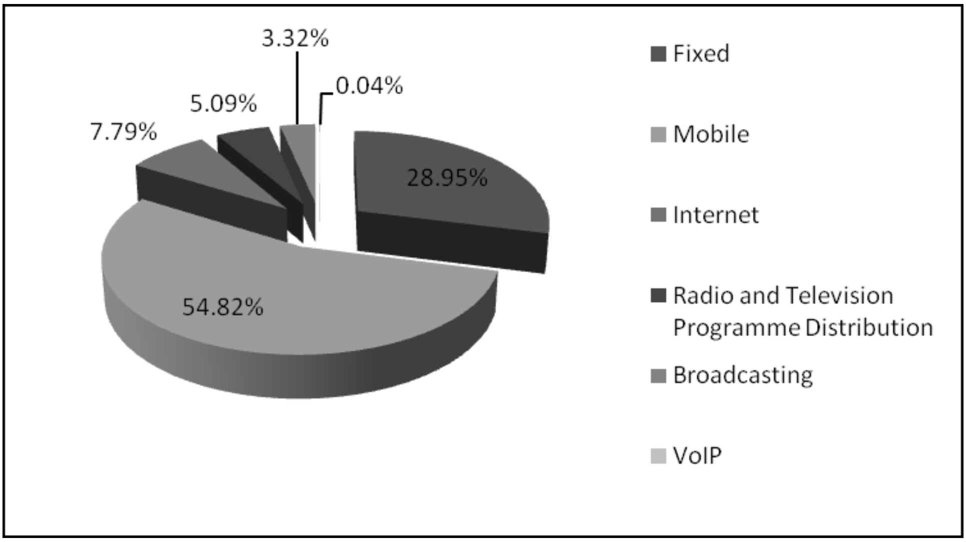


Figure 1. Revenue structure for 2009 by services [4].

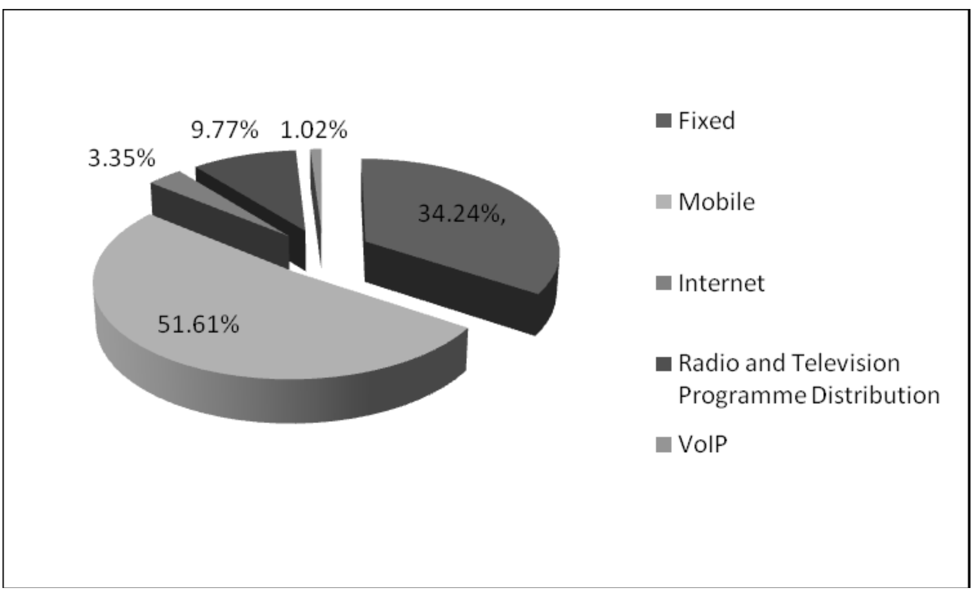


Figure 2. Investment structure by services in 2009 [4].

The estimates of the IBRD and the World Bank suggest that for every 10% increase in the investments made in the development of modern telecommunications, we can expect an average of 1.4% growth in the economic development of the society [3].

An overview of the development in the telecommunications market in the Republic of Serbia in the previous several years will be given further in this chapter, through the following segments:

- Public fixed telecommunications networks and services,
- Public mobile telecommunications networks and services,
- Internet services,
- Radio and television programme distribution services,
- Broadcasting.

3.1. Public Fixed Telecommunications Networks and Services

Public fixed telecommunications networks and services involve the following types of services provided via fixed network:

- Local calls (origination and termination) – telephone service providing telephone connection between two users in the same code/network area via switching system(s);
- National calls (origination and termination) – telephone service providing telephone connection between two users in different code/network area via switching or transmission systems;
- International calls (origination and termination) – telephone service providing telephone connection between two users, where one is a user of a national operator in the Republic of Serbia and the other a user of a foreign operator, via switching and transmission systems;
- Calls to value added services – service offered by operators to users, making available value added services (by dialling numbers from the special numbering reserved for this type of service), which are different premium rate services or voting;
- Special services (telephone directory, abbreviated dialling, call-unit count device, call diverting, conference calls, etc.);
- Calls to emergency numbers – police, medical assistance and fire department.

The development and activities of the public fixed telecommunications network and services in the previous period were marked by the activities of the incumbent operator Telekom Srbija Joint Stock co., which had been the only public fixed telecommunications service operator for a long time. Since 2003, Telekom Srbija has been in the ownership of two shareholders: Public Company of PTT traffic “Srbija” (80%) and OTE from Greece (20%). Telekom Srbija is a large company / a closed joint stock company with more than 10 000 employees and annual revenues of approximately one billion euros, encompassing four dependent legal entities. In addition to fixed telephony service provision, Telekom Srbija also provides mobile telephony, internet, RTV programme distribution, interconnection and leased lines services. Telekom Srbija is the owner of complete cable duct network and telecommunications infrastructure in the whole territory of the Republic of Serbia.

The Government of the Republic of Serbia launched a tender for the sale of the 51% of the capital in Telekom Srbija company on the 20 October 2010 and the procedure is underway. The effects of the privatisation of the incumbent telecommunications operator are analysed in detail in the reference [6]. In the situation where the market has practically been liberalised by regulatory measures, it is necessary for the incumbent operator to increase significantly the business efficiency, raise the productivity per employee, stimulate innovation and use new technologies. This requires considerable investments but the incumbent operators, especially in countries in transition, lack necessary financial resources. The existing tariff system is a particularly challenging issue, since charges for subscription and local calls in fixed network are extremely low, whereas the international call charges are unreasonably high.

Telekom Srbija held a monopoly for years in the area of fixed telephony, Internet wholesale, interconnection and leased lines. Specifically, the incumbent operator used the market advantage and the infrastructure to impose prices that were not based on real costs or market situation. Therefore, before the adoption of the Rules on Interconnection in 2008 only a small number of interconnection agreements between the telecom operators had existed, which has now significantly changed. Also, the introduction of VoIP operators in the Serbian market has enabled international calls at considerably lower prices.

The issue of monopoly in this telecom segment could only be solved only by introducing new operators in the market. Therefore, in May 2009 RATEL launched the public bidding procedure for the issuance of two licences for public fixed wireless telecommunications network (FWA) in the 411.875-418.125/421.875-428.125 MHz frequency bands and voice services, data package transmission and simultaneous voice and data transmission for the territory of Serbia, using CDMA technology. After the public bidding procedure had been carried out, Telekom Srbija and Orion telekom Ltd. were pronounced the best bidders, while the amount achieved through bidding was 540 000 euros, paid by each of the winning bidders. The operators Telekom Srbija and Orion telekom were awarded licences in June 2009, with the obligation to begin with the commercial service provision within six months. In addition to the bidding procedure, in October 2009, the Ministry of Telecommunications and Information Society adopted the Rules on the number of licences for public fixed telecommunications networks and services and the period for which the Licence is issued, minimum conditions for Licence issuance and minimum amount of the one-off Licence issuance fee. Pursuant to these Rules, RATEL launched the public bidding procedure for the issuance of one licence for public fixed telecommunications network and services for the territory of Republic of Serbia. The procedure was completed in January 2010 and the licence for public fixed telecommunications network and services was awarded to “Telenor” d.o.o, with the obligation to start with the public fixed service provision within 12 months. In this way, the monopoly in the retail segment has been fully eliminated. The wholesale market regulation is laid down under the Law on Electronic Communications and, according to RATEL’s plan it should be implemented during 2011.

Fixed network traffic in billions of minutes during the previous years is shown in Table 3 below. As can be observed, this traffic has been stagnating, which is in line with the tendency of the EU countries. The trend is mainly characterized by the traffic transfer to mobile telephony and the Internet.

Table 3. Fixed network traffic in billions of minutes [4]

Year	2004	2005	2006	2007	2008	2009
Traffic	16.7	17.3	15.8	12.9	15.9	12.6

The number of users of fixed networks in the Republic of Serbia in the past 7 years in constant growth as shown in Table 4. The absolute growth is not large; therefore stagnation can be expected in the following years.

Table 4. Number of fixed telephony users [4]

Year	2003	2004	2005	2006	2007	2008	2009
Number of users	2,332,432	2,481,457	2,527,328	2,719,400	2.854,500	3,084,870	3,145,920

The fixed network penetration rate, as measured according to the number of inhabitants in the past 7 years has been, is given in Table 5. The penetration rate in regard to the number of households is 125%, which is a rather significant data for the further development of telecommunications in the Republic of Serbia.

Table 5. Fixed telephony penetration rate [4]

Year	2003	2004	2005	2006	2007	2008	2009
Penetration	31.1%	33.1%	33.7%	36.3%	38%	41.14%	41.96%

The quality of service in the fixed telephone network has been improving over years. Table 6 shows the number of pending requests for new telephone lines in the past years.

Table 6. Number of pending requests for fixed lines [4]

Year	2003	2004	2005	2006	2007	2008	2009
Pending requests	313499	284048	419405	428576	350685	314628	249486

From a rather small percentage of digital exchanges in the beginning of this century (80%) the digitalization rate has grown to 97%. The number of party-lines, which amounted to 390 thousand in 2005, became three times less over five years, amounting to 143 thousand today. With the introduction of CDMA technology, this number is expected to drop significantly.

The revenue from the fixed network is given in Table 7. The compound annual growth rate (CAGR) in the past 7 years is only 2.8%, measured in the national currency, which brings us to the conclusion that the existing services in this segment will not produce growth in revenues and could even stagnate. The reason lies in the transfer of traffic from the fixed to other types of network. If the revenues are measured in euros, the situation is even worse due to negative trend in exchange rates.

The investments in the telephone service are shown in Table 8. It should be noted that the decrease in investments that can be observed from the table is the result of a number of factors. First of all, the fixed telephony sector is becoming less significant due to the transfer of traffic to mobile networks and the Internet. Also, the incumbent operator had no incentive to invest in the telecommunications network development, since the extremely low prices of fixed telephony could hardly cover the costs. Finally, the absence of competition results in poor investments, since Telekom Srbija felt no pressure to increase the quality of services or make them more available especially in rural areas.

Table 7. Revenue from the fixed telephony services [4]

Year	2003	2004	2005	2006	2007	2008	2009
Revenue in bn. RSD	28.764	30.604	31.624	35.070	33.146	33.827	34.913
Revenue in bn. EUR	443	405	381	416	414	413	371

Table 8. Investments in the fixed network in millions of euros [4]

Year	2003	2004	2005	2006	2007	2008	2009
Investments	69	222	81	59	70	65	51

Taking into account the fact that Telekom Srbija was the only public fixed telecommunications service operator, in March 2006, according to the Law on Telecommunications [11], RATEL declared the company Telekom Srbija the operator with significant market power. In this regard a procedure for accounting separation of costs, revenues and fixed investments for operators with significant market power was initiated. Once the appropriate model had been developed in less than three years, the application of cost-based principle began as the only valid basis of the special tariff regime for the public fixed telephony services. In other words the obligation was imposed on the SMP operator to apply the cost-based model in the telecommunications services pricing, subject to subsequent verification by the regulatory authority i.e. RATEL. Special tariff regime for SMP operators was laid down in accordance with the Rules on the application of the cost-accounting principle, separate accounts and reporting of telecommunications operator with significant market power (*Official Gazette of RS*, number 103/08), regulating the basic principles, models and methodology of cost and performance calculation, calculation of cost price and sale prices for the services of the SMP operator. Essentially, the Rules stipulate “cost plus” method or average costs of services increased by appropriate return on assets employed in service production and sales. The Rules stipulate the application of Historical Cost Accounting (HCA) according to top-down method, based on the functional principle of Fully Distributed Cost (FDC) or Activity Based Costing – (ABC). HCA has been implemented in the process of sales pricing for the services of the SMP operator for the past two years, while RATEL’s preparations for introducing a more advanced model, namely Current Cost Accounting (CCA), are underway. The implementation of Long Run Incremental Cost (LRIC) will begin after the successful application of CCA model by RATEL and the SMP operator, which requires 3-5 years according to estimates.

Monthly subscription without VAT for analogue telephone connection in 2005 was 0.53 euros for residential users (natural entities) with 150 call-units free of charge, and 0.53 euros for business users (legal entities) with no free call-units. The price of a 3-minute local call in peak traffic was 0.63 eurocents. This price was extremely low and was maintained as such due to inherited political reasons. Obviously, fixed phone in the Republic of Serbia had been considered a social category for a long time. Therefore, a tariff rebalance needed to be carried out, since it was evident that Telekom Srbija was bearing losses in this business segment. On the other hand, the total profitability of the firm was kept at a high level, leading to a conclusion that the losses in the fixed segment were covered from other revenues of the operator, mainly from the mobile telephony, international calls and wholesale services. According to the regulatory obligation, imposed by the regulator in order to make the work of other operators using the infrastructure and wholesale services of Telecom Srbija easier, the retail prices had to reflect the real costs. As a result, the retail prices of services have been corrected several times during the past 5 years in accordance with the Rules on application of cost-based principle. The current price of monthly subscription charge for analogue telephone connection without VAT is 3.68 euros for residential users with 150 call-units free of charge, and also 3.68 euros for business users with no free call-units. The price of a 3-minute local call in peak traffic was 1.15 eurocents. Finally, it should be noted that currently the prices are significantly lower compared with other countries in the region, as can be seen from Figure 3. Therefore, further harmonization can be expected once Telekom Srbija is privatised.

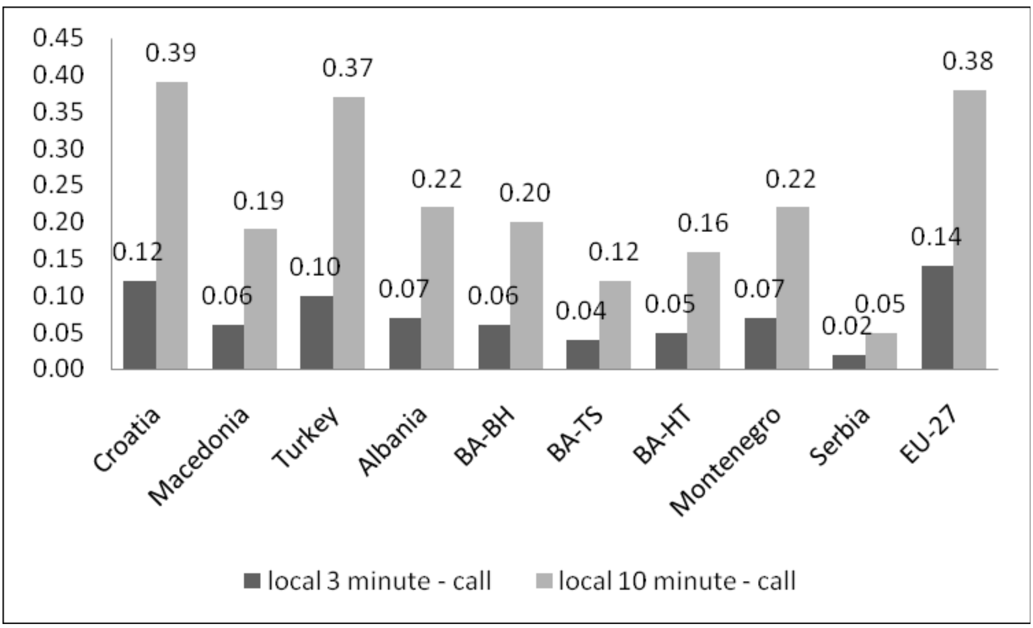


Figure 3. Price of a 3-minute and a 10-minute Local Call (€) (VAT included) [13].

The problem of disproportion between the prices for local and international calls needs to be addressed. Namely, the prices of these two types of traffic are disproportionate and do not produce the same return on invested capital, which is an unfeasible practice. Therefore, the prices need to be adjusted so as to provide identical yield to the operator. However, it should be kept in mind that the price structure of the incumbent operator is usually conditioned by social policy. The governments that controlled the incumbent operator thought that it would be politically correct to subsidize the local calls in order to provide affordable access to the fixed services to everyone. It should also be noted that the subsidies were usually generated from the boosted revenues from the international traffic and leased lines. Consequently, disproportionate prices of services have a number of negative effects on the economic efficiency of the market. Firstly, they are seen as a warning or even discouragement to new investors. Secondly, high prices of some services will lead to discouraging the usage of those services, whereas the services offered at extremely low price will be used more than necessary due to the low price. This creates a market imbalance that is not likely to foster market liberalization [7].

Telekom Srbija remains, however, the dominant operator in the fixed telephony market since it owns the infrastructure which is hard to duplicate, but it can be rather quickly modified at low cost to provide other services. Its dominance will continue in the following period, especially in the area of access networks and national calls provision. There is a possibility of competition, but it will certainly not be enough to modify the power distribution in the fixed market anytime soon. Significant competition can be expected only when other fixed operators, landline and wireless and VoIP service providers conquer the market.

The new Law on Electronic Communications [12] in addition to relevant market of fixed telephony, also identifies the wholesale market as relevant, so that along with the retail tariff rebalance price regulation for wholesale services of the dominant operator is also required. Furthermore, the new Law defines the markets related to the telephone service and telecommunications network, specifically retail PSTN access market, wholesale PSTN call origination market, wholesale PSTN call termination market, wholesale market (physical) access to network elements and accompanying assets (including shared access and fully unbundled access to local loop), wholesale broadband market and wholesale leased lines market.

The regulation of these markets will lead to fulfilment of conditions for further development of telecommunications in the fixed network segment. The analysis of relevant markets began in late 2010 and the first results can be expected in mid 2011. More details about the analyses procedures can be found in the reference [2].

3.2. Public Mobile Telecommunications Network and Services

The mobile telephony has been present in the Republic of Serbia since 1994 with the establishment of Mobtel company. When once the GSM technology was introduced the usage of mobile services became more diffused. By 2006, there were two operators with the licence for public mobile telecommunications network service provision, Telecom Srbija and Mobtel, both in majority state ownership. In order to provide real competition, the Government of the Republic of Serbia launched a tender for the sale of Mobi63 company (which took over the infrastructure and the employees of Mobtel), leading to the purchase by the Norwegian operator Telenor in July 2006 for 1.513 billion euros. The price included the GSM/UMTS licence issued by RATEL for a period of 10 years with the possibility of 10 year-extension without additional charge. This procedure is the biggest direct investment in the territory of Republic of Serbia. The licence holder is required to begin with the commercial provision of UMTS service within 6 months, to cover 25% of population within 2 years and 60% of population within 5 years following the licence issuance.

During the same year RATEL launched the tender for the third licence, won by Mobilkom Austria in a public bidding procedure for 320 million euros. The licence was issued for the territory of Republic of Serbia for a period of 10 years with the possibility of 10 year-extension. The licence holder is required to begin with the commercial provision of GSM service within 6 months, to cover 20% of population and main roads within 12 months, 50% of population within 24 months and 80% of population and 90% of the territory within 4 years following the licence issuance. In addition to 320 million euros paid for the licence, during 2008, Mobilkom Austria made considerable investments in the development of infrastructure and employed a large number of professionals, thus making the biggest greenfield investment in Serbia so far. Mobilkom Austria became operational in July 2007, under the name of VIP Mobile d.o.o.

The following operators have been present in the public mobile telecommunications network market in the Republic of Serbia since 2006:

- Telecommunications company Telekom Srbija - Mobilna telefonija Srbije MTS, owned by Public company for PTT Traffic Srbija (80%) and OTE, Greece (20%) (licence replaced on 01.08.2006).
- Telenor Belgrade, 100% owned by Sonofon, Denmark, (licence issued on 01.09.2006).
- Vip mobile, owned by Telekom Austria Group, Austria (licence issued on 01.12.2006).

All three operators were granted licences for public mobile telecommunications networks and public mobile telecommunications network services in accordance with GSM/GSM1800 and UMTS/IMT-2000 standards, issued by the Republic Agency for Electronic Communications.

Apart from the basic voice transmission service via public mobile telecommunications network, the mobile operators in the Republic of Serbia offer a wide range of additional services to users. These services include the following: voice mail, call divert, call waiting, conference call, SMS text messaging, data transmission, incoming call identification, hidden identification, regular monthly itemized bill, connection/disconnection upon request, change of tariff package, replacement of a damaged or lost SIM card, WAP, MMS service, etc. The introduction of the 3G network meant launching new services: real time video calls, video streaming, video clips, etc.

Table 9 shows the number of base stations for all three operators in the past 5 years. As can be noted, this period was marked by large investment cycles for all three operators, resulting in as many as 3279 new base stations on the territory of Republic of Serbia built in the last two years.

Table 9. The number of base stations of the mobile operators [4]

Year	2005	2006	2007	2008	2009
Telekom Srbija	770	1085	1274	1798	2041
Telenor	831	879	1280	1820	2703
VIP	-	-	173	727	1262

Since 2005, mobile telephony has been the most profitable telecommunications branch in Serbia. Its share in the total business volume has grown from 45% in 2005 to 55% at the end of 2009. Also, up to the beginning of the great economic crises, the revenues from the mobile telephony have shown remarkable growth, as can be observed in the Figure 4 below. Furthermore, the average annual growth rate of the revenues from the mobile telephony in the past 8 years has reached 15.8%. The revenues from the mobile market in the Republic of Serbia in the past years are given in Figure 4.

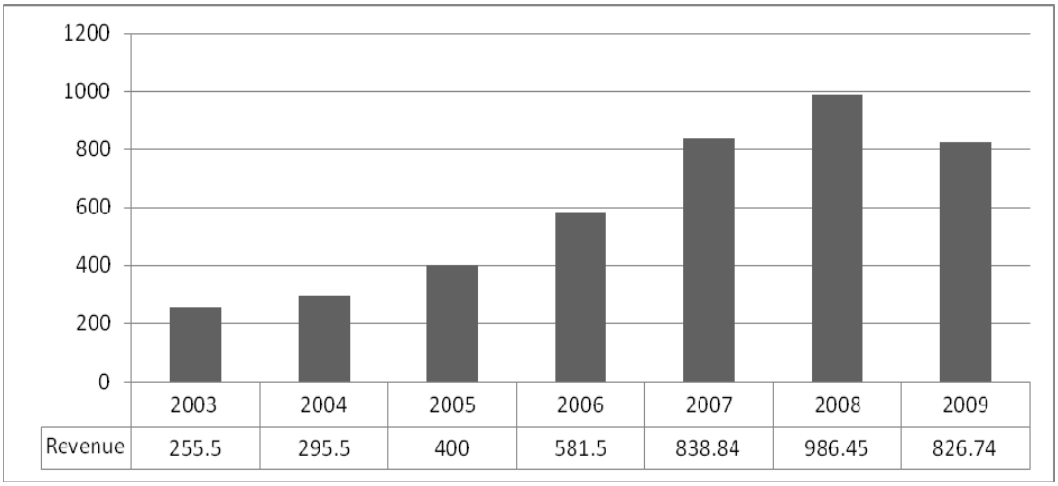


Figure 4. The revenues from the mobile services in millions of euros [4].

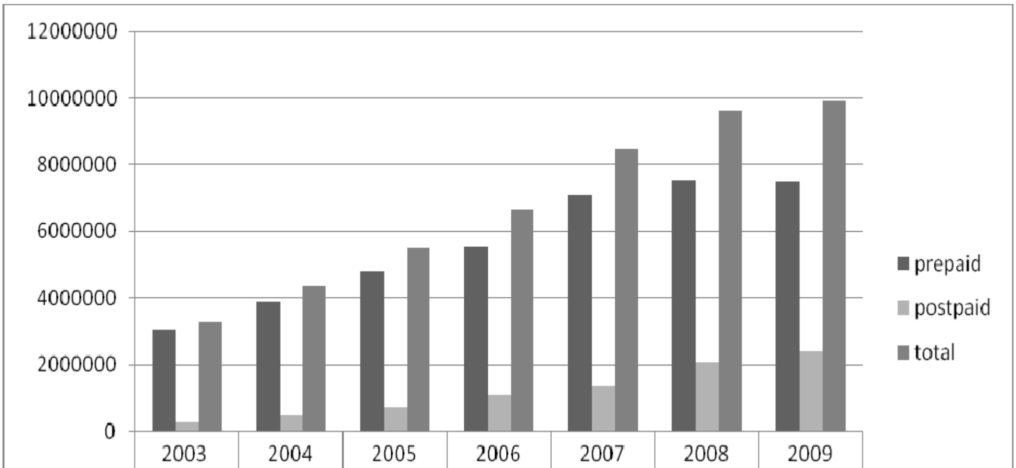
It should be noted that the decrease in the revenues from the mobile telephony in 2009 is, to a certain extent, a consequence of calculation, i.e. the difference in exchange rates, since in 2009, euro had the average annual value of approximately 94 dinars, compared to 82 dinars in 2008. When expressed in RSD currency, the total revenues from mobile telephony service amounted to 77.8 billion dinars, which in an increase of about 4% in respect to the previous year. As one of the temporary measures aimed at diminishing the effects of the World Economic Crisis, the Government of the Republic of Serbia introduced an additional tax rate in the mobile telephony sector on 1 June 2009. The 10% tax applies to all calls, standard SMS and MMS messages, data transmission and additional services in the country and abroad, while this tax is not included in the calculation of VAT. Such a measure considerably decelerated the growth of the mobile market, even though the operators continued to compete trying to win over the users by cutting down on prices and introducing a wide range of new services, such as SMS directory, favourite numbers, money transfer services, tickets purchasing via mobile phone, downloading music from WAP portals, combination of prepaid and postpaid packages, Facebook profile SMS notification, etc.

The economic crises and the 10% tax have also had significant impact on the cuts in fixed investments compared with 2007, which was the peak year in the investment cycle, marked by the entrance of two global mobile operators in the market. Table 10 shows the investments in the mobile network in the past 4 years.

Table 10. Investments in the mobile network in millions of euros [4]

Year	2006	2007	2008	2009
Investments	99	319	241	114

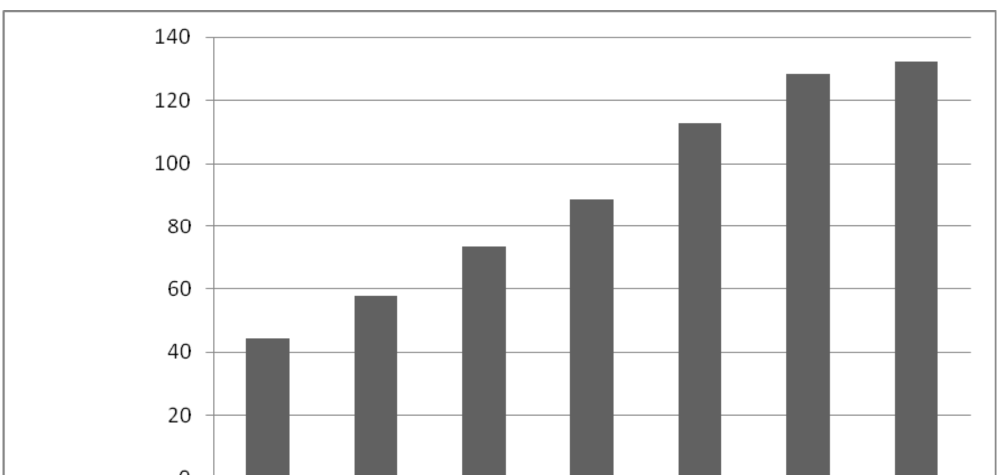
The entrance of new operators, price cuts, increase in the quality of services, rather agressive marketing and affordable technology have enabled the increase in the number of users of mobile telephony in Serbia in the period 2003-2009 at an average annual rate of 20%. Even though the crises was reaching its peak, growth was seen even during 2009. The number of users (prepaid, postpaid and total) is given in Figure 5.



	2003	2004	2005	2006	2007	2008	2009
prepaid	3020000	3865000	4800000	5560000	7070000	7530000	7500000
postpaid	260000	475000	710700	1080000	1380000	2090000	2410000
total	3280000	4340000	5510700	6643722	8452642	9618767	9912339

Figure 5. The number of mobile users in the Republic of Serbia [4].

From a rather low penetration of only 45% in 2003, in 7 years the penetration rate has exceeded 130%, which is a remarkable achievement. Details can be seen in Figure 6.



	2003	2004	2005	2006	2007	2008	2009
Penetration rate	44.52	57.94	73.5	88.6	112.7	128.28	132.2

Figure 6. Mobile penetration rate in percentages [4].

The growth of the number of users was accompanied by the average annual growth of outgoing traffic per user in minutes, as shown in Table 11. The increase in the number of minutes is the result of competitive and affordable prices for national calls, transfer of traffic from fixed to mobile network and high QoS.

Table 11. The average growth of outgoing traffic per user in mobile network in minutes [4]

Year	2004	2005	2006	2007	2008	2009
Outgoing traffic	270	330	390	475	622	828

The total outgoing traffic is given in Figure 7, where a significant growth over the past few years can be observed.

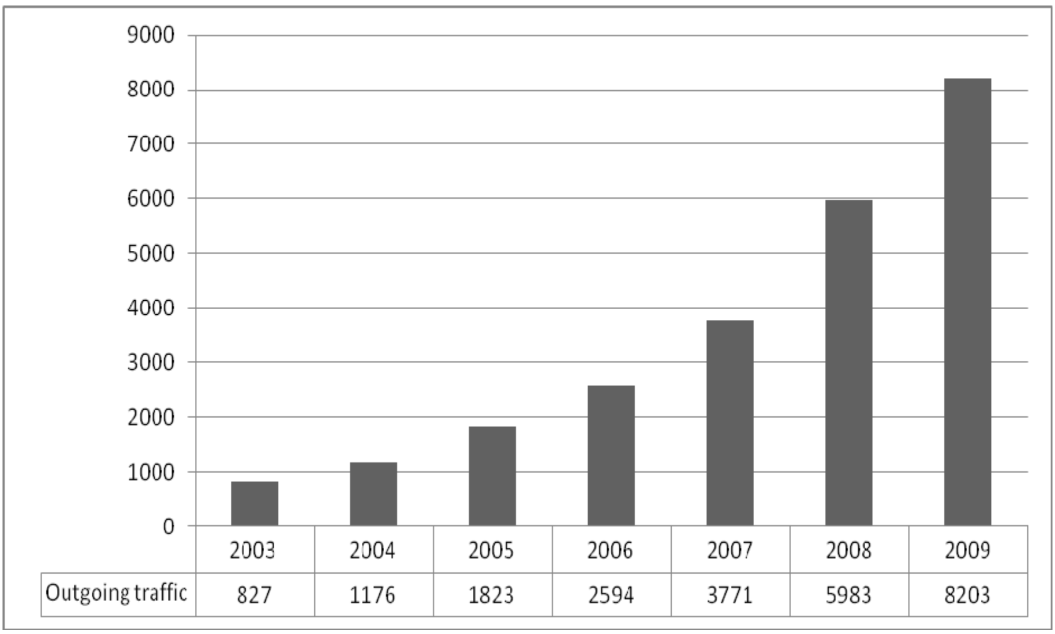


Figure 7. Outgoing traffic in mobile network in millions of minutes [4].

The number of SMS and MMS followed the same trend, as shown in Table 12.

Table 12. The number of SMS and MMS in millions [4]

Year	2003	2004	2005	2006	2007	2008	2009
SMS	1192	1602	2093	3208	4480	7981	9290
MMS	0	5.75	11.5	9.4	16.8	21.6	33.2

The calculation of OECD basket for mobile services led to a conclusion that the market arrived from extremely high prices in 2006 to the prices of national mobile traffic which are among the lowest in Europe, which is further examined in reference [1]. The main findings are as follows. The average low usage basket offered by the operators in Europe for 30 minutes of outgoing traffic and 33 SMSs in line with the OECD standard in 2007 amounted to 12.33€. Three out of four cheapest operators in Europe (among the countries comprised by the analyses) are from Serbia. According to this analysis, low usage basket in Serbia was 5.62€ which is more than 50% below the average price. The average high usage basket offered by the operators in Europe for 30 minutes of outgoing traffic and 33 SMSs in line with the OECD standard in 2007 amounted to 33.36€. Three out of five cheapest operators in Europe (among the countries comprised by the analyses) are from Serbia, as the high usage basket in Serbia was 14.62€ which is over 55% below the European average.

The competition in the Serbian mobile market was measured by the Herfindahl – Hirschman- Index (HHI), which is an indicator used for determining the degree of concentration of a given market and it is defined as the sum of the squares of the market shares whose value ranges from 1 to 10000. It is important to say that smaller index values indicate greater market competition. Table 13 clearly shows that the entry of new operators in the mobile market in the Republic of Serbia has led to a continual increase in competition, which was the ultimate goal.

Table 13. The value of index in the mobile market [4]

Year	2006	2007	2008	2009
HHI index	5332	4759	4684	4520

Finally, it should be noted that there are no structural barriers in the retail market, since every new operator can obtain access to the network of an existing operator in the wholesale market and begin with the service provision in the retail market. This market is characterized by intense competition reflected in price cuts, introduction of new services, a variety of packages offered, free choice of operator thanks to their presence throughout the country and wide GSM signal coverage. Therefore, the market features are such as to indicate further intensification of competition between the operators to be expected in near future. This is especially likely to happen in view of the fact that all three operators are large enterprises, that they are both network operators and service providers, that they are using the same technology and that, soon, compliant with the obligation under the licences, their respective networks will cover the entire population and territory of Serbia thus enabling them to offer the network to a new entrant. Also, the operators are selling their services throughout the territory of Serbia both through their own sales centres and authorized dealers. All operators offer the same basic set of services, the operators compete for end-users through special offers in terms of prices and different promotions, in order to keep the old and attract new customers. The users can choose between different packages and different operators. The mobile market is constantly growing, both in terms of the number of users and in terms of traffic volume and the number of services provided. According to relevant previsions, this trend is expected to continue in the following period. An additional enhancement of competition can be expected with the introduction of number portability in mobile networks, which should enable the mobile users to change the operator while maintaining the same number. The implementation of number portability in Serbia is expected to begin in the first quarter of 2011.

3.3. Internet Services

The retail broadband market in the Republic of Serbia consists of the following segments:

1. ADSL with direct network connection,
2. Access via cable network of the cable distribution operators,
3. Wireless access (Wi-Fi, etc.),
4. Access via mobile network (EDGE, UMTS, HSDPA),
5. Access via leased lines,
6. Access via fibre-optic cables (FTTH).

On the other hand, the wholesale broadband market consists of the following segments:

- Bitstream service provided over copper twisted pair, optical cable or hybrid cable. Access points can be located on:
  - IP level
  - Ethernet level
  - DSLAM/OLT level
- ADSL access via copper twisted pair provided by Telekom Srbija for their own needs.

The retail broadband market had not been regulated until the setting up of the regulatory authority, which soon adopted the Rules on the terms and conditions for the Internet service provision regulating the manner in which these services are to be provided. The rules stipulate basic technical and other requirements necessary for Internet service and other data transmission services provision. On the other hand, the monopoly in the wholesale Internet market was more striking since all providers were forced to use Telekom Srbija's infrastructure in order to reach the end-users. Moreover, Telekom Srbija was the only one with substantial capacities and links connected to the surrounding countries. Until December 2008, it was the only holder of the authorization for the international interconnection (Verat company had the authorization for wireless connection only), which is one of the main reasons for such situation in the market. Once the Rules on terms and conditions and the procedure for the issuance of authorization to a public telecommunications operator for interconnection of national telecommunications network with a telecommunications network of another country (*Official Gazette of* RS, no. 94/08) were adopted in December 2008, conditions were met for opening up the broadband market. So far RATEL has issued 11 authorizations. As for the access, it is expected that the significant market power of Telekom Srbija in this segment will be put under control and subsequently reduced by imposing of the regulatory measures, following the marked analyses. This primarily regards obligation to grant access to the "last mile", by offering local loop unbundling product on market terms. Also, by introducing other technologies, such as cable Internet and wireless systems, providers can access end-users directly.

The Rules on classes of radio stations for which radio station licence is not required (*Official Gazette,* no. 26/07) referring to 2.4 and 5.5 GHz, made it possible for ISPs to freely use these frequencies to provide the Internet of non-guaranteed quality of service. Following the adoption of these Rules, 133 ISPs providing this service have been registered.

The Internet market in the Republic of Serbia was developing rather slowly in the early days. At first, narrowband dial-up access provided by ISPs was available. It was the ADSL technology, offered by ISPs in 2005, that set off a more significant development of the Internet. Until then the number of users was extremely low. In the meantime, other technologies were developed, first the cable modem, which enabled very high bit-rates by the end of 2009 with the introduction DOCSIS 3.0 standard, followed by the wireless access and the Internet. Table 14 provides a chronological overview of the number of users according to different technologies.

Table 14. Number of users according to different technologies [4]

	Dial up	ADSL	Cable modem	Wireless access	Mobile Internet	3G network subscribers	Other
2004	545164	-	8372	2918	-	-	898
2005	708226	9530	23956	1049	-	-	13.914
2006	882611	26126	54598	21968	-	9687	10210
2007	692905	132359	87731	36059	-	257379	2276
2008	397202	267876	151154	48130	25489	738401	1135
2009	252195	351252	187923	45864	100628	762307	5547

Along with the growth of the number of users, the revenues from the Internet services grew as well, as shown in Figure 8. Interestingly, regardless of the economic crises, the revenues continued to grow in 2009, which can be accounted for by a considerable expansion of different technologies offering broadband services.

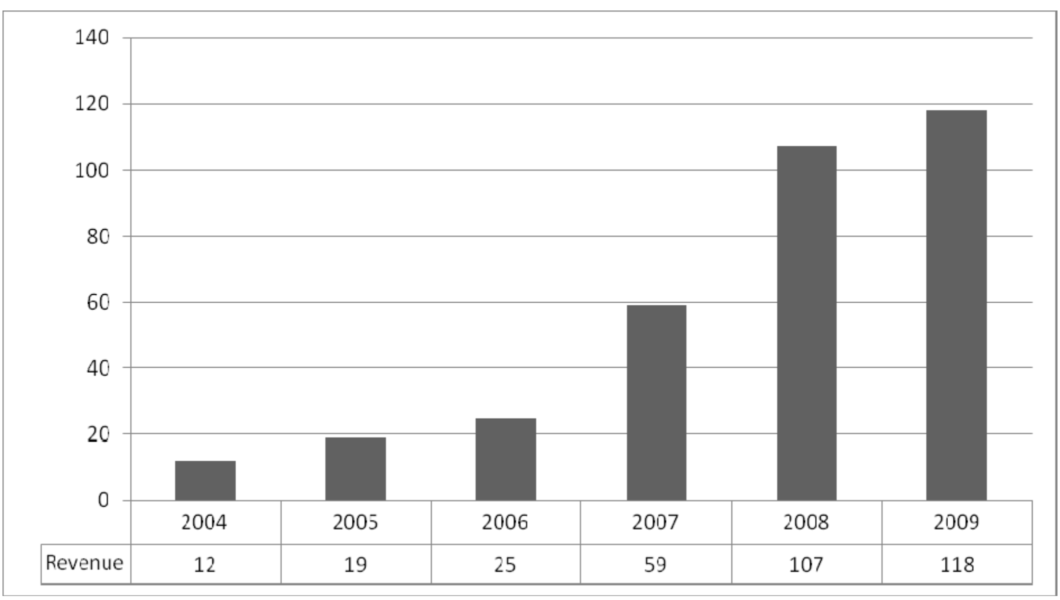


Figure 8. Revenues from the Internet service provision in millions of euros [4].

Once the regulatory measures defining the business environment of ISPs were introduced, the increase in revenues and the increase in the number of users led to the increase in the number of providers as well, as shown in Table 15.

Table 15. Number of ISPs [4]

Year	2004	2005	2006	2007	2008	2009
Number of ISPs	61	66	109	159	197	199

The growth of the Internet sector in Serbia is evident in terms of the total revenues, cumulative number of users (especially broadband access users) and market penetration. Moreover, the choice of access technology revealed a striking tendency towards the use of broadband services which was reflected in the increase of ADSL, cable and mobile Internet access users (the number of mobile Internet access users has quadrupled in respect to 2008), as well as in the continuous decrease in the number of dial-up users (which decreased by 70% over the 2006-2009 period).

Within the services provided to end-users, ISPs offer a variety of bundled services among which the best selling one include permanent Internet connection and upload/download speeds of 1024/128 kbps, accounting for approximately 300,000 connections or about 50% of all broadband connections.

Comparing the amounts of the monthly subscription fees of various Internet packages as well as the structure of Internet connections in 2009 with the ones recorded during previous years, we may observe significant changes. Specifically, the considerable decrease in prices, as a result of increased competition and increasingly demanding end-users, had considerable impact on the improvement of Internet service provision which was evident from the constant rise in the number of high-speed Internet connections. For example, in 2006, the monthly subscription fee for permanent Internet access for 512/128 kbps speeds amounted to approximately 70.5 euros. Moreover, it should be noted that the highest Internet access speed offered by Internet service operators in 2006 was 768/192 kbps. Currently, the monthly charge for the 1024/128 kbps Internet is 11.5 euro, whereas the highest speed commercially offered in the retail market to residential users is 61440/2048.

All operators compete for end-users through special offers in terms of prices and different promotions, in order to keep the old and attract new customers. The users can choose between different packages with the possibility of substitution between ADSL and cable access. The Internet market is constantly growing both in terms of the number of users and traffic volume. The growth of these parameters is expected to continue in the following period.

3.4. Cable Distribution of Radio and Television Programmes

Pursuant to the Rules on terms and conditions for radio and television programme distribution service (Official Gazette of RS, no. 26/09), on the radio and TV programme distribution market in the Republic of Serbia, the service may be provided via the following public telecommunications networks:

- radio and TV programme distribution via cable network (coaxial, hybrid and optical) - CATV which includes analogue and digital CATV and IPTV
- radio and TV programme distribution via satellite (Direct to Home – DTH)
- radio and TV programme distribution via MMDS (Multichannel Multipoint Distribution System) and LMDS (Local Multipoint Distribution System)

For the provision of radio and TV programme distribution service, RATEL issues an authorization in line with the abovementioned Rules, whereby an additional licence for radio frequency usage is issued for the MMDS and LMDS platforms.

There were 78 operators registered for providing these services in 2009, among which 75 operators had the authorization for radio and TV programme distribution service via cable distribution network (including 2 operators which provide IPTV service), and 3 operators had the authorization for providing the service via satellite distribution network (DTH).

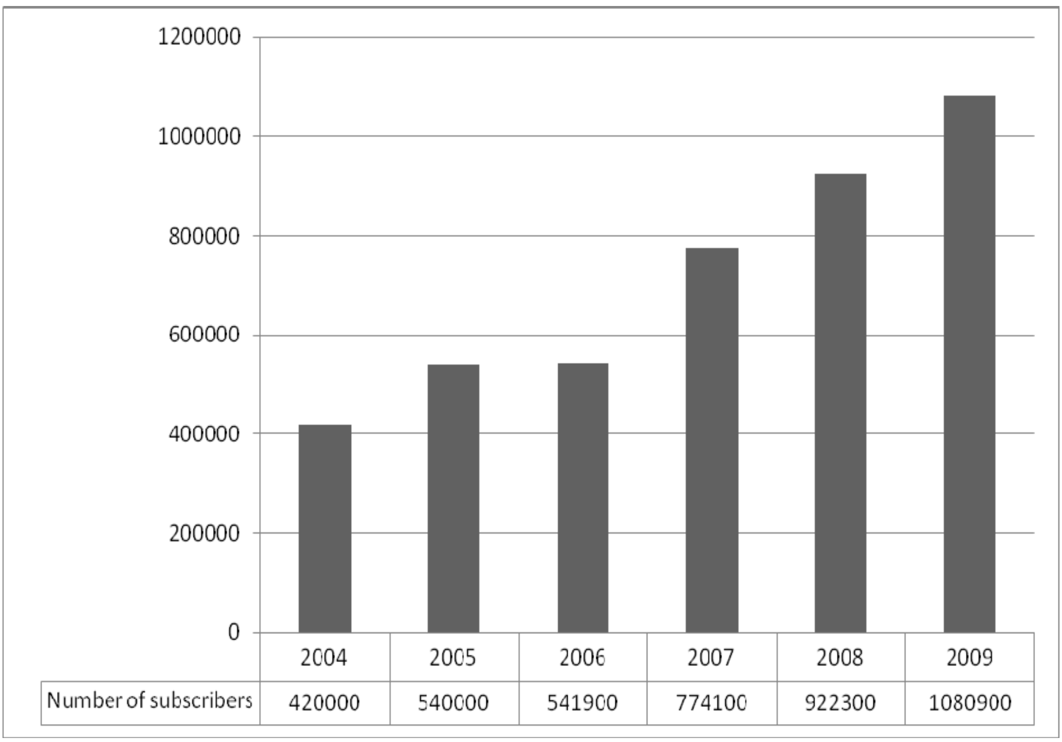


Figure 9. The number of CATV subscribers in the Republic of Serbia [4].

According to the number of subscribers, Serbia Broadband – Srpske kablovske mreže (SBB) still represents the leading operator within radio and TV programme distribution services, with a 52% market share. Consequently, in accordance with the Law, RATEL's Managing Board passed the Decision on Designating the Public Telecommunication Operator with Significant Market Power for the Radio and Television Programme Distribution via Cable Distribution Network, declaring the company SBB the operator with significant market power. Accordingly, the regulation of an SMP operator's services requires the application of a special tariff regime for radio and TV programme distribution services along with the obligation to apply the cost-based model in forming the prices of telecommunications services. Hence, SBB is required to observe the rules and conditions set out in the Rules on the application of the cost-accounting principle, separate accounts and reporting of a telecommunications operator with significant market power and the Law.

The total number of subscribers of radio and TV programme continues to grow, exceeding one million in 2009, which is by 17% more compared with 2008. The average growth rate of the number of radio and TV programme distribution service subscribers in the period from 2004 to 2009 was 21%. Penetration rate amounted to 14.4%, or 42.9% in terms of the number of households. The number of CATV subscribers is given in Figure 9.

The cable radio and television distribution penetration rate is given in Table 16 below.

Table 16. CATV penetration rate [4]

Year	2004	2005	2006	2007	2008	2009
Penetration rate	5.6%	7.2%	7.3%	10.3%	12.0%	14.4%

The peculiarity of the Serbian market lies in the fact that a major part of broadcasting infrastructure was destroyed in the 1999 bombardments, which left a great portion of the Republic of Serbia without the radio or television signal. For this reason, the development of the cable networks was more prominent in respect to other European countries. Due to mergers in the CATV market, SBB acquired a dominant position that needed to be regulated, making this provider the only cable operator in Europe whose prices are regulated according to the cost-based principle. These measures were a consequence of its vastly spread out network and its influence over the other operators, leading to RATEL's decision on price regulation in accordance with the Law on Telecommunications in force at the time.

Another serious problem was the fact that the networks had been rolled out often disregarding the technical standards. Therefore, RATEL had to publish a series of technical requirements applying to the roll out of cable telecommunications systems, which additionally contributed to a greater market regulation in this area.

The expansion of the cable operators and the increase in the number of users in this area has been accompanied by a steady growth of revenues in the past years, as shown in the Figure 10 below.

In 2005, more that 80% of networks were mainly coaxial with characteristics such as to enable solely one-way service (distribution of radio and TV programme), whereas in 2009, 90% of networks were hybrid HFC networks. Also, digital television has been introduced along with DOCSIS 3.0 standard enabling high bit-rates necessary for HDTV and high-speed Internet.

The basic feature of this market in the past period was the absolute dominance of the radio and TV programme distribution over cable network of both analogue and digital television. Since cable operators are already positioned according to geographic areas we can expect further market development to be marked by mergers and acquisitions.

IPTV appeared in the CATV market in late 2008, and in terms of the number of users and revenues it still has not acquired a significant market share. Considering the fact that fixed telephony subscribers, which add up to over 3 million, can be regarded as potential IPTV users, an expansion of this platform can be expected in the following period. Also, since fixed telephony and ADSL broadband Internet services are already being delivered through this infrastructure, the conditions for triple play services should be met in the following period.

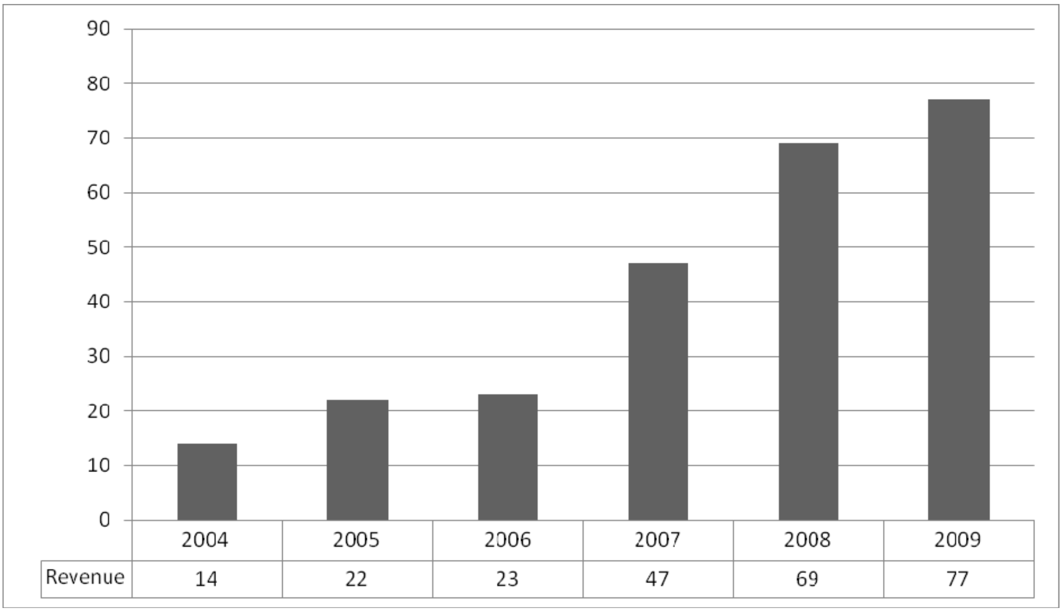


Figure 10. Revenues from CATV in millions of euros [4].

Unlike the radio and TV programmes distribution via cable network, there are no structural barriers for DTH service provision since the satellite used as a distribution network can cover the whole territory of the Republic of Serbia, which is particularly interesting for rural areas.

Since the users have a choice when it comes to access technology (cable, DTH or IPTV), involving adequate quality and prices, they can choose between different radio and TV providers, which improves market competition.

3.5. Broadcasting

Broadcasting refers to wireless terrestrial transmission of radio and television signal to end-users. It has already been mentioned that the major part of infrastructure was destroyed during the 1999 bombardments, which was also underlined in the 126-ITU Resolution in Marrakesh in 2002.

The situation in the broadcasting in 2005 in the Republic of Serbia can be described as unregulated and facing a variety of problems. Excessive number of broadcasting stations (FM and TV) with poor technical broadcasting quality caused interference, both on the national and the international level. Disregard of technical standards and regulations, especially in larger towns and cities, resulted in congestion and interferences, especially when it came to radio programmes. A large number of radio and TV stations were working illegally with inadequate equipment.

Following the public invitation of the Republic Broadcasting Agency (RBA) and based upon the data received from the responsible office of the Ministry of Capital Investments, in 2005 RBA made record of 755 broadcasters – where 543 were solely FM radio stations, 73 solely TV stations and 139 stations were broadcasting both radio and TV programme. Since many of these stations had several networked transmitters, the number of active frequencies (transmitters) in Serbia during 2005 was even higher.

The usage of radio frequency spectrum is regulated by the international Radio Regulations, international agreements, technical standards and legislation pertinent to telecommunications. Accordingly, two basic documents regulating broadcasting have been adopted: Radio Frequency Assignment Plan (Official Gazette of RS, no. 112/04) and Frequency/Location Allocation Plan for Terrestrial Analogue FM and TV Broadcasting Stations for the Territory of the Republic of Serbia (Official Gazette of RS, no. 6/06). The launching of the Public competition for the issuance of radio and television programme licence in 2006 marked the beginning of regulation of the broadcasting sector. In the past several years 5 national licences for TV signal coverage, 5 national licences for radio signal coverage, 135 regional and 288 local licences for radio and TV programme were issued.

Since the full application of the terrestrial digital broadcasting (T-DAB) in VHF (174-230 MHz) and DVB-T in VHF and UHF (470-862 MHz) frequency bands would not be possible in the European broadcasting without the revision of the international plan in force (Stockholm 61), Regional Radio Conference RRC-06 was held. In the second session (May, 2006), DVB-T and T-DAB terrestrial broadcasting digital plan was adopted which will remain if force for the next 40-50 years.

One of the final results of the Conference was the coordination and international recognition of the digital Plan GE06, namely allotment/assignment plan for digital broadcasting, DVB-T, for the territory of the Republic of Serbia, as well as the digital allotment/assignment plan for digital broadcasting, T-DAB, for the territory of the Republic of Serbia. This implies the following:

- eight coverages of the whole territory of the Republic of Serbia for digital TV (one coverage in the VHF band and seven in the UHF band) for portable and mobile reception,,
- two coverages of the whole territory of the Republic of Serbia for digital radio (in the VHF band) for portable and mobile reception,,
- six coverages for digital TV in the UHF band for the area of the city of Belgrade.

In view of the presented results of the Conference, it is important to underline the fact that the current level of technology is adequate to form a multiplex of 4-6 channels with a studio quality image. This means that, due to the results achieved at the Conference by the delegation of the Republic of Serbia, the coverage of the whole territory of the Republic of Serbia with 32-18 different TV programmes and 12-16 radio programmes has been enabled. Also, a possibility of the coverage of the wider area of the City of Belgrade with 56-84 programmes has been ensured. As the deadline for the switchover to the new digital plan is 2015, it can be expected that the further development of technology will lead to even greater number of TV programmes with the same frequency resources.

Having regard to the results of the RCC-06 Geneva Conference, further development of broadcasting, whether satellite, terrestrial or cable, will be based solely on digital technologies with the following standards already in use: DRM (Digital Radio Mondiale) for short-wave and medium-wave audio broadcasting, , T-DAB (Terrestrial Digital Audio Broadcasting) for audio broadcasting and DVB (Digital Video Broadcasting) suite of standards for digital television (S-Satellite, C-Cable, T-Terrestrial and H-Handheld). Digital technologies used for terrestrial broadcasting enable better capacity utilization rate of the existing frequency recourses (multiple programmes per channel) and higher resistance in terms of reception quality degradation, whereby the digital service is much closer to a studio level of quality in respect to the analogue service. When it comes to digital terrestrial broadcasting, Serbia has opted for T-DAB and DVB-T2 standards.

The Government of the Republic of Serbia adopted the Strategy for Switchover from Analogue to Digital Broadcasting in the Republic of Serbia, which aims at defining the framework and providing fundamental strategic guidelines for the introduction of digital radio and TV programme broadcasting in the Republic of Serbia. The Strategy enumerates, inter alia, the basic advantages of the digitalization for the users (audio and image resolution, content variety, more radio and TV programmes, new services for the disabled and elderly, etc.), for service providers (the possibility of adapting the content to suit the needs of different target groups, interactivity, etc.) as well as for the state itself primarily by enabling the more efficient radio frequency spectrum usage. The switchover from analogue to digital television signal broadcasting will result in digital dividend, which will surely contribute to further development of telecommunications. The specific allocation of the freed out RF spectrum has not yet been decided upon, but a part of it will most certainly be used for the commercial usage of the broadband systems. The date set as the final deadline for the digital switchover for terrestrial TV broadcasting in the Republic of Serbia is 4 April 2012.

For the purpose of a more rational signal delivery to the end-user, the Government of the Republic of Serbia has established public company for broadcasting equipment and communications ("Emisiona tehnika i veze") whose task will be to form multiplexes from the national, regional and local television broadcasters and deliver them to the homes on the whole territory or part of the territory of the Republic of Serbia.

3. Analytical and Technical Framework of the Market Regulation

Having regard to the situation encountered in the telecommunications market in the Republic of Serbia, the declared orientation of the state administration to liberalize the market and to eliminate monopoly and in view of the adopted legal solutions, RATEL was determined to enable good business practice to enter this segment of Serbia's economy. Due to numerous problems and their complexity, which had been overlooked, in good faith, by the legal framework regulating the telecom market, RATEL's activities were rather delicate at the very beginning. In order to approximate the common practice to the experiences of developed countries, RATEL's Managing Board sought grounds in the following mechanisms:

- Encouragement to introduce advanced technology so as to remove the existing barriers to entry in respect to the new operators and the provision of new services,
- Continual market analysis, benchmark analysis, annual publication of reports on the tendencies of principal indicators,

- Introduction and application of the cost-based model according to principles of accounting separation of costs, revenues and fixed investments.

Before we further explain each mechanism, let us first underline that insisting on continual market analysis, ensuring transparent presentation of the tendencies of principal indicators to the wide public and introducing and applying appropriate cost-based model is of particular importance, since RATEL literally found itself in the middle of uncultivated land in terms of practical experience in these areas in the Republic of Serbia. Prior to RATEL there was hardly any meaningful market analysis, much less a developed cost accounting which had been cast out of the companies with the first appearance of socialist economy.

The experience of many countries reveal that the development of an open and equal market can be provided by applying new technologies which was largely used by RATEL in the given conditions. Specifically, the fact that modern technologies allow for telecommunications systems to be only partially or entirely built by applying new technologies, led to a significant cost cuts, both concerning the initial installation and the subsequent maintenance. This is especially applicable to FWA and wireless signal transmission systems, or high-capacity microwave links. Since their application does not involve classical digging and roll-out of wires or cables, the costs of building and maintenance are significantly reduced.

The application of new technologies and protocols in the packet transmission and routing have led to a completely new network architecture in all layers. Voice transmission has been digitalized (VoIP) and this service is by far less expensive if the new generation networks (NGN) are applied. Moreover, the NGNs enable high quality transmission of a large number of services over the same network. Therefore, on top of the more affordable investments and maintenance, the variety of services available to a single user is multiplied. Regulatory support in the NGN usage creates new possibilities for a level playing field in the telecommunications market.

Also, in the area of cable distribution the application of the new technologies has led to the switchover from analogue to digital broadcasting. This digitalization has enabled direct to home digital transmission of the TV signal via satellite. Moreover, digital transmission of the TV signal via Internet Protocol broke the barrier of territorial or geographic monopoly present on the cable distribution.

The process of the terrestrial television digitalisation opens the possibility of a considerably higher number of frequency channels and hence a greater number of programmes and broadcasters. Naturally, it is of particular importance that the digitalization liberates a significant portion of the RF spectrum, which will be newly allocated.

Along with the implementation of the solutions stipulated under the law and the creation of conditions for the application of the advanced technology, RATEL carries out a detailed and continual analysis for the application of advanced technology and compares movements in the carefully selected set of indicators with the countries in the region and developed countries for the purpose of identifying the SMP operators whose prices are subject to a strict control in order to prevent monopolistic behaviour.

The previous Law on Telecommunications recognized four relevant markets, namely: fixed telephony market (including the infrastructure), mobile telephony market, leased lines market and interconnection market. Considering the fact that the telecom market had already been split into numerous segments hard to encompass with the abovementioned markets, in 2008 RATEL's Managing Board passed the Decision on the monitoring and analysis of the telecommunications market by applying the four identified markets with the addition two more markets, namely the cable distribution systems service provision market and Internet service provision market. The Decision also stipulates the preparation of the telecommunications market overview and periodic detailed analysis pertaining to the identified markets with the purpose of examining the existing and identifying any potential SMP operators. Pursuant to the Decision, a multidisciplinary team has been formed during 2009, consisting of economists, engineers and lawyers, with the task of a continual monitoring and analysis of the telecommunications market.

For the purpose of this paper, the results of these analyses will be briefly presented. First, the analyse has revealed that Telekom Srbija owns the infrastructure which includes the access network to over 3 million users, transmission systems, fibre-optical network at the transit, local and subscriber level, and the duct system containing the entire cable infrastructure. Since this is fundamental for the provision of leased lines, it can be concluded that, when it comes to the leased lines market, Telekom Srbija controls the infrastructure that cannot be easily duplicated any time soon. Based upon these facts, the results of the analysis have provided sufficient indicators showing that the regulation of market and prices in the segment of leased lines is feasible only if Telekom Srbija is declared an SMP operator until the conditions for a full and equal competition are met.

Second, the analysis has shown that Telekom Srbija has the most favourable position in the wholesale broadband Internet market, also due to the ownership of the mentioned infrastructure. Also, it should be noted that until December 2008, this company was the only holder of the authorization for the international interconnection (Verat company had the authorization for wireless connection only), which is one of the main reasons for such situation in the market.

Third, based upon the results of the analysis it can be concluded that all three operators present in the mobile network interconnection market have a dominant position in the area call termination service, whereas as regards all interconnection services in fixed network, including national calls termination, international call termination and national and international calls transit, Telekom Srbija holds a dominant position.

Fourth, the market share of Telekoma Srbija in the fixed telephony market was 100%, according to the number of users and revenues, since it was the only operator in fixed operator. It owns infrastructure which is hard to duplicate, but it can be rather quickly modified at low costs to provide other services. Its dominance will continue in the following period, especially in the area of access networks and national calls provision.

Fifth, the analysis of the radio and television programme distribution market has shown that a large number of operators are present in the market, while the dominance of one particular operator is evident, since SBB hold a market share of over 50% in terms of the number of users. It provides the services over HFC network and also via satellite system. With the introduction of IPTV platform greater competition can be expected primarily by the dominant operator Telekom Srbija.

Sixth, the mobile market can be divided into wholesale and retail market. As for the retrial market there is a possibility of substitution both on demand and supply side. On the demand side, the users can choose between all three operators, and between prepaid and postpaid packages. As for the supply side, since the same network is used for the provision of different services, the operators can easily start with the provision of new services. Thanks to the network access service in the wholesale market, a new entrant can quickly enter into the retail market, which enables full competition. The following markets have been identified on the wholesale level: network access and the possibility of outgoing calls and international roaming. When it comes to the network access and the possibility of outgoing calls, there are three operators on the supply side with their own networks covering the major part of territory and population. Considering the market size, the entry of a new infrastructure operator can hardly be expected, however virtual operators may show interest in the market. The Agency should open up this market by imposing on the operators the obligation to provide network access and the possibility of outgoing calls to any operator who submits a reasonable request for the usage of this service. As for the international roaming, foreign operators can close agreements with all three operators, however, despite the competition present in the market, the prices for end-users are rather high, exceeding the prices offered to the subscribers in the EU. The EU practice has shown that the competition alone is not enough to reduce the prices, but the intervention of the regulatory authority is necessary.

In addition to these analyses, the regulatory authority has regularly been publishing an annual overview of the situation in the telecommunications market in the Republic of Serbia [4]. This comprehensive and detailed analysis offering specific statistical data has proved to be extremely useful for all existing and any potential telecom market players. It should be noted that this publication has been published every year since 2006 and it is a rare example of a transparent monitoring and publishing of data relevant for a particular branch of the Serbian economy.

With the purpose of additional regulation of prices and QoS, the regulatory authority currently also applies benchmarking method in respect to other countries in the region. Benchmarking is the process of systematic and continual measuring and comparison of the organized processes in respect to the processes in other countries [14]. Each benchmarking aims at the reduction of differences between the compared processes. In order to achieve this objective, first of all the area of comparison has to be carefully selected. In our case, the selection was based exclusively on the country size, per capita income and the size and situation in the relevant countries' telecom market. Subsequently, specific services and/or networks whose performances are to be compared during a certain time-sequence are selected. After that, specific values of the chosen parameters from data bases are compared, taking into account all peculiarities of the countries relevant for the comparison. Naturally, the most important part concerns the analysis of differences between the obtained results to which RATEL pays particular attention, because the objective of the benchmarking is not a mere comparison, but identification of causes and possible consequences of the differences between the results obtained. Subsequently, relevant conclusions can be drawn on the basis of this analysis, which should enable further development of the market and eventual reaching of the EU level of development in the long run.

Finally, the most complex activity RATEL was engaged in as a part of the telecommunications market regulation was indeed the introduction and application of the cost-based model according to the accounting separation of costs, revenues and fixed investments. Thanks to enormous efforts, the regulator managed to implement this model in less than three years, thereby making the telecommunications the only economic sector where such model is applied. The first step in the adoption of the mechanism for price control of the SMP operator was the adoption of Rules on the application of cost-based principal by RATEL's Managing Board, regulating the basic principles, models and methodology of cost and performance calculation, calculation of cost price and sale prices for the services of the SMP operator [15].

When fulfilling the obligation to select a cost accounting model and method, RATEL considered the feasibility of its application with the SMP operators in Serbia and has chosen Historical Cost Accounting – HCA model according to top-down method, based on the functional principle of Fully Distributed Cost – FDC. The selected model was chosen as the most appropriate for the level of development of the existing accounting information system in the Serbian companies, and it is only the initial solution on the road to the implementation of other two accounting models, namely, Current Cost Accounting - CCA and Long Run Incremental Cost - LRIC. Through the introduction of this model, RATEL has ensured a greater social role i.e. the role of an educator of operators in the area of cost accounting application, in which domestic operators lag seriously.

Model HCA is based upon real costs se expressed in the bookkeeping records of the operators or expenditures and revenues of the company as a whole. This model is accompanied by the principle of Fully Distributed Cost, which means that the costs generated in the company are allocated to final products and services launched in the market. The cost allocation can be direct or indirect – applying appropriate cost drivers, which may differ according to the type of cost or service it applies to. The major advantage of this model is a high degree of SMP's data verifiability, whereas the principal setback is that it applies only to the existing networks and services. This model is accompanied by top-down method, which includes the primary cost accounting by types for the company as whole; internal allocation and reallocation of capital and operational costs by company segments; allocation of capital and operational costs by types and service units.

Pursuant to the Rules, the SMP operator is required to prepare unit cost calculations for all market services and submit them to RATEL in writing, by the set deadline. In addition to the official financial reports (Balance Sheet, Income Sheet, Cash Flow Statement, Report on Changes in Capital and Explanatory notes to Financial Statements), SMP operators are required to prepare and submit Internal Statements prepared according to the recommendations under Rules, at least twice a year.

The controlled prices of the telecom services are formed by means of a consistent application of the cost-based model, which may be accompanied by a comparative analysis of prices in the region, while making sure that both user protection and the assets for the further development of the SMP operator are ensured. A particular value of these Rules lies in the openness and flexibility towards other cost accounting methods, such as Activity Based Costing – ABC. This means that, irrespective of the primarily applied FDC principle, the operators are free to apply the ABC methodology as well. The Rules also stipulate the application of CCA to begin in mid 2011, whereas LRIC is planned to start being applied later on.

The significance of the price control applied to the SMP operators in the Serbian telecom market is manifold since on the one hand it enables a transparent monitoring of cost prices and sales prices of the services within the special tariff regime, and on the other hand it stimulates a gradual increase of the corporate social responsibility of the operators. In this way, the SMP operators are given the possibility to constantly upgrade their business efficiency and capital and assets management.

CONCLUSION

The paper provides an overview of RATEL's five-year work and the results accomplished in the Serbian telecommunications market. The principal task of the regulator is to create conditions for competition in all segments of the telecoms sector, in order to provide a liberal and open market, and hence ensure the future development of the entire sector, greater number of services and improved quality of service available to users at a lower price. Having regard to these principles laid down under the Law on Telecommunications, appropriate bylaws, decisions and decrees were adopted leading to the creation of conditions for competition in all segments of the telecommunications market: public fixed telecommunications networks and services, public mobile telecommunications networks and services, Internet services, cable distribution and broadcasting.

The analysis of the public fixed telecommunications network market showed that the number of users of fixed telephony services and the intensity of its traffic were becoming saturated. The main reason was the transfer of telephone services from fixed to mobile network, accompanied by the fall in revenues and investments in this market segment. Since Telekom Srbija was the only operator, it was declared by RATEL an SMP operator required to keep separate accounting for each organizational unit (fixed, mobile, Internet, IPTV, leased lines, interconnection). This prevented cross-subsidies in terms of costs, revenues and investments, thus averting further monopolistic practices and uncompetitive impact on particular segments of the telecom market. Since the price of telephone call was extremely law, there was no interest for the entry of new operators. With the tariff rebalance according to the cost-oriented model, conditions were created for other operators to see an interest in the provision of this type of service. It was only then that the tender for the national licence for fixed wireless access using CDMA technology could be launched, opening the doors to competition and leading to price-cuts.

Modern technology has enabled the convergence of the fixed network service transmission, which made it possible for the user to simultaneously make telephone calls and use high-speed Internet. Furthermore, IPTV and other numerous e-services are becoming more available. The result is the increase in the number of users and, more importantly, the multiplication of the number of services per user. Also, the possibilities of revenues from fixed networks have significantly increased. In order to take advantage of the new possibilities, great fixed investments are required in the reconstruction of the existing and the roll-out of the fixed new generation networks. In view of this trend, RATEL initiated and carried out the issuance of the national licence for fixed wireless access. As a result, there are now three operators in the fixed network market: Telekom Srbija, Telenor and Orion. The first two operators will be able to provide broadband services in the future, specifically service packages consisting of VoIP, high-capacity Internet, IPTV and other services. All this suggests that the conditions have finally been met for the beginning of the full liberalization of the fixed telecommunications network market.

In the public mobile telecommunications networks the competition is present, since three large operators are in the game: Telekom Srbija, Telenor and VIP. This resulted in a significant increase in the number of users. Along with the growth in the number of users, the improvement in the QoS and the increase in the number of services leads to a greater traffic volume and hence to the rise of the revenues. There is also a considerable increase in the investments in these networks. The introduction of the number portability will lead to additional competitiveness among the mobile network operators.

The Internet service market is displaying growth in the number of users accompanied by the increase in revenues from the services. The initial situation was rather unsatisfactory, with a low number of users and the wholesale monopoly held by Telekom Srbija. In order to increase the number of users, RATEL first liberalized the usage of the 2.4 and 5.5 GHz, utilized by a great number of ISPs to provide the wireless Internet of non-guaranteed quality of service. The This was followed by the regulator’s adoption of the Rules on the authorization for the international interconnection of ISPs, whereby the Internet wholesale was liberalized. New network technologies have created new possibilities of convergence and globalization of the access, especially in the Internet services, both in fixed and in mobile networks. The user is given a whole range of possibilities concerning the choice of quality and comfort, but also regarding the price of the service provided. This will indeed increase the revenues from these services, however it also involves considerable investments and the creation of conditions for the application of the new technologies.

The cable distribution market is characterized by the growth in the number of users, largely due to damages suffered by the terrestrial radio and TV broadcasting systems during the bombardments. Since there is a territorial monopoly in this segment, SBB was declared SMP by RATEL and adequate measures were taken, TV programme broadcasting service price regulation according to the Rules on the price-oriented principle. With the purpose of reducing this monopoly, the Rules on TV broadcasting service provision using DTH technology was adopted. Also, the regulator took measures to enhance IPTV usage, as the new type of competition in this market. Recent technological developments led to the convergence of services in the distribution systems, thus making possible the provision of the Internet service together with all its advantages, in addition to the TV programme broadcasting, which resulted in the significant increase of the number of users. A considerable improvement of the quality of the technologies applied in these networks was also noted. Indeed, this led to the increase in the investments and the growth in revenues.

The broadcasting was an unregulated sector facing a series of problems, such as the excessive number of broadcasting stations (FM and TV) with poor technical broadcasting quality and the disregard of technical standards and regulations, which led to interference and illegal operation of many stations. The market was largely in the area “grey economy”.

With the implementation of the Broadcasting Law and the La on Telecommunications, optimal usage of the RF spectrum and the legal operation in this market was ensured, ultimately resulting in the stabilization in the revenue inflow. The process of the digitalization of the terrestrial TV broadcasting is underway, which will free a portion of spectrum and enable it to be used for other modern services. This should lead to new investments and new revenues in the telecommunications market.

As shown in the discussed segments of the telecommunications market, the regulator made an effort to create conditions for market competition through the regulatory measures. The analyses show that the competition in the mobile networks is rather effective, whereas in other telecom market segments the necessary regulatory conditions are in place. Therefore, the real market competition is still to come. An important factor of guarantee is the new Law on Electronic Communications, based on the 2007 EU Regulatory Framework. Also, the regulation will cover a greater field divided into the following segments of the telecommunications market:

1. Retail access to the public telephone network at a fixed location,
2. Wholesale call origination on the public telephone network provided at a fixed location,
3. Wholesale call termination on the public telephone network provided at a fixed location,
4. Wholesale physical access to telecommunications infrastructure,
5. Wholesale broadband access,
6. Wholesale terminating segments of leased lines,
7. Wholesale call termination on mobile network.

The 2007 EU Regulatory Framework leaves it up to the national regulatory authorities to analyze additional markets according to the peculiarities of the respective countries.

Considering the results achieved and the new regulatory challenges of the Serbian telecommunications market, it can be concluded that the implementation of the new regulatory framework is the logical next step to be taken by RATEL in order to fulfil the goal of an effective market liberalization of the Serbian telecommunications, feasible in the long-run.

REFERENCES

[1] Vladica Tintor, Vlade Milićević, Milan Janković and Jovan Radunović, Liberalization of the Mobile Telephony Market in the Republic of Serbia, *Technology in Society*, Volume 31, Issue 4, November 2009, Pages 384-398, Elsevier Ltd.

[2] V. Tintor, M. Janković, V. Milićević, Legal and economic framework of EU telecom market regulation, *Economic Annals*, Volume 185, April-June 2010.

[3] *Telecommunication Regulation Handbook*, IBRD, World Bank, 2011.

[4] RATEL's yearbook publication on telecom market, available on www.ratel.rs.

[5] Jonathan E. Nuechterlein, *Digital Crossroads: American Telecommunications Policy in the Internet Age*, MIT Press, 2005.

[6] Kjell A. Eliassen, Johan From, *The privatisation of European telecommunications*, Ashgate, 2007.

[7] Ian Walden, *Telecommunication Law and Regulation*, Oxford University Press, 2009.

[8] M. L. Dukic and M. Jankovic: A Proposal for Telecommunications Strategy in Serbia, *IEEE Communications Magazine*, August 2005, Vol. 43, No. 8.

[9] Milan Jankovic, Jovan Radunovic, Regulation of the Telecommunication Market in the Republic of Serbia: Start-up, Global Communications Newsletter, *IEEE Communication Magazine*, vol. 46, No. 9, September 2008, 22. 2,6.

[10] Jankovic M., Fixed Communications in the Republic of Serbia are Entering competition, Global Communications Newsletter, *IEEE Communication magazine*, Vol. 48, no. 11, November 2010.

[11] Law on Telecommunications (*Official Gazette of RS*, number 44/03, 36/06 and 50/09-US).

[12] Law on Electronic Communications (*Official Gazette of RS*, number 44/10).

[13] Enlargement Country Monitoring Report, available on www.cullen-international.com.

[14] Dejan Malinić, Financial benchmarking as a tool for improving corporate performance, *Proceedings of the XIV Symposium of Accountants and Auditors, Banja Vrućic*a, 2010.

[15] Rules on the application of the cost-accounting principle, separate accounts and reporting of a telecommunications operator with significant market power, (*Official Gazette of RS*, number 103/08).

[16] Carl W. Stern, *Boston Consulting Group on Strategy*, John Wiley&Sons, 2006.

In: Advances in Communications,Volume 8
Editor: Anthony V. Stavros
ISBN: 978-1-61324-794-5

Chapter 4

THE IMPACT OF THE TELECOMMUNICATIONS ACT OF 1996 IN THE BROADBAND AGE

Hanlong Fu
Department of Communication Sciences
University of Connecticut, Storrs, Connecticut, U. S.

Abstract

The Telecommunications Act of 1996 is arguably the single most important piece of legislation since the Communications Act of 1934, one that affects the telecommunications industry, consumers, and ultimately the balance of political power. The 1996 Act was designed to usher in competition to telephony and cable by breaking down the cross-entry barriers that were put in place by the Communication Act of 1984. Yet immediately after the passage of the 1996 Act, telecommunication industries witnessed a deluge of mega mergers and acquisitions. This unprecedented merger wave resulted in a handful of conglomerates dominating industries that were previously separated by telecommunications regulations. The present chapter assesses the structural impact of the 1996 Act by analyzing the most recent industry trends and statistics of media ownership. We discuss, in particular, the implications of the 1996 Act for the National Broadband Plan.

INTRODUCTION

The Telecommunications Act of 1996 is arguably the single most important piece of U.S. legislation since the Communications Act of 1934 (e.g., Atkin, Lau, & Lin, 2006; Howard, 1998; Pickard, 2007), one that affects the telecommunications industry, consumers, and ultimately the balance of political power (McChesney, 2000). The 1996 Act was designed to usher in competition to telephony and cable by breaking down the cross-entry barriers reified by the Communication Act of 1934 and reified under the Telecommunication Act of 1984 (Krattenmaker, 1996). Architects of the legislation envisioned that local telephony would evolve from a natural monopoly to an open, competitive marketplace (Bates, Jones, &

Washington, 2002). Yet, immediately after the passage of the 1996 Act, telecommunication industries witnessed a deluge of mega mergers and acquisitions (e.g., Atkin, Hallock, & Lau, 2006; Bates, Washington, & Jones, 2002), which resulted in a handful of conglomerates dominating industries that were previously separated by telecommunications regulations.

A decade and a half later, merger mania continues to run rampant in the telecommunication industry (Abelman& Atkin, 2011). The 1996 Act failed to deliver the competitive stimulus that it was designed to offer. Instead, the telecommunications industry was steeped in $165 billion of debt and plagued by plummeting stock values only a few years after the passage of the 1996 Act (Dibadj, 2003). McChesney (2000) argues that these results are predictable consequences, originating from a corrupt and anti-democratic process that shirked public input. The public interest, a supposed goal of the Act, was undermined by corporate lobbyists and the undue influence they brought to bear on the political process determining the 1996 Act. Although much has been written about of the impact of the 1996 Act (Atkin et al., 2006; Huntemann, 1999), scholarly attention is needed to monitor the ongoing impact of the 1996 Act in this rapidly evolving media landscape. This chapter first briefly retraces major telecommunications legislation in the U.S., including the 1996 Act. Second, we provide an up-to-date analysis of the impact of the 1996 Act on the telecommunications industry. In light of newly proposed National Broadband Plan (NBP), the chapter concludes with an analysis of the likely impact of the 1996 Act on the NBP.

BACKGROUND: HISTORICAL OVERVIEW

Communications policy has historically been shaped by the emergence of a new communication technology. Each new technology necessitates the formation of a different policy regime, defined by the new attributes associated with the technology. As a result, print media, common carriers, broadcasting, cable TV, and the Internet were the leading technological forces determining the contours of communications policy in the U.S. (BerQuist, 2010).

Modern telecommunications policy making can be traced back to the early diffusion of the telephone, a technology invented in 1876 by Alexander Bell. As the telephone began to diffuse, a state of laissez-faire oversight led to the rise of several thousand independent telephone operators that were neither interconnected nor efficient (e.g., Dizard, 2000). Established originally for regulating railroad transportation, the Transportation Act of 1920 was later extended to classify telegraph and telephone as “common carrier,” an archaic term used to describe the duties associated with operators of public utilities (Spector, 2002). By law, common carriers are required to offer interconnection and other services to customers without discrimination (BerQuist, 2010; Spector, 2002).

The Radio Act of 1927 was put in place amidst a chaotic environment in which scarce electromagnetic spectra were overcrowded with too many broadcasters (McChesney, 2000). In 1927, Congress passed the Radio Act to restore the order of over-the-air broadcasting. The Federal Radio Commission (FRC) was established to regulate the burgeoning broadcasting industry and to reallocate the electromagnetic spectra. As Krattenmaker (1996) notes, the telecommunications industry was operated under the auspices of strong governmental regulation from the very beginning, as there were no comparable industry-specific agencies at a federal level for other industries.

The FRC used a "public interest" criterion to reallocate electromagnetic spectra, that is, stations were permitted to broadcast as long as they promised to serve the public interest (Hazlett, 1998; Krattenmaker, 1996, Napoli, 2001). The FRC favored those stations which best served "public interest, convenience, or necessity" (McChesney, 2000, p. 192). In the absence of a clear definition for those terms, some thought that the FRC would initially favor non-profit and educational stations. The FRC instead argued that commercial broadcasters would best serve the public interest because they were driven by profits and less likely to broadcast "propaganda" than other types of stations (McChesney, 2000; 2004). In that vein, commercial broadcasters were already in a privileged position, given that they were better financially and materially prepared to "serve" the public interest. The result of spectrum reallocation was that educational and nonprofit broadcasters lost their broadcasting rights to commercial broadcasters such as NBC and CBS (McChesney, 2000). Nevertheless, the "public interest" criterion has become thestandard for assigning broadcasting licenses (Hazlett, 1998).

Dizard (2000) notes that bans on telephone company entry into electronic media were debated when Congress established the FRC. But the separation between common carrier and broadcast media was not formalized until 1934, when Congress supplanted the 1927 Act with new legislation. As McChesney (2002) pointed out, the 1927 Act was an overnight bill designed as a temporary measure to resolve the chaos in the radio industry. The Communications Act of 1934 was a follow-up bill engineered to bring long-term order to over-the-air broadcasting. The 1934 Act formalized the separation between broadcasting and telephony by assigning different sections to broadcasting and telephony in the Act (Atkin, 1999). As part of the 1934 Act, the Federal Communications Commission (FCC) was established through the merger of the Interstate Commerce Commission (ICC) and the FRC (Bednarski, 2003). Backed by the broadcasting industry, the 1934 Act was established to preserve the growing broadcasting industry and protect the interests of existing stations (Bates & Chambers, 1999).

One of the continuities in the 1934 Act was that the public interest was still considered to be an important determinant of communications regulation, despite the fact that the Act was premised on the fact that the radio spectrum would remain primarily commercial (Bednarski, 2003). In particular, the common carrier principle was reaffirmed in the 1934 Act. The public interest standard, poorly defined in both the 1927 and 1934 Acts and often sidelined by commercial interests, remained an important aspect of policymaking rhetoric during 1930s. The 1927 and 1934 Acts are often framed in the context of the trusteeship model in the literature (Hazlett, 1998; Bednarski, 2003), one in which broadcasters were granted free licenses in return for commitments to serving the public interest. As a case in point, the 1934 Act recognized AT&T as a "natural monopoly" in telephony, and in return, AT&T was expected to deliver universal service in order to comply with the public interest mandate (Bates & Chambers, 1999). These rather paternalistic measures reflected prevailing beliefs – forged during the depths of the Great Depression – that telecommunication channels should serve the public and perhaps even help grease the wheels of commerce (Atkin, 1999: McChesney, 2000).

The trusteeship model remained the de facto model for telecommunications policymaking through the early 1980s (Bednarski, 2003). The model was called into question after 1965, when AT&T's monopoly attracted increasing criticism as new market players such as MCI sought access to AT&T's long-distance telephone lines and were denied entry (Melody, 1999). In 1974, the Department of Justice (DOJ) proposed a motion to disintegrate AT&T on the grounds that it "foreclosed the equipment market with a bias toward the Western Electric subsidiary;engaged in predatory pricing, particularly in the intercity service area;denied interconnection of specialized common carriers with the Bell network; anddenied interconnection of non-Bell equipment to the AT&T network" (Atkin, Lin, & Lau, 2006, pp. 81-82). Public outcry and regulatory interference led to a divestiture settlement with AT&T in 1982 (see US vs. AT&T, 1982), which broke AT&T into smaller parts based on their geographic locations, known as Regional Bell Operating Companies (RBOCs) or "Baby Bells" (Brock, 1994). Although RBOCs were originally heavily regulated, they were gradually allowed to experiment with tapping into cable and other information service industries (Dizard, 2000). Nevertheless, the Baby Bells were under strong regulatory scrutiny, which preempted them from growing into monopolies in their respective markets.

At about the same time as the AT&T divestiture, the regulatory model of the telecommunications industry in the U.S. shifted from that of a trusteeship model to a marketplace model, which accentuates the role of market forces in meeting the public interest (Bates & Chambers, 1999; Bednarski, 2003). Although the model still has a public interest mandate, it assumes that broadcasters will automatically act in the public interest by adjusting their content to meet the audience's needs (Zaragoza, Bodorff, &Emord, 1988). Under this model, content regulation is harmful to the public interest because any content regulations would deny the average consumers maximum satisfaction from the medium (Zaragoza et al., 1988). In short, the marketplace model relies on market forces to serve the public interest.

Without defining the public interest, market forces within a laissez-faire economic system were assumed to be in natural harmony with the public interest. This new model of telecommunications regulation significantly influenced the FCC's major policy initiatives since that time (Bednarski, 2003; Zaragoza et al., 1988). Additionally, the shift in regulatory approaches meant the abolition of the Fairness Doctrine, which required broadcasters to provide reasonable amount of opportunities for the public to access to rivaling or conflicting viewpoints (Bednarski, 2003). Consequently, as the rise to prominence of conservative talk radio suggests, the diversity of voices was undermined by this shift of regulatory models (Abelman& Atkin, 2011).

Cable TV was originally designed for sending signals to underserved households and did not become a commercial enterprise until the 1950s (Guillory, 2006). After the initial trial, cable TV enjoyed exponential growth and reached 10 million households by the mid-1970s. As cable adoption and commercialization increased, broadcasters complained about cable operato retransmission of their signals on copyright grounds (BerQuist, 2008, 2010). Although the courts did not support the FCC's attempt to regulate cable, Congress passed legislation that required cable operators to pay royalty fees to broadcasters (Guillory, 2006). As the number of cable systems was rapidly increasing, the disoriented practice of franchising precipitated consumer complaints, prompting Congress to take measures to restore order in the cable industry (BerQuist, 2008; Shaw, 2002). The Cable Communications Policy Act of 1984 was instituted to put the cable industry under the regulatory jurisdiction of the FCC (Einstein, 2004) and to formalize the franchising process at local levels (BerQuist, 2008). In response to skyrocketing cable rates, Congress passed the Cable Television Consumer Protection and Competition Act of 1992 to reaffirm the FCC's authority to regulate basic cable rates (BerQuist, 2008; Shaw, 2002). The 1992 Act mandated the "must carry" and "retransmission consent" rules to allow broadcasters to negotiate with cable operators for carriage (Meadows, 2010). These two rules stipulate that cable operators must carry signals from local stations and the arrangement must be reached by consent between cable operators and local broadcasters (Meadows, 2010). The latter provision was joined by several others that were included in a major overhaul of the Communication Act four years later.

THE TELECOMMUNICATIONS ACT OF 1996

As a product of the marketplace model, the Telecommunications Act of 1996 was designed to "promote competition and reduce regulation in order to secure lower prices and higher qualities for American telecommunications consumers and encourage the rapid deployment of new telecommunications technologies" (see the 1996 Act). The 1996 Act was signed by Bill Clinton on February 8, 1996 and was touted by the former president as "a truly revolutionary legislation that will bring the future to our doorsteps" (Clinton, 1996). Commentators have generally been less sanguine about the impact of the 1996 Act (e.g., Dibadj, 2003; Huntemann, 1999; McChesney, 2004; Napoli, 1999; Warf, 2007; Wirth, 2002). Having brought colossal debt and hyperinflationary rates (Atkin, Lau, & Lin, 2006), the 1996 Act has yet to deliver its key promises, i.e., increased diversity and competition, lower prices, and higher service quality. This section profiles the 1996 Act by analyzing its key stipulations and institutional and technological antecedents.

Antecedents

Mauritsen (2002) noted that the 1996 Act was designed to address two basic structural issues in telecommunications: technological convergence and legal balkanization. To address these issues, Congress expected the FCC to formulate policies that could eliminate cross-entry barriers between relevant industries. Congress also anticipated that the FCC would loosen its grip on market entry standards which could retard technological innovations. Additionally, the potentially harmful effects of competition within and across industry segments had to be constrained (Krattenmaker, 1996). The 1996 Act was envisioned to bring competition to monopolized markets so that consumers could access new technologies at reasonable costs through a competitive marketplace (Economides, 1998; Mauritsen, 2002).

Technological Convergence

As Krattenmaker (1996) observed, everything seemed much simpler when the Communications Act was established in 1934, because electronic communications occurred via either the air or wires. However, telecommunications policymaking was complicated by the emergence of a host of new communication technologies such as satellites, microwaves, computers, television, fiber optics, and the Internet (see also Dizard, 2000). As these new technologies defy the traditional delineations between air and wire transmission, regulatory regime based on traditional classification schemes are often rendered obsolete. For instance, the century-old telephone industry is being transformed by digital transmission of voice, video, and data services via wireless and fiber optics networks (Dizard, 2000; Atkin et al., 2006; Warf, 2007). Fiber optics technology blurs the traditional dichotomy between telephony and cable by allowing cable operators to offer telephone and Internet services and telephone operators to offer cable and Internet services. In the meantime, Internet service providers (ISPs) are capable of offering traditional voice, video, and data services through digital technologies such as Voice over Internet Protocol (VoIP), and Internet Protocol TV (IPTV). In short, technological convergence equips traditionally separate industries with new capabilities, allowing them to offer services that were previously exclusive to other industries.

Galperin (2002) notes that technological convergence calls for new regulatory regimes built upon new technological and economic dynamics. The Communications Act of 1934 was modeled upon a media landscape that differed drastically from the early 1990s. The 1934 Act proved increasingly inadequate in dealing with the challenges presented by a rapidly-evolving media landscape. As the section to follow details, older regulatory schemes governing the pricing, interconnection, and last-mile rental charges for traditional cable and telephonybeca-me obsolete.

Legal Balkanization

Legal balkanization is a state of affairs that materialized as a consequence of governmentally imposed entry barriers separating older telecommunications industries (Krattenmaker, 1996). The problem of legal balkanization loomed large in the context of technological convergence, as regulators struggled to override legal barriers to facilitate cross-industry entries. In that sense, breaking down legal balkanization became a necessary condition for establishing new regulatory policies in the context of technological convergence.

The rationales behind legal balkanization were the fear of predation and the disruption of pro-social cross-subsidies (Krattenmaker, 1996; Krattenmaker& Salop, 1986). Put differently, entry barriers were considered necessary to prevent predatory activities between telephone and cable industries. For example, telephone companies, advantaged by their more extensive networks, could deny cable companies' equal access to necessary facilities. Telephone operators could also engage in more subtle forms of predation such as discriminatory pricing and cross-subsidization (Krattenmaker, 1996).

Key Stipulations

In essence, the 1996 Act was designed to facilitate cross-media competition by allowing: 1) local and long distance telephone companies to enter each other's markets, and 2) cable and telephone companies to enter each other's markets (Bates & Chambers, 2002; Dizard, 2000; Telecommunications Act, 1996). The 1996 Act contains detailed provisions that affect many telecommunications industries.

For radio, the 1996 Act abolishes the ownership cap on the number of stations that any single entity can control nationwide (Krattenmaker, 1996). Although anti-trust law remained applicable, this meant that the FCC was willing to give up its essential regulatory measures for the so-called “invisible hand.” In the television domain, the 1996 Act raised the ownership cap of stations that a single entity can own nationwide to 35% (Chan-Olmsted & Kang, 2004; Zhang, 2003). The FCC further lifted the cap to 45%, which caused Congress to issue a resolution of disapproval (e.g., Atkin, Lau, & Lin, 2006). These industry-friendly measures, together with the FCC’s early orders eliminating prohibitions on broadcast-newspaper cross-ownership, were the major culprit behind the merger mania to follow (Huntemann, 1999; Warf, 2007). Additionally, the 1996 Act also hastened rate deregulation in cable systems–a process initiated in the Cable Communications Act of 1984 – where cable rate regulation was authorized only when there is no “effective” competition in the cable market (Krattenmaker, 1996). When telephone operators enter the cable market, effective competition, by definition, is achieved and both cable and telephone operators are then freed from rate regulation.

At its core, the 1996 Act is predicated upon breaking the monopoly of Baby Bells and thereby promoting an open local loop (Dizard, 2000; Krattenmaker, 1996). Codified in the 1996 Act, sections 251 and 271 stipulate how other industries can enter an RBOCs’ markets and how RBOCs can enter long-distance markets, respectively (Dibadj, 2003). Section 251 dictates that incumbent local exchange carriers (ILECs) “provide, at just and reasonable rates, interconnection with their networks for the transmission and routing of telephone exchange service and exchange access at any feasible point within the ILECs’ networks” (Telecommunications Act, 1996). In addition, all ILECs must “provide access to their poles, ducts, conduits, and rights of way to competing providers of telecommunications services” (Telecommunications Act, 1996).

Simply put, RBOCs must open up their markets to new entrants and provide facilities for interconnection at reasonable prices (Dibaj, 2003; Krattenmaker, 1996). On the other hand, section 271 mandates that RBOCs can enter long-distance telephone markets when a series of four conditions are met. Briefly, an RBOC must enter into binding agreements with other new entrants for access to the ILEC’s network. The FCC must determine whether the requested authorization for long-distance services is consistent with the “public interest, convenience and necessity,” a rhetorical meme dating back to the Radio Act of 1927. Finally RBOCs must meet 14 requirements of a checklist designed to ensure RBOC’s compliance with section 251 and 252.

As Dizard (2000) notes, long-distance telephone markets were already very competitive when the 1996 Act was passed. As a case in point, AT&T’s dwindling profit margins led the company to stop marketing its long-distance services to consumers in 2004 (Atkin, Hallock, & Lau, 2006). Regulators assumed that the crux of breaking local telephone monopolies was to shatter the dominance of RBOCs in their respective local markets (Economides, 1998; Mauritsen, 2002). Without proper incentives, local telephone monopolies continued to retard the entry of long-distance service providers by protesting the FCC’s decisions every step of the way (Mauritsen, 2002).

Local telephone providers were thus poised to gain more from the 1996 Act than were long-distance providers, as nearly 40%of long-distance telephone company revenues were derived from calls that originated and terminated within RBOC territories (Mauritsen, 2002). The 1984 divestiture of AT&T helped create multiple regional monopolies. In a sense, the 1996 Act reversed the process by freeing these regional monopolies from proper regulation and turning them into bigger monopolies. In the end, the 1996 Act negated the 1984 settlement by allowing the Baby Bells to merge with each other. This highly deregulatory measure thus created favorable conditions for the emergence of another monopoly, one of the many ramifications for telecommunication media that are explored in the section to follow.

THE IMPACT OF THE 1996 ACT

Howard (1998) noted that the broadcasting industry witnessed a burst of merger and acquisition (M&A) activities, evidenced by the fact that 81.2% of television stations became group-owned in the top 100 media markets one year after the passage of the 1996 Act. In the radio industry, Wirth (2002) found that direct format competition in large media markets decreased 12% from 1995 to 2000. Wirth argued that the 1996 Act's claim to contribute to format diversity in radio programming cannot be supported. Similarly, the consolidation frenzy was found to suppress local news and music programming, as top 10 radio groups controlled the majority of the country's commercial radio stations shortly after the Act's passage (Huntemann, 1999).

Telephony

*Wired Telephony.*While there are some small differences among income brackets, the telephone penetration rate on average has reached a level of saturation. By the end of 2009, 95.7% of all U.S. households had telephone service (FCC, 2010). Wired telephone operators provide fixed local services, fixed long-distance services, carrier services, private network services, and other ancillary services, among which fixed local services are the major source of revenue (IBISWorld, 2010a). Due to the intense competition from such upstart rivals as VoIP and wireless telephony, the industry has been shrinking in terms of revenue. In 2011, industry revenue is projected to decline by almost 10% to $141.4 billion, and this trend is likely to persist for the next decade, further paring down the revenue to $97.0 billion by 2016 (IBISWorld, 2010a).

The 1996 Act brought about a slew of M&A activities that left 60% of America's 104 million households controlled by Verizon and Southwestern Bell Company (Atkin, Lau, & Lin, 2006). By early 2011, the two telecom giants – AT&T and Verizon – commanded 53.9% of industry revenue, and the top four firms including Qwest and Sprint Nextel account for 60.4% (CR4[*] = 60.4%) of industry revenue (IBISWorld, 2010a). Based on this index, the industry is considered to be moderately concentrated. The current level of concentration represents a slight increase from CR4 of 59.5% in 2002 (IBISWorld, 2010a). The slow increase in concentration can be explained by the dwindling domestic consumer base for wired telephony. Nevertheless, this concentration level is expected to increase during the next five years, as major players are poised to benefit from the economies of scale by acquiring smaller and less competitive players in the market (IBISWorld, 2010a).

As Bates et al. (2002) note, the major obstacle to robust competition in telephony lies in the local loop. In 2006, the local telephone loop was dominated by the RBOCs, as they

[*] CR4 stands for the concentration ratio of top four firms, which can be calculated by summing the percentages of the top four firms in the market.

controlled 98% of local telephone revenues and 90% of the 155 million phone lines (Atkin, Lin, & Lau, 2006). By 2011, the ILEC are expected to have a loosened grip on the number of phone lines. IBISWorld (2010a) estimates that ILECs will control 83% of access lines, a 7% drop from 2006. Nevertheless, ILECs are poised to benefit more from the deregulatory incentives that the 1996 Act provides in the next five to ten years.

At present, a number of entry barriers retard competitive local exchange carriers (CLECs) from gaining significant leverage against ILECs; they include regulatory, market, and access barriers. To compete, new entrants must obtain permission from the FCC and regulatory authorities at local levels. Even if they do obtain the license to compete, they must build and maintain capital intensive access networks (IBISWorld, 2010a). Further, CLECs will need to compete directly against the entrenched local giants and their assets in branding, business relations, and facilities. When unable to compete effectively with these regional giants, for instance, long-distance operators such as AT&T opted to merge with RBOCs (SBC). In so doing, RBOCs regained the territories lost to them in the1984 divestiture, by merging with each other or their former parenting company. In the end, the local loop remains under the tight grip of regional giants, leaving few prospects for enhanced competitiveness in the near term.

Wireless Telephony. Wireless telephony represents a robust growing sector. In 2010, the industry boasted 296 million wireless subscribers and revenues of $205 billion dollars (IBISWorld, 2010b). With the steady decrease in landline subscribership, it is projected that cell-only households (currently around 20% of U.S. homes) will keep increasing and become the primary medium for personal communication (Karlis&Tellen, 2010). Technological improvements such as 3G and 4G promise to deliver faster wireless services to customers, providing a boost to this growing potential. In concert with broadband technology, wireless operators are now able to offer a variety of multimedia services, which allows them to compete with other multimedia service providers. Similar to wired telephony, a few major operators dominate the wireless telephone market. As of 2010, Verizon Wireless and AT&T own 32.2% and 27.4% of the market share respectively. Sprint Nextel owns 14.4% of the market share and T-Mobile owns 11% of the market share (IBISWorld, 2010b). Together the top four firms take up 84% of the entire wireless market. With a CR4 of .84, the wireless market is considered highly concentrated. As wireless operators are vying for bigger market shares, the rate of concentration is unlikely to decrease in the next few years.

VoIP. The public switched telephone network (PSTN) is designed for voice-only traffic (Wei & Lee, 2008). Under PSTNs, a phone connection is operated in a closed circuit, which disallows other information from travelling simultaneously along the same line. As a result, traditional wired telephony can be inefficient and costly for long-distance connections. Voice over Internet Protocol (VoIP) offers several advantages over traditional PSTN-based wired telephony. First, telephone connections based on VoIP are much cheaper due to the elimination of toll charges associated with PSTN service (Wei & Lee, 2008). Efficiency is significantly increased because a channel can be used to send multiple packages of voice and data at the same time. Additionally, users can enjoy greater mobility as they're no longer tethered to a landline. Users can make phone calls wherever there is an Internet connection.

In 2010, the VoIP industry had 29.8 million subscribers and revenues of $12.3 billion (IBISWorld, 2010c). Some 70% of the VoIP services are currently provided by cable operators (Karlis&Tellen, 2010). The top four VoIP providers are all cable operators. With 31.5% of the market, Comcast surpasses other players in the VoIP market by a wide margin.

Time Warner Cable, Cox Communications, and Vonage own 16.2%, 9.6%, and 8.3% of the VoIP market respectively (IBISWorld, 2010c). With a CR4 of 65%, the VoIP market is considered to be concentrated at a medium level.

Broadcasting

Cable TV. The proliferation of cable TV was sped up by its commercialization. In 1964, there were only 1 million cable households. The number of cable households reached 10 million in ten years (Guillory, 2006). By the end of 2005, 60% of all TV households subscribed to cable TV (Guillory, 2006). In recent years, cable subscribership has been on the decline due to competition from other multichannel video programming distributors (MVPD). Nevertheless, there were still 64.7 million cable households in 2010 (NCTA, 2010).

Dating back to 1970, the cable TV industry has been exhibiting a comparable trajectory toward concentration. As Howard (1986) noted, the top ten cable operators were in possession of 42.5% of the industry revenues by 1985. In 2005, the top four cable firms controlled more than 60% of the industry revenues (Guillory, 2006). As McChesney (2005) noted, the concentration level would have been 90% if the FCC had approved Comcast and Time-Warner Cable's initiative to purchase Adelphia. The CR4 of cable TV industry has jumped to 70.6% (see Table 1), a net 10% increase since 2005. Although the decrease in cable subscriptions can partially account for the increase in CR4, cable markets today are much more concentrated than before. Similarly, the Herfindahl–Hirschman Index (HHI)—calculated by adding the squared market share of each firm--now reaches 1830, indicating a highly concentrated market.

As Guillory (2006) noted, cable providers' primary growth of revenue has not been in traditional programming services, but in digital video recording (DVR), VoIP, Internet, and business services. Bundling is an effective strategy for cable operators to aggregate services to gain profit margins. Crawford and Cullen (2007) estimated that consumers' surplus benefits would increase by 65.6% when cable operators offer a la carte rather than bundling services.

Bundling allows cable operators to tap into other telecommunication markets while maintaining a competitive edge at home. For instance, top cable operators currently are also major players in VoIP and Internet services. The vision that the 1996 Act's architects had for greater competition and content diversity by allowing cross-entry has failed to materialize in the domain of cable TV. Although cross-entry occurred in Internet telephony and Internet services, cable operators, and telephone companies have primarily held firm on their respective ground.

Radio. The 1996 Act eliminated radio ownership caps, prompting intensive consolidation in the radio industry. For instance, Clear Channel Communications at one time owned 1200 radio stations that reached 100 million listeners on a daily basis (Pitts, 2006). As of 2010, CC Media Holdings, the parent company of Clear Channel Communications, owns 894 radio stations and has a 17.3% of the market share (IBISWorld, 2010e). The second largest company, Sirius XM Radio Inc. – which consists of four satellites and 700 ground facilities,now controls 15.5% of the market. The third largest radio broadcaster, CBS Radio, operates 137 radio stations and has 8.9% of the market share. The fourth largest provider of radio services, Citadel Broadcasting Corporation, is in possession of 3.5% of the market share

(IBISWorld, 2010e). Additionally, Entercom Communications and Cox Enterprises Inc. control 2.5% and 2.3% of the market share respectively (IBISWorld, 2010e). As of 2010, CR4 for the radio industry is 44.2% and roughly 50% of the market is controlled by the top six companies.

Conglomeration in radio broadcasting did not lead to increases in productivity, but rather to severe financial failures. Both Clear Channel Communications and Citadel, the largest and third largest radio companies of the early 2000s, were steeped in huge debt (Dealbook, 2009; De La Merced, 2009). Recent economic analysis of deregulation on the broadcasting industry finds no evidence of enhanced efficiency and productivity. Due to market concentration, increased market power has contributed to increased return on sales, but operating efficiency and return on assets have dropped substantially after the 1996 act (Zhong, Cao, &Ning, 2008).

The Internet

In 2010, 75% of all Americans used the Internet and household penetration reached 72.8% (IBISWorld, 2010d). Internet access (72.2%) constitutes the bulk of the total $ 40.3 billion industry revenue. Equipment rentals, Website hosting, backbone services, and VoIP are among the primary services that ISPs are currently offering. Fueled by newer technologies, industry revenue is expected to keep growing in tandemwith increased adoption of broadband.

Internet service provision is one of the few domains in which telephone companies and cable operators are competing for subscribership. However, the competition has been limited to a few “big guns” in the telephone and cable industry. The Internet service provider (ISP) industry is highly concentrated (CR4 = 79.1%). As of 2010, AT&T occupies 34.2% of the ISP market. The second largest ISP provider, Comcast Corporation, had 20.3% of the market. With 12.6% and 12.1% of the market share respectively, Verizon Communications and Time Warner cable was the third and fourth largest internet provider (IBISWorld, 2010d).

The Internet has largely been unregulated in the U.S. (Brown & Halter, 2010). Scholars have argued for a common carrier approach to Internet interconnection (e.g., Specta, 2002). This would require ISPs to provide services to interconnection and end-users on a non-discriminatory basis. However, recent court decisions challenge this notion of common carriage. In 2009, Comcast’s delay of Web access to peer to peer (P2P) file sharing services prompted consumers to file complaints with the FCC. After holding hearings, the FCC subsequently ordered Comcast to stop the practice (BerQuist, 2010). However, the U.S. Court of Appeals for the D. C. Circuit stayed the order on the basis that the FCC had no statutory grounds to regulate ISP’s network practices (Comcast vs. FCC, 2010). This court decision emphasizes existing concerns over network neutrality, as this precedent opens up the possibility of discriminatory pricing by ISPs, leaving the openness of the Internet in question.

Broadband. The 1996 Act demands that the FCC facilitate the deployment of advanced telecommunications capabilities to all Americans (Telecommunications Act of 1996). Advanced telecommunications essentially means broadband Internet connection (Lombardi & Meadows, 2010). As broadband technologies have evolved, the definition of broadband has been changing accordingly. The most recent FCC definition of broadband sets the download speed at 4 Mbps and upload speed at 1 Mbps (FCC, 2010). The Organization for Economic

Co-Operation and Development (OECD) defines broadband as connections with a download speed of 256 Kbps (OECD, 2010a). With an advertised download speed at 105 Mbps and 101 Mbps, respectively, Japan and Portugal currently are leading OECD countries in broadband speed (OECD, 2010b). The U.S. is behind many OECD countries in advertised broadband speed, averaging 14 Mbps (OECD, 2010b).

As an umbrella term, broadband encompasses a variety of allied technologies including digital subscriber line (DSL), cable modem, satellite, fiber, broadband over power line (BPL), and wireless (Lombardi & Meadows, 2010). The OECD estimates that 58% of broadband connection is through DSL and 29% of broadband connection is through cable modem (OECD, 2010b). In the U.S., cable operators currently have the upper hand in their competition against telecoms, as 55% of wired broadband connection is through cable, 38% through DSL, and 6% through fiber optic cable (IBISWorld, 2010d). As of June 2010, Wireless broadband offered by telecommunications carriers had 136 million subscriptions and wired broadband had 83 million subscriptions (OECD, 2010b). In the meantime, wired broadband penetration has reached 27.1 subscribers per 100 inhabitants in the U.S. (OECD, 2010b).

The American Recovery and Reinvestment Act (ARRA) of 2009 appropriates $10.5 billion for communications, information, and security technologies, of which $ 7.5 billion were designated for deploying broadband and wireless Internet services. As part of ARRA, Congress urged the FCC to develop a comprehensive plan for broadband deployment and adoption (BerQuist, 2010). In response to Congress' request, the FCC introduced the National Broadband Plan (NBP) in 2010 after evaluating studies on broadband adoption. The primary goal of the plan is to provide 100 million U.S. households with access to the Internet at download speeds of 100 Mbps or more and upload speeds of 50 Mbps or more by 2020 (FCC, 2010a).

To encourage competition, the plan includes twenty recommendations about broadband networks, devices, and applications. In the domain of broadband networks, the plan recommends that the FCC take measures to make more radio spectrum available to wireless providers and market entrants, enabling them to compete with wireline broadband operators; the plan also recommends that the FCC collect more data to allow more detailed analyses of market competition and concentration (FCC, 2010a). As the plan notes, the FCC has only 50 free megahertz in reserve, only a fraction of the planned 500 megahertz free radio spectrum (FCC, 2010a). In terms of devices, the plan recommends that all MVPDs adopt a gateway device with equivalent functionality to replace existing set-top boxes (STBs) starting December 31, 2010.

The current competitive regime in the broadband industry is very fragile, if it exists at all, as 96% of the population is served by at most two wireline providers (FCC, 2010a). First, the broadband industry is a high fixed cost industry. As building broadband networks requires immense financial investment, new market entrants are unlikely to be able to compete with existing broadband providers. This is particularly true when existing broadband providers finish their roll-out of capital intensive broadband networks based on fiber optics such as fiber-to-the-home (FTHN) and fiber-to-the-node (FTTN). Second, the consumer cravings for speed is likely to eliminate the slower broadband technologies such as BPL and DSL, as it is already happening in some markets (Meadows, 2008). As a result, the market could be dominated by one or two companies who promise to provide top speeds (FCC, 2010a). In this vein, cable providers are in a favorable position, as their upgraded data over cable service interface specifications (DOCSIS 3.0) could offer up to 320 Mb/s downstream speeds. Third, wireless broadband based on 3G/4G services is unlikely to shrink the demand for wired broadband services (IBISWorld, 2010d). Wireless broadband suffers from several disadvantages such as the screen size, radio spectrum, programming, and technology specifications.

The NBP offers no remedies for these hurdles to competition. Even if they were to do so, the guidelines wouldonly be a recommendation without any regulatory or statutory jurisdiction. Given the existing court cases, broadband providers are likely to challenge the FCC on issues in which they have a high stake if the FCC decides to institutionalize their recommendations. Additionally, recommendations under the NBP plan are based on an isolated classification scheme rather than cross-industry dynamics. Owing to the 1996 Act, telecoms are offering both wireless and wired broadband services. In light of this, the first recommendation, in some ways, contradicts its purpose to facilitate cross-industry competition; because, both wireless and wireline broadband are provided by telecoms. Freeing up additional radio spectrum could not help new market entrants, but existing telecoms or market giants may benefit from such a policy.

Table 1. Concentration in Telephony, Broadcasting and Internet

Industry	Telephony			Broadcasting		Internet	
	Wireless	Wireline	VoIP	Cable	Radio	Wireline	Broadband
Top Four	Verizon	AT&T	Comcast	Comcast	CC Media	AT&T	Cable
	AT&T	Verizon	Time Warner	Time Warner	Sirius Radio	Comcast	DSL
	Sprint Nextel	Qwest	Cox	Cox	CBS Radio	Verizon	Fiber Optics
	T-Mobile	Sprint-Nextel	Vonage	Charter	Citadel	Time Warner	Other
CR4	84%	60.4%	65%	70.5%	45%	79%	59%

Source: Atkin et al., 2006; NTIA, 2010; IBISWorld, 2010a, 2010b, 2010c, 2010d, 2010e

Further Analysis

The 1996 Act was heralded by former president Bill Clinton as a revolutionizing force in bringing new technologies to our doorsteps (Clinton, 1996). However, a brief survey of thetelecommunications, broadcasting, and ISP industriesreveals a convergent pattern of consolidation after the enactment of the Telecommunications Act of 1996. As of 2010, CR4 readingsfor these industries were all above 60% except for the radio industry, indicating highly concentrated markets. Heightened concentration restrains competition and undermines the interests of consumers. Intensive concentration could severely disrupt existing business models and lead to business melt-down. Additionally, the diversity of information sources could be undermined by intensive concentration, weakening the basis of a representative democracy (e.g., McChesney, 2000).

The radio industry is not as concentrated as telecommunications and other broadcasting industries.A few years after the 1996 Act, uncurbed M&A activities led to immense debt in the radio broadcasting industry, including Clear Channel Communications and Citadel Broadcasting (Armstrong, 2010; Jeffrey, 2010). In the end, unsustainable business models characterized by expansive buyouts drove radio conglomerates to shrinkthe sizes of theirradio station fleets, resulting in a lesser degree of concentration. L. Lowry Mays, the chief executive officer of Clear Channel Communications, once argued before Congress that deregulation and economies of scale enabled his company to diversify radio offerings and contributed to a healthier, more robust industry (Lee, 2003). However, shortly after his testimony, the industry was steeped in turbulent financial waters. Research also suggests that consolidation led to a decrease rather than increase in radio offerings (Huntemann, 1999). The 1996 Act contributed to the radio industry melt-down by dismantling the ownership cap, which ultimately undermines the public interest by encouraging further consolidation.

Consumers' interests were more blatantly violated in the cable industry. Since the passage of the 1996 Act, cable rates have been rising at hyper-inflationary rates (Atkin et al., 2006). For instance, cable rates increased by 59% from 1996 to 2004 and another 21% between 2006 and 2008 (FCC, 2009). Average cable users watch only about six of the channels available to them, which poses the question of whether the increasedrates have produced a commensurate rise in value from the medium (Neuendorf,Jeffres, & Atkin, 2000). Additionally, concentration leads to limited options for consumers in a given media market. As a result, consumers often have to settle for bundling packages even if a la carte pricing has been shown to offer greater benefits (e.g., Crawford & Cullen, 2007).Market concentration typically increases a company's pricing power, leaving leading firms to freely increase rates while reducing the quality of products and services offered (Albarran&Dimmick, 1994; Chan-Olmsted &Litman, 1988). In this vein, the 1996 Act defeats its own stated raison d'etreof protecting the interests of consumers; that is, cable conglomerates–not consumers--have been benefiting from the deregulation.

In particular, media merger and acquisition activity reached record levels in the wake of the 1996 Act during late 1990s, contributing to unprecedented levels of within-industry concentration. Following Comcast cable's 2009 merger with AT&T Broadband, the combined company serves 25% of multichannel viewers. This level of concentration, in and of itself, would have been sufficient to trigger Justice Department scrutiny in days past (e.g., Atkin et al., 2006; Chan-Olmsted &Litman, 1988), leaving the top three firms in command of two-thirds of the nation's cable market. Commentators (McChesney, 2004) observe that the impact of these mergers was greatest in cable distribution, where HBO and the three next largest channels command nearly 90% of pay cable revenue. Abelman and Atkin (2011, p. 151) concluded that this merger activity, in concert with acquisitions in telephony, leaves "the audience's essential communication lines under the control of two powerful oligopolies".

Similar to cable TV, telephony, particularly wireless telephony, is highly concentrated. The 1996 Act does not provide enough incentives tobreak the closed local loop. Instead, deregulation helped ROBCs to regain their territories lost in the 1984 divestiture. Cross-entry as the architects of the 1996 Act envisioned did not occur as telecoms and cable operators have held firm to their respective territories. In the end, consumers did not benefit from deregulation, as increased media offerings were stymied by failed cross-entry.

Few would dispute that the free flow of information and diversity of voices is critical to the maintenance of a representative democracy (Gilens & Hertzman, 2000; Huntemann, 1999; McChesney, 2004; Napoli, 1999). As Napoli (1999) noted, there are three kinds of diversity: source, content, and exposure. Merger mania facilitated by the 1996 Act has produced mega-corporations owning numerous media outlets across different industries and markets. Source and content diversity could be severely undermined if this consolidation is not properly harnessed. Research already shows that source and content diversity has been eroded by the growing influence of a handful of industry potentates. For instance, newspapers owned by media conglomerates are more likely to slant news in which their parent companies have a stake (Gilens & Hertzman, 2000).

Content diversity is negatively influenced when media conglomerates, such as Clear Channel, cut costs by sharing nationally syndicated programs across their networks rather than airing local educational and music programming (Huntemann, 1999). Source diversity is also hampered by both Congressional directives and the 1996 Act. Ethnic radio formats are shrinking as stations switch to formats that cater to listeners with high consumption habits (e.g., Huntemann, 1999; Wirth, 2002). In 1978, Congress eliminated the tax certificate program, which dampened the transfer of ownership to minorities (Huntemann, 1999). The program allowed sellers of broadcast properties to defer payment of the capital gains tax when they sold their license to minorities.

The only competitive domain is that of the ISP, where telecoms and cable operators are already competing to deliver Internet services. Cable operators have an upper hand in high-speed internet services while telecoms have a bigger market share of the traditional wirelineISP business (IBISWorld, 2010d). Broadband is becoming the battleground between cable and telecom operators. Cable operators' DOCSIS 3.0 technology promises to offer much faster download speeds than telecoms' FTTN and FTTH. As such, some argue that telecoms may not be able to compete effectively with cable operators in providing wireline broadband once they finish deploying DOCSIS 3.0 networks (IBISWorld, 2010d). However, telecoms could counter the competition from cable operators by expanding their wireless broadband services, which depends eventually on the progress of 3G and 4G networks rollout.

The relative competitiveness of the ISP industry seems to provide a counter-factual case for understanding the 1996 Act's impact. However, competition in the ISP industry is limited to a few "big guns" from the cable and telephone industries. Although cable and telecom giants are competing to deliver Internet services, the diversity of sources remains stagnant as the majority of consumers are served, at most, by two providers in a given market (FCC, 2010a). It should be noted that the 1996 Act was introduced when the Internet was in its infancy. As a result, there remains a regulatory delay on the issues pertainingto Internet oversight. Because major telephone companies are also among the top ISP providers, they are offering two services that are under different regulatory regimes. For instance, telecommunications services are regulated under the 1996 Act, while ISP services are largely unregulated (IBISWorld, 2010d). This is a significant result because this differential allows ISP providers to avoid contribution to universal service funds and other access charges (IBISWorld, 2010d).

As broadband essentially means high-speed Internet, the regulatory lag looms large in the age of broadband.The aging regulations have proven ineffective in coping with the emerging challenges due to convergence and technological innovations. The debates on broadband regulation have focused, of late, on the issue of net neutrality. A recent court decision implies that broadband operators may keep fighting over the rights to regulate net traffic (Comcast v. FCC). The FCC, to date, has not provided any regulatory incentives to facilitate competition in broadband-related industries. The National Broadband Plan does not offer any concrete suggestions regarding how to enhance competition between telecoms and cable operators. As a result, the broadband industry will be as concentrated as the ISP market unless the FCC and Congress are willing to weigh in on these issues. After fourteen years, technological innovations and failures in the existing regulatory scheme make it necessary to revisit the 1996 Act. In particular, Congress and the FCC need to ensure that regulatory measures are taken to safeguard consumers' interests while maintaining a competitive marketplace.

CONCLUSION

In sum, the 1996 Act contributed to a heightened level of M & A activities in the telephone,broadcasting, and ISP industries. As Noam (1999) noted, the intense concentration dynamic cuts across major media industries The 1996 Actthus represents a seminal piece of legislation, one that was designed to break local monopolies in local telephony and bring competition to cable and telephony by allowing them cross-enter each other's market. However, the 1996 Act not only failed to deliver its major promises,but it has actually dampened competition by dismembering the competitive mechanisms that were put in place by the Communications Act of 1934. Additionally, the wide diffusion of the Internet, along with broadband technologies, brings home new regulatory challenges. Aging regulatory schemes governing the cable and telecom industries have becoming increasingly ineffective in meeting these new challenges. The ISP broadband industries remain largely unregulated. Currently, the ISP industry is highly concentrated. Broadband, as an extension of the ISP industry, is likely to replicate the existing business model. If left unchecked, concentration in broadband, the backbone of knowledge-based economy, could jeopardize consumer interests, competition, and ultimately the diversity of voices upon which we depend for the vibrancy of our democracy.

REFERENCES

Armstrong, R. (2010). Clear Channel's looming debt crunch is a confidence test. Retrieved February 01, 2011, from http://blogs.wsj.com/deals/2010/12/30/clear-channels-looming-debt-crunch-is-a-confidence-test/

Abelman, R. & Atkin, D. (2011).*The televiewing audience: The art and science of watching television* (2nded.). New York: Peter Lang.

Albarran, A., & Dimmick, J. (1994).Concentration and economies of multiformity in the communication industries. *Journal of Media Economics*, 9, 41-62.

Atkin, D. (1999). Videodialtone reconsidered: Prospects for competition in the wake of the Telecommunications Act of 1996. *Communication Law and Policy Journal*, 4, 35-58.

Atkin, D., Hallock, J. & Lau, T.Y. (2006).Telephony.In A. Grant & J. Meadows (Eds.). *Communication Technology Update*. Boston: Focal, pp. 273-283.

Atkin, D., Lau, T.Y. & Lin, C.A. (2006). Still on hold: Prospects for competition in the wake of the Telecommunication Act of 1996 on its 10 year anniversary. *Telecommunications Policy*, 30, 80-95.

Bates, B. J., & Chambers, T. (1999).The economic basis for radio deregulation. *Journal of Media Economics*, 19, 12-22.

Bates, B., Jones, K. A., & Washington, K. D. (2002). Not your plain old telephone: New services and new impacts. In C. Lin, & D. Atkin (Eds.), Communication technology and society: Audience adoption and uses (pp. 91-124). Cresskill, NJ: Hampton.

Bednarski, A. (2003). From diversity to duplication: Mega-mergers and the failure of the marketplace model under the Telecommunications Act of 1996. Retrieved April 24, 2010, from http://www.law.indiana.edu/fclj/pubs/v55/no2/bednarski.pdf.

BerQuist, L. (2008). Communication policy & technology.In A. Grant & J. Meadows (Eds.). *Communication technology updates* (pp. 66-76). Boston: Focal Press.

BerQuist, L. (2010). Communication policy & technology.In A. Grant & J. Meadows (Eds.). *Communication technology updates* (pp. 71-80). Boston: Focal Press.

Brock, G. (1994).*Telecommunication policy for the information age: From monopoly to competition*. Cambridge, MA: Harvard University Press.

Brown, T., & Halter, H. (2010).The Internet.In A. Grant & J. Meadows (Eds.). *Communication technology updates* (pp. 281-291). Boston: Focal Press.

Chan-Olmsted, S. M. (1998). Mergers, acquisitions, and convergence: The strategic alliances of broadcasting, cable television, and telephone services. *Journal of Media Economics,* 11(3), 33-46.

Chan-Olmsted, S., & Kang, J. W. (2004).Theorizing the strategic architecture of a broadband television industry. *Journal of Media Economics*, 16, 3-16.

Chan-Olmsted, S., & Litman, B. (1988).Antitrust and horizontal mergers in the cable industry. *Journal of Media Economics*, 1, 3-24.

Clinton, B. (1996). Remarks by the president at the signing ceremony for the Telecommunications Act conference report. Retrieved February, 10, 2010, from http://www.whitehouse.gov/WH/eop/op/telecom/release.html.

Comcast V. *FCC*. No. 08-1291 (D.C. Cir., 2010). Retrieved February, 10, 2010, from http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-297351A1.pdf.

Crawford, G. S., & Cullen, J. (2007). Bundling, product choice, and efficiency: Should cable television be offered a la carte. *Information Economics and Policy*, 19, 379-404.

Dibadj, R. (2003). Competitive debacle in local telephony: Is the 1996 Telecommunications Act to blame. *Washington University Law Quarterly, 81*, 1-65.

De la Merced, M. J. (2009). Citadel broadcasting plans prearranged bankruptcy. http://dealbook.nytimes.com/2009/12/10/citadel-broadcasting-plans-prearranged-bankruptcy-filing/?ref=clearchannelcommunicationsinc.

Dealbook. (2009). Clear Channel Outdoor to Sell $750 Million Notes http://dealbook.nytimes.com/2009/12/11/clear-channel-outdoor-to-sell-750-million-notes/?ref=clearchannelcommunicationsinc.

Economides, N. (1998). The Telecommunications Act of 1996 and its impact.*Japan and the World Economy*, Retrieved April 20, 2010 from http://www.stern.nyu.edu/networks/98-08.pdf.

Federal Communications Association.(2010). Sixth broadband deployment report. Retrieved February 10, 2011, from http://www.fcc.gov/Daily_Releases/Daily_Business/2010/db0720/FCC-10-129A1.pdf.

Federal Communications Association. (2010a). The national broadband plan. Retrieved February 10, 2011, from http://download.broadband.gov/plan/national-broadband-plan.pdf.

Galperin, H. (2002). Can the US transition to digital TV be fixed? Some lessons from two European Union cases. *Telecommunications Policy*, 26, 3-15.

Gilens, M., & Hertzman, C. (2000). Corporate ownership and news bias: Newspaper coverage of the 1996 Telecommunications Act. *The Journal of Politics*, 62(2), 369-386.

Grant, A. E. (2010). The structure of the communication industries.In A. Grant & J. Meadows (Eds.).*Communication technology updates* (pp. 57-70). Boston: Focal Press.

Guillory, B. (2006). Cable television.. In A. Grant & J. Meadows (Eds.). *Communication technology updates* (pp. 70-82). Boston: Focal Press.

Hazlett, T. W. (1998). Assigning property rights to radio spectrum users: Why did FCC license auctions take 67 years? *Journal of Law and Economics*, 41(2), 529-575.

Howard, H. H. (1998). The 1996 Telecommunications Act and TV station ownership: 1 year later. *Journal of Media Economics*, 11(3), 21-32.

Huntemann, N. (1999). Corporate interference: The commercialization and concentration of radio post the 1996 Telecommunications Act. *Journal of Communication Inquiry*, 23, 39-407.

IBISWorld. (2010a). Wired telecommunications carriers in the U.S. (No. 51331b): IBISWorld.

IBISWorld. (2010b). Wireless telecommunications carriers in the U.S. IBISWorld.

IBISWorld. (2010c). Cable, Internet & Telephone providers in the U.S. (No. 51322): IBISWorld.

IBISWorld. (2010d). Internet service providers in the U.S. (No. 51419a): IBISWorld.

IBISWorld. (2010e). Radio broadcasting in the U.S. (No. 51311): IBISWorld.

Jeffrey, D. (2010). Citadel broadcasting wins bankruptcy plan approval (update2). Retrieved February 01, 2011, from

http://www.businessweek.com/news/2010-05-18/citadel-broadcasting-wins-bankruptcy-plan-approval-update2-.html.

Karlis, J., &Telleen, M. (2010).Telephony.In A. Grant & J. Meadows (Eds.). *Communication technology updates* (pp. 264-280). Boston: Focal Press.

Krattenmaker, T.G. (1996). The Telecommunications Act of 1966.*FederalCommunications Law Journal*, 49, 1-49.

Krattenmaker, T. G., & Salop, S. (1986). Anticompetitive exclusion: Raising rivals' costs to achieve power over price. *The Yale Law Journal*, 96, 209-293.

Lee, J. (2003). Radio giant defends its size at senate panel hearing. Retrieved February 10, 2010, from http://www.nytimes.com/2003/01/31/business/media/31RADI.html.

Lombardi, J. J., & Meadows, J. H. (2010). Broadband & home networks. In A. Grant & J. Meadows (Eds.).*Communication technology updates* (pp. 305-321). Boston: Focal Press.

Mauristen, R. H. (2002). The Telecommunications Act of 1966: Relationship to functional theory. *Perspectives Journal, Proceedings of AABBS.*

McChesney, R. W. (2000). *Rich media, poor democracy: Communication politics in dubious times*. New York: New Press.

McChesney, R. W. (2004). *The problem of the media*. New York: Monthly Review Press.

McChesney, R. W. (1997). *Corporate media and the threat to democracy*. New York, Seven Stories.

Meadows, J. H. (2010).Multichannel television services.In A. Grant & J. Meadows (Eds.). *Communication technology updates* (pp. 101-116). Boston: Focal Press.

Meadows, J. H. (2008).Broadband & home networks.In A. Grant & J. Meadows (Eds.). *Communication technology updates* (pp. 303-320). Boston: Focal Press.

Melody, W. H. (1999). Telecom reform: Progress and prospects. *Telecommunications Policy*, 23, 7-34.

Neuendorf, K., Jeffres, L., & Atkin, D. (2000). The television of abundance arrives: Cable choice and interest maximization. *Telematics & Informatics*, 17, 169-197.

Napoli, P. M. (1999). Deconstructing the diversity principle. *Journal of Communication,* 49(4), 7-34.

Noam, E. (2009). *Media ownership and concentration in America*. New York: Oxford University Press.

OECD. (2010a). OECD broadband subscriber criteria. Retrieved February 10, 2010, from http://www.oecd.org/document/46/0,3746,en_2649_33703_39575598_1_1_1_1,00.html.

OECD. (2010b). OECD broadband portal. Retrieved February 10, 2010, from http://www.oecd.org/document/54/0,3746,en_2649_33703_38690102_1_1_1_1,00.html.

Pickard, V. (2007). In T. M. Schaefer & T. A. Birkland (Eds.), *The encyclopedia of media and politics.* Washington, DC: CQ Press.

Pitts, G. (2006). Radio broadcasting.In A. Grant & J. Meadows (Eds.). *Communication technology updates* (pp. 139-154). Boston: Focal Press.

Shaw, J. (2002). *Telecommunication deregulation and the information economy* (2nd ed.). Boston: Artech.

Specta, J. B. (2002). A common carrier approach to Internet interconnection.*Federal Communications Law Journal*, 54, 225-280.

Telecommunications Act of 1996, 104 Pub. L. 104, 110 Stat. 56, 111. (1996) (codified as amended in 47 C.F.R. S. 73.3555).

U.S. v. AT&T, (1982).552 F. Supp. 131, 195.(D.D.C. 1982).aff'd sub nom. Maryland v. US, 460 US 1001 (1983).

Warf, B. (2007). Oligopolization of global media and telecommunications and its implications for democracy. *Ethics, Place & Environment*, 10(1), 89-105

Wei, R., & Lee, Y. (2008).Telephony.In A. Grant & J. Meadows (Eds.). *Communication Technology Updates* (pp. 245-267). Boston: Focal Press.

Wirth, D. L. (2002). Direct format competition on the radio dial and the Telecommunications Act of 1996: A five-year trend study. *Journal of Radio Studies, 9*, 33-50.

Zaragoza, R. R., Bodorff, R. J., & Emord, J. W. (1988). The public interest concept transformed: The trusteeship model gives way to a marketplace approach. In J. T. Powell & W.Gair (Eds.), *Public interest and the business of broadcasting* (pp. 30-45). London: Quorum Books.

Zhang, B. (2003). Understanding the impact of convergence on broadband industry regulation: A case study of the United States. *Telematics & Informatics*, 19, 37-53.

Zhong, K., Cao, Y., &Ning, Y. (2008).The deregulatory effects of the Telecommunications Act of 1996 on the broadcasting industry: Expectations vs. reality. *Journal of Accounting and Public Policy 27*, 238-261.

In: Advances in Communications,Volume 8
Editor: Anthony V. Stavros
ISBN: 978-1-61324-794-5

Chapter 5

E-COMMERCE SECURITY IN THE HASHEMITE KINGDOM: CALIBRATING JORDAN'S ELECTRONIC TRANSACTIONS LAW

Stephen E. Blythe[1]
School of Management
New York Institute of Technology
CERT Technology Park, P.O. Box 5464
Al Muroor Road, Abu Dhabi
United Arab Emirates

Abstract

Jordan's economy has been growing in recent years due in no small part to a most-favorable trade treaty with the United States. E-commerce emerged after the enactment of the Electronic Transactions Law ("ETL") in 2001. The ETL covers most E-commerce transactions, but is inapplicable to wills, real property deeds, powers of attorney, termination of insurance contracts and utility services, court documents and securities. The statute recognizes the enforceability and admissibility of electronic documents and signatures. The electronic form may be used to comply with statutory requirements pertinent to writing, originality, retention and signing. The ETL contains E-contract rules concerning attribution, acknowledgement of receipt, and the time and place that an electronic message can be assumed to have been sent and received. In an unusual but admirable move, the statute includes rules regarding the transfer of electronic notes and electronic funds transfers. The ETL contains third-generation E-signature rules; the utilization of stringent security procedures distinguishes secure E-signatures from insecure ones. The ETL lists several computer crimes: fraudulent use of a certificate; submission of false information to a Certification Authority ("CA"); offenses by CA's; and commission of other crimes using electronic means. Is the ETL up-to-date according

[1] E-mail Address: itlawforever@aol.com

to current trends in international E-commerce law? Not quite. The most significant of the recommended amendments are to: (1) eliminate the list of exclusions and recognize the legal validity of electronic wills; (2) add E-contract rules pertinent to automated and carriage contracts; (3) add consumer protections for E-commerce buyers; (4) establish Information Technology tribunals as a court-of-first-instance for E-commerce disputes; (5) claim "long arm" jurisdiction over foreign parties in E-commerce transactions; (6) designate the Jordanian Post Office as a licensed CA; and (7) add several new computer crimes.

Keywords: Jordan, E-Commerce, E-Signature, Transactions, Law.

AN INTRODUCTION TO THE HASHEMITE KINGDOM OF JORDAN

Recent History

Jordan attained independence from Britain in 1946 and was ruled by King Hussein from 1953 until his death in 1999. King Hussein was an intelligent ruler and successfully handled attempts to influence him from the United States, the Soviet Union, Israel, his Arab neighbors, and a large contingent of Palestinians residing in his country. Parliamentary elections were resumed in 1989 and political liberalization began. A peace agreement was reached with Israel in 1994. King Abdullah II succeeded his father in 1999 and has implemented substantial economic reforms. Jordan became a member of the World Trade Organization ("WTO") in 2000 and the European Free Trade Association in 2001,[1] and a bilateral free trade agreement was signed with the U.S. in 2001.[2] Despite the fact that 92% of its residents are Sunni Muslims, Jordan supported the Coalition's invasion of Iraq and the overthrow of Saddam Hussein in 2003; the country also accepted thousands of displaced Iraqis after the violent insurgency began. In the municipal elections of 2007, 20% of the city council seats were reserved for women, a dramatic move toward equal gender rights in the Middle East. After Prime Minister Nader Al-Dahabi was appointed in 2007, King Abdullah directed him to focus on implementation of economic reforms, the improvement of healthcare and housing for both civilians and the military, and strengthening the educational system.[3]

The Economy

Recently, Jordan has experienced chronic unemployment (over 12%), poverty (more than 14%) and inflation (almost 15%). However, its relationships with the World Trade

---

[1] U.S. Central Intelligence Agency ("CIA"), *The World Factbook*, "Jordan," 2009, pp. 1, 4, 5; http://www.cia.gov/library/publications/the-world-factbook/geos/countrytemplate_jo.html.

[2] This free trade agreement will phase out duties on almost all goods and services by January, 2010. U.S. Department of State, Bureau of Near Eastern Affairs, *Background Note: Jordan*, February, 2009, p. 3; http://www.state.gov/r/pa/ei/bgn/3464.htm . *See* Bashar H. Malkawi, "E-Commerce in Light of International Trade Agreements: The WTO and the United States-Jordan Free Trade Agreement," 15:2 *International Journal of Law and Information Technology* 153-169 (2007); http://ijlit.oxfordjournals.org/cgi/content/abstract/15/2/153 .

[3] Note 1 supra.

Organization, its advantageous trade treaties with the United States and Europe, and its ongoing economic reforms justify hope for a brighter future. The country's gross domestic product was estimated to be $31.6 billion in 2008, an increase of 5.6% over the previous year. Jordanians are relatively well educated, with an average of thirteen years of schooling. The country successfully completed an International Monetary Fund ("IMF") program in 2002 and has since adhered to IMF guidelines pertinent to monetary funds. Privatization of formerly public entities has continued at a steady pace, and in recent years the Jordanian government has successfully applied privatization proceeds (in conjunction with the ending of government subsidies to consumer goods and the oil industries) to achieve a reduction of its national debt. All of these positive measures have increased productivity and have made Jordan a more attractive target for foreign direct investment. In the future, Jordan will be challenged to continue to create jobs, reduce the budgetary deficit, attract foreign investment and reduce dependence on foreign grants. Jordan's banking sector seems to have incurred little damage from the worldwide financial crisis, but 2009 has seen a slowdown in its tourism and real estate markets.[1]

Jordan is hampered by a scarcity of oil, water and other natural resources. The country's traditional sources of income were from exports of phosphates and potash, overseas remittances and foreign aid. However, during the past eight years, exports of garments and textiles that enter the U.S. and Europe with little or no tariffs have been the primary source of economic growth.[2] Many of those goods are manufactured in the Aqaba Special Economic Zone, a "liberalized, low tax duty-free and multi-sector development zone…in a strategic location on the Red Sea…"[3] Other industries include fertilizers, pharmaceuticals, petroleum refining, cement, inorganic chemicals light manufacturing and vegetables.[4] Tourism, a burgeoning industry, is also expected to be a significant source of future income; this trend was given a boost when the ruins of Petra, an ancient city, were designated one of the "New Seven Wonders of the World" in 2007.[5]

The Internet and E-Commerce

To date, the rate of internet penetration has been modest. In a population of about 6.3 million, only 1.5 million use the internet, amounting to an internet penetration rate of only 24%. However, there has been a 300% growth in internet penetration during the past three years. Jordan has 28,896 internet hosts, a 9-fold increase during the past four years. This

---

[1] Note 1 supra at 5, 7, 8.
[2] Note 1 supra at 7.
[3] Aqaba Special Economic Zone Authority, home page of website; http://www.aqabazone.com/ .
[4] Note 1 supra at 9, and Note 2 supra at 1.
[5] "New Seven Wonders of the World," *Wikipedia: The Free Encyclopedia*;
http://en.wikipedia.org/wiki/New_Seven_Wonders_of_the_World. The other six are: Chichez Itza in Mexico; Christ the Redeemer in Rio de Janeiro, Brazil; the Colosseum in Rome, Italy; the Great Wall of China; Machu Picchu in Peru; and the Taj Mahal in Agra, India. Id. Tourism is on the rise, notwithstanding the fact that the "threat of terrorism remains high in Jordan." U.S. Department of State, *Jordan: Country Specific Information*, 6 May 2009, p. 2;
http://travel.state.gov/travel/cis_pa_tw/cis/cis_1149.html .

positive trend is expected to continue due to governmental promotion of internet usage and its development of the information technology sector.[1]

As the degree of internet penetration grows, so does E-commerce. During 2008, retail E-commerce in Jordan was an estimated US $181 million.[2] Jordan is especially conducive to the growth of E-commerce because of its absence of legal restrictions, no E-commerce taxes and a lenient bankruptcy statute which provides an incentive for entrepreneurs to undertake the risk of creation of an online business.[3] In an empirical study in 2005 of all Jordanian firms having a website, the authors concluded that "Jordan has adequate and efficient e-commerce requirements" and that two possible scenarios could lead to a breakout of E-commerce and its widespread adoption by Jordanian firms: either the private sector could take the lead and the government could follow, or the government could take the lead and the private sector could follow.[4] The article predicted the latter would happen because of the Jordanian government's announced plan to open an E-government website by 2007,[5] and the article's prediction has come true; since 2007, more and more private firms have gone online.[6] Despite the dramatic growth in E-commerce during the past several years, Jordanian E-commerce and internet banking remain hampered by security issues, a problem that could and should have been solved with adoption and use of electronic signatures early in this decade.[7]

ELECTRONIC SIGNATURES

Contract law worldwide has traditionally required the parties to affix their signatures to a document.[8] With the onset of the electronic age, the electronic signature made its appearance. It has been defined as "any letters, characters, or symbols manifested by electronic or similar

[1] Note 1 supra at 2, 3, 9, 12. For the results of a comprehensive online survey of Jordanian internet users (with 931 respondents), *see* Arab Advisors Group, *Jordan Internet Users and E-Commerce Survey* 2008; http://www.arabadvisors.com/consulting-report/11695 .

[2] International Telecommunications Union, *ICT Statistics Newslog*, 2 December 2008; http://www.itu.int/ITU-D/ict/newslog/ .

[3] Samir Ghawl, "Chances of Success in E-Commerce Higher, US Expert Tells Jordanians," *Jordan Times*, 29 August 2007; http://www.jordanembassyus.org/08292007003.html . Successful E-commerce entrepreneurs in Jordan have employed a variety of approaches. *See* Gaith Seqer, "E-Commerce in Jordan: Different Approaches From 3 Startups and From Jabbar and Aramex," *Arabcrunch*, 29 August 2009;
http://arabcrunch.com/2009/08/e-commerce-in-jordan-different-approaches-from-3-startups-and-from-jabbar-and-aramex.html .

[4] Rifat O. Shannak and Mu'taz M. Al-Debei, "The Current State of E-Commerce in Jordan: Applicability and Future Prospects – An Empirical Study," *Proceedings of the Fifth Ibima Conference on the Internet and Information Technology in Modern Organizations*, Cairo, Egypt, November 13-15, 2005, p. 473;
http://bura.brunel.ac.uk/handle/2438/2888 .

[5] Id.

[6] Note 1 supra at 12.

[7] Ahmad Khasawneh, Iyad Al Azzam and Mohammad Bsoul, "A Study on E-commerce Security in Jordan," 3:2 *International Journal of Electronic Finance* 166-176 (2009);
http://portal.acm.org/citation.cfm?id=1552230 .

[8] See, e.g., U.C.C. Sect. 2-201, 2-209 (1998).

means and executed or adopted by a party with the intent to authenticate a writing,"[1] or as "data in electronic form which are attached to or logically associated with other electronic data and which serve as a method of authentication."[2] An electronic signature may take a number of forms: a digital signature, a digitized fingerprint, a retinal scan, a pin number, a digitized image of a handwritten signature that is attached to an electronic message, or merely a name typed at the end of an e-mail message.[3]

A well-known U.S. consumer group has stated, "Given the current state of authentication technology, it's much easier to forge or steal an e-signature than a written one."[4] This statement seems to assume that all E-signatures offer an equal degree of security. However, such an assumption would be erroneous; some electronic signatures offer more security than others. It is prudent for E-Commerce participants to use the more secure types of electronic signatures, notwithstanding their greater degree of complexity and expense.

Online Contracts: Four Levels of Security

When entering into a contract online, four degrees of security are possible.

1. The first level would exist if a party accepted an offer by merely clicking an "I Agree" button on a computer screen.[5]
2. The second level of security would be incurred if secrets were shared between the two contracting parties. This would be exemplified by the use of a password or a credit card number to verify a customer's intention that goods or services were to be purchased.[6]
3. The third level is achieved with biometrics. Biometric methods involve a unique physical attribute of the contracting party, and these are inherently extremely difficult to replicate by a would-be cyber-thief. Examples include: a voice pattern, face recognition, a scan of the retina or the iris within one's eyeball, a digital reproduction of a fingerprint,[7] or a digitized image of a handwritten signature that is attached to an electronic message. In all of these examples, a sample would be taken from the person in advance and stored for later comparison with a person purporting to have

[1] Thomas J. Smedinghoff, "Electronic Contracts: An Overview of Law and Legislation," 564 PLI/P at 125, 162 (1999).

[2] European Union Directive 1999/93/Ec of the European Parliament and of the Council Ïf 13 December 1999 On A Community Framework For Electronic Signatures, (1999/93/EC) – 19 January 2000, OJ L OJ No L 13 p.12.

[3] David K.Y. Tang, "Electronic Commerce: American and International Proposals for Legal Structures," in *Regulation and Deregulation: Policy and Practice in the Utilities and Financial Services Industries* 333 (Chrisopher McCrudden ed., 1999).

[4] Michael Dessent, "Browse-Wraps, Click-Wraps and Cyberlaw: Our Shrinking (Wrap) World," 25 T. *Jefferson L. Rev*. 1, 4 (2002).

[5] Jonathan E. Stern, Note, "Federal Legislation: The Electronic Signatures in Global and National Commerce Act," 16 *Berkeley Tech. L.J.* 391, 395 (2001).

[6] Id.

[7] In the highly successful Hong Kong Identity Card, the two thumb prints are used as a biometric identifier. See, Rina C.Y. Chung, "Hong Kong's 'Smart' Identity Card: Data Privacy Issues and Implications for a Post-September 11th America," 4 *Asian-Pacific L. & Pol'y J.* 442 (2003).

the same identity.[1] For example, if a person's handwriting was being used as the biometric identifer, the "shape, speed, stroke order, off-tablet motion, pen pressure and timing information" during signing would be recorded, and this information is almost impossible to duplicate by an imposter.[2]

Biometrics, despite its potential utility as a form of electronic signature, has at least two drawbacks in comparison with the digital signature: (1) The attachment of a person's biological traits to a document does not ensure that the document has not been altered, i.e., it "does not freeze the contents of the document;"[3] and (2) The recipient of the document must have a database of biological traits of all signatories dealt with in order to verify that a particular person sent the document.[4] The digital signature does not have these two weaknesses and most seem to view the digital signature as preferable to biometric identifiers.[5] Many also recommend the use of both methods; this was the course taken by the Hong Kong government in designing its identity card.[6]

4. The digital signature is considered the fourth level because it is more complex than biometrics. Many laypersons erroneously assume that the digital signature is merely a digitized version of a handwritten signature. This is not the case, however; the digital signature refers to the entire document.[7] It is "the sequence of bits that is created by running an electronic message through a one-way hash function and then encrypting the resulting message digest with the sender's private key."[8] A digital signature has two major advantages over other forms of electronic signatures: (1) it verifies authenticity that the communication came from a designated sender; and (2)

---

[1] Note 6 supra at 395-96; and "The Legality of Electronic Signatures Using Cyber-Sign is Well Established," CYBER-SIGN, at http://www.cybersign.com/news news.htm

[2] Id.

[3] K.H. Pun, Lucas Hui, K.P. Chow, W.W. Tsang, C.F. Chong & H.W. Chan, "Review of the Electronic Transactions Ordinance: Can the Personal Identification Number Replace the Digital Signature?," 32 *Hong Kong L.J.* 241, 256 (2002).

[4] Id. at 257.

[5] Id. However, one of the experts in computer law and technology—Benjamin Wright—is a notable exception. Wright contends that biometrics is a more preferable authentication method in the case of the general public, although he concedes that digital signatures using PKI are preferable for complex financial deals carried out by sophisticated persons. In PKI, control of the person's "private key" becomes all-important. The person must protect the private key; all of the "eggs" are placed in that one basket, and the person carries a great deal of responsibility and risk. With biometric methods, the member of the general public would be sharing the risk with other parties involved in the transaction, and the need to protect the "private key" is not so compelling. See, Benjamin Wright, "Symposium: Cyber Rights, Protection, and Markets: Article, 'Eggs in Baskets: Distributing the Risks of Electronic Signatures,'" 32 *West L.A. L. Rev.* 215, 225-26 (2001).

[6] Note 23 supra.

[7] The Hong Kong E-commerce law typically defines a digital signature as follows: "an electronic signature of the signer generated by the transformation of the electronic record using an asymmetric cryptosystem and a hash function such that a person having the initial untransformed electronic record and the signer's public key can determine: (a) whether the transformation was generated using the private key that corresponds to the signer's public key; and (b) whether the initial electronic record has been altered since the transformation was generated." Hong Kong Special Autonomous Region, *Electronic Transactions Ordinance*, Ord. No. 1 of 2000, s 2.

[8] Note 17 supra at 146.

it verifies the integrity of the content of the message, giving the recipient assurance that the message was not altered.[1]

Digital Signature Technology: Public Key Infrastructure

The technology used with digital signatures is known as Public Key Infrastructure, or "PKI."[2] PKI consists of four steps:

1. The first step in utilizing this technology is to create a public-private key pair; the private key will be kept in confidence by the sender, but the public key will be available online.
2. The second step is for the sender to digitally "sign" the message by creating a unique digest of the message and encrypting it. A "hash value" is created by applying a "hash function" – a standard mathematical function – to the contents of the electronic document. The hash value, ordinarily consisting of a sequence of 160 bits, is a digest of the document's contents. Whereupon, the hash function is encrypted, or scrambled, by the signatory using his private key. The encrypted hash function is the "digital signature" for the document.[3]
3. The third step is to attach the digital signature to the message and to send both to the recipient.
4. The fourth step is for the recipient to decrypt the digital signature by using the sender's public key. If decryption is possible the recipient knows the message is authentic, i.e., that it came from the purported sender. Finally, the recipient will create a second message digest of the communication and compare it to the decrypted message digest. If they match, the recipient knows the message has not been altered.[4]

Advantages of the Digital Signature

Unlike biometric and other forms of electronic signatures, the digital signature will "freeze" the contents of the document at the time of its creation. Any alterations to the document's contents will result in a different hash value. Furthermore, the encryption of the hash value with the signatory's private key "links uniquely the digital signature to the signatory, i.e., the owner of the private key." [5] Although a handwritten signature is only

---

[1] Christopher T. Poggi, "Electronic Commerce Legislation: An Analysis of European and American Approaches to Contract Formation," 41 *Va. J. Int'l L.* 224, 250-51 (2000).

[2] Susanna Frederick Fischer, "California Saving Rosencrantz and Guildenstern in a Virtual World? A Comparative Look at Recent Global Electronic Signature Legislation," Association of American Law Schools 2001 Annual Meeting, Section on Law and Computers, 7 *B.U. J. Sci. & Tech. L.* 229, 233 (2001).

[3] Note 26 supra at 249.

[4] Jochen Zaremba, "International Electronic Transaction Contracts Between U.S. and E.U. Companies and Customers," 18 *Conn. J. Int'l L.* 479, 512 (2003).

[5] Note 26 supra at 250.

"signatory-specific," the digital signature is both "signatory-specific" and "document-specific." [1]

The digital signature is the only form of electronic signature which satisfies all three of the UNCITRAL evaluation factors, i.e., that an electronic signature should:

(1) authorize; (2) approve; and (3) protect against fraud.[2] Authorization is achieved because the digital signature will accompany the document, which allows for confirmation of the identity of the signatory. Approval is attained via computation of the hash value of the electronic document, which freezes the contents of the document at the time of its creation, and allows for detection of any subsequent alterations. Finally, there is protection against fraud because it is extremely unlikely – virtually impossible – for anyone to determine a signatory's private key with only the public key as a starting point.[3]

Disadvantages of the Digital Signature

The digital signature has at least two drawbacks. Firstly, since the private key of each person is rather difficult to memorize, they are most often stored in computers. If the computer is not kept in a secure location, the contents of the private key may be vulnerable. This heightens the necessity of maintaining the security of the private key and protecting it from intruders. However, it should be noted that this weakness of the digital signature is also common to most other forms of electronic signatures. The password or the PIN face similar security problems. Therefore, with good security policies and procedures, this disadvantage can be minimized.[4]

The other disadvantage of the digital signature pertains to the digital certificate, which must be issued by a Certification Authority ("CA"). Obtaining the certificate and having to interact with the CA is somewhat inconvenient and costly for the user, but over time this disadvantage should be alleviated as digital signatures become more popular, easier to use, and cheaper.[5] Because the CA plays such a vital role in the viability of the digital signature, it is essential for the user to understand exactly what the CA does.

The Critical Role of the Certification Authority

In order for PKI to realize its potential, it is crucial that the user be able to ensure the authenticity of the public key (available online) used to verify the digital signature. If Smith and Jones are attempting to consummate an online transaction, Smith needs an independent confirmation that Jones' message is actually from Jones before Smith can have faith that Jones' public key actually belongs to Jones. It is possible that an imposter could have sent Jones his public key, contending that it belongs to Smith. Accordingly, a reliable third party –

[1] Id.
[2] Note 26 supra at 243.
[3] Note 26 supra at 252.
[4] Note 26 supra at 253.
[5] Id.

the Certification Authority[1] – must be available to register the public keys of the parties and to guarantee the accuracy of the identification of the parties.[2]

The most important job of the CA is to issue certificates which confirm basic facts about the subscriber, the subject of the digital certificate. Of course, the certificate is a digitized, computer-held record containing the most pertinent information about a transaction between two transacting parties. Typical information contained in a certificate includes the following: the name and address of the CA that issued the certificate; the name, address and other attributes of the subscriber; the subscriber's public key; and the digital signature of the CA.[3] Sufficient information will be contained in the certificate to connect a public key to the particular subscriber.[4]

In making an application to a CA for a certificate, the prospective subscriber must provide some sort of photo I.D., e.g., a passport or a driver's license. If the application is approved and the certificate is issued, the CA will issue a private key to its new subscriber which corresponds to the public key. This is done, however, without disclosing the specifics of the private key.[5] The steps in this application procedure vary somewhat from CA to CA, according to the type of certificate being offered by the CA. Ordinarily, however, once the CA has verified the genuine connection between the subscriber and the public key, the certificate will be issued.[6]

In order to indicate the authenticity of the digital certificate, the CA will sign it with her digital signature. Ordinarily, the public key corresponding to the subscriber's private key will be filed in the CA's online repository which is accessible to the general public and to third parties who have need of communication with the subscriber. Additionally, the online repository contains information pertaining to digital certificates which have been revoked or suspended by the CA due to lost or expired private keys. This is an important positive aspect of PKI technology: the general public has access to the status of digital signatures, and relying third parties are kept informed, allowing them to judge whether they should place reliance on communications signed with a certain private key.[7]

One of the recurring problems for digital signature lawmakers is in trying to fairly apportion the liability for risk of computer fraud between the CA and the subscriber. Nations around the world, and the state laws of the United States, have arrived at different conclusions regarding this apportionment. The problem is compounded if each CA is required to modify its practices every time it issues a certificate pertaining to a transaction affecting another jurisdiction which happens to have dissimilar digital signature laws.[8]

---

[1] Certification Authority ("CA") is the term used in this article because it seems to be the most commonly used designation around the world.

[2] Tara C. Hogan, Notes and Comments – Technology, "Now That the Floodgates Have Been Opened, Why Haven't Banks Rushed Into the Certification Authority Business?," *4 N.C. Banking Inst.* 417, 424-25 (2000).

[3] A. Michael Froomkin, "The Essential Role of Trusted Third Parties in Electronic Commerce," 75 OR. L. REV. 49, 58 (1996).

[4] Note 43 supra at 425-426.

[5] Note 17 supra at 149.

[6] Note 17 supra at 150.

[7] Note 43 supra at 426-27.

[8] Andrew B. Berman, Note, "International Divergence: The 'Keys' To Signing on the Digital Line – The Cross-Border Recognition of Electronic Contracts and Digital Signatures," 28 *Syracuse J. Int'l L. & Com*. 125, 143-44 (2001).

A digital certificate is only as reputable as the CA who issued it. If the CA is unreliable and untrustworthy, the digital certificate is also unreliable and untrustworthy. In the final analysis, a party contracting with an unknown stranger must rely upon the CA's registration expertise and its judgment that the subscriber's identification is accurate.[1]

THREE GENERATIONS OF ELECTONIC SIGNATURE LAW

The First Wave: Technological Exclusivity

In 1995, the U.S. State of Utah became the first jurisdiction in the world to enact an electronic signature law.[2] In the Utah statute, digital signatures were given legal recognition, but other types of electronic signatures were not.[3] The authors of the Utah statute believed, with some justification, that digital signatures provide the greatest degree of security for electronic transactions. Utah was not alone in this attitude; other jurisdictions granting exclusive recognition to the digital signature include Bangladesh,[4] India[5], Malaysia,[6] Nepal,[7] New Zealand[8] and Russia.[9]

Unfortunately, these jurisdictions' decision to allow the utilization of only one form of technology is burdensome and overly-restrictive. Forcing users to employ digital signatures gives them more security, but this benefit may be outweighed by the digital signature's possible disadvantages: more expense because of the fee paid to the certification authority; lesser convenience due to being forced to use a certification authority; forcing users to use one type of technology to the exclusion of others when another type of technology might be

---

[1] David Hallerman, "Will Banks Become E-commerce Authorities?," 12 *Bank Tech. News*, June 1, 1999.

[2] *Utah Code Ann*. 46-3-101 *et seq.*, 1995. This first-generation statute was repealed in 2000 and replaced with the Uniform Electronic Transactions Act, a second-generation model law. *Utah Code Ann*. 46-4-101 *et seq.* (2000); http://le.utah.gov/~code/TITLE46/46_04.htm. *See* Note 61 infra, first citation.

[3] Id.

[4] Bangladesh, *Information Technology (Electronic Transaction) Act* ("ITA") 2000 (Draft); http://www.bangladeshgateway.org/lawit.pdf .

[5] Republic of India, *The Information Technology Act* ("ITA"), 9 June 2000; http://www.mit.gov.in/itbillionline/itbill2000.asp. *See* Stephen E. Blythe, "A Critique of India's Information Technology Act and Recommendations for Improvement," 34 *Syracuse Journal of International Law and Commerce* 1 (2006), a publication of the College of Law, Syracuse University, Syracuse, New York USA.

[6] Republic of Malaysia, *Digital Signature Act* ("DSA"), 1997; http://www.mycert.org.my/bill/digisign/digi1.html .

[7] Federal Democratic Republic of Nepal, *Electronic Transactions Ordinance No. 32 of yhe Year* 2061 B.S. (2005 A.D.), s 60-71. An official English version was released by the Nepal Ministry of Law, Justice and Parliamentary Affairs and was published in the *Nepal Gazette* on 18 March 2005; http://www.hlcit.gov.np/pdf/englishcyberlaw.pdf. *See* Stephen E. Blythe, "On Top of the World, and Wired: A Critique of Nepal's E-Commerce Law," 8:1 *Journal of High Technology Law* (2008), a publication of Suffolk University School of Law, Boston, Massachusetts USA.

[8] New Zealand, *Electronic Transactions Act* 2000 ("ETA"); http://www.med.govt.nz/templates/MultipageDocumentPage_9779.aspx.

[9] Russian Federation, *Electronic Digital Signature Law*, Federal Law No. 1-FZ, 10 January 2002. *See* Note 18 supra at 234-37.

better suited to a particular type of transaction; use of a more complicated technology which may be less adaptable to technologies used in other nations, or even by other persons within the same nation; inappropriate risk allocation between users if fraud occurs; and the potential disincentive to invest in development of alternative technologies.[1]

The Second Wave: Technological Neutrality

Jurisdictions in the Second Wave overcompensated. They did the complete reversal of the First Wave and did not include any technological restrictions whatsoever in their statutes. They did not insist upon the utilization of digital signatures, or any other form of technology, to the exclusion of other types of electronic signatures. These jurisdictions have been called "permissive" because they take a completely open-minded, liberal perspective on electronic signatures and do not contend that any one of them is necessarily better than the others. In other words, they are "technologically neutral." Permissive jurisdictions provide legal recognition of many types of electronic signatures and do not grant a monopoly to any one of them. The United States of America[2] is a member of the second wave; the overriding majority of its jurisdictions (forty-five states, the District of Columbia, and the Territories of Puerto Rico and Virgin Islands) have enacted the Uniform Electronic Transactions Act (either in its entirety or with minor amendments), a permissive second-generation model law.[3] Australia has also enacted a second-generation statute.[4]

The disadvantage of the permissive perspective is that it does not take into account that, in fact, some types of electronic signatures *are* better than others. A PIN number and a person's name typed at the end of an E-mail message are both forms of electronic signatures, but neither is able to even approach the degree of security that is provided by the digital signature.

---

[1] Amelia H. Boss, "The Evolution of Commercial Law Norms: Lessons To Be Learned From Electronic Commerce," 34:3 *Brooklyn Journal of International Law* 673, 689-90 (2009). It is debatable as to whether technological-neutrality or technological-specificity is the correct road to take. *See* Sarah E. Roland, Note, "The Uniform Electronic Signatures in Global and National Commerce Act: Removing Barriers to E-Commerce or Just Replacing Them with Privacy and Security Issues?" 35 *Suffolk U. L. Rev*. 625, 638-45 (2001).

[2] For analysis of American law, *see* "E-Commerce and E-Signature Law of the United States of America," *The Ukrainian Journal of Business Law*, Kiev, Ukraine, November, 2008. For concise coverage of American, British, E.U. and U.N. law, *see* Stephen E. Blythe, "Digital Signature Law of the United Nations, European Union, United Kingdom and United States: Promotion of Growth in E-Commerce With Enhanced Security," 11: 2 *Richmond Journal of Law and Technology* 6 (2005).

[3] United States of America, National Conference of Commissioners on Uniform State Laws, *Uniform Electronic Transactions Act*, 7A U.L.A. 20 (Supp. 2000);

http://www.law.upenn.edu/bll/archives/ulc/fnact99/1990s/ueta99.htm. The State of Washington is the only U.S. jurisdiction presently having a first-generation statute, and these states have third-generation statutes: Alabama, Georgia, Florida and Ohio. *See also* United States of America, *Electronic Signatures in Global and National Commerce ACT* ("E-Sign"), Public Law 106-229, 15 U.S.C. 7001, 114 Stat. 464, 30 June 2000;

http://www.esignrecords.org/resources/esign.pdf.

[4] Commonwealth of Australia, ELECTRONIC TRANSACTIONS ACT 1999;

http://www.austlii.edu.au/au/legis/cth/consol_act/eta1999256/. *See* Note 18 supra at 234-37.

The Third Wave: A Hybrid

Singapore was in the vanguard of the Third Wave. In 1998, this country adopted a compromise, middle-of-the-road position with respect to the various types of electronic signatures. Singapore's lawmakers were influenced by the UNCITRAL Model Law on Electronic Commerce.[1] In terms of relative degree of technological neutrality, Singapore adopted a "hybrid" model – a preference for the digital signature in terms of greater legal presumption of reliability and security, but not to the exclusion of other forms of electronic signatures. Singapore did not want to become "hamstrung" by tying itself to one form of technology. The Singapore legislators realized that technology is continually evolving and that it would be unwise to require one form of technology to the exclusion of others. The digital signature is given more respect under the Singapore statute, but it is not granted a monopoly as in Utah. Singapore allows other types of electronic signatures to be employed. This technological open-mindedness is commensurate with a global perspective and allows parties to more easily consummate electronic transactions with parties from other nations.[2]

In recent years, more and more nations have joined the Third Wave. They recognize the security advantages afforded by the digital signature and indicate a preference for the digital signature over other forms of electronic signatures. This preference is exhibited in several ways: (1) utilization of a digital signature using a PKI system is explicitly required for authentication of an electronic record; (2) utilization of a digital signature with PKI seems to be necessary in order for an electronic record to comply with any statutory requirement that a record be in paper form; and (3) in order for a signature in electronic form to comply with a statutory requirement that a pen-and-paper signature be affixed, it must be a digital signature created with PKI. Nevertheless, the Third Wave jurisdictions do not appear to be as technologically-restrictive as those in the First Wave. They do not compel the E-commerce participant to use only the digital signature, *in lieu* of other forms of electronic signatures, as the State of Utah did in its original statute of 1995.

The moderate position adopted by Singapore has now become the progressive trend in international electronic signature law. The hybrid approach is the one taken by the European Union's E-Signatures Directive,[3] Armenia,[1] Azerbaijan[2] Barbados,[3] Bermuda,[4] Bulgaria,[5]

---

[1] United Nations Commission on International Trade Law ("UNCITRAL"), *Model Law on Electronic Commerce with Guide to Enactment* (Hereinafter "MLEC"), G.A. Res. 51/162, U.N. GAOR, 51st Sess., Supp. No. 49, at 336, U.N. Doc. A/51/49 (1996). *See* Stephen E. Blythe, Note 42 supra, first citation.

[2] Republic of Singapore, *Electronic Transactions Act* (Cap. 88) ("ETA"), 10 July 1998; Although granting legal recognition to most types of electronic signatures, the Singapore statute implicitly makes a strong suggestion to users – in two ways – that they should use the digital signature because it is more reliable and more secure than the other types of electronic signatures: (1) digital signatures are given more respect under rules of evidence in a court of law than other forms of electronic signatures, and electronic documents signed with them carry a legal presumption of reliability and security – these presumptions are not given to other forms of electronic signatures; and (2) although all forms of electronic signatures are allowed to be used in Singapore, its electronic signature law established comprehensive rules for the licensing and regulation of Certification Authorities, whose critical role is to verify the of authenticity and integrity of electronic messages affixed to electronic signatures. Id. *See* Stephen E. Blythe, "Singapore Computer Law: An International Trend-Setter with a Moderate Degree of Technological Neutrality," 33 *Ohio Northern University Law Review* 525-562 (2006).

[3] Note 3 supra; *see* Stephen E. Blythe, Note 42 supra. For concise coverage of European Union law, *see* Stephen E. Blythe, "E-Signature Law and E-Commerce Law of the European Union and its

Burma,[6] China[1] Colombia,[2] Croatia,[3] Dubai,[4] Finland,[5] Hong Kong,[6] Hungary,[7] Iran,[8] Jamaica,[9] Japan,[1] Lithuania,[2] Pakistan,[3] Peru,[4] Slovenia,[5] South Korea,[6] Taiwan,[7] Tunisia,[8]

Member States," *The Ukrainian Journal of Business Law*, pp. 22-26, May, 2008, Kiev, Ukraine. In an assessment of the effectiveness of its E-Signature Directive in 2006, the European Commission concluded that contracting parties had been slow to use digital signatures, but that "many other simpler electronic signature applications had become available." Reasons advanced by the Commission for the slow rate of adoption of digital signatures include: "technical problems in the marketplace, a lack of criteria for certification and mutual recognition, a lack of interoperability at national and cross-border levels, and the existence of isolated areas where certificates were used for a single purpose." Overall, the primary reason advanced was an economic one, caused by a typical user's decision to eschew development of a multi-application digital signature in favor of an E-signature which is applicable to its own industry, e.g., the banking sector. *Report on the Operation of Directive 1999/93/Ec on a Community Framework for Electronic Signatures*, s 5.2, COM (2006), cited in Boss, Note 59 supra at 695-96. Despite the less than enthusiastic reception of the digital signature in Europe and elsewhere, that rate of acceptance is expected to be given a "shot in the arm" felt worldwide by the "United Nations Convention on Contracts for the International Carriage of Goods Wholly or Partly by Sea (hereinafter "Rotterdam Rules");"

http://www.unis.unvienna.org/unis/pressrels/2008/unisl125.html . The Rotterdam Rules became effective on 23 September 2009 and recognize the legal validity of electronic bills of lading. In order to comply with the security requirements of Article 38 of the Rotterdam Rules, it will apparently be necessary to employ a digital signature. Felix W.H. Chan, "In Search of a Global Theory of Maritime Electronic Commerce: China's Position on the Rotterdam Rules," 40 *Journal of Maritime Law and Commerce* 185 (2009). *See also* Manuel Alba, "Electronic Commerce Provisions in the UNCITRAL Convention on Contracts for the International Carriage of Goods Wholly or Partly by Sea," 44 *Texas International Law Journal* 387 (2009). Accordingly, *a la* Mark Twain's rumored death, any notion that the digital signature is passé appears to have been "greatly exaggerated." The digital signature appears to have a bright future because, presently at least, it is the epitome of security.

[1] Republic of Armenia, *Law on Electronic Document and Electronic Signature*, 2002; http://www.gipi.am/?i=223 . *See* Stephen E. Blythe, "Armenia's Electronic Document and Electronic Signature Law: Promotion of Growth in E-Commerce via Greater Cyber-Security," *Armenian Law Review*, May, 2008, a publication of the Department of Law, American University of Armenia, Yerevan, Republic of Armenia.

[2] Republic of Azerbaijan, *The Law of the Azerbaijan Republic on Digital Electronic Signature*, 2003; http://unpan1.un.org. *See* Stephen E. Blythe, "Azerbaijan's E-Commerce Statutes: Contributing to Economic Growth and Globalization in the Caucasus Region," 1:1 *Columbia Journal of East European Law* 44-75 (2007), a publication of Columbia University School of Law, New York NY USA.

[3] Barbados, *Electronic Transactions Act, Cap*. 308B, 8 March 2001;

http://www.barbadosbusiness.gov.bb/miib/Legislation/Acts/investment_acts.cfm.

See Stephen E. Blythe, "The Barbados Electronic Transactions Act: A Comparison with the U.S. Model Statute," 16 *Caribbean Law Review* 1 (2006), a publication of the Faculty of Law, The University of the West Indies, Barbados.

[4] Commonwealth of Bermuda, *Electronic Transactions Act* 1999 ("ETA");

http://www.bakernet.com/ecommerce/bermuda-eta.doc . *See* Note 18 supra at 234-37.

[5] Republic of Bulgaria, *Law on Electronic Document and Electronic Signature* ("EDL"), 2001; http://www.csd.bg/news/law/E-CommercePublE.htm. *See* Stephen E. Blythe, "Bulgaria's Electronic Document and Electronic Signature Law: Enhancing E-Commerce With Secure Cyber-Transactions," 17:2 *Transnational Law and Contemporary Problems* 361 (2008), a publication of the University of Iowa College of Law, Iowa City, Iowa USA.

[6] The Union of Myanmar, *Electronic Transactions Law* ("ETL"), The State Peace and Development Council Law No. 5/2004, The 12 Waxing of Kason 1366 M.E., 30 April 2004; http://ibiblio.org/obl/docs/Electronic-transactions.htm. *See* Stephen E. Blythe, "Rangoon Enters the Digital Age: Burma's Electronic Transactions Law As a Sign Of Hope For a Troubled Nation," 3:1

International Business Research __ (2010), a publication of the Canadian Center of Science and Education, Toronto, Canada; http://ccsenet.org/journal/index.php/ibr/ .

[1] People's Republic of China, Order No. 18 of the President, *Law of the People's Republic of China on Electronic Signature*, Adopted at the 11th Meeting of the Standing Committee of the Tenth National People's Congress of the People's Republic of China, Promulgated 28 August 2004, Effective 1 April 2005. The statute was translated into English by the Beijing University School of Law, Beijing, China, and is available (by subscription only) at their website:

http://www.lawinfochina.com/dispecontent.asp?db=1&id=3691.

See Stephen E. Blythe, "China's New Electronic Signature Law and Certification Authority Regulations: A Catalyst for Dramatic Future Growth of E-Commerce," 7 *Chicago-Kent Journal of Intellectual Property* 1 (2007), a publication of Chicago-Kent College of Law, Illinois Institute of Technology, Chicago, Illinois USA. *See also* Felix W.H. Chan, "E-Commerce All at Sea: China Welcomes Digital Bills of Lading Under the Electronic Signature Law 2005," 3 *Oklahoma Journal of Law and Technology* 31 (2006).

[2] Republic of Colombia, *Law Regulating Data Messages, Electronic Trade, Digital Signatures and Certification Entities* ("ETL"), 13 January 1999, Official Translation No. 7 by Maria del Pilar Mejia de Restrepo; http://www.qmw.ac.uk/~t16345/colombia_en_final.htm. *See* Stephen E. Blythe, "Computer Law of Colombia and Peru: A Comparison With the U.S. Uniform Electronic Transactions Act," a book chapter in *Internet Policies and Issues*, Frank Columbus, Editor, Nova Science Publishers, Inc., New York NY USA, 2009.

[3] Republic of Croatia, *Electronic Signature Act* ("ESA"), 17 January 2002;

http://www.ehrvatska.hr/sdu/en/Zakonodavstvo/RH/categoryParagraph/00/document/eSignatureActOG10_2002.pdf. *See* Stephen E. Blythe, "Croatia's Computer Laws: Promotion of Growth in E-Commerce Via Greater Cyber-Security," 26: 1 *European Journal of Law and Economics* 75-103 (August, 2008), a publication of Springer Netherlands Ltd., Amsterdam.

[4] Emirate of Dubai, *Law of Electronic Transactions and Commerce* No. 2/2002 ("ETL"), 12 February 2002; http://www.tecom.ae/law/law_2.htm. *See* Stephen E. Blythe, "The Dubai Electronic Transactions Statute: A Prototype for E-Commerce Law in the United Arab Emirates and the G.C.C. Countries," 22:1 *Journal of Economics and Administrative Sciences* 103 (2007).

[5] Republic of Finland, Ministry of Justice, *Act on Electronic Signatures*, 2003; http://www.finlex.fi. *See* Stephen E. Blythe, "Finland's Electronic Signature Act and E-Government Act: Facilitating Security in E-Commerce and Online Public Services," 31:2 *Hamline Law Review* 445-469 (2008), a publication of Hamline University School of Law, St. Paul, Minnesota USA.

[6] Hong Kong Special Autonomous Region, People's Republic of China, *Electronic Transactions Ordinance*, Ordinance No. 1 of 2000. Before amending its original digital signature law, Hong Kong only recognized digital signatures and was therefore a member of the First Wave. After amendments were enacted, Hong Kong joined the Third Wave. *See* Stephen E. Blythe, "Electronic Signature Law and Certification Authority Regulations of Hong Kong: Promoting E-Commerce in the World's 'Most Wired' City," 7 *North Carolina Journal of Law And Technology* 1 (2005), a publication of the University of North Carolina School of Law, Chapel Hill, NC USA.

[7] Republic of Hungary, *Act XXXV of 2001 on Electronic Signature*, 2001; http://www.techlawed.org. *See* Stephen E. Blythe, "Hungary's Electronic Signature Act: Enhancing Economic Development With Secure E-Commerce Transactions," 16:1 *Information and Communications Technology Law* 47-71 (2007), a publication of Routledge Publishing Co., a member of the Taylor & Francis Group. Executive Editor: Prof. Indira Carr, Centre for Legal Research, Middlesex University, London, U.K.

[8] Islamic Republic of Iran, *Electronic Commerce Law of the Islamic Republic of Iran* ("ECL"); http://irtp.com/laws/ec/IR%20Iran%20E-Commerce%20Law.pdf. *See* Stephen E. Blythe, "Tehran Begins to Digitize: Iran's E-Commerce Law as a Hopeful Bridge to the World," 18 *Sri Lanka Journal of International Law* (2006), a publication of the University of Colombo Faculty of Law, Colombo, Sri Lanka.

[9] Jamaica, *Electronic Transactions Act*, 2005. *See* Stephen E. Blythe, "Internet Law As A Potential Catalyst For Growth Of Caribbean E-Commerce: Jamaica's Statute As A Model," a paper to be

United Arab Emirates,[9] Vanuatu[1] and in the proposed statute of Uganda.[2] Many other nations are either currently using the hybrid approach or are considering the adoption of it; Jordan is one of them.[3]

---

presented and published in the *Proceedings of the Second Annual Academy of Business Administration Global Trends Conference*, Cancun, Mexico, December 18-23, 2009.

[1] Japan, *Law Concerning Electronic Signatures and Certification Services*, promulgated 24 May 2000, effective 1 April 2001;

http://www.meti.go.jp/english/report/data/gesignconte.html. *See* Stephen E. Blythe, "Cyber-Law of Japan: Promoting E-Commerce Security, Increasing Personal Information Confidentiality and Controlling Computer Access," 10 *Journal of Internet Law* 20 (2006), a publication of Aspen Publishers, Inc., New York, NY USA.

[2] Republic of Lithuania, *Law On Electronic Signature*, No. VIII – 1822 (July 11, 2000), As Amended, No. IX—934 (June 6, 2002);

http://www3.lrs.lt/cgibin/preps2?Condition1=204802&Condition2. *See* Stephen E. Blythe, "Lithuania's Electronic Signature Law: Providing More Security in E-Commerce Transactions," 8 *Barry Law Review* 23 (2007), a publication of Dwayne O. Andreas School of Law, Barry University, Orlando, Florida USA.

[3] Islamic Republic of Pakistan, *Electronic Transactions Ordinance*, 2002; http://unpan1.un.org/groups/public/documents/apcity/unpan010245.pdf. *See* Stephen E. Blythe, "Pakistan Goes Digital: the Electronic Transactions Ordinance as a Facilitator of Growth for E-commerce," 2:2 *Journal of Islamic State Practices In International Law* 5 (2006), a publication of ElectronicPublications.org Ltd., Stockport, U.K. Editors: Prof. Javaid Rehman, School of Law, Brunel University, West London, U.K.; and Dr. Amir Ali Majid, School of Law, London Metropolitan University, London, U.K.

[4] Republic of Peru, *Law Regulating Digital Signatures and Certificates*, 28 May 2000, translated by National Law Center for Inter-American Free Trade; http://natlaw.com/interam/ar/ec/tn/tnarecl.htm. *See* Note 54 supra.

[5] Republic of Slovenia, Centre for Informatics, *Electronic Commerce and Electronic Signature Act*, 2000; http://e-uprava.gov.si/eud/e-uprava/en/ECAS-Act-in-English.pdf. *See* Stephen E. Blythe, "Slovenia's Electronic Commerce and Electronic Signature Act: Enhancing Economic Growth With Secure Cyber-Transactions," 6: 4 *The I.C.F.A.I. Journal of Cyber La*W 8-33 (2007), a publication of ICFAI University Press, Institute of Chartered Financial Analysts of India, Hyderabad, India.

[6] Korean Legislation Research Institute, *Digital Signature Act* No. 5792, *Statutes of the Republic of Korea,* Vol. 16 (II), pp. 1217-1220 (1999). The statute has been amended two times: (1) Act No. 6360 of 16 January 2001; and (2) Act. No. 6585 of 31 December 2001. *See* Stephen E. Blythe, "The Tiger on the Peninsula is Digitized: Korean E-Commerce Law as a Driving Force in the World's Most Computer-Savvy Nation," 28: 3 *Houston Journal Of International Law* 573-661 (2006), a publication of the College of Law, University of Houston, Houston, Texas USA.

[7] Republic of China, *Electronic Signatures Act* ("ESA"), 2002;

http://law.moj.gov.tw/Eng/Fnews/FnewsContent.asp?msgid=944&msgType=en&keyword.

See Stephen E. Blythe, "Taiwan's Electronic Signature Act: Facilitating the E-Commerce Boom With Enhanced Security," a paper presented and published in the *Proceedings of the Sixth Annual Hawaii International Conference on Business*, Honolulu, Hawaii USA, May 25-28, 2006.

[8] Republic of Tunisia, *Electronic Exchanges and Electronic Commerce Law*, 9 August 2000; http://www.bakernet.com.org. *See* Stephen E. Blythe, "Computer Law of Tunisia: Promoting Secure E-Commerce Transactions With Electronic Signatures," 20 *Arab Law Quarterly* 317-344 (2006), a publication of Brill Academic Publishers, Leiden, The Netherlands.

[9] United Arab Emirates, Federal Law No. (1) of 2006 o*n Electronic Commerce Andtransactions* ("ECL"), 30 January 2006;

http://www.tra.ae/pdf/legal_references/Electronic%20Transactions%20%20Commerce%20Law_Final%20for%20May%203%202007.pdf. *See* Stephen E. Blythe, "The New Electronic Commerce Law

JORDAN'S ELECTRONIC TRANSACTION LAW

Jordan's Electronic Transaction Law ("ETL")[4] was enacted in 2001 in order to support the utilization of contracts consummated in electronic form.[5] The ETL does not override other relevant laws and is subject to those laws.[6] The ETL should be applied in view of international business practices and technological innovations.[7] The Prime Minister and Ministers are charged with the responsibility of implementation of the statute,[8] and the specific governmental departments to be executing the ETL will be decided by the Cabinet.[9]

Applicability of the ETL

The ETL should be applied to: (1) "electronic transactions,[10] electronic records,[1] electronic signatures[2] and any data messages;"[3] (2) electronic transactions between

---

of the United Arab Emirates: A Progressive Paradigm for Other Middle Eastern Nations to Emulate," a paper presented and published in the *Proceedings of the Annual International Conference on Global Business*, Dubai, United Arab Emirates, May 10-13, 2009.

[1] Republic of Vanuatu, *Electronic Transactions Act* (Act. 24 of 2000); http://www.paclii.org/cgi-paclii/disp.pl/vu/legis/num%5fact/eta2000256.html. The E-commerce law of the Commonwealth of Bermuda was used as a model for this statute. "Vanuatu E-commerce," LOWTAX, p. 1; http://www.lowtax.net/lowtax/html/jvaecom.html. For a discussion of the ETA by the Prime Minister of Vanuatu—the person who introduced the bill in Parliament – *see* Hon. Prime Minister Barak T. Sope Maautamate, MP, Government of the Republic of Vanuatu, "The e-Business Act of 2000, The International Companies (E-Commerce Amendment) Act of 2000, The Companies (E-Commerce Amendment) Act of 2000: A Plain English Explanation," pp. 3-7;

http://www.vanuatu.gov.vu/government/library/Explanation%20of%20the%20ecommerce%20acts.htm.

See also Stephen E. Blythe, "South Pacific Computer Law: Promoting E-Commerce in Vanuatu and Fighting Cyber-Crime in Tonga," 10: 1 JOURNAL OF SOUTH PACIFIC LAW (2006), a publication of the School of Law, University of the South Pacific, Emalus Campus, Port Vila, Republic of Vanuatu.

[2] Republic of Uganda, *Electronic Signatures Act*, Draft, 2004;

http://www.sipilawuganda.com/downloads/electronic%20signatures%20bill%202004.pdf.

See Stephen E. Blythe, "The Proposed Computer Laws of Uganda: Moving Toward Secure E-Commerce Transactions and Cyber-Crime Control," a paper to be presented and published in the *Proceedings of the Tenth Annual Conference of the International Academy of African Business and Development*, Kampala, Uganda, May 19-23, 2009.

[3] Note 93 infra.

[4] Hashemite Kingdom of Jordan, *Electronic Transactions Law* No. 85 OF 2001 (hereinafter "ETL"); http://www.cbj.gov.jo/uploads/Electronic_Transactions_Law.pdf .

[5] ETL art. 3(A). An electronic contract is defined as "An agreement concluded in whole or in part by electronic means." ETL art. 2.

[6] Id.

[7] ETL art. 3(B).

[8] ETL art. 41.

[9] ETL art. 39. Cabinet decisions will determine the ETL's implementation provisions, to include the following: (1) fees charged by government departments for carrying out electronic transactions; and (2) rules pertinent to Certification Authorities, their issuance of certificates, and the fees they may charge for doing so. ETL art. 40.

[10] Electronic transactions are defined simply as "Transactions conducted by electronic means." ETL, art. 2.

government departments;[4] but (3) only to those parties that have agreed to use the electronic form in a specific transaction.[5] The ETL does not apply to documents controlled by a specific statute in which designated procedures and forms have been mandated, including but not limited to, the following: (1) wills and codicils; (2) sale of real property (not leases), transfer of title to real property, and real property deeds; (3) powers of attorney; (4) notices of termination of life insurance, health insurance or public utilities services; (5) court-related documents;[6] and (6) securities (except for "special instructions issues").[7]

Legal Recognition of the Electronic Form

Ordinarily, "electronic records, contracts, messages, and signatures" are just as enforceable and admissible as paper records, contracts, messages and signatures.[8] Their legal enforceability and admissibility cannot be denied based on the mere fact of their electronic form, so long as they are in compliance with the ETL.[9]

Compliance With Originality Requirement

If a statute mandates that the original of a legal document must be presented in order to exercise a legal right pertinent to that document, presentation of an electronic record will be deemed to have complied with that mandate, provided: (1) the information in the electronic record can be stored and readily accessed at a later time; (2) the record is stored in the original form it was generated, sent or received and the accuracy of the information in the record can

---

Transactions are defined as "actions occurring between two parties or more for establishing obligations upon one party, or mutual obligations between more than one party relating to the conduct of business, a civil obligation, or any government department." Id. Electronic is inclusively defined as "The technology utilizing electrical, magnetic, optical or electro-magnetic means or any other similar means in the interchange and storage of information." Id.

[1] An electronic record is defined as a "record, contract, or data message generated, sent, received or stored by electronic means." ETL art. 2.

[2] An electronic signature is defined as "electronic, numeric or photic data or others taking the shape of letters, numbers, symbols, or signs, or the like in a data message or added or related thereto, having a shape identifying the person who timed or distinguished it from others for reasons of the person's signature and the approval of content." ETL art. 2.

[3] ETL, art. 4(A). A data message is defined as "Information generated, sent, received or stored by electronic or similar means, including Electronic Data Interchange (EDI), electronic mail, telegram, telex or telecopy. ETL, art. 2. EDI is the "Electronic transfer of information from one person to another using information processing systems." ETL art. 2.`

[4] ETL, art. 4(B).

[5] ETL, art. 5(A). However, just because a party has agreed to use the electronic form in one transaction does not imply that she will use the electronic form in other transactions. ETL art. 5(B).

[6] ETL, art. 6(A).

[7] ETL, art. 6(B).

[8] ETL, art. 7(A).

[9] ETL, art. 7(B). However, in some situations, electronic documents will not enjoy the same legal rights as paper documents. This section of the article (pertaining to ETL art. 7 – 11) is inapplicable in the following situations: (1) if a statute mandates the sending of information to a "related person" in writing but allows other forms of communication if the parties agree; and (2) the parties have agreed to use a particular form of mail, e.g., "first class mail, express mail or regular mail." ETL, art. 12.

easily be verified; and (3) the record shows the identity of the sender and receiver,[1] and the date and hour the message was sent and received.[2]

Compliance With Writing Requirement

If a statute mandates a transaction must be effected in writing on paper to be enforceable, utilization of the electronic form will suffice to meet this requirement if the receiver has the ability to print and retain the information and can readily access the information at a later time.[3] However, the sender must do anything to prevent the receiver from being able to print or to retain the information. If she does, her communiqué will have no legal effect.[4]

Compliance With Signature Requirement

If a statute mandates that an agreement be signed by a party in ink on paper, the signatory's electronic signature attached to an electronic record will have the same legal effect.[5] An electronic signature will be attributed to the signatory, and will be considered legally enforceable, provided: (1) a means of identification has been used to confirm the sender; and (2) some reliable means of ensuring that the sender has consented to the information in the electronic record has been applied.[6]

Compliance With Retention Requirement

If a statute mandates that a paper document that be kept in storage, it may be kept in electronic form unless a statute enacted after the ETL states otherwise.[7]

Electronic Contracts[8]

Contractual offers and acceptances may be made electronically.[9]

---

[1] The ETL refers to the sender as the "originator,' and to the receiver as the "addressee." ETL, art. 2. An originator is defined as "A person by whom or whose behalf, a data message is generated or sent prior to its receipt or storage by the addressee. Id. An addressee is defined as "The person who is intended by the originator to receive the data message." Id.

[2] ETL, art. 8(A). These conditions may be created through a third party. ETL, art. 8(C). It is not necessary for the electronic record to contain information pertinent to its transmission or receipt. ETL, art. 8(B).

[3] ETL, art. 9(A).

[4] ETL, art. 9(B).

[5] ETL, art. 10(A).

[6] ETL, art. 10(B). The degree of reliability of the method of confirmation that the signatory approves the information in the electronic record will depend upon all of the circumstances, including whether the parties agreed to use of that method. Id.

[7] ETL, art. 11.

[8] These are defined as "An agreement concluded in whole or in part by electronic means." ETL, art. 2.

[9] ETL, art. 13. In a conference paper presented in 2006, the authors questioned whether the ETL's E-contract rules replaced the pre-existing Jordanian civil contract law, or whether the civil contract law continued to be applicable to electronic contracts despite the presence of the ETL. *See* Marwan Al Ibraheem and Hisham Tahat, "Regulating Electronic Contracting in Jordan," *Proceedings of the 21st*

Attribution To Sender

An electronic message is deemed to have been sent by the sender if it was transmitted by: (1) the sender herself; (2) the sender's agent; or (3) an electronic agent[1] programmed by the sender or by the sender's agent, to operate automatically.[2]

Often, a receiver may not know for sure whether the sender actually sent the message. However, "life goes on" even though the receiver is uncertain. A receiver may assume that a received message was sent by the sender if: (1) an information processing system[3] procedure – previously agreed to by the sender – was applied by the receiver to ascertain whether the sender actually sent the message, and the procedure confirmed that was the case; or (2) the data message received was transmitted by the sender's subordinate or agent, and was empowered by the sender to utilize the method of identification of the sender.[4]

The rules in the preceding paragraph are inapplicable if: (1) the receiver was in receipt of a timely notice from the sender that said electronic record did not belong to her or her agent;[5] (2) the receiver either knew, or should have known if reasonable care or a specifically-agreed procedure had been employed, that the electronic record did not belong to the sender or her agent.[6]

Acknowledgment of Receipt

When Sender Has Requested Acknowledgement. If the sender, who has requested acknowledgement of receipt, does not specify the form or method to be used by receiver, then the receiver may employ: (1) any type of communication, manual or automated; or (2) conduct sufficient to show the sender that the reception has occurred.[7]

If the sender has instructed the receiver that the electronic record will have no legal impact[8] until the sender is in receipt of the acknowledgement, then the electronic record is assumed to not have been sent until the sender is in receipt of the acknowledgement from the receiver.[9] On the other hand, if the sender has not placed conditionality upon her receipt of the acknowledgement, and the sender has not received an acknowledgment within a reasonable time or within a previously-specified time, then the sender: (1) after informing the receiver that no acknowledgement has been received, give the sender a reasonable time deadline by which the acknowledgement must be received; and (2) if the sender is not in

---

Bileta Conference; Globalisaton and Harmonisation in Technology Law, Malta (April 2006), British and Irish Law, Education and Technology Association.

[1] An electronic agent is defined as "A computer program or other electronic means used independently to implement an action or respond thereto for the purpose of generating, sending, or receiving a data message without review or action by an individual." ETL, art. 2.

[2] ETL, art. 14.

[3] An information processing system is defined as "An electronic system used for generating, sending, receiving, processing or storing data messages or for handling data messages in any other respect. ETL, art. 2.

[4] ETL, art. 15(A).

[5] ETL, art. 15(B)(1). In this case, the receiver is not allowed to assume that the received message was transmitted by the sender in question. However, the sender is liable for any damages incurred by the receiver prior to the receiver's receipt of notice from the sender. Id.

[6] ETL, art. 15(B)(2).

[7] ETL art. 16(A).

[8] The ETL refers to this situation of no legal impact as "conditional." ETL, art. 16(B).

[9] ETL art. 16(B).

receipt of the acknowledgement by the deadline, then the sender may take the legal position as if the electronic record had never been sent and exercise her rights accordingly.[1]

Legal Effect of Receipt of Acknowledgement. Once the acknowledgement has been received by the sender, the sender may assume (in the absence of evidence to the contrary) that the receiver has received her message. However, the acknowledgment is insufficient legal evidence for the sender to assume that the content of the electronic message received by the receiver is identical to what was sent by the sender.[2]

Time and Place of Transmission and Reception[3]

Time of Dispatch. An electronic record will be assumed to have been sent when it enters a computer information system outside the control of the sender, or one outside the control of the sender's agent.[4]

Time of Reception: Receiver Has Designated Location. If the receiver has indicated a specific computer information system for the electronic record to be sent to: (1) the assumed time of reception is when it enters the designated computer information system; or (2) if it enters another computer information system of the receiver that is not the designated one, then the assumed time of reception is when the receiver retrieves the electronic record from the computer information system.[5]

Time of Reception: Receiver Has Not Designated Location. If the receiver has not designated a specific computer information system for the electronic record to be sent to, then the assumed time of reception is when the electronic record enters any computer information system of the receiver.[6]

Place of Transmission and Reception. The electronic record is assumed to have been sent from the sender's place of business, and the electronic record is assumed to have been received at the receiver's place of business.[7] If either the sender or the receiver has more than one place of business, then the place will be assumed to be the one that has the closest connection to the transaction in question; if this cannot be determined, then the principal place of business will be the applicable one.[8] If either the sender or the receiver does not have a place of business, then the assumed location will be the sender or receiver's place of residence.[9]

---

[1] ETL art. 16(C).
[2] ETL art. 16(D). Let's be realistic: there may have problems in the transmission of the message.
[3] If the parties have made an agreement as to the assumed time and place of dispatch and reception, that agreement will be controlling. ETL art. 17(A).
[4] ETL art. 17(A).
[5] ETL art.17(B).
[6] ETL art.17(C).
[7] ETL art. 18(A).
[8] ETL art. 18(B).
[9] ETL art. 18(A).

Transferable Electronic Notes

Requirements for Transferability

In order to be transferable, electronic notes: (1) must comply with all requirements for "negotiable instruments" pursuant to the Commercial Code, except for the writing requirement; and (2) must have the agreement of the drawer that the note is negotiable.[1]

If a statute requires for a financial institution to maintain a copy of a check, use of the electronic form is allowed provided both sides of the cancelled check are accessible for reference at a later time.[2] In order for electronic checks to be transferable, the Central Bank of Jordan must approve and issue criteria for attainment of the approval.[3]

Rights and Obligations of Holders

In order for the holder of a transferable electronic note to enjoy full rights, the computer information system she uses to generate and transfer the note can record the transfer of the rights and confirm the identity of the transferee-beneficiary.[4] The computer information system must be able to create an "authoritative copy" of the note that is: (1) "uniquely identifiable and unalterable;"[5] and (2) states the names of the beneficiary and transferee as well as the transferability status of the note.[6] The authoritative copy must be transmitted and retained by the party that assumes control or her agent.[7]

If all criteria are complied with, the holder of an electronically transferred note is able to assert control of it and enjoys the same rights and privileges as a holder of a written note.[8] Similarly, the obligor of an electronic note has the same privileges and defenses as the obligor of a written note.[9] If a party makes an objection to payment of the electronic note, the party controlling the note has a duty to show "reasonable proof" that she controls the note.[10]

Electronic Transfer of Funds

Electronic transfer of funds ("ETF") is legally sanctioned by the ETL. However, the ETL does not override other relevant laws.[11] Banks engaged in ETF must adhere not only to the

---

1 ETL art 19(A).

2 ETL art. 19(B). The retention requirement is covered in ETL art. 8. Id.

3 ETL art. 19(C). The rules relating to transferable electronic checks are covered in ETL art. 20—24. Id.

4 ETL art. 20.

5 ETL art. 21(A)(1). In order to maintain uniqueness, the authoritative note or the revised note can be reproduced only with the consent of the party assuming control over the note. If copies are made, each one must be so labeled. Every time the authoritative copy is reproduced, there must be an indication on the copy that it is an "identical copy of the authoritative copy." ETL art. 21(C).

6 ETL art. 21(A)(2).

7 ETL art. 21(B).

8 ETL art. 22.

9 ETL art. 23.

10 ETL art. 24. Proof may take the form of a showing of the authoritative copy or of pertinent business records, or other evidence which confirms the details of the note and the name of the party controlling the note. Id.

11 ETL art 25.

ETL, but also to the Central Bank of Jordan Law,[1] the Banks Law, and all implementation regulations pertinent to those statutes. Furthermore, banks participating in ETF have an obligation to take measures to ensure the security of the transfer and the confidentiality of the customer's private information.[2]

After the bank's client informs the bank that an unlawful entry to her account has been made as a result of an ETF, and that a hold should be placed on all ETF's, the client is not responsible for any losses. The client should have informed the bank that the account's security has been compromised. A possible reason is that the access card has been lost or its PIN number has been disclosed.[3]

Secure Electronic Records and Secure Electronic Signatures

"Secure" Status

An electronic record will be assumed to have not been modified from the time of its creation if, on the date in question, the confirmation process: (1) uses accredited security procedures; (2) uses "commercially acceptable" security procedures;[4] or (3) uses security procedures that were previously-agreed-to by the parties.[5] In order for an electronic signature to attain "secure" status, it must utilize a security procedure meeting the standards of the preceding sentence,[6] and must also have the following characteristics: (1) be unique to the person possessing its private key; (2) identify the holder of the private key; (3) be created by the holder of the private key; and (4) be attached[7] to an electronic record and not allow the record to be modified[8] without first modifying the electronic signature.[9]

---

[1] The Central Bank of Jordan promulgates regulations relating to ETF's. These include: information to the conveyed to the Central Bank by the bank making the ETF, procedures for correcting mistakes, uncovering of ETF's which have been used illegally, and other issues arising from electronic payments. ETL art. 29.

[2] ETL art. 26.

[3] ETL art. 27. However, the client cannot escape liability for loss if her account and the ETF were used for an illegal purpose, provided the primary reason for the illegal use was the client's negligence and the bank has exercised due diligence to prevent illegal use of the account. ETL art. 28.

[4] In order to "commercially acceptable" security procedures, they must take into account the specific situation faced by the parties, to include these factors: (1) the purpose and type of the transaction; (2) the degree of sophistication of the parties; (3) the experience of the parties in consummating similar transactions; (4) whether other procedures were available that could have been used by the parties, and their cost; and (5) the procedures that are most commonly used in similar types of transactions. ETL art. 30(B).

[5] ETL art. 30(A).

[6] Id.

[7] There is a legal presumption that a secure electronic signature was attached by the purported signatory, and that the signatory intended to show her approval of the electronic record's contents. ETL art. 32(A)(2). An insecure electronic signature does not enjoy such a presumption. ETL art. 32(B).

[8] There is a legal presumption that a secure electronic record was not changed after the security procedures became effective. ETL art. 32(A)(1). An insecure electronic record does not enjoy such a presumption. ETL art. 32(B).

[9] ETL art. 31.

Certificates

If an electronic record is attached to an electronic signature which happens to be a digital signature, there is another requirement for that record to achieve "secure" status: the digital signature must have been created during the period of validity of its "accredited[1] security certificate"[2] and the digital signature must correspond to the identification code expressed in the certificate.[3]

Computer Crimes

Fraudulent Use of Certificate

It is a crime for a party to create, present or use a certificate in order to commit fraud or to carry out another illicit purpose. The convicted offender will be subject to: (1) imprisonment in the range of 3 months to 2 years; (2) a fine in the range of 3,000 to 10,000 Jordanian Dinars;[4] or (3) both.[5]

Submission of False Information to CA

It is a crime for a party to submit false information to a CA for the purpose of attaining issuance, suspension or revocation of a certificate. The convicted offender will be subject to: (1) imprisonment in the range of 1 to 6 months; (2) a fine in the range of 1,000 to 5,000 Jordanian Dinars;[6] or (3) both.[7]

Offenses by CA's

It is a crime for a CA or prospective CA to: (1) submit false information to the governmental authority in an application for a CA license; (2) allow the confidentiality of its clients' private information to be compromised; or (3) violate the ETL or its implementation regulations. The convicted offender will be subject to fine of at least 5,000 Jordanian Dinars.[8]

---

[1] In order to be "accredited," the certificate containing the identification code must have the following attributes: (1) be issued by a "licensed or competent entity" (i.e., a Certification Authority ("CA")); (2) be issued by a recognized foreign CA; (3) be issued by an authorized Egyptian government department, institution or body); or (4) be issued by a body that is accepted by all parties to the transaction. ETL art. 34.

[2] Egyptian law uses this three-word designation. Many jurisdictions around the world simply use the one-word designation ("certificate") and that is used in this article.

[3] ETL art. 33.

[4] This corresponds to a range of approx. U.S. $4,228 to $14,094. XE.com, 24 November 2009.

[5] ETL art. 35.

[6] This corresponds to a range of approx. U.S. $1,409 to $7,047. XE.com, 24 November 2009.

[7] ETL art. 36.

[8] ETL art. 37. This is approx. U.S. $7,047. XE.com, 24 November 2009.

Commission of other Crimes

Using Electronic Means

If a person uses electronic means to carry out another crime that is defined by other acts, the convicted offender will be subject to: (1) imprisonment in the range of 3 months to 1 year; (2) a fine in the range of 3,000 to 10,000 Jordanian Dinars;[1] or (3) both. If the other act defining the crime provides for a more stringent penalty, the other penalty will apply.[2]

OTHER DIGITAL ACCOMPLISHMENTS OF THE JORDANIAN GOVERNMENT

E-Government

Unlike the electronic transactions statutes of many other nations, Jordan's ETL does not mention E-government. Nevertheless, the Jordanian government has created a website and encourages its citizens to take advantages of the online services offered there.[3]

Jordan's national E-government strategy has these objectives: "deliver high-quality services to consumers, businesses and organizations; improve government performance and efficiency; enhance Jordan's competitiveness; ensure public sector transparency and accountability; reduce costs and increase ease of interaction with government; promote development of Jordan's ICT sector; develop skills within the public sector; boost e-commerce activities; and improve information security."[4]

Jordan's government portal is first-rate and conveniently includes the following breakdown of its services: Government-to-Citizens; Government-to-Business; and Government-to-Government. The services most used by citizens are: issuance of documents; issuance or renewal of National ID Card; payment of tax in full or in installments; initial issuance of passports; and acquisition of debit clearance.[5]

The portal provides potential or existing businesses with information pertinent to: the start-up and operation of a business in Jordan; investment incentives offered by the government of Jordan; Jordan's privatization program; investment benefits in free zones; qualified industrial zones; trade agreements; laws and regulations; and selling to the government of Jordan.[6] The government-to-government section of the Portal is accessible

[1] This corresponds to a range of approx. U.S. $4,228 to $14,094. XE.com, 24 November 2009.

[2] ETL art. 38.

[3] The Jordanian Government Portal is located at: http://www.jordan.gov.jo/wps/portal. For a discussion of the launching of Jordan's E-government, its planning and implementation features and the main obstacles, *see* Hussein Al Omari, "E-Government Architecture in Jordan: A Comparative Analysis," 2:11 *Journal of Computer Science* 846-852 (2006); www.scipub.org/fulltext/jcs/jcs211846-852.pdf .

[4] Id., first citation, at home page.

[5] Hashemite Kingdom of Jordan, Government Portal; http://www.jordan.gov.jo/wps/portal/G2C/?New_WCM_Context=/wps/wcm/connect/gov/eGov/Home+Citizen/.

[6] Hashemite Kingdom of Jordan, Government Portal;

only by employees of relevant government ministries of Jordan and is not accessible by the general public.[1]

National ID Card

Jordan adopted a National ("smart") ID card in 2001. It contains several types of personal information, including voter registration.[2] Application and other information pertinent to the National ID Card is available at the Government Portal.[3] Only a handful of other jurisdictions have adopted an ID card; they include Belgium[4] and Hong Kong.[5] In those jurisdictions, the ID Card's computer chip can serve as the E-signature of the cardholder.[6] This idea is recommended for adoption in Jordan.[7]

RECOMMENDATIONS FOR IMPROVEMENT OF THE ETL

Jordan has made a commendable beginning toward attainment of a first-class electronic transactions law. Although the ETL is a significant accomplishment, it has not gone far enough. The following amendments should be considered.

---

http://www.jordan.gov.jo/wps/portal/G2B/?New_WCM_Context=/wps/wcm/connect/gov/eGov/Home+Business/.

[1] Hashemite Kingdom of Jordan, Government Portal; http://www.jordan.gov.jo/wps/themes/html/warning.html.

[2] Privacy International, PHR2006: THE HASHEMITE KINGDOM OF JORDAN, 18 December 2007, p.2; http://www.privacyinternational.org/article.shtml?cmd%5B347%5D=x-347-559523 .

[3] Hashemite Kingdom of Jordan, Government Portal; http://images.jordan.gov.jo/wps/wcm/connect/gov/egov/home+citizen/services+for+citizens/most+used+services/issuance+of+personal+id+cards+for+the+first+time%2C+renewal%2C+in+lieu+of+lost+%2C+in+lieu+of+destroyed.

[4] By 2010, Belgium will have issued an electronic ID card to each of its 9 million inhabitants, becoming the first European nation to carry out this achievement. Each resident will pay approximately EU 10 for his card. These cards contain two E-signatures; one will be used for identification of the holder, and the other will be used to sign E-documents. Already, the electronic ID card is being used for: access to the Belgian government website and its E-government services; signing of legal documents in digital form (e.g., tax declaration, VAT declaration, and social security affirmations); access to community container parks; parking tickets; signing of registered mail; signing of Flemish Parliament Decrees; requests for official documents and access to National Register records; and access to the E-library service. Additionally, Dell, HP and Siemens computers are now able to read the Belgian ID card and to process its E-signature data. Interdisciplinary Centre for Law & Information Technology, *The Legal and Market Aspects of Electronic Signatures*, 2003, pp. 177-178; http://ec.europa.eu/information_society/eeurope/2005/all_about/security/electronic_sig_report.pdf.

[5] Rina C.Y. Chung, Note 23 supra. For information pertinent to the Hong Kong I.D. card, refer to the Hong Kong Government Portal; http://www.smartid.gov.hk/. The list of other nations which have adopted national ID cards includes: Austria, Bahrain, Belgium, Hong Kong, Israel, Italy, Spain and the United Arab Emirates.

[6] Notes 173 and 174 supra.

[7] See Recommendations, infra.

The ETL Should Be Comprehensive

All of the laws pertinent to electronic transactions should be included under the umbrella of the statute. If a number of other laws pertinent to electronic transactions have been previously enacted, they should be consolidated in the new statute. A comprehensive statute is easier for all affected parties to research and to comprehend. For example, all of the E-government laws and regulations should be made a part of the ETL.

Make the ETL Supreme In All Things Electronic

Amend Article 2 to state that, if the ETL is in conflict with another law or statute, the ETL will prevail.

Add: A List of Other Laws Affected by the ETL

There should be a list of other statutes and regulations that are modified or affected by the ETL. Additionally, there should be a list of the names of all other statutes currently in force (and the applicable provisions in each) which can be complied with using the electronic form instead of the paper form.

Delete: All Exclusions in Article 6

Article 6 of the ETL excludes several types of documents from coverage. The result is that the following types of documents must be in paper form to have legal validity: wills and codicils; [1] real estate deeds and documents relating to sale or purchase of real estate; powers of attorney; notice of termination of insurance and public utility services; and court documents.[2]

All of the exclusions should be eliminated. This would recognize the legal validity of electronic documents in all of those categories. This would firmly tell the world that Jordan sees virtually no limits to the utilization of the electronic form and would hasten the adoption of the electronic form by its citizens and residents. Only a few nations have completely eliminated exclusions in their E-commerce statutes,[3] and none of them are in the Middle East.

---

[1] The aversion to electronic wills is beginning to dissipate. In 2005, the U.S. State of Tennessee became the first American jurisdiction to recognize the legal validity of a will that is executed with an electronic signature. *See* Chad Michael Ross, Comment, "Probate – Taylor v. Holt – The Tennessee Court of Appeals Allows a Computer Generated Signature to Validate a Testamentary Will," 35 *University Of Memphis Law Review* 603 (2005).

[2] ETL art. 6.

[3] For example, Azerbaijan's statute contains no exclusions from coverage; it states that electronic documents "can be used (applied) in *all activity spheres* where software and technical equipment could be applied to create, use, store, transmit and receive information." Republic of Azerbaijan, Electronic Document Law, 2003, art. 1(1) (emphasis added), note 67 supra. Montenegro and New Zealand also

Add: Legal Validity of Electronic Form to Comply with Several Additional Requirements of other Statutes

The ETL should state a general presumption that the electronic form may be used to satisfy requirements contained in other statutes which are prerequisite to incurrence of a legal right. Those requirements include, but are not limited to, the following: the witnessing of a handwritten signature or seal; a paper document's notarization, certification, acknowledgement, verification, attestation, or being made under oath; production of multiple copies of a paper document (where production of one electronic copy is sufficient); communication by registered or certified mail (provided that the electronic message is transmitted thorough the sender's Certification Authority and confirmed by him); and seller's provision of a notice to a consumer in writing. For a comprehensive list of such electronic compliance allowances, refer to the New Zealand statute.[1]

Add: Two New Types of E-Commerce Contract Rules

Jordan needs E-commerce contractual rules pertaining to automated contracts and carriage contracts. For automated contracts, the U.S. Uniform Electronic Transactions Act contains a good model.[2] For attribution, refer to South Korea's Electronic Commerce Act.[3] For acknowledgement of receipt, look to Singapore's Electronic Transactions Act.[4] For time and place, use Holland's Electronic Commerce Act.[5] For carriage contracts, Colombia's Electronic Trade Law has a commendable paradigm.[6]

Add: More Consumer Protections for E-Buyers

Jordan needs to enact a general consumer protection statute applicable to all internet consumers.[7] The Republic of Tunisia can be used as a model for good consumer protections.

---

have no exclusions from coverage; *see* Republic of Montenegro, Electronic Signature Law, 2003; www.mipa.cg.yu; and Commonwealth of New Zealand, *Electronic Transactions Act*, Note 57 supra.

[1] Note 57 supra.

[2] Note 61 supra. *See* Stephen E. Blythe, Note 60 supra.

[3] Republic of South Korea, Korean Legislation Research Institute ("KLRI"), *Framework Act on Electronic Commerce*, *Statutes of the Republic of Korea,* Vol. 13, pp. 395-400 (1999). The KLRI is an independent non-profit organization funded by the government of the Republic of South Korea. The KLRI's charge is to translate all of the Korean federal statutes into English. They do an admirable job of this and the *Statutes'* twenty volumes, in loose-leaf form, are continually updated. This is one of the Korean government's globalization thrusts. Of course, the "official" statutes are the ones in Korean Language as originally enacted. However, given that the KLRI's work is financed by the Korean government, the English-Language versions of the statutes used in research for this article could be described as "*quasi*-official." *See* Stephen E. Blythe, Note 86 supra.

[4] Note 64 supra (first citation), s 14.

[5] Kingdom of the Netherlands, *Act on Information Society Services* (30 June 2004), Art. 11. *See* Stephen E. Blythe, Note 60 supra, third citation.

[6] Note 73 supra.

[7] In their 2006 conference paper, two Jordanian professors also recognized the need for E-commerce consumer protections in Jordan. *See* Note 117 supra.

The Tunisian E-commerce statute gives consumers: (1) a “last chance” to review an order before it is entered into; (2) a 10-day window of opportunity to withdraw from an agreement after it has been made; (3) a right to a refund if the goods are late or if they do not conform to specifications; and (4) no risk during the 10-day trial period after goods have been received. Tunisian E-consumers enjoy some of the best protections in the world.[1]

Add: I.T. Courts for E-Commerce Disputes

Because of the specialized knowledge often required in the adjudication of E-commerce disputes, Information Technology Courts should be established as a court-of-first-instance for them. The I.T. Courts would be tribunals consisting of three experts. The chairperson would be an attorney versed in E-commerce law, and the other two persons would be an I.T. expert and a business management expert. The attorney would be required to hold a law degree and be a member of the bar with relevant legal experience; the I.T. person would be required to hold a graduate degree in an I.T.-related field and have experience in that field; and the business management expert would be required to hold a graduate degree in business administration and have managerial experience. The E-commerce law of Nepal can be used as a model.[2]

Add: Long-Arm Jurisdiction Against Foreign E-Commerce Parties

Because so many of the E-commerce transactions incurred by the residents of Jordan will be with parties outside the borders of Jordan, it would be prudent for the ETL to explicitly state its claim of “long arm” jurisdiction against any E-commerce party who is a resident or citizen of a foreign jurisdiction, so long as that party has established “minimum contacts” with Jordan. The Kingdom of Tonga can be used as a model.[3]

Minimum contacts will exist if a cyber-seller outside of the country makes a sale to a person in Jordan. In that situation, the laws of Jordan should be applicable to the foreign party because that party has had an effect upon the country through the transmission of an electronic message that was received in Jordan. The foreign party should not be allowed to

---

[1] Note 88 supra. One of the few nations that may offer better consumer protections is Korea. That country has enacted a separate statute specifically for E-commerce consumer protections—the E-Commerce Transactions Consumer Protection Act. *See* Korean Legislation Research Institute, Act on the Consumer Protection in the Electronic Commerce Transactions (“CPA”), *Statutes of the Republic of Korea*, Vol. 13, pp. 481 to 485-30. Originally enacted by Law No. 6687 (30 March 2002), and amended by Act Nos. 7315 and 7344 of 31 December 2004 and 27 January 2005, respectively. The CPA recently underwent a major overhaul with substantial amendments in Act No. 7487 of 31 March 2005; those amendments became effective on 1 April 2006. For an analysis of the CPA, *see* Stephen E. Blythe, Note 84 supra. Iran also provides good consumer protections, including a window of opportunity to withdraw from an E-commerce transaction previously entered into; however, the window in Iran is only seven days, as opposed to Tunisia’s ten days. *See* Stephen E. Blythe, Note 79 supra.

[2] Note 56 supra.

[3] The Republic of Tonga explicitly states its claim of long-arm jurisdiction over foreign E-commerce parties. *See* Stephen E. Blythe, Note 90 supra.

evade the jurisdiction of the Jordanian courts merely because he is not physically present in the country. After all, E-commerce is an inherently international and multi-jurisdictional phenomenon.

Add: Jordan's Post Office To Become a Licensed Certification Authority

In order to promote the utilization of E-signatures among the general public and to make them cheaper and more accessible, Jordan's Post Office should be designated as a licensed CA. For a model, look to the Belgian Post Office, which has implemented a promotional campaign to educate the general public about E-signatures and their availability through the Post Office.[1]

Add: Registration Agents

In order to increase the number of locations offering CA services throughout Egypt, the office of Registration Agent ("RA") should be created. An RA is employed by a CA and works under the authority granted in the CA's license; an RA does not need a separate license. The RA is able to operate branch locations of the CA. The RA's responsibilities include the processing of applications for certificates and confirming the identification documents submitted by those applicants. Several nations have experimented with the RA idea, including Peru[2] and the Slovak Republic.[3]

Add: Digital Signature in the National ID Card

As mentioned, Jordan adopted a National ID card in 2001; each resident is issued a "smart" card containing a computer chip. Each card should be allowed to serve as an E-signature of the cardholder after the computer chip has been activated by a CA.

Add: Several New Computer Crimes

The list of computer crimes needs to be expanded. The following computer crimes, with appropriate penalties, should be recognized: (a) Unauthorized Tampering with Computer Information; (b) Unauthorized Use of a Computer Service; (c) Unauthorized Interference in the Operation of a Computer; (d) Unauthorized Dissemination of Computer Access Codes or

---

[1] Kingdom of Belgium, *Legal Framework for Electronic Signatures and Certification Services* ("ESA"), 9 July 2001. This statute was supplemented by the *Royal Decree Organizing the Supervision And Accreditation Of Certification Service Providers Issuing Qualified Certificates*, 6 December 2002.

[2] Republic of Peru, *Digital Signature Law*, Art. 13. *See* Stephen E. Blythe, Note 73 supra.

[3] Slovak Republic, *Electronic Signature Act*, Art. 21. *See* Stephen E. Blythe, Note 60 supra, third citation.

Passwords; and (e) Injection of a Virus into a Computer. The Singapore Computer Misuse Act can be used as a model.[1]

SUMMARY AND CONCLUSIONS

Introduction

King Abdullah II succeeded his father in 1999. In response to chronic unemployment, poverty and inflation, he has implemented substantial economic reforms and advantageous trade treaties with the United States and Europe; these moves justify hope for a brighter future. To date, the rate of internet penetration has been modest. However, there has been a 300% growth in internet penetration during the past three years and Jordan has incurred a 9-fold increase in internet hosts during the past four years. This positive trend is expected to continue due to governmental promotion of internet usage through its commendable E-Government Portal and its development of the information technology sector.

Electronic Signatures

An E-signature is used to sign an electronic document. There are several types of E-signatures: a digital signature, a digitized fingerprint, a retinal scan, a pin number, a digitized image of a handwritten signature that is attached to an electronic message, or merely a name typed at the end of an e-mail message. The most secure of these is the digital signature because it will freeze the contents of the message to which it is attached and will indicate if the message has been altered since its creation. Because of the high degree of security it offers and its assurance that an attached document has not been altered, the digital signature is the most preferred and is given the highest degree of legal status. However, biometric E-signatures (e.g., a retinal scan) are also very useful and are often employed in conjunction with the digital signature.

Three Generations of E-Signature Law

There have been three generations of E-signature law since the world's first E-signature statute was enacted in 1995. These three successive generations emphasized, respectively: exclusive recognition of public key infrastructure ("PKI") technology and the digital signature; technological neutrality, with all types of E-signatures and technologies recognized; and a hybrid perspective which recognized all types of E-signatures, with a preference shown for PKI in admission of E-signatures and electronic documents ("E-documents") into evidence.

---

[1] Republic of Singapore, *Computer Misuse Act* (Cap. 50A), 30 August 1993; http://agcvldb4.agc.gov.sg/non_version/cgi-bin/cgi_gettopo.pl?actno=1998-REVED-50A. *See* Stephen E. Blythe, Note 59 supra, second citation.

Jordan's Electronic Transaction Law

Jordan enacted its Electronic Transaction Law ("ETL") in 2001. It applies to most E-commerce transactions, but it is inapplicable to: wills; real property deeds and sales; powers of attorney; notice of termination of insurance or public utility services; court documents; and securities.

The basic assumption of the ETL is that an electronic record, contract, message or signature is just as enforceable and admissible in court as their paper counterparts. If a statute mandates that an original paper document must be presented in order to verify the existence of a legal right, that mandate is deemed met if an electronic document is produced. However, there are certain requirements pertaining to storage in the original form, accessibility at a later time, verification of the document's accuracy, and the identity of the sender and addressee and the date and hour the document was transmitted.

If a statute mandates that a record of a transaction must be in writing on paper to be enforceable, that mandate is deemed met if the transaction is recorded in an electronic document. However, there are certain requirements relating to the ability to print a hard copy of the record and its accessibility at a later time.

If a statute mandates that a paper document containing an agreement must be signed in ink to be enforceable, that mandate is deemed met if an electronic signature is attached to the electronic document which contains the details of the agreement. Once again, however, there are certain requirements concerning confirmation of the sender of the document, and use of a reliable means of ensuring that the sender is in agreement with the information contained in the electronic document.

If a statute mandates that a paper document must be retained in storage for a length of time, that mandate is deemed met if the document is retained in electronic form.

The ETL has adopted a number of specific rules pertinent to the creation and implementation of E-contracts. These rules are concerned with: whether an electronic message that has been received from a purported party can be legally assumed to have actually emanated from that party; the acknowledgement of receipt from the receiver to the sender, in different types of situations (i.e., when the sender has requested acknowledgement, and when the sender has not so requested); and the time and place that an electronic message can be assumed to have been sent and received.

In a somewhat unusual move, Jordan decided to include rules pertinent to the transfer of electronic notes and electronic funds transfers in its ETL. Electronic notes are transferable if all requirements for "negotiable instruments" under the Commercial Code are complied with, and the drawer of the note agrees that it is negotiable. Electronic funds transfers are sanctioned by the ETL as well. However, the participating financial institution must do everything reasonably necessary to ensure the security of the transfer and to maintain the confidentiality of the customer's private information.

The ETL contains third-generation electronic signature provisions. In order for an electronic record to achieve "secure" status, it must use stringent security procedures or procedures that have been agreed to by all relevant parties. In order for an electronic signature to attain "secure" status, it must have met those same requirements, and must also have the following characteristics: be unique to the signatory, identify the signatory, be created by the signatory, and be attached to an electronic document and not allow the document to be changed without first changing the electronic signature. The ultimate requirement of a secure electronic signature is for it to have been created during the period of validity of a certificate issued by a licensed CA.

The ETL lists several computer crimes: fraudulent use of a certificate; submission of false information to a CA; offenses by CA's; and commission of other crimes using electronic means. All of them are punishable by a prison term, a fine, or both.

Recommendations for Improvement of the ETL

The ETL needs to be refined. Recommended amendments are to: (1) make the ETL comprehensive; (2) make the ETL supreme in all things electronic; (3) add a list of other laws affected by the ETL; (4) delete all exclusions from coverage; (5) recognize the legal validity of the electronic form in order to comply with several additional requirements of other statutes; (6) add rules for electronic automated contracts and electronic carriage contracts; (7) add consumer protections for E-buyers; (8) add I.T. Courts for resolution of E-commerce disputes; (9) add long-arm jurisdiction over foreign E-commerce parties; (10) add Jordan's Post Office as a licensed Certification Authority; (11) add Registration Agents to assist Certification Authorities; (12) add a digital signature to the National ID Card; and (13) add several computer crimes.

In: Advances in Communications,Volume 8
Editor: Anthony V. Stavros
ISBN: 978-1-61324-794-5

Chapter 6

ENHANCEMENT AND EXPANSION OF LANGUAGE EDUCATION IN THE INTERNET ERA ROLE OF WEB 2.0 TECHNOLOGIES

C. K. Cheung and Aditi Dubey
Faculty of Education
University of Hong Kong, PR of China

Abstract

The increasing public engagement in Internet communication and new media, especially in Web 2.0 technologies, has generated new interest regarding the effects of technological mediation in the area of language learning. The changes brought about due to the multimodal nature and stylistic variations of new media technology have resulted in increased theoretical reflection on language within a mediated communication reality (Rowe & Ryss, 2009). These theories imply the need for today's pedagogy to move beyond the traditional written and oral texts and to embrace digital text formats. Knobel & Wilber (2009) affirm that digital formats don't mesh well with traditional language learning practices such as book reports, comprehension questions, reading tasks, and weekly spelling tests. Thus, the Internet era with its Web 2.0 media tools, questions the dominant ideas around language and language learning. This in turns prompts the need for language learning to be rethought in the new media context.

Web 2.0 refers to web applications that facilitate a user-centred design and interactive information sharing. Most "digital natives" (Prensky, 2001) or learners, who grew up using computers, are well-versed with Web 2.0 tools such as blogs, wikis, podcasts, and social networking sites. Unlike other applications, since the focal point of web 2.0 tools is the user, this paper aims to shed light into the potential of Web 2.0 tools to enhance the language ability of students, and to expand or broaden the range of language learning opportunities. The growth and development of Web 2.0 tools, which has been a result of the rapid advancement in new media and Internet technologies, can contribute to language learning by promoting interaction, exposure, feedback, reflection, empowerment and learner autonomy. They also provide an environment for language learners to reflect, comment, question and review their progress even outside the classroom in a more authentic environment (Pinkman, 2005).

To conclude, the paper argues that the user-friendliness of Web 2.0 applications, when combined with appropriate guidelines, training and adequate knowledge, can enable teachers or “digital immigrants” (Prensky, 2001) to integrate Web 2.0 tools to assist students’ language learning in the classroom. Web 2.0, thus, offers many opportunities for language learners and teachers. This not only has tremendous potential for the enhancement and improvement of language learning, but it is also more conducive for meeting the expanding language needs of the Internet era.

Keywords: Web 2.0, blogs, social networking sites, wikis, podcasts, language education.

INTRODUCTION

This paper explores the ways in which Web 2.0 technologies can enhance language learning and expand on the existing notion of language education. As the student demographic gets increasingly accustomed to Web 2.0 technologies, the changes brought about due to the multimodal and stylistic variations of new media technology have resulted in a reflection on language learning within a mediated communication reality (Rowe & Ryss, 2009). Language instructors can now leverage Web 2.0 tools such as blogs, wikis, podcasts, social networking sites and other participatory technologies that allow access to authentic sources of language for today’s students and provide more opportunities for input, social interaction and collaboration with native speakers of the target language. Since language is all about communication, being able to use their newly acquired language skills in an authentic environment can be highly motivating for students. Thus, the impact of delivery mode on students’ language skills and performance is an area that has been receiving significant attention in the past few years, especially as the use of technology for language instruction has expanded rapidly (McLoughlin, Hutchinson & Koplin, 2002).

While the world witnesses a change in social practices and language use due to a proliferation of communicative Web 2.0 tools, language teachers are beginning to ask themselves if they have been successful in integrating these tools in their classrooms to enhance language learning. Moreover, since the user-friendliness of Web 2.0 applications enables language teachers to integrate these tools easily into the class setting, scholars and researchers have pondered the potential of Web 2.0 technologies to produce a transformation in pedagogy, particularly in relation to foreign language and second language acquisition. As a result, this paper synthesizes and scrutinizes empirical evidence on the patterns of use of Web 2.0 tools such as blogs, social networking sites, wikis and podcasts in language education. It draws on existing information from reports, conference papers and journal articles to identify trends, experiences and challenges. Consequently, it examines the pedagogical potential and realities of Web 2.0 in language education, and proposes steps for the integration of Web 2.0 tools in the language classroom.

WEB 2.0 TECHNOLOGY

Although it is not easy to define Web 2.0 precisely because there are no clear or set standards for a Web 2.0 application, it is a term attributed to Tim O' Reilly (2005) to describe recent developments on the World Wide Web. It refers to the emergence of a set of online applications which facilitate a more socially connected web where everyone is able to add and edit information (Anderson, 2007). The term is often used to describe the increased use of the web for user-generated content, collaboration and social networking (Dalsgaard & Sorensen, 2009). Echoing this view, Knobel & Wilber (2009) state that "A Web 2.0 ethos values and promotes three interlocking functions or practices: participation, collaboration, and distribution." Consequently, with Web 2.0 data, the Internet also becomes a platform for social software that enables groups of users to socialize, work and share information with each other. Franklin & van Harmelen (2007) offer the following definition of Web 2.0 technologies:

- Web 2.0 encompasses a variety of different meanings that include an increased emphasis on user generated content, data and content sharing and collaborative effort, together with the use of various kinds of social software, new ways of interacting with web-based applications, and the use of the web as a platform for generating, re-purposing and consuming content (p. 4).

Web 2.0 can also be understood by contrasting it with the former "Web 1.0". In Web 1.0 a few content authors provided content for a wide audience of relatively passive readers, unlike Web 2.0 that allows greater user participation. However some, such as Tim Berners-Lee (2006), founder of the Web, argue that Web 2.0 is simply a "piece of jargon" because if Web 2.0 is meant to refer to blogs and wikis – that is people to people – then that was what the web was supposed to be all along. Despite this line of reasoning, there is widespread acceptance that the Internet has become dominated by a group of applications that share the following characteristics:

- Power to the user: Whereas Web 1.0 was dominated by content provided in static pages, Web 2.0 applications have democratized the web by prioritizing user-generated content, ownership and social connectivity.
- Harnessing collective intelligence: O'Reilly (2005) points out the importance of interactivity in Web 2.0 technologies, which empowers the masses rather than the media elite.
- Web as a platform: O'Reilly (2005) emphasizes that Web 2.0 users are able to source information and are able to run rich internet applications in their browsers. These applications have a participative element that encourages users to add, edit or rehash content.

Based on these definitions and criteria, a typology of Web 2.0 tools and activities can be done as follows for the purpose of this paper (see Table 1 below):

Table 1. A Typology of Web 2.0 Tools and Activities

Technology	Activity	Examples
Media sharing	Creating and exchanging media with peers or wider audiences.	Flickr, YouTube, Slideshare
Media manipulations & mash ups	Using web-accessible tools to design and edit digital media files and combining data from multiple sources to create a new application, tool or service.	Geotagged photos on maps, Voice mashups
Instant messaging, chat, web 2.0 forums	One-to-one or one-to-many conversations between Internet users.	MSN, Paltalk, Arguementum
Online games and virtual world	Rule-governed games or themed environments that invite live interaction with other Internet users.	WorldofWarcraft, SecondLife
Blogging	An Internet-based journal or diary in which a user can post text and digital material while others can comment.	Worldpress, Edublog, Twitter
Social networking	Websites that structure social interaction between members who form subgroups of 'friends'.	Facebook, Myspace, Linkedin, Elgg
Social bookmarking	Method for Internet users to share, organize, search, and manage bookmarks of web resources. Unlike file sharing, the resources themselves aren't shared, merely bookmarks that reference them.	Del.icio.us, Citeulike, Zotero
Recommender systems	Websites that aggregate and tag user preferences for items in some domain and thereby make novel recommendations.	Digg, LastFm, Stumbleupon
Wikis and collaborative editing tools	Web-based services that allow users access to create edit and link pages.	Wikipedia, GoogleDocs, Bubbl.us
Syndication / RSS feeds	Users can ‘subscribe’ to RSS feed enabled websites to be automatically be notified of any updates.	Bloglines, Podcast, Googlereader

THE WEB 2.0 WAY OF LANGUAGE EDUCATION

Language instruction has always had a close relationship with technology. In the early days of foreign language pedagogy, the first computers turned out to be excellent for language drills. Nowadays, in the discussion about the potential of Web 2.0 technologies to produce a pedagogical transformation, Web 2.0 applications are portrayed as technologies that will enhance collaboration and participation in the classroom, and enable students to develop their language skills. Simon (2008) notes that along with teaching experience and excellent scholarship, foreign language job candidates are asked by an increasing number of higher education institutions to have proficiency in instructional technologies. He points out that a keyword search for the word "technology" and "computer" in the Modern Language Association (MLA) job list returns several ads among the job postings (as of November 20, 2007), which state some of the following requirements: "familiarity with teaching-related technologies" (tenure track in Spanish, Missouri); "experience with technology in the classroom" (tenure track in French, Michigan); "ability to use technology effectively in teaching and learning" (tenure track in Japanese, South Carolina) etc. Although the wording varies slightly from one ad to the next, it is clear that the candidates applying for the job of a language teacher need to have an answer ready when asked how they use technology in the classroom.

However, integrating technology into teaching requires the combination of adequate technical skills and sound pedagogical foundations because if not done right, bad experiences feed the argument of technophobic educators who believe that computers and other electronic gadgets do not belong in the classroom, and it leads skeptics to turn their backs on educational technology. This can have unfortunate consequences for language learning because many Web 2.0 applications serve as powerful socialization and communication tools. As such, the subsequent discussion aims to analyze the potential of some of the major Web 2.0 technologies, namely blogs, social networking sites, podcasts and wikis, in the area of language instruction.

Blogs in the Language Classroom

Short for weblog, a blog is an easy to maintain online journal that can provide foreign language learners a venue in which they can reflect, comment, question, review, and communicate – outside the classroom in an authentic environment. Shihab's (2009) research points out that teachers found "blogging to be the most powerful tool for journal writing and sharing ideas" (p. iii). Research by Thorne and Payne (2005) suggests that language students prefer blogging to traditional journals or weekly essays. In this research, students also reported frequently looking back over their own and other students' blog postings, and most of them noticed a progress in their writing over time. According to Solomon & Schrum (2007) blogs are "natural tools for writing instruction" (p. 81) leading to greater number of English language teachers to start experimenting with blogs in their classes. In a blog students can not only write, revise, review the writings of peers, and get feedback from their teachers, they can also share their writing on a global level, and see a permanency of their words that usually can never happen when a paper is submitted in a traditional class. Similarly, Pinkman's (2005) findings suggest that learner-perceived benefits of using blogs included increased interest and motivation to use English because of interaction with, and feedback from, classmates and teachers. Blogs have been known to be successful in the language classroom primarily because of significant teacher input / feedback and interaction with peers. This reinforces Campbell's (2003) three potential ways of using blogs in the language classroom:

- The tutor blog: It is posted by the tutor as a resource for students' self-study. It is intended for the daily reading practice of learners and for encouraging exchange of comments online.
- The learner blog: It is posted by students and not only provides them with an opportunity to get writing practice, but also to develop their sense of ownership for their writing.
- The class blog: It is created by students and can be like a free-form bulletin board or a project-based language learning exercise. It can promote interaction between students and promote language exchange.

Moreover, in order to encourage learners' input, focus and motivation, blogs can be structured around specific activities or topics which can stretch their language output beyond their comfort level while at the same time allowing students to take their time over postings in a low-pressure environment. Since writing is a core skill for bloggers, Penrod (2007) too considers blogs to be effective for teaching composition because besides offering ease of use and publishing, blogs can potentially empower students who are often marginalized in the classroom. The power of saying words in writing is achieved in blogs. Students who are shy or who have specific learning styles might prefer to express themselves on blogs rather than during open discussions. Blogging is very helpful for deaf students to communicate with others. Based on the collaborative nature of blogs, Penrod (2007) argues that blogging also encourages fluency in writing, cooperative learning, critical thinking, and performance-based learning.

Furthermore, the uses of microblogging in language education have also been considered. MicroBlog allows users to share, disseminate and access information within individual communities with about 140 characters long text updates and instant sharing. The earliest and most famous is the United States microblog Twitter. Castro (2009) identifies several uses of microblogs and notes that besides organizing language class activities based on twitter messages, the teacher can also use microblogs to get in touch with students for sending memos, sharing resources, providing quick feedback and for keeping students' motivation level high. One of the most important advantages of the use of microblogging such as Twitter in class is that the community continues to be active even after the class or the entire course has ended, which contributes to members' life-long learning (Castro, 2009).

However, Lenhart et al. (2010) show through two Pew Internet Project surveys of teens and adults that teens in the age range of 12-17 years do not use Twitter in large numbers. This puts Twitter far down the list of popular online activities for teens and stands in stark contrast to their record of being early adopters of nearly every online activity. Moreover, the project also indicates that "blogging has declined in popularity among both teens and young adults since 2006" (p. 2) and attributes the rapid decline in blogging among younger internet users to changes in social network use by teens and young adults.

Social Networking Sites (SNSs) in the Language Classroom

Lenhart et al. (2010) point out that many of the functions that blogging served for teens in the mid-2000s for communicating about their lives and updating their activities for their friends have now become central activities on SNSs. She adds that status updating on social networks have replaced old-style 'macro-blogging' for many teens and adults. Social networking allows an individual to create a profile for themselves on the service and share that profile with other users with similar interests to create a social network. Harrison and Thomas (2009) define SNSs as follows:

- At the core of SNSs are the profiles and network of 'friends' that users create. After a user signs up to an SNS, s/he will be prompted to create a profile that varies in sophistication depending on the networking site concerned. Facebook, for instance, provides a sophisticated profiling system that allows users to create very detailed information about themselves and also fine tune the level of privacy by determining what information is to be made public (p.111).

Users can usually post notes, photographs, music and videos on their site and receive comments. The majority of SNSs such as Facebook, which reportedly has 175 million users, also have group or community functions that allow users to create or join groups within the SNS based on a particular theme.

The popularity of social software, such as MySpace and Facebook, has led to a huge increase in the number of language-related social networking sites and Carroll (2008) notes that by 2005 social networking had been recognized as a Web 2.0 technology with the potential to be used for structured language learning purposes. Whereas early internet-mediated language exchange was text-based, developments in software mean that students can now also talk (and see) their language partner, who are much easier to find due to social networking sites. Typically, language learners sign up to a site (most of which are free), and specify their native language and the language(s) they are learning. Students can either find a partner for language practice on a one-to-one basis outside the classroom, or a teacher can set up a class exchange. Some examples are:

- Mixxer: Developed by Dickinson College in the US, it offers student or class exchange.
- xlingo: Uses Skype as the main form of communication.
- LiveMocha: Language-learning site along with social-networking.
- Worldia: Social-networking site aimed at international users. Include a service to provide and receive
- language lessons.
- My Language Exchange: Offers voice and text chat, lesson plans.

Language practice with native speakers through social networking sites offers authentic interaction, and the ability to put learning into practice in a real context can be a powerful motivational tool. Kern (cited in Thorne & Payne 2005) found through a quantitative assessment of a chat programme – 'Daedalus Interchange' – with a group of French students that the Interchange led to more language production, and was of an overall greater level of sophistication than face-to-face discussion. Also, there was evidence that the environment reduced communication anxiety, as students felt free to communicate in what they considered to be a more informal atmosphere. Another language-related development of social software is the advent of the social dictionary such as 'Word Source' and 'Lingoz', which can offer teachers ideas for creating vocabulary-building exercises. The forums can also be a useful resource for query resolution for the independent learner.

Social bookmarking tools such as del.icio.us make use of human-generated tags to create an online repository of websites, which can also serve as good sources of knowledge for both the language learner and the teacher. A quick search for sites marked 'Spanish' on del.icio.us returns a wealth of useful resources: dictionaries, language-learning sites, Spanish newspapers. Within a class context, a unique class-specific tag can be used to share resources with other members of the class. There can also be plenty of opportunities for educators to network online with peers. The 'ning' platform, which enables users to create social networks around specific subjects, is particularly popular. 'Classroom 2.0' ning and 'EFL Classroom' 2.0 ning are two examples of 'ning' interaction in language education. However, a study conducted by the European Commission (2009) on the impact of new media on language learning found the following (p.12):

- Evidence from all areas of the study suggests that the use of English is growing in informal use and across social networking sites. It creates a threat to the learning of other languages, of more formal English and even to learners' first languages where these are minority languages.

Similarly, it has been argued that since many language teachers are daunted by the speed of technological development and the popularity of new social networking sites, teachers' knowledge and confidence levels in this area, need to be raised through training (The Open University, 2010).

Wikis in the Language Classroom

A wiki is a collaborative website that anyone within the community of users can contribute to or edit. A wiki can be open to a global audience or can be restricted to a select network or community. Wikis can cover a specific topic or subject area. Wikis also make it easy to search or browse for information. Although primarily text, wikis can also include images, sound recordings & films. Wikipedia the free internet encyclopedia is the most well known wiki. Wikis represent a particular type of collaborative learning environment where collaboration can result in aggregated, collective products. Zorko's (2009) research confirms that wikis can be used to enhance effective collaboration in a constructivist approach to language learning in a classroom if a user friendly software is used, the wiki is organized clearly, all necessary resources are provided, students are taught how to use wiki features, students are made aware that their contributions in the wiki are part of the assessment, prompt feedback is given, the group's work is made visible and easily accessible, and a group identity is built. Thus, wikis can be used by teachers to enrich and supplement (online) learning (Choy & Ng, 2007). The literature points to the following wiki characteristics that make them appropriate for uses in education, including language learning:

- User friendliness*:* Wikis are robust, simple, easy to use, constantly updated by the providers, require no technical knowledge and can be used by a very large number of users.
- Flexibility: The information and resources in wikis are available and easily accessed at any time from anywhere (Augar et al., 2004; Nicol et al., 2005). Moreover, the time lag in communication offered by them and other such asynchronous Web 2.0 tools. This temporal delay gives the reader time to reflect before posting a reply. It thus connects students regardless of time and space (Arnold & Ducate, 2006).
- Low cost: There are minimal costs in terms of purchasing, maintenance, support and the teacher's time.

- Digital information Storage: In wikis information such as class instruction documents, reports, web links, ideas, sketches etc. is easily stored (Nicol et al., 2005), and wikis can function as archives of a community's practices (Choy and Ng, 2007) that are later used as repositories of knowledge by new students (Godwin-Jones, 2003).
- Information sharing (Augar et al., 2004; Nicol et al., 2005): Wikis allow public visibility of students' work and facilitates sharing of information, resources and ideas;
- Progress monitoring: Teachers can monitor student participation and assess their work more easily (Augar et al., 2004; Nicol et al., 2005).
- Democratic participation (Schwartz et al., 2004): The nature of wikis permits very easy and free access to any student who has been entrusted with the password, which in turn encourages the participation of all students.
- Student empowerment: The freedom to create, change or delete anything empowers students with more responsibility (Thorne & Payne, 2005).
- Collaboration and social interaction (Augar et al., 2004, p.95; Choy and Ng, 2007; Godwin-Jones, 2003; Schwartz et al., 2004; Thorne and Payne, 2005; Wei et al., 2005; Zorko, 2009): The possibility of the students' use of wikis to fundamentally change their own knowledge constructions and develop a deeper understanding of concepts through transformative dialogue and the construction of new knowledge makes wikis transformative in nature. Oblinger (2005) argues that collaborative online learning environments are particularly suitable for 'Net Generation learners' who have grown up with and used the Internet all their lives, i.e. the 'digital natives' (Prensky, 2001) or the 'NGeners' (Thorne & Payne, 2005).What they expect and want from learning technology is interactivity, be it with a computer, a teacher or a classmate (McNeely, 2005).

Despite the several benefits of using wikis in the language classroom, Lund's (2008) study examines the potential problems facing language production practices in school where the individual learner's output is often the focus of attention and points out that students' typical concerns about using wikis is regarding inexpert editing, their texts getting deleted, or / and someone changing what has been written by them.

- Such concerns testify to the notions of ownership and individual accountability, although learners were aware that previous versions could be restored. These notions are characteristic of how school work traditionally has been linked to individual assessment and grades. Thus, we see the tensions between historical and institutional practices on the one hand, and on the other the emerging practices that in our introduction were linked to knowledge advancement and complex problem solving (Lund, 2008, pp. 48-49).

This goes to show that using wikis in the classroom can create tensions between individual and collective, institutional and novel practices, just as it can hold immense potential for language development.

Podcasts in the Language Classroom

Podcasting is a way of making audio or video files available on the internet that can either be listened to or viewed on a PC or downloaded to a hand-held device such as an iPod or mp3 player. A podcast will be treated as a sound recording (audio podcasts) or a film (video podcasts, as known as vodcasts). Podcasts may also include images, including PowerPoint presentations. If there is a script or lecture notes for the podcast, they are protected as a literary work. There may be multiple layers of copyright in a podcast, depending on the content of the podcast. If there is a presenter or a subject being interviewed, they will not only own copyright in their presentation or interview but they will have performers' rights. There will also be a separate copyright in the actual recording itself. Besides being aware of digital rights management, language teachers must ensure that the podcast content is screened and evaluated for its appropriateness, acceptability, credibility and clarity before it is used in the classroom.

There are various ways in which podcasts can be integrated into the curriculum. Williams (2007) writes that teachers can either use ready-made podcasts or produce their own during their sessions with students or while preparing for their classes; students may also create their own podcasts for poetry and plays; teachers could use podcasts to archive their lectures for future reference and for offering remedial help to students who miss classes because of illness. Students just need to use their computer or portable media players to listen to lectures and review lesson content. Slow students can benefit a lot from podcasts of archived lectures especially if instructors speak faster than students can understand at one time. Podcasts can be located with regular search engines and thus they can be considered as resources that students can listen to and use in their research.

The ubiquity of the MP3 player among students means that podcasting can be an effective tool for language learning. Consequently, there has been a proliferation of teaching material available online. Podcasting can be helpful for language learners / teachers because it provides them:

- Authentic input: Since podcasts are generally created with native speakers in mind, Rosell-Aguilar (2007) argues that podcasts do not only provide access to authentic materials, but also enable users to gain knowledge about the culture, history and politics of the countries in which the target language is spoken. For instance, radio broadcasts (e.g. BBC radio podcasts) for English learning or El Pais podcast for Spanish can be used by language learners of English and Spanish.
- Semi-authentic language: Robin (2007) shows that 'semi-authentic language' can be created specifically for language learners who find authentic texts too difficult or challenging to comprehend.
- Opportunity to create podcasts: Students can be motivated by the idea of creating their own podcasts in the target language either individually or through collaboration. It stretches student users by prompting them to pay greater attention to details in their use of the target language, especially because students are aware of a wider audience (Stanley, 2006).
- Opportunity to subscribe to language courses: Students can listen to any language online through the available language courses anytime, anywhere.

In view of the numerous resources available online, Williams (2007) makes the following observation about the use of podcasts in the classroom:

- Podcasts, presented in the proper context in a learning environment, are a great way to deliver information, especially for auditory (or visual if it's a video podcast) learners. The ability to stop, start and replay also makes for a tool appropriate for students with special earning needs or challenges—as well as everyone else who wishes to learn by repetition. As you listen to and review podcasts for use in your classroom, take special care to think about all the tools available and resist the "just because it's cool" temptation (p. 48).

Thus, podcasts can enrich the language learning environment only when they are selected as the right tool for the right job.

Current Challenges Facing the Use of Web 2.0 Tools in the Language Classroom

Students now have at their disposal a range of Web 2.0 authoring forms such as blogging, social bookmarking, social networking, podcasting, and wiki writing. Many language teachers are interested in enabling students to demonstrate their learning by creating content in these forms. However, incorporating Web 2.0 technologies for language learning is not without its difficulties.

Tackling New Trends

As the discussion in the previous sections show, the rapidly declining number of bloggers and the small number of microbloggers among teens and young adults may limit the of potential of blogs in language learning. Recent literature (discussed before) also indicates that the language used on SNSs can pose a significant threat to the proper learning of any language. Moreover, a closer examination has revealed that wikis can present tensions between the individual student and the collective, and between institutional and novel practices. Thus, in order to be successfully implemented, Web 2.0 applications such as podcasts should not be used simply because it is "cool" to do so, but instead should be chosen to support instruction only after a careful evaluation of their appropriateness and acceptability in the language classroom. Therefore, after an in-depth exploration of the prospects and problems facing the integration of blogs, SNSs, wikis and podcasts, the paper proposes that educators must go through the following stages or take the following steps when incorporating a Web 2.0 tool in the language classroom (see Figure 1 below).

Assessment Methods

Another area that scholars are giving serious thought to is regarding the assessment of Web 2.0 use in the language classroom. The promise of Web 2.0 is that "Learning progress and achievements become visible not only in tests but rather in the learning process documented in portfolios (for example in wikis or weblogs), learning products and social interactions" (Ehlers, 2009, p. 304).The interactivity and social interaction that it encourages cannot be assigned or marked to full effect by using assessment strategies that academics may have used previously, for written reports, essays, examinations or class presentations. In this respect, Web 2.0 activities are different even from earlier forms of online learning activities such as uploading files for assessment or contributing to discussion boards. However, the design and conduct of assessment for such student-created content is a challenging task (Gray et al., 2010). Although including students' use of Web 2.0 authoring in academic learning seems to have educational merit, student Web 2.0 authoring is substantially different from earlier forms of assessable student work. Furthermore since Web 2.0 activities extend the nature not only of individual student work but also of group work “The Web 2.0 tension to be managed is one of deciding how to balance the private and the social within the experience of learning.... But it is also a matter of protecting the realistic demands of assessment" (Crook, Fisher, Graber, Harrison, Lewin, Logan et al., 2008, p. 39).

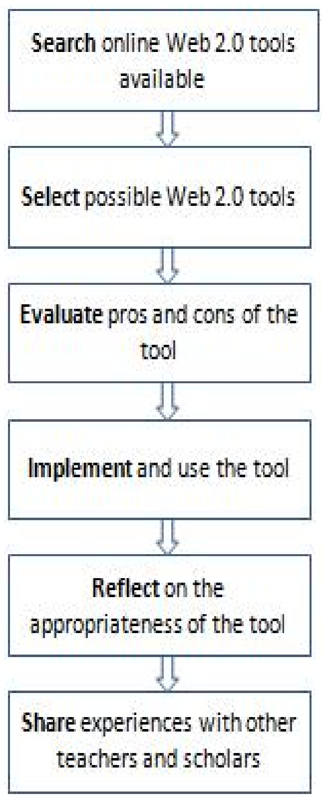


Figure 1. Recommended Steps for Using Web 2.0 Tools in the Language Classroom.

Students’ Perceptions

The digital age presents a paradox and the raises the question regarding the connection between the formal writing teens do and their informal e-communication and written exchanges because despite the fact that a vast majority of teens have eagerly embraced written communication with their peers as they share messages through various web 2.0 technologies, they do not think that the material they create electronically is real writing and dissociate e-communication with “writing” (Lenhart et al. 2008). This casts a doubt on the notion that Web 2.0 applications that rely primarily on written means, such as blogs, SNSs and wikis, can be truly useful for language education. If students remain unconvinced of the benefits of online writing for education, they are likely to regard its use in the language classroom as frivolous. Educators and policy makers can play an important role in showing students the connection between online writing and language learning. But they can only do so if they are well versed and comfortable with the use of Web 2.0 tools in the classroom. However, since ‘old world’ teachers or ‘digital immigrants’ (Prensky, 2001) may not have the necessary skills to combine new technology with their pedagogy, they may feel forced to use unfamiliar tools and work and in unfamiliar ways. Hence, these teachers would benefit from training in teaching with Web 2.0 applications.

Ethical Issues

The introduction of Web 2.0 systems into language learning is not without problems. Web 2.0 opens systems up to much wider and more open use and highlights concerns over student protection and cyber-bullying. Thus, the use of Web 2.0 technologies brings important issues of personal security to the fore and also presents problems with network and IT systems security. Besides security, issues of accessibility to Web 2.0 applications, visibility of content and privacy of users (Anderson, 2007) are other areas which are to be carefully thought of before teachers incorporate Web 2.0 tools into language education. Moreover, since much of Web 2.0 based student work is about content sharing and repurposing, students can easily succumb to the copy-and-paste culture that runs counter to traditional notions of plagiarism. In addition, Web 2.0 raises a variety of issues in relation to intellectual property rights such as that of ownership of content, its re-use and control (Franklin & van Harmelen, 2007). Therefore, it becomes pertinent for policymakers, educational institutions and language teachers to consider current practices and determine how best to address them to enhance language learning.

CONCLUSION

Today’s learners are the primary citizens in this new user-generated participatory Web, also known as the “read-write Web” in contrast to the previous metaphor of “Internet as information superhighway”. The metaphor used for Web 2.0, i.e. the “read-write Web”, itself shows its significance for language education. Since language acquisition today is seen as a fundamentally communicative process catalyzed by students' exposure to authentic language, material, and audiences, for those students who are learning a foreign language and cannot travel to the target country to immerse themselves in the culture and the language, the Internet is a valuable alternative that not only delivers authentic content but also communication capabilities. Emerging Web 2.0 applications continue to expand on traditional notions on literacy and provide affordable ways to bring the target language and culture to the classroom.

Moreover, the participatory nature of Web 2.0 technologies allows students to manage aspects of their own language learning by themselves. This becomes especially important in today's era of lifelong learning in which studying usually continues beyond students' academic lives.

The potential of these tools for enhancing language learning is clear, but a greater body of research is required to inform the creation of guidelines to ensure these tools are integrated appropriately. The role of language technologists is not limited to telling students about blogs or social networking sites. It includes sorting through novel technologies, evaluating their instructional potential, researching current educational uses, and sharing findings with educators. The most promising applications available today were not designed for instructional use and do not come with an instruction manual. Furthermore, many educators concede that student Web 2.0 authoring in higher education raises significant challenges for assessment, posing a barrier to further adoption. This highlights the need to create a systematic process of evaluation, feedback and measurement of outcomes. To use them in the classroom requires the ability to redirect their intended purpose and, more importantly, to think through possible consequences of doing so. Therefore, training teachers in the art of integrating Web 2.0 technologies in language learning programmes will inculcate more confidence among professional language educators in the use and application of these tools. Indeed, as Web 2.0 applications become more mainstream, new and exciting developments in the area of language learning are likely to continue apace.

REFERENCES

Anderson, P. (2007). *What is Web 2.0? Ideas, technologies and implications for education.* JISC techwatch report. Retrieved from http://www.jisc.ac.uk/media/documents/techwatch/tsw0701b.pdf

Arnold, N., & Ducate, L. (2006). Future foreign language teachers' social and cognitive collaboration in an online environment. *Language Learning & Technology*, 10(1), 42-66.

Augar, N., Raitman, R., & Zhou, W. (2004). Teaching and learning online with wikis. In *Beyond the comfort zone: Proceedings ASCILITE Perth 2004* (pp. 95-104). Perth, 5-8 December. Retrieved from http://www.ascilite.org.au/conferences/perth04/procs/augar.html

Berners-Lee, T. (2006). *Linked Data - Design Issues.* Retrieved from http://www.w3.org/DesignIssues/LinkedData.html

Campbell, A. (2003). Weblogs for Use with ESL Classes. The Internet TESL Journal, 9(2). Retrieved from http://iteslj.org/Techniques/Campbell-Weblogs.html

Carroll K. (2008). *Skype, social networks and learning.* Retrieved from http://ken-carroll.com/2008/01/14/skype-social-networks-and-language-learning/

Castro, M. (2009). The use of microblogging in the classroom. Proceedings of the *Third International Wireless Ready Symposium.* Retrieved from http://wirelessready.nucba.ac.jp/castro2009.pdf

Choy, S.O., & Ng, K.C. (2007). Implementing wiki software for supplementing online learning. *Australasian Journal of Educational Technology*, 23(2), 209-226.

Crook, C., Cummings, J., Fisher, T., Graber, R., Harrison, C., & Lewin, C. (2008). Web 2.0 technologies for learning: The current landscape – opportunities, challenges and tensions. Retrieved from http://partners.becta.org.uk/uploaddir/downloads/page_documents/research/web2_technologies_learning.pdf

Dalsgaard, C., & Sorensen, E. (2008). A typology for Web 2.0. Proceedings of *The 7th European Conference on e-learning, ECEL 2008*. Retrieved from http://www.cil.au.dk/index.php/component/content/article/15-publikationer/38-artikel-a-typology-for-web-20

Ehlers, U. (2009). Web 2.0 - e-learning 2.0 - quality 2.0? Quality for new learning cultures. *Quality Assurance in Education,* 17(3), 296-314.

European Commission (2009). *Study on the impact of ICT and new media on language learning*. Executive Agency: Education, Audiovisual and culture. Retrieved from http://eacea.ec.europa.eu/llp/studies/study_impact_ict_new_media_language_learningen.php

Franklin, T., & van Harmelen, M. (2007). *Web 2.0 for content of teaching and learning in higher education.* Retrieved from http://ie-repository.jisc.ac.uk/148/1/web2-content-learning-and-teaching.pdf

Godwin-Jones, R. (2003). Blogs and wikis: Environments for on-line collaboration. *Language Learning & Technology*, 7(2), 12-16. Retrieved from http://llt.msu.edu/vol7num2/pdf/emerging.pdf

Gray, K., Thompson, C., Sheard, J., Clerehan, R., & Hamilton, M. (2010). Students as Web 2.0 authors: Implications for assessment design and conduct. *Australasian Journal of Educational Technology*, 26(1), 105-122.

Harrison, R., & Thomas, M. (2009). Identity in online communities: Social networking sites and language learning. *International Journal of Emerging Technologies and Society*, 7(2), 109-124.

Knobel, M., &Wilber, D. (2009). Let's talk 2.0. Educational Leadership, 66(6), 20-24. Retrieved from http://www.ascd.org/publications/educational_leadership/mar09

Lenhart, A., Purcell, K., Smith, A., & Zickuhr, K. (2010). *Social media and young adults.* Pew Internet. Retrieved from http://www.pewinternet.org/Reports/2010/Social-Media-and-Young-Adults.aspx

Lenhart, A., Arafeh, S., Smith, A., & Macgill, A. (2008). *Writing, technology and teens.* Pew Internet. Retrieved from: http://www.pewinternet.org/Reports/2008/Writing-Technology-and-Teens.aspx

Lund, A. (2008). Wikis: A collective approach to language production. *ReCALL,* (20) 1, 35-54.

McLoughlin, C., Hutchison, H., & Koplin, M. (2002). Different Media for Language Learning: Does Technology Add Quality? *International Conference on Computers in Education, ICCE'02* (p. 881). New Zealand, Auckland, 3-6 December.

McNeely, B. (2005). Using technology as a learning tool, not just the cool new thing. In D. Oblinger (Ed.), *Educating the Net Generation.* EDUCAUSE e-book. Boulder: Colorado. Retrieved from http://www.educause.edu/educatingthenetgen

Nicol, D., Littlejohn, A., & Grierson, H. (2005). The importance of structuring information and resources within shared workspaces during collaborative design learning. *Open Learning: The Journal of Open and Distance Learning*, 20(1), 31-49.

Oblinger, D. (2005). Learners, learning and technology: The Educause learning initiative. *EDUCAUSE Review*, 40(5), 67-75.

O'Reilly, T. (2005). *What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software*. Retrieved from http://oreilly.com/web2/archive/what-is-web-20.html

Penrod, D. (2007). *Using blogs to enhance literacy: The next powerful step in 21st century learning*. Lanham, MD: Rowman & Littlefield Education.

Pinkman, K. (2005). Using Blogs in the Foreign Language Classroom: Encouraging Learner Independence. The JALT CALL Journal, 1(1), 12-24. Retrieved from: http://jaltcall.org/journal/articles/1_1_Pinkman.pdf

Robin, R. (2007). Commentary: Learner-based Listening and Technological Authenticity. Language Learning & Technology, 11(1), 109-115. Retrieved from http://llt.msu.edu/vol11num1/robin/default.html

Rosell-Aguilar, F. (2007). Top of the Pods-In Search of a Podcasting "Podagogy" for Language Learning. *Computer Assisted Language Learning*, 20(5), 471-492.

Rowe, C., & Wyss, E.L. (2009). *Language and new media*. Creskill, NJ: Hampton Press.

Prensky, M. (2001). Digital natives, digital immigrants. *On the Horizon*, 9(5), 1-6. Retrieved from http://www.marcprensky.com/writing/prensky%20-%20digital%20natives,%20digital%20immigrants%20-%20part1.pdf

Schwartz, L., Clark, S., Cossarin, M. & Rudolph, J. (2004). Educational wikis: Features and selection criteria. *The International Review of Research in Open and Distance Learning*, 5(1).

Shihab, M.M. (2009). *Web 2.0 Tools Improve Teaching and Collaboration in High School English Language Classes.* A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Computing Technology in Education. Graduate School of Computer and Information Sciences, Nova Southeastern University.

Simon, E. (2008). Foreign language faculty in the age of Web 2.0. EDUCAUSE Quarterly, 31 (3).

Solomon, G., & Schrum, L. (2007). *Web 2.0: New tools, new schools.* Eugene, OR: International Society for Technology in Education (ISTE).

Stanley, G. (2006). Podcasting: Audio on the Internet comes of age. Teaching English as a Second or Foreign Language, 9(4). Retrieved from http://www-writing.berkeley.edu/TESL-EJ/ej36/int.html

The Open University (2010). *Social networking threat to language learning.* Retrieved from http://www3.open.ac.uk/media/fullstory.aspx?id=18269

Thorne, S. L., & Payne, J. S. (2005). Evolutionary Trajectories, Internet-mediated Expression, and Language Education. The CALICO Journal, 22(3), 371-397. Retrieved from http://language.la.psu.edu/%7Ethorne/thorne_payne_calico2005.pdf

Varvel, V. (2005). Honesty in online education. *Pointers and Clickers,* 6(1), 1-20. Retrieved from http://www.ion.uillinois.edu/resources/pointersclickers/2005_01/VarvelCheatPoint2005.pdf

Wei, C., Maust, B., Barrick, J., Cuddihy, E., & Spyridakis, J.H. (2005). Wikis for supporting distributed collaborative writing. *Proceedings of the Society for Technical Communication 52nd Annual Conference.* Seattle, WA. May 8-11, 2005. The Society for Technical Communication: Arlington, VA, USA. Retrieved from http://faculty.washington.edu/farkas/TC510/Wei-WikisForSupportingCollabWriting.pdf

Williams, B. (2007). *Educator's podcast guide*. Eugene, OR: International Society for Technology in Education.

Zorko, V. (2009). Factors affecting the way students collaborate in a wiki for English language learning. *Australasian Journal of Educational Technology*, 25(5), 645-665.

In: Advances in Communications, Volume 8
Editor: Anthony V. Stavros
ISBN: 978-1-61324-794-5

Chapter 7

APPROACH OF IMPROVED TOPOLOGY DEVELOPMENT PROTOCOL IN AD HOC NETWORK MINIMIZING THE NUMBER OF HOPS AND MAINTAINING CONNECTIVITY OF MOBILE TERMINALS WHICH MOVE FROM ONE TO THE OTHERS

***Kohei Arai*[*] *and Lipur Sugiyanta*[†]**
Department of Information Science
Faculty of Science and Engineering
Saga University
Saga, Japan

Abstract

Wireless ad-hoc mesh network is a special kind of network, where all of the nodes move in time. Node is intended to help relaying packets of neighboring nodes using multi-hop routing mechanism in order to solve problem of dead communication. Wireless mesh network which engages broadcasting and contains multiple hops become increasingly vulnerable to problems such as routing problem and rapid increasing of overhead packets. During this progress, the delay on account of multi hop characteristics and redundant packets caused by communication nature potentially existed during communication. Typically, delay will increase in linearity with number of hops. There is a certain minimum level of delay that will be experienced due to the time it takes to transmit a packet through a link. Topology development holds a significant point prior to the data transmission. Without improved topology development protocol, this problem can decrease network's performance in overall data transmission. We analyze the delay

[*] E-mail address: arai@is.saga-u.ac.jp
[†] E-mail address: lipurs@gmail.com

performance of a multi-hop wireless network with a dynamic route between each source and final destination pair. There are fluctuate interference constraints on the set of links that impose a fundamental delay performance of any instant network topology. At first, we present a similar Link State Routing network simulation to derive such referential lower bounds. We conduct extensive simulation studies to suggest that the average delay of multi-hop transmission policy can be made lower compared to the referential bound by using appropriate functions of network metrics. This chapter provides a broadcast framework that engages various network metrics and at the same time maintaining connectivity of nodes (mobile terminals). The framework captures the essential features of the wireless network metrics, i.e. bandwidth, throughput, network buffer, direction, and round trip time. This research is useful since, in many cases, it find that the throughput is the most important parameter in reduction of delay transmission. This result is confirmed with another composite simulation result. Most of network hop delay is impacted with this composite metric, particularly in delay minimization on the longer hops. The reduction achievement on average delay by this algorithm is 0.577% and the total average delay reduction for this simulated network is 0.683%. This research will be further designed primarily for achieving maximum throughput in the multiple wireless network area.

Keywords: multi-hop, delay, connectivity, metrics.

1. INTRODUCTION

A Wireless (Ad-Hoc) Mesh Network consists of mobile nodes platforms which are free to move in the area. Node is referred to a mobile device which equipped with built-in wireless communications devices attached and has capability similar to autonomous router. The nodes can be located in or on airplanes, ships, cars, or on people as part of personal handheld devices, and there may be multiple hosts among them. Each node is autonomous. The system may operate in isolation, or have gateways to a fixed network. In the future operational mode, multiple coverage of the network is expected to operate as global “mobile network” connecting to legacy “fixed network”.

At each time and every node’s positions, a wireless connectivity in the form of a random, single-hop, multi-hop path may exist among nodes. This topology may change as the nodes move or adjust their parameters. Among networks, Wireless (Ad-Hoc) Mesh Network has several characteristics:

1) Dynamic topologies,
2) Bandwidth-constrained,
3) Energy-constrained operation, and
4) Limited physical security.

These characteristics create a set of underlying assumptions and performance considerations for protocol design which extend beyond static topology of the fixed network. The design should reacts efficiently to topological changes and traffic demands while maintaining effective routing in a mobile networking context.

All nodes in Wireless (Ad-Hoc) Mesh Network rely on batteries or other exhaustible energy modules for their energy. As a result of energy conservation or some other needs, nodes may stop transmitting and/or receiving for arbitrary time periods. A routing protocol should be able to accommodate such sleep periods without overly adverse consequences. If this is done intelligently, it can utilize network energy and bandwidth resources more efficiently, at the cost of increased route discovery delay. For this network, the design of routing algorithm which able to adapt to the traffic pattern on a demand or need basis is needed for overall network performance.

Routing schemes differ in their delivery semantics: (a) unicast delivers a message to a single specified node; (b) broadcast delivers a message to all nodes in the network; (c) multicast delivers a message to a group of nodes that have expressed interest in receiving the message; and (d) anycast delivers a message to any one out of a group of nodes, typically the one nearest to the source. Broadcast is the dominant form of message delivery on the wireless network.

Research on multi-hop wireless networks have been devoted to routing protocol, system stability, network performance, and throughput maximization. The delay performance of these systems, however, has largely been an open problem. This problem is enormously difficult even in the context of legacy wire networks, primarily because of complex network interactions that complicate the multi-hop routing mechanisms. The chapter presents an analysis delay performance network to obtain an improvement average delay of packet transmission from source to the destination in the dynamic wireless network environment. Furthermore, we re-engineer a routing metric topology development to achieve good delay performance while at the same time maintaining connectivity to other nodes.

In this chapter, the proposed framework provides the broadcast schemes in Wireless (Ad-Hoc) Mesh Network that engages various network metrics inside and at the same time maintaining connectivity of nodes (mobile terminals). Various formation options of nodes and their potential overheads and impacts on reduction of delay performance are evaluated via simulation study. We analyze a multi-hop wireless network with multiple source final destination pairs, given routing and traffic information. The remainder of this chapter is organized as follows: Section 2 gives preliminaries and our model. Section 3 discusses the detail design of the simulation model, its notations, and assumptions. A performance evaluation of generic algorithms and comparison to a similar Link State Routing broadcast-based network that uses distance weight as the link cost unit are presented in Section 4. Section 5 concludes the chapter.

2. Preliminaries

Wireless mesh network is generally set up with a centralized access point for provide high level of connectivity in certain area. The access point has knowledge of all devices in its area and routing to nodes is done in a table driven manner [1], [2], [5]. Nemoto introduced a technical review of wireless mesh network technology products that implemented IEEE802.11 standard through experiments of fixed wireless mesh network nodes [2]. In terms of review the network performance at this stage, it will be represented as the view of use and evaluation of outdoors Muni-WiFi devices in accordance to applying the legacy LAN technology inside the corporate network. Performance of network access layer, i.e. performance of voice and TCP data transmission in terms of throughput, response time between mesh nodes, and communication delay in multi-hop transmission are presented.

However, Nemoto intended to operate in static topology network [2]. With recent performance in computer and wireless communications technologies, advanced mobile wireless is expected to see increasingly widespread use and application. The vision of future mobile ad hoc networking is to support robust and efficient operation in mobile wireless networks by incorporating routing functionality such that networks are capable to be dynamic, rapidly-changing with random, multi-hop topologies which are likely composed of relatively bandwidth-constrained wireless links. Supporting this form of host mobility requires address management, protocol interoperability enhancements and the likes.

Broadcasting plays a critical role especially in vehicular communication where a large number of nodes are moving and at the same time sending a large size of packet. Broadcast is essential in order to either transmit packet data or build topology prior to data transmission. In wireless network where nodes communicate with each other using broadcast messages, the broadcast environment works as receivers collect information from all transmitting nodes within its coverage pattern's neighborhood, and then allowing receivers to aware of immediate surrounding respond before re-transmitting packet. Several transmissions may be unnecessary during broadcasting mechanism. These redundant cause the broadcast storm problem [8], in which redundant packets cause contention and collision consume a significant delay of the aimed transmission time. Thus, routing protocols should be capable to respond these changes using minimum signaling and taking into account the routing topology metric as a parameter distributed in the network. To address these challenges, a delay-efficient routing must satisfy the following properties; (1) ensure high throughput and (2) allocate resources equitably. Multiple link hop paths in the network have to be scheduled such that selected link paths are not starved for service. Starvation leads to an increase in the average delay in the network.

Minimizing the delay involves applying a path selection mechanism of routing metric to multiple routes, in order to select (or predict) the best route. The metric is computed by a routing algorithm, and can cover such information as throughput, round trip time, hop count, interface buffer, load, MTU, direction, and communication cost. The routing table stores the best possible routes and other potential routes.

3. Simulation Model, Notations, and Assumption

Relays are intermediate nodes between source and final destination which help relaying packets using multi hop routing mechanism. The appearance of relays is required to avoid dead communication if the distance is not in the proximity of each node. It can extend the mobility and expand the coverage area, but in the same time increase the delay time. In general, the more relay nodes (hop), the longer the delay time. Energy consumption is also affected. Direct transmission is seemed to have more aggregate energy required than indirect communication. Thus the adjustment of relay nodes will influence the balance of delay time and energy consumption.

The proposed framework assumes that nodes are capable of dynamically adjusting their relay nodes on per move step base. It attempts to minimize the number of relay nodes between source and final destination pairs and at the same time maintain the node's energy level required. This behavior is almost similar to MANET routing protocols (e.g., AODV, DSR and TORA). One common property of these routing protocols is that they discover routes using broadcast flooding protocols whose value of distance metric in order to minimize the number of relay nodes between any source and final destination pair.

The approach is initiated from broadcast mechanism and propagated through node-to-node based routing topology metrics approach. During propagation, it takes into account all topology development, route discovery, and data transmission. Each source injects single big packet which fragmented into multiple packets in the network, which traverse through the network until those reach the final destination. Packets are queued at each node in its path where it waits for an opportunity to be transmitted. This model is not only applicable in direct communication (one hop transmission) but it can also work in multi-hop transmission. In this situation, when the source and final destination nodes are located outside the maximum transmission range, source node is capable to discover multiple hop routing efficiently thus maintain the energy level required in comparison to standard flooding based ad hoc routing designs.

The Model

Simulation describes that antenna module installed in each node is capable of dynamically adjusting the transmission energy used to communicate with other nodes. Industrial standard of antenna module supports a management for controlling this energy consumption. Simulation assumes that the energy consumption required to transmit a packet between nodes A and B is similar to that energy required between nodes B and A if and only if the distance and the size of packet are same. The coverage distance range of the nodes is a perfect symmetric unit disk (omni-directional). If $d_{x,y} \leq r_x \rightarrow$ x and y can see each other. This assumption may be acceptable in the condition that interference in both directions is similar in space and time; which is not always the case. Usually interference-free Media Access Control (MAC) protocol such as Channel Sense Multiple Access (CSMA) may exist. In addition, wireless link channel is assumed to have no physical noise; i.e., the errors in packet reception due to fading and other external interferences are not considered as a serious problem. Packets from sender to receiver will be transmitted as long as the bandwidth capacity is sufficient and the received signal to noise ratio (SNR) is above a certain minimum value. Thus every packet successfully received is acknowledged at the link layer and de-encapsulate at the higher layer. Each node is capable of measuring the received SNR by analyzing overheard packet. A constant bit error rate (BER) is defined for the whole network. Whenever a packet is going to be sent, a random number is generated and compared to the packet's CRC. If the random number is greater, the message is received, otherwise it is lost. The default value for the BER is 0, which means there is no packet loss due to physical link error.

Simulation cover a single area of homogeneous nodes that communicate with each other using the broadcast services of IEEE 802.11. There are nodes with different roles simulated in this simulation, namely initiator node/source node, receiver node, sender node, destination node, and final destination node. Initiator node/source node is node that initiates transmission of packet. Packet can be either route discovery or data transmission. Like other nodes, initiator is always moving with random direction, speed, and distance. At the time it is moving, initiator node is always sensing its neighbor to maintain connectivity. Receiver node is node that can be reached by source/sender node. Nodes are defined as neighbors if it located within its distance radius range. At initial time, node senses its neighbors before packet data is required to be transmitted. Coverage neighbor nodes always receive packets that are broadcasted from sender. Destination node is selected receiver node in multi hop transmission that should relay packets to the next receiver node. Final destination node is node that became the end destination of packets.

The layered concept of networking was developed to accommodate changes in local layer protocol mechanism. Each layer is responsible for a different function of the network. It will pass information up and down to the next subsequent layer as data is processed. Among the seven layers in the OSI reference model, the link layer, network layer, and transport layer are 3 main layers of network. The framework is configured in those layers. Genuine packets are initiated at Protocol layer, and then delivered sequentially to next layer as assumed that fragmented packets to be randomly distributed. Simulation models each layer owned with finite buffers. Limited buffer makes packets are queued up according to the drop tail queuing principle. When a node has packets to transmit, they are queued up provide the queue contains less than K elements ($K \geq 1$). To increase the randomization of the simulation process, simulation introduces some delay on some common processes in the network, like message transmission delay, processing delay, time out, etc. This behavior will result that at each instance of a simulation would produce different results. The packets exchanged between sender and receiver is of a fixed rate transmission λ based on a Poisson distribution. Nodes that have packet queued are able to transmit it out using in each available bi-directional link channel.

Energy is power kept in each node. Heinzelman et al. assumed that the radio dissipates E_{elec} = 50 nJ/bit to run the transmitter or receiver circuitry and ε_{amp} = 100 pJ/bit/m^2 for the transmit amplifier [6]. Thus, to transmit a k-bit message a distance d using this radio model, the radio expends:

$$E_{Tx}(k,d) = E_{elect} * k + \epsilon_{amp} * k * d^2 \tag{1}$$

and to receive this message, the radio expends:

$$E_{Rx}(k) = E_{elect} * k \tag{2}$$

The energy model included in simulation was based on the following formulas, taken from [5]:

$$E_{TXBit} = E_{elect} + \left(\varepsilon_{amp}.(\pi r^2)\right) \tag{3}$$

$$E_{RXBit} = E_{elect} \tag{4}$$

The energy behaviors of node are defined as follow:

- During the idle time, a node does not spend energy. Even though this assumption has been proven untrue because being idle might be as costly as receiving data, this is still an assumption that can be done in most experiments, since the most important factor is the overhead in terms of message exchange and its associated cost.
- The nodes are assumed to have one radio for general messages. The main radio is used in all operations when the node is in active mode, and to send and receive

control packets. When this radio is turned off, then no messages will be received and no energy will be used.

- Energy distribution among nodes can either be constant value, normally distributed, Poissonly distributed, or uniformly distributed.

Improved Topology Algorithm

The goal of this research is to find out the essential metric of wireless network parameters which required during topology development prior to data transmission in the network and maintaining connectivity to the others. The parameters analyzed are distance, round trip time, interface buffer/network buffer, throughput, energy, and combination of those.

The core algorithm is developed from static mode (e.g., sensor networks). The enhancement algorithm for serving mobility then detailed in support of topology building, topology maintenance, and routing maintenance. We show our methodology on a tree network. The tree topology decomposes the paths between source and final destination into several route paths. The algorithm underestimates the interference between the route paths. The algorithm starts to operate with building the network topology. Establishment of the next hop in this multi-hop environment is conducted with topology development mechanism. The role of the route maintenance algorithm is to make sure that a minimum flow of packets is transmitted in order to maintain the route when there are no data packets available to send at the transmitter.

Network topology must be executed before data transmission takes place. Topology development is proactive, it uses Topology Control (TC) messages to discover and disseminate link state information. It involves transmit and receives of HELLO packets, REPLY packets, CONFIRM packets, and so on; mostly redundant. These are packets that successfully received by link layer and will update an entry in the neighbor table which cache information about surrounding nodes exists. HELLO packets and corresponding REPLYs have contents of [ID, hop, energy, time, initialTime], where ID is a unique neighbor node (IP address), hop is a number which increment each time packet reach at relay node, energy is current available energy level needed to ensure the communication with the neighbor node, time is current time at which this event is executed, and initialTime is time from which this event was generated.

The routing maintenance algorithm is responsible for performing the route optimization operation that leads to the discovery of routes changes. The route maintenance algorithm performs two basic operations: initiate broadcast maintenance packets, which computes whether a route optimization between two nodes is needed and sets up broadcast mechanism; and executes maintenance packets, which determines when to transmit routing maintenance packets. The framework optimizes routes through sequence of steps to converge to an optimum route. The step refers to the event in which a packet initiates a source node to transmit a Hello request for the first time. The network will converge as fast as the transmission speed of data transmitted by node.

We built network simulator to evaluate this performance. The simulator supports physical, link and routing layers for single/multi hop ad-hoc networks. We assume that IEEE 802.11 Distributed Coordination Function (DCF) or MAC protocol which uses Channel Sense Multiple Access with Collision Avoidance (CSMA/CA) already deployed.

Successfully received packet by receiver's interface is packet whose SNR is above a certain minimum value otherwise the packet cannot be distinguished from background noise/interference. Packets are transmitting through physical layer in accordance with Poisson distribution. Communication between two nodes in IEEE 802.11 uses RTS-CTS signaling before the actual data transmission takes place. Simulation simulates this with random hearing to link's condition.

The simulator uses two-steps propagation model to simulate interactive propagation in the operation of the protocol in dynamic environment. As a future research, the appropriate propagation model that best matches to this environment should replace the simple two-steps model presented here. The two-steps propagation model is appropriate for outdoor environments where a line of sight communication existed between the transmitter and receiver nodes and when the antennas are omni-directional.

When a node first starts, it only knows of its immediate neighbors, and the direct cost involved in reaching them. (This information, the list of destinations, the total cost to each, and the next hop to send data to get there, makes up the routing table, or distance table.) Each node, on a regular basis, sends broadcast packets to neighbors to get all costs of destinations. The neighboring node(s) examine this information, and compare it to what they already 'know'; anything which represents an improvement on what they already have, thus update their own routing table(s). Over time, all the nodes in the network will discover the best next hop for all destinations, and the best total cost. When one of the nodes involved goes down, those nodes which used it as their next hop for certain destinations discard those entries, and create new routing-table information. They then pass this information to all adjacent nodes, which then repeat the process. Eventually all the nodes in the network receive the updated information, and will then discover new paths to all the destinations which they can still "reach".

During this sequence, relay node is determined by relevant information gathered from neighbor nodes. After omitted redundant packets and based on calculation metric value, relay node is set (i.e., a small set of nodes that potentially forward the broadcast packet) to achieve high delivery ratio with certain metric consideration. It means that only selected neighbors able to forward the packet to the next neighbors. The selected neighbor or new relays added to a route during iteration are very much dependent on the relay found in the previous iteration. The set can be selected dynamically (based on both topology and broadcast state information). This relay node set forms a connected dominating set (CDS) and achieves full coverage of connected network. It is possible that the first iteration, which seemed as most optimum value of metric value is not the route achieving the optimum topology with optimum delay path.

Real data transmission is triggered by source node which injects one packet into the protocol layer. The packets either fragmented or not, flow through node layers at every time-slot. The length of the active periods (denoted by random variable) is distributed randomly according to Mersenne Twister algorithm. The mean of transmission rate and arrival rate of packets can be controlled by changing the value of p, a Poisson distribution. Upon receiving a packet, neighbors create reply packet which contain its condition (i.e. trip time, interface buffer) and sent it back to sender. The arrival process is defined as the arrival packets stream at each node is a series of active and idle periods. The received packet is then processed by the layering module with the result that one of the following actions is taken: (i) the packet is passed to the higher layers if both MAC and IP addresses match; (ii) the packet is dropped if neither MAC nor IP addresses match; or (iii) the packet is forwarded to another node when only the MAC address matches. In the latter case, it searches the routing table to find the next route node with the higher metric calculation to reach next destination node. When receiver receives a packet data from the higher layers it searches the routing table to see if a route toward the destination node exists. If this is not the case, node searches the neighbor table to see if information regarding the destination node is available. After the neighbor node replies with a packet of its own then route optimization follows as described previously. When nodes are mobile and no data packets are available for transmission, a source node required to transmit explicit signaling packets to maintain a topology.

Because several relay nodes may exist between source and final destination, the source node will choose the one providing a highest metric value. Multiple packets are sent to that single (next) relay node. From the simulation, it noted that transmission of multiple route-redirect packets wastes bandwidth and network resources (overhead packets increased). For sparsely populated networks, this may not be a problem. However, this is an issue in the case of densely populated networks where several potential nodes can be chosen. The framework addresses this issue by giving priority for the execution of an update routing maintenance packet to the potential neighbor node that computes highest route metric energy-distance values first. After receiving an update routing maintenance packet, a node modifies its routing table, putting the source of the received packet as the next hop node for the specific sender-destination route path. To execute preferential event in sequentially distributed events, we used a simple approach that consists of applying a different time-event execution after the triggering event takes place. The lower and upper bound of the queuing interval are set such that they do not interfere with predefined timers used by the other events for layers and modification events.

4. Performance Evaluation

In this section, an evaluation of the framework is discussed and followed by a number of performance issues associated with routing metrics and route maintenance. Much of the analysis for multi-hop wireless networks has been limited to establishing the performance of the network. In this chapter, we have taken an important step towards the expected delay analysis of network.

The general research on the delay analysis has progressed in the following main directions:

1. Heavy traffic using fluid models: used to either establish stability of the system or to study the workload process in the heavy traffic environment.
2. Large Deviations: to calculate queue-overflow probability.

Here, we have taken a different approach to reduce the wireless network to single topology tree_network which are then analyzed to construct the delay performance for data transmission. This technique captures the essential features of the wireless network and is useful since, in many cases, we can also find that the proposed framework performs higher than the similar Link State Routing (LSR) network. The similar LSR network is selected because it is simple to deploy and can be used for analyzing a large scale of packets processes using known network topology. We compare the framework and similar LSR network to best understand the various tradeoffs and limitations of the algorithm.

A similar (LSR) network would generate full routing tables in advance where, all nodes in the network would be aware of distance level and routes to all other nodes in the network. This network can compute the optimum metric with shortest distance to a next relay node by listening replies of topology construction and topology maintenance packets transmitted by the neighbors. This network operation requires each node in the network to broadcast a routing packet. The broadcast packets contain information about the distance metric of all known destinations. Each node floods the network with information about what other nodes it can connect to, and the received packets may require to be forwarded by other nodes to propagate the entire network. After collecting packets from all nodes of the network, any node should be capable of computing optimum routes to any other node in the network. Each node then independently assembles this information into a tree. Using this tree, each node then independently determines the least-cost path from itself to every other node using a standard shortest paths (distance) algorithm. The iteration of propagation events to be entirely flooded mainly depends on the density of nodes in the network. The result is a tree rooted at the source node such that the path through the tree from the root to any other node is the least-cost path to that node. This tree then serves to construct the routing table, which specifies the best next hop to get from the current node to any other node.

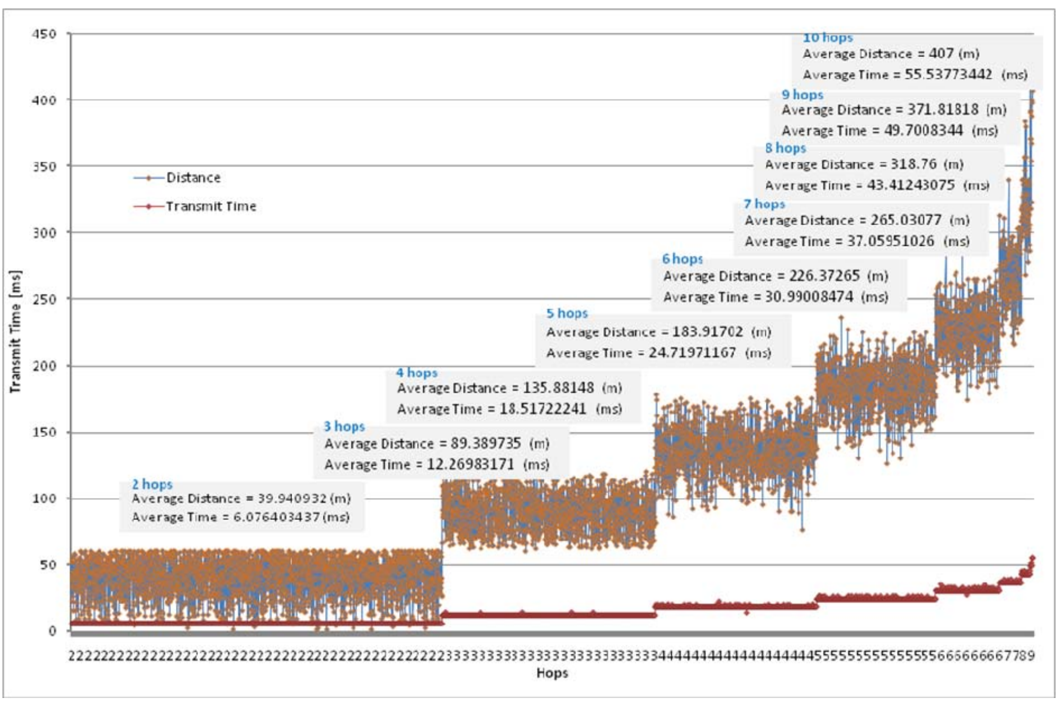


Figure 1. The delay performance of similar LSR network.

We consider a network composed of 20 nodes located within transmission range of each other. In this example, we analyze the wireless tree topology with randomly generated flows initially described in Section 3. There are several metrics in this system which investigated with each topology development under the 20 nodes network model. We studied the network for several scenarios with different initiator nodes load. We find that depending on the input topology metric, the algorithm computes different hop/relay connectivity type for the packets flow in the system. We discuss five representative topology metrics to evaluate the impact of the topology development made in the analysis. The mean arrival rate of fragmented packets of completed 10000 bytes packet, a Poisson distribution, is set to 100.

Case 1. For the given round trip time metric weight, the algorithm computes the decomposition of topology into maximum 10 tree hops. 37.37% is two hops, 21.98% is three hops, 17.68% is four hops, 12.71% is five hops, and rest is for other hops. Note that this metric impacts on the number of hops from two to four significantly, but this effect is not resulting in larger delay minimization on the longer hops. Also note from Figure 2 that the average distances for different hops are almost similar. The higher reduction on average delay by this metric algorithm is 0.183% compared with Figure 1 respectively. The total average delay reduction for this simulated network is 0.344%.

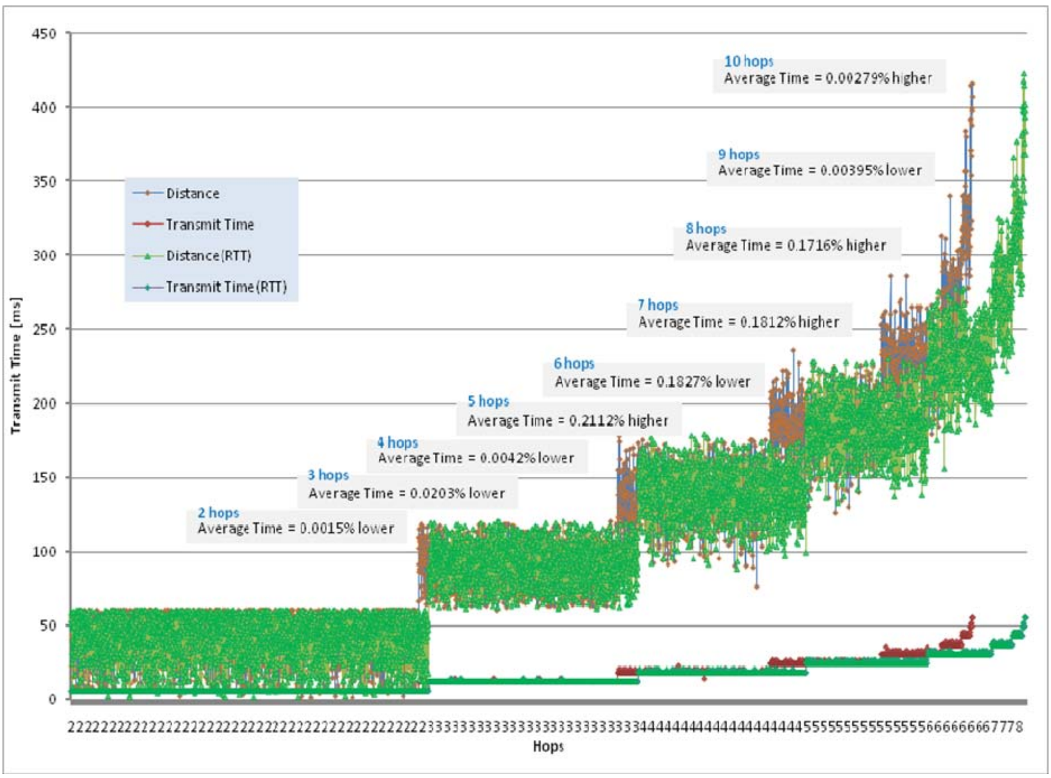


Figure 2. Comparison between improved Topology construction with RTT metric and similar LSR network.

Case 2. For the given throughput metric weight, the algorithm computes the decomposition of topology into maximum 12 tree hops. 37.25% is two hops, 21.58% is three hops, 18.57% is four hops, 12.88% is five hops, and rest is for other hops. Almost each hop delay is impacted with this metric from two to eight hops, but this effect is not resulting in larger delay minimization on the longer hops. Also note from Figure 3 that the average distances for different hops are almost similar, except for 11 and 12 hops. The higher reduction on average delay by this metric algorithm is 0.359% compared with Figure 1 respectively. The total average delay reduction for this simulated network is 0.542%.

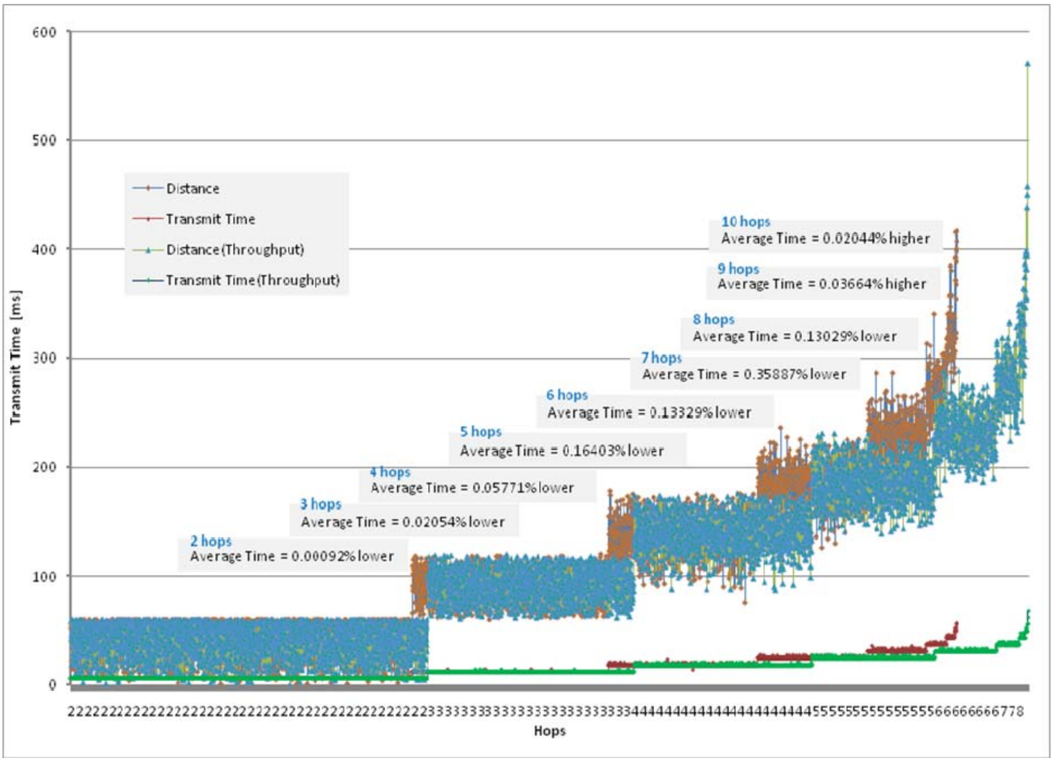


Figure 3. Comparison between improved Topology construction with Throughput metric and similar LSR network.

Case 3. Direction can be distinguished as approaching and away. For the given direction metric weight, the algorithm computes the decomposition of topology into maximum 10 tree hops. 37.22% is two hops, 22.32% is three hops, 18.91% is four hops, 12.58% is five hops, and rest is for other hops. Only small portion of network delay is impacted with this metric, particularly from five to seven hops, but this effect is not resulting in delay minimization on the shorter hops. From Figure 4, the average distances for different hops are almost similar. The higher reduction on average delay by this metric algorithm is 0.388% compared with Figure 1 respectively. The total average delay reduction for this simulated network is 0.522%.

Case 4. A combination of distance metric and energy metric is used. For the given metric weight, the algorithm computes the decomposition of topology into maximum 12 tree hops. 36.18% is two hops, 22.01% is three hops, 18.13% is four hops, 11.65% is five hops, and rest is for other hops. With the average distances for different hops is almost similar, less portion of network delay is impacted with this metric as shown in Figure 5. The higher reduction on average delay by this metric algorithm is 0.253% compared with Figure 1 respectively. The total average delay reduction for this simulated network is almost zero.

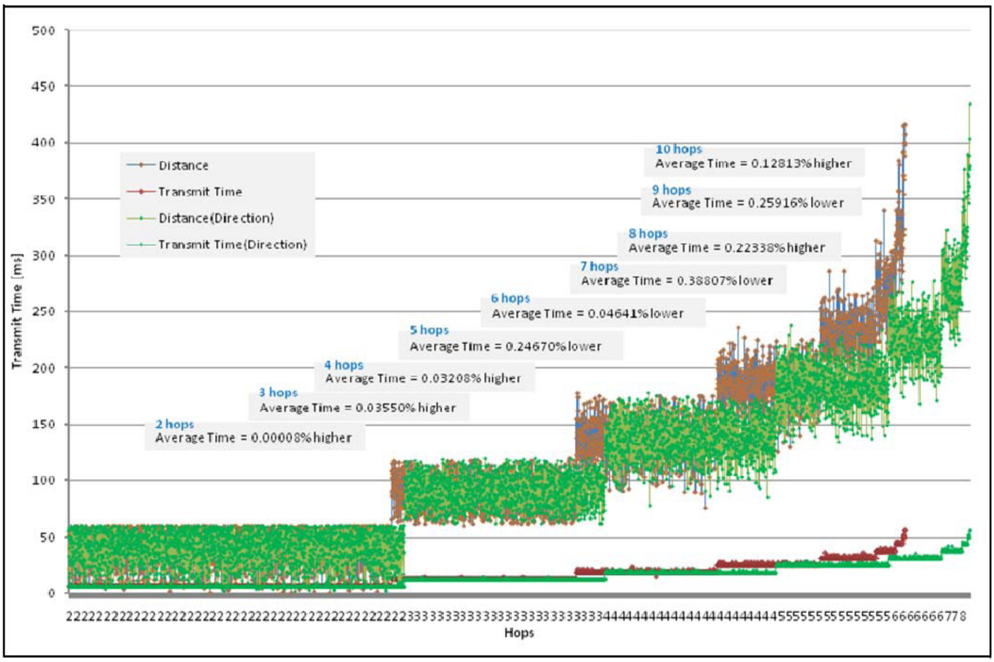


Figure 4. Comparison between improved Topology construction with Direction metric and similar LSR network.

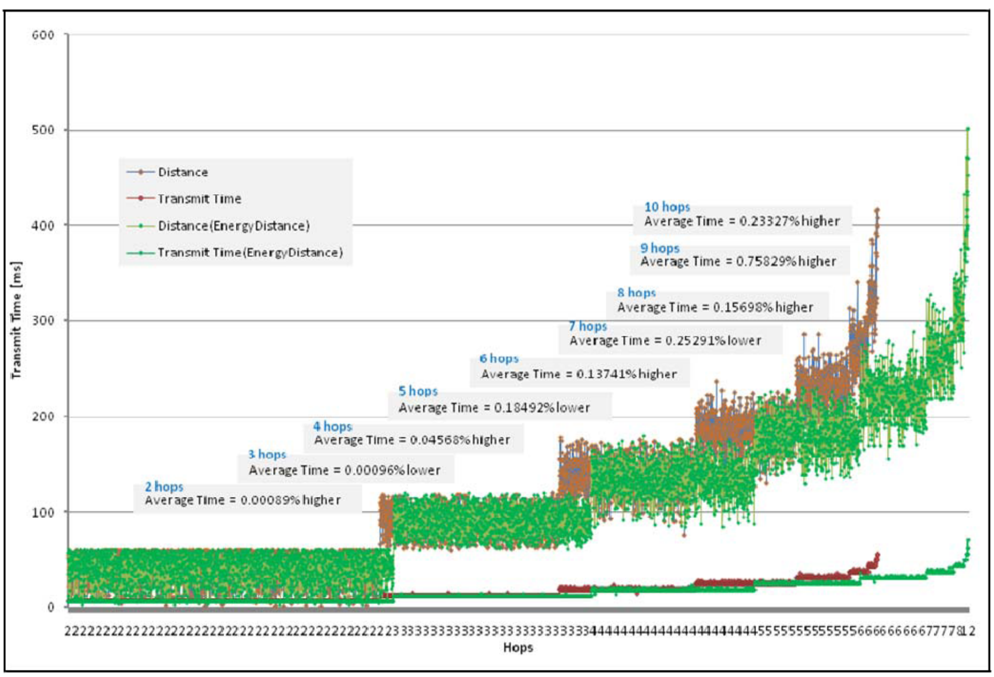


Figure 5. Comparison between improved Topology construction with Energy+Distance metric and similar LSR network.

Case 5. A composite metric consists of distance, energy, interface buffer, round trip time, and direction is set. For this given metric weight, the algorithm computes the decomposition of topology into maximum 10 tree hops. 36.56% is two hops, 21.46% is three hops, 18.02% is four hops, 12.18% is five hops, and rest is for other hops. Most of network hop delay is impacted with this metric, particularly in delay minimization on the longer hops. From Figure 6, the average distances for different hops are almost similar. The higher reduction on average delay by this metric algorithm is 0.577% compared with Figure 1 respectively. The total average delay reduction for this simulated network is 0.683%.

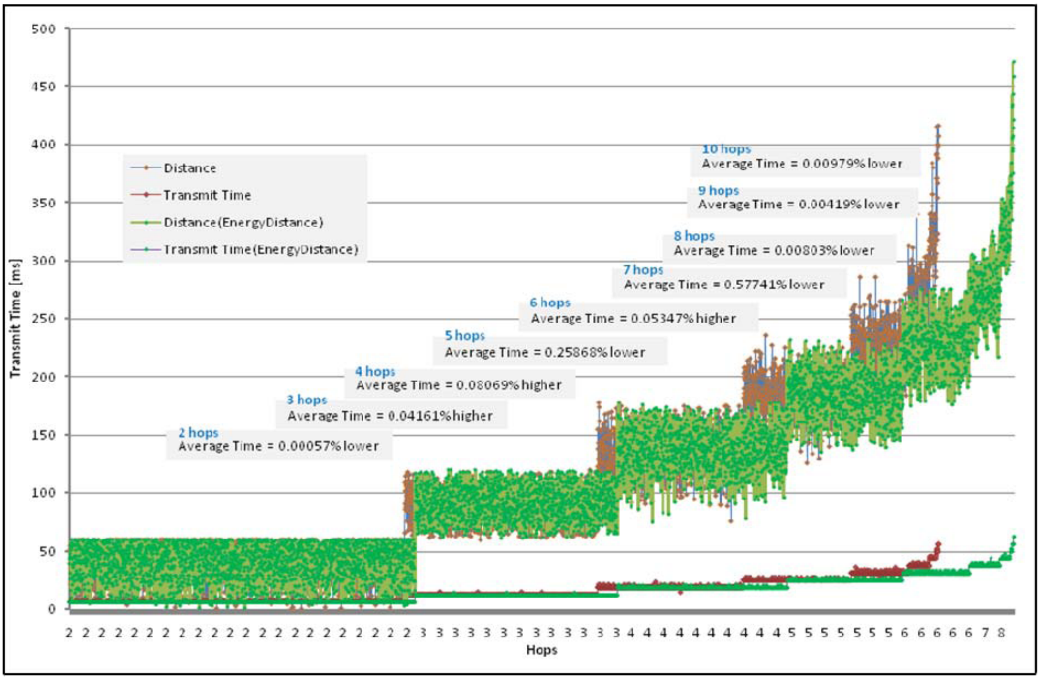


Figure 6. Comparison between improved Topology construction with Composite metric and similar LSR network.

CONCLUSION

In this chapter, we have described a simple approach to reduce the delay in a multi-hop wireless topology construction routing prior to carry out the data transmission analysis. The analysis is taken generally and admits a large class of arrival processes. The analysis already includes handling multiple topology type combination. The main difficulty however is in identifying the bottlenecks in the network. The result obtained in this simulation is compared against the similar LSR network with the initial average distances for different hops are kept similar. The delay performance in similar LSR is important to identify reference optimum relay hop delay and help in the design of a delay-efficient policy of comprehensive network simulation. Here, we have taken a different approach to reduce the wireless network delay in multi hop environment through the analysis of network routing metric in charge of the topology creation. This algorithm captures the essential wireless network performance parameters, i.e. bandwidth, throughput, network buffer, direction, and round trip time. This result is useful since, in many cases, we can also find that the throughput is the most important parameter in reduction of delay transmission. This result is confirmed with another simulation result using composite network metric. It is interesting to note that the routing metric policy, which was designed primarily for achieving maximum throughput in the single wireless network area, can also be engineered to achieve good delay performance in multiple wireless network area.

REFERENCES

1. Masato, Tsuru. *"Simulation-based Evaluation of TCP Performance on Wireless Networks"*. Journal of the Japan Society for Simulation Technology: 67-73, 2009.
2. Nozomu, Nemoto. *"Consideration and Evaluation of Wireless Mesh Network"*. Nomura Research Institute (NRI) Pacific Advanced Technologies Eng.: 70-85, 2006.
3. Javier G., Andrew T. C., Mahmoud N., Chatschik B. *"Conserving Transmission Power in Wireless Ad Hoc Networks"*. Network Protocols Ninth International Conference on ICNP: 24-34, Nov 2001
4. Chang-Woo Ahn, Sang-Hwa Chung, Tae-Hun Kim, Su-Young Kang. *"A Node-Disjoint Multipath Routing Protocol Based on AODV in Mobile Adhoc Networks"*. In: Proceeding of Seventh International Conference of Information Technology ITNG2010: 828-833, April 2010.
5. Prasanthi. S, Sang-Hwa Chung. *"An Efficient Algorithm for the Performance of TCP over Multi-hop Wireless Mesh Networks"*. In: Proceeding of Seventh International Conference of Information Technology ITNG2010: 816-821, April 2010.
6. Heinzelman, W., Chandrakasan, A., Balakrishnan, H. *"Energy-efficient communication protocol for wireless microsensor networks"*. In: Proceedings of the 33rd International Conference on System Sciences (HICSS): 1-10, 2000.
7. Feldmann, N. Kammenhuber, O. Maennel, B. Maggs, R. D. Prisco, and R. Sundaram. "*A methodology for estimating interdomain web traffic demand*". In IMC, 2004.

In: Advances in Communications, Volume 8 ISBN: 978-1-61324-794-5
Editor: Anthony V. Stavros

Chapter 8

APPLYING MULTIMODAL ANALYSIS TO MYSPACE: AN INSTRUCTIONAL FRAMEWORK TO DEVELOP STUDENTS' DIGITAL LITERACY OTHERS

Kirsty Young
Faculty of Education
Kuring-gai Campus University of Technology
Sydney, NSW, Australia

Abstract

The use of popular online social networking sites (eg. MySpace, Facebook) has resulted in an abundance of authentic texts where words, graphics, pictures and music are combined by users to project a desired image of oneself. This Chapter proposes that these online profiles can be utilised for instructional purposes to develop students' digital literacy. This can be achieved by using social semiotic multimodal analysis to deconstruct online social networking profiles. The process of applying multimodal analysis to online social networking profiles develops student understanding of the impact of text, graphics and sound on the readers' interpretation which, in turn, facilitates students becoming critical readers and producers of online content. This Chapter presents an analytical framework that educators can use with students to deconstruct online content, with an example drawn from MySpace.

Keywords: online social networking, MySpace, multimodal analysis, digital literacy, adolescents.

INTRODUCTION

Adolescents regularly make use of words, pictures, graphics and music to create an online identity using online social networking sites such as MySpace. Marketers, advertisers, and graphic designers have undertaken extensive education to learn how to use language and media to great effect to promote a person, place, object, experience or event while adolescents, in trying to promote themselves online are generally self-taught but not taught-well (Considine, Horton & Moorman, 2009).

To understand the digital literacy experiences of adolescents, educators must engage with the media used by them in their daily life and recognise the semiotic systems being used (Considine, et al, 2009; Jewitt, 2005). The recent phenomenon of online social networking provides an authentic context for examining digital texts produced by adolescents. This Chapter presents an analytical framework to demonstrate how social semiotic multimodal analysis can be applied. In this instance, the framework is applied to the MySpace (www.myspace.com) context to gain insights into the communicative modes being exploited by MySpace users to create an online identity and socialise with others. The premise being that this instructional famework provides an appropriate tool to aid educators in developing students' digital literacy with respect to both the consumption and production of online content.

This Chapter commences by describing the communicative affordances enabled by the Internet and particularly social networking sites. A description of social semiotic multimodal analysis is then presented, including a framework for its application to deconstruct online social networking profiles. A single case study example is then presented wherein a MySpace profile is analysed using the framework developed in this Chapter. The result demonstrates the potential for applying this framework to enhance students' digital literacy.

COMMUNICATION AND THE INTERNET

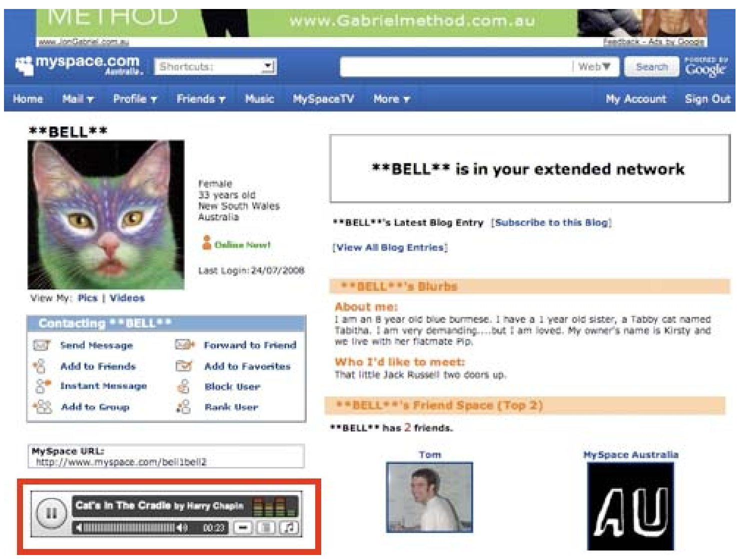
Historically discussion about the communicative functions of the Internet has focussed on pragmatic issues of speed, reach, storage and privacy (van Dikj, 1999). There has also been interest in the nature of language used online, particularly acronyms, emoticons and unconventional spellings in online material ("netspeak") (Crystal, 2001, 2005). These areas of focus emerged at the time of Web 1.0 where communications via the Internet were primarily limited to text-based email, discussion boards and chatrooms. Today, Web 2.0 enables Internet users greater opportunities to engage multimodalities, including graphics and sound, to communicate in more engaging and collaborative ways. As a result, formal education must include instruction on the social communicative acts enabled by the multimodal presentation of oneself online.

Daley (2003) argues that the multimedia language of the screen is the current vernacular which enables construction of complex meanings, beyond that possible through text. She suggests that to be truly literate today one must master reading and writing in the multimedia language of the screen. This expanded notion of literacy also incorporates 'remixing'. This term is used to describe the technical practices required to blend text, images, video, audio and games in the creation and maintenance of an online profile (Perkel, 2006). Importantly, the production of online content no longer requires expert HTML knowledge or use of complex publishing platforms which makes it possible for anyone with Internet access to contribute online (Rosen and Nelson, 2008).

Today we must acknowledge that adolescents are both consumers and creators of media (Considine et al. 2009). Knowing how to re-use media (both socially and technically) is essential for communication and creative expression over the Internet. Rodriguez-Illera (2004, p. 49) suggests a view of literacy that is not simply linguistic or cognitive, but requires a communicative competence and takes into consideration the cultural and interpersonal context in which it is produced. In this Chapter it is argued that MySpace is an informal learning environment that fosters digital literacy through authentic online communicative acts. This proposition is based on a model that tries to reconcile the social and technical perspectives on literacy (Perkel, 2006).

MYSPACE AS A PLATFORM FOR DIGITAL LITERACY DEVELOPMENT

MySpace has specifically been chosen as the platform for multimodal analysis because it allows its users more design freedom compared with than other social networking sites. An example of a MySpace profile is presented in Figure 1 below.



Screenshot 1. MySpace profile example*.

MySpace was also chosen for analysis because of its popularity as an online social networking site. Global membership to MySpace grew from 4.9 million in November 2004 to 26.7 million in November 2005 (Read/Write Web, 2007), with over 70 million registered users in 2006 (Snyder, Carpenter & Slauson, 2006). MySpace peaked as the second most popular web property in the world with 87 million accounts and was attracting 115 million

* This is a mock MySpace profile created by the researcher for demonstration purposes.

people to the site each month (Techcrunch, 2008). MySpace is currently available in 15 languages and as at February 2010 it is reportedly the 12th most popular website globally (Alexa, 2010). The popularity of MySpace in different countries is varied but it is regularly reported in the Top 25 sites accessed:

Country	MySpace Ranking*
Australia	13th
Canada	25th
Germany	19th
Ireland	24th
Mexico	13th
United Kingdom	15th
United States	5th

MySpace provides a platform for individuals to create an online identity. The site enables the user to present a desired impression of him/herself which stands in for their physical self around the clock. In the digital age this form of online identity creation and self-promotion has also been referred to as 'building a personal brand'. There is a growing body of literature around issues of online identity creation (see, for example, Boyd, 2006; Boyd & Heer, 2006; Chan, 2006; Miller & Arnold, 2003; Simpson, 2005; Valkenburg, Schouten & Peter, 2005). For the purposes of this Chapter, however, the focus is on the functional communicative elements being used to construct an online profile. Specifically, the use of text, graphics, sound and layout within the context of the MySpace platform and the limitations and restrictions associated with this particular website. Further, how these communicative acts can be deconstructed to examine the literacy elements present in this online environment and their effect on online text construction and interpretation.

The argument presented is that those individuals creating profiles on MySpace "*are developing new technical and social skills that enable their participation in a variety of social activities using a new medium. They are learning to incorporate the Web in multiple overlapping facets of their lives*" (Perkel 2006, p. 13). In order to understand the technical and social skills that require mastery a framework for analysis of websites has been developed. This process enables understanding of the elements of online identity creation relevant to adolescents' social activity. The development of an analytical framework can be subsequently used as an instructional tool to develop students' digital literacy through deconstruction and reconstruction of authentic online texts. As shall be demonstrated below, social semiotic multimodal analysis provided the necessary principles to develop this instructional framework.

* MySpace Ranking as at 12 August 2009 determined by Alexa (2009) based on website traffic statistics.

SOCIAL SEMIOTIC MULTIMODAL ANALYSIS

Social semiotics attempts to make meaning from the social aspects of communication. It is an approach to investigating human signifying practices in specific social and cultural circumstances that recognises that languages and other systems of socially accepted meanings evolve over time, as we are currently seeing with the Internet, particularly Web 2.0 technologies. Social semiotics focuses all systems for meaning-making including speech, writing, images, gestures and music and seeks to understand how these elements work together to create meaning. The 'author' and 'reader' is encouraged to consider the placement of elements, colours, styles of typeface, and so on, to understand meaning potential (Machin, 2007).

Social semiotics has been developed extensively by Kress & van Leeuwen to analyse multimodal forms of communication in both traditional and digital media, including, women's magazines, advertisements, websites, paintings, diagrams and instructional texts. Multimodal analysis has since been more broadly applied to print media, electronic texts, moving image and 3D objects (see, for example, Levin & Scollon, 2004; O'Halloran, 2004; Royce & Bowcher, 2007). Given the demonstrated effectiveness of multimodal analysis in research contexts where it has been applied to explore meaning within various multimodal texts, the principles of multimodal analysis are considered useful to underpin a pedagogical framework for teaching digital literacy.

In developing an instructional framework for analysis of online social networking sites two primary sources are being drawn upon. The work of Bateman & Delin (2001) is applied to identify the constraints of the medium and specific structural elements available for analysis, while the work of Kress & van Leeuwen (2006) is used to understand the meaning of composition. In what follows, a brief summary of the key elements being drawn from these two bodies of work is presented, before moving to the instructional framework that has been developed as a result of these influences.

As defined in Table 1 below, Bateman & Delin (2001) identify three areas of overlapping constraint imposed by various media – canvas constraints, production constraints and consumption constraints.

Moving beyond the constraints of the media, Kress & van Leeuwen (2006, p. 177) highlight the importance of examining text composition and the ways in which these parts interact and affect each other. Two compositional elements particularly relevant in the analysis of MySpace are 'salience' and 'framing', as defined in Table 2.

Given the nature of online social networking sites (such as MySpace) to facilitate communication, the work of Bateman & Delin (2001) has been further drawn upon to extract the communicative elements based on text content, expression and delivery. Five communicative elements are outlined in Table 3 below.

Taking the ten elements (Tables 1-3 above) into account the Chapter now moves to describe how aspects of multimodal analysis have be applied to create an analytical framework to be used as an instructional tool to facilitate students' digital literacy. That is, to guide student understanding of the ways in which text, graphics, music and images are used in an authentic online context (ie. an online social networking site). To demonstrate, the framework is then applied to deconstruct a MySpace profile to uncover the individual technical and symbolic elements being used by the online author, and then reconstructed to explore the presentation of self in the online context.

Table 1. Bateman & Delin (2001) Constraints of Media

1. Canvas constraints	Constraints arising out of the physical nature of the object being produced eg. paper or screen size, fold geometry, columns.
2. Production constraints	Constraints arising out of the production technology eg. limit on page numbers, colours, size of included graphics, availability of photographs, constraints arising from the micro- and macro-economy of time or materials eg. deadlines, expense of using colour, necessity of incorporating advertising.
3. Consumption constraints	Constraints arising out of the time, place, and manner of acquiring and consuming the document. Such as method of selection at purchase point or web browser sophistication and the changes it will make on downloading; also constraints arising out of the degree to which the document must be easy to read understand or otherwise use; fitness in relation to task (read straight through? Quick reference?); assumptions of expertise of the reader etc.

Table 2. Kress & van Leeuwen (2006) Composition Elements

4. Salience	The elements are made to attract the viewer’s attention to different degrees eg. placement in the foreground or background, relative size, contrasts in colour, etc. Regardless of where elements are placed, salience can create a hierarchy of importance among elements. Salience reflects the relationship between a number of factors: size, sharpness of focus, tonal contrast; colour contrast; placement in the visual field; perspective; specific cultural factors.
5. Framing	The presence or absence of framing devices (realised by elements which create dividing lines, or by actual frame lines) disconnects or connects elements of the image, signifying that they belong or do not belong together in some sense. The stronger the framing of an element, the more it is presented as a separate unit of information. Context and colours in a more precise nature of this “separation”. In the absence of framing elements of the spatial composition are connected, the more they are presented as belonging together, as a single unit of information.

Table 3: Bateman & Delin (2001) Five Communicative Elements

6. Content structure	The structure of the information to be communicated. The 'raw' data out of which documents are constructed.
7. Rhetorical structure	The rhetorical relationships between content elements; how the content is 'argued'.
8. Layout structure	The nature, appearance and position of communicative elements on the page
9. Navigational structure	The ways in which the intended mode(s) of consumption of the document is/are supported.
10. Linguistic structure	The structure of the language used to realise the layout elements.

ANALYTICAL FRAMEWORK FOR MULTIMODAL ANALYSIS OF MYSPACE

Literacy instruction should always be carefully and purposefully planned to maximise student learning. Digital literacy similarly requires logical, sequential instruction. Presented here is an analytical framework that incorporates multimodal analysis for instructional development of digital literacy using authentic online texts (in this instance, MySpace profiles). This framework involves a four-step process:

(a) identify the communicative function/s of a MySpace profile;
(b) identify the contextual constraints (canvas, production, consumption);
(c) deconstruct the MySpace profile into its component parts to identify the compositional and communicative elements being used;
(d) reconstruct the MySpace profile to view the elements as a whole – a social semiotic interpretation of the profile which explores the creation of an online identity.

(a) Identify the communicative function/s of MySpace

The first step in the analysis process is for students to define their interpretation of the communicative function of MySpace. The interpretation of the sites' purpose will be informed by the students' out-of-school exposure to MySpace. For instance, one perspective is that MySpace presents an avenue for individuals to create an impression of him/herself for the purpose of self-promotion. So, the communicative function is to establish an online presence that presents the profile creator in a desired light and subsequently elicits contact from, and engagement with, others.

(b) Identify the contextual constraints of MySpace

Next, the students use multimodal analysis to identify the constraints of the MySpace platform. The design of the MySpace website imposes significant limitations on the user and these must be considered in any analysis of profiles which exist on the site. After instruction on the relevant work of Kress & van Leewuen (2006) and Bateman & Delfin (2001) students use the questions suggested in Table 4 below to identify contextual constraints.

Table 4. Identification of contextual constraints of MySpace

1. What are the canvas constraints of MySpace?
Student responses could include: Width of screen Long scrolling page Set response boxes (with strong frames) Limitations on ability to move response boxes around page (without HTML knowledge)
2. What are the production constraints of MySpace?
Student responses could include: Extensive but not exhaustive background selections available Limited to typewritten text (although extensive fonts, colours available) Designated headings which can only be altered, deleted or relocated with HTML knowledge (although colour and font are determined by the user) Extensive but finite number of songs available for inclusion on one's profile Advertising (can not be removed but noted to 'adapt' to reflect user's interests) Logos which are automatically inserted when user adds an 'application'[*] Links provided by MySpace (eg. at top of page – home, browse, find people, etc and at bottom of page – about, FAQ, Terms, Privacy, etc)
3. What are the consumption constraints of MySpace?
Student responses could include: Assumption that desired audience has Internet access Assumption computer/web browser used to 'read' the page is capable of supporting all features/applications. Assumption reader can understand 'textspeak' and other non-standard language and punctuation used Issues associated with reading across and down the page simultaneously Unable to control what 'friends' comment/write/ask on author's profile Issues associated with control of visibility of one's profile (privacy settings) Issues associated with the inability to permanently remove one's profile from the website
4. To what extent/effect is 'framing' used in MySpace?
Student responses could include: The developers of the MySpace website have made substantial use of 'framing'. The framing presents elements of the 'profile' as isolated rather than viewing the page as the 'whole' identity of the 'creator'.

Addressing questions 1-4 directs student attention toward both the construction and visual elements that influence text presentation. The potential for digital literacy development occurs when students independently recognise the contextual constraints of the online space being examined.

[*] MySpace applications include online games, activities, content, gift services, etc which are posted on a profile. Anyone can develop an 'application' for use by others so long as it complies with the MySpace guidelines.

(c) Deconstruct the MySpace profile (identification of compositional and communicative elements)

Once students have extracted the contextual constraints that impact MySpace users' text production they move to identify the elements that can be used by creators of MySpace profiles to present his/her online self. That is, identification of the compositional and communicative features available for analysis. Educators can guide students by having them respond to the questions presented in Table 5.

Table 5. Deconstruction of MySpace (compositional and communicative elements)

5. How is 'salience' used to create online identity on MySpace?
Student responses could include: Salience is the one area where the MySpace user can override some of the limitations placed on them by the MySpace designers. MySpace users can make use of colour and size in relation to graphics and photographs. MySpace users can also the use different fonts and unique punctuation to add salience text. Salience is interrelated with elements of layout and navigational structure below.
6. How is 'content' used to create online identity on MySpace?
Student responses could include: The content included would be the person's internal/desired perception of him/herself. Other content would be external to the author – being posted by an online 'friend' (a co-constructed profile). Content is the communication between online 'friends' (eg. wall postings)
7. How is 'rhetoric' used to create online identity on MySpace?
Student responses could include: Picture of 'self' argued through use of background, colours, photo/s, text (personal details, about me, interests, blurb, blog), graphics, applications
8. How is 'layout' used create online identity on MySpace?
Student responses could include: Positioning of controllable elements (eg. insertion of own photograph) Use of colour, font, paragraphs and punctuation to add emphasis (aligned with salience) Inclusion of particular types of 'applications'.
9. How is the 'navigational structure' used to create an online identity on MySpace?
Student responses could include: Limited by MySpace design features. Sense of author's identity dispersed across long, scrolling page. MySpace profile can use salience to direct viewer's attention to certain elements and thus influence navigation toward components valued by author.
10. What is 'linguistic structure' used to create an online identity on MySpace?
Student responses could include: Use of 'textspeak' and/and traditional spelling/grammar/punctuation Headings, colour, graphics, font to 'speak' to the reader of the author's identity.

In Table 4, digital literacy is developed through identification of the contextual constraint of the MySpace website. In Table 5, students are focused on the ways in which compositional and communicative elements are used in a digital space for a specific purpose (in this case to create an online identity). Students' digital literacy is developed through the act of deconstruction to examine elements in isolation.

Following students' analysis of contextual constraints and deconstruction of the compositional and communicative elements the students are then in a position to list the specific components of a MySpace profile that are available for analysis. An example of the analysable components of a MySpace profile that could be developed after application of questions in Table 4 and Table 5 is presented in Table 6 below.

Table 6. Example of specific components of a MySpace profile available for analysis

1.	MySpace name
2.	MySpace URL name
3.	Profile picture
4.	Mood status updates
5.	Demographic information provided
6.	Date of last login
7.	Statement about networking purpose
8.	Opening statement/message
9.	Background/graphic design choice
10.	Song choice
11.	Blog
12.	Personal details (status, here for, sexual orientation, hometown, zodiac sign, smoke/drink habits, occupation)
13.	'About me' information provided: Who I'd like to meet; Interests (general, music, films, television)
14.	Additional photos inserted by author
15.	Videos inserted by author
16.	'Applications' inserted by author

It is important to note that as the MySpace site continues to evolve so will these analysable elements. Also, student learning is most effective when they develop the list of elements for themselves; when this list emerges from their own exposure to the MySpace website and level of proficiency in using the site. The list of analysable components can vary from student to student. For instance, students with HTML knowledge (who can manipulate site with greater expertise) may have a different range of analysable elements than those students without such understanding.

(d) Reconstruct the MySpace profile (social semiotic interpretation of the online identity created)

Having identified the constraints of the MySpace platform and, subsequently, the analysable components, students then move to analyse individual MySpace profiles in detail. A single case study of such analysis is now presented. This MySpace profile has been purposefully selected because its creator is not clearly identifiable from the profile picture. To further protect online identity this profile has been given the pseudonym * Yam A *. This pseudonym shares characteristics which resemble the name originally chosen by the profile creator. The analysed profile can be seen in Screenshot 2 below.

Table 7. *Yam A* MySpace profile summary analysis[*]

1.	MySpace name	Creative /ambiguous
2.	MySpace URL name	Contrast between girly (barbie) and cool/tough (007)
3.	Profile picture	Level of anonymity. Creative/abstract picture. Salience used to draw attention to the eyes
4.	Mood status updates	Used – 'sleepy'
5.	Demographic information provided	Appears factual – no obviously outrageous claims
6.	Date of last login	Active profile at time of analysis
7.	Statement about networking purpose	Exploring creative avenues/networks
8.	Opening statement/message	Tough (expletive used) – shock value
9.	Background/graphic design choice	Contrast between girly (pink hearts) and tough (black with skull and crossbones)
10.	Song choice	Contrast between cute (Muppets' tone) and tough (about pornography)
11.	Blog	Not used
12.	Personal details (status, here for, sexual orientation, hometown, zodiac sign, smoke/drink habits, occupation)	Appears factual - no outrageous statements made
13.	'About me' information provided: Who I'd like to meet; Interests (general, music, films, television)	Does not make use of 'netspeak'. Creative punctuation – random use of capitalisation no use of fullstops. Three vertical lines ||| used as para marker. Strong association to people she cares for. List of favourites are non-mainstream/alternative. Describes her personal values.
14.	Additional photos inserted by author	Not applicable
15.	Videos inserted by author	Not applicable
16.	'Applications' inserted by author	Tough – Mob application

The 16 elements identified in Table 6 above are firstly examined (as far as possible) in isolation. Then, taking account of the questions in Tables 4 and 5 the profile is reconstructed and viewed in context, where the integrated use of text, graphics, pictures and music is examined to interpret *Yam A*'s online identity. It is important to note that this is one

[*] Due to publication restrictions this is an extremely brief summary overview of a very detailed and in-depth analysis of the MySpace profile.

interpretation of the profile and student learning (ie. digital literacy development) comes from their ability to justify their own criteria for analysis and the subsequent interpretation based on the analytical criteria which they have developed.

Screenshot 2. *Yam A*'s MySpace profile.

After deconstruction of individual elements of the profile in isolation it is necessary to review the profile in context. Based on a holistic analysis of *Yam A*'s MySpace profile one interpretation is that she has projected *tough*, *girly* and *creative* sides to her online identity. Visual analysis of the deconstruction/reconstruction of *Yam A*'s profile highlights the elements she has used to present three sides to her online identity:

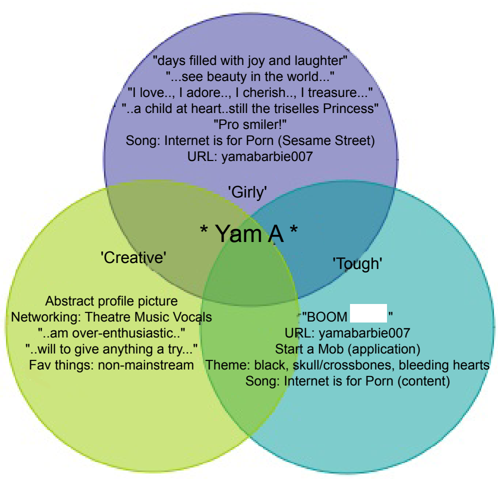


Figure 1. Interpretation of *Yam A*'s online identity.

One interpretation of Figure 2 is that that different semiotic modes are being used to represent three sides to *Yam A*'s online identity:

- **Descriptive text/words** have been used to create a girly image
- Detail of her **activities/actions** establishes her creative image
- **Visual/Graphics** have been used to develop tough image

The purpose of the instructional multimodal analysis framework presented in this Chapter is to develop student digital literacy through the deconstruction/reconstruction process. However, given the texts being analysed are authentic and presented in the context of an online social networking site important social issues are also raised and can be explored. For example, in the case of *Yam A* interpretation of her profile suggests that she has managed to capture different facets of her identity, as far as possible, in the two-dimensional arena. As humans we often present different sides of our personality in various, family, work and social situations. We adapt our behaviour in accordance to the people and contexts in which we find ourselves. Here the interpretation presented of *Yam A*'s profile suggests that has integrated three different sides to her character through the use of, and integration of, different semiotic modes. This emphasises the use of digital texts to communicate in complex ways.

It is important to note that the intention is not to direct students to find a single, common interpretation of the online profile being analysed. The students' social and cultural differences should be reflected in their own interpretation. The aim of the framework is to cause students to justify their analysis and interpretations. Application of the instructional multimodal framework to MySpace draws students' attention to the structure of the website and the subsequent constraints imposed on the user. It encourages students to consider how the creator of the online profile has made use of text, colour, graphics and music to present an online identity. This approach need not be limited to MySpace but the framework could be used in the analysis of many other authentic online texts.

CONCLUSION

The use of online social networking profiles to explore text production is not unique (see, for example, Dowdall, 2009). It is necessary, however, to develop instructional frameworks which can be used as tools to facilitate digital literacy; not only in terms of online text production, but also online text consumption. As suggested by Hansford and Adlington (2009):

> *"*[to] *initiate student learning, effective digital texts can be co-examined to determine the multimodal components that combine to result in effective texts. The design components such as linguistic, spatial, gestural, visual and audio elements can be analysed both for their separate and synergistic qualities. Such an exploration would be a strong starting point to scaffold students in composing effective contemporary texts."* (p. 64)

The permeation of digital media across all facets of students' lives dictates that effective digital literacy learning must be high on the educational agenda. This Chapter deems the use of online social networking profiles to develop digital literacy as particularly pertinent because of the popularity and authenticity of such websites.

Digital literacy education based around analysis of social networking sites presents an environment where students naturally create and consume digital text. Importantly, one glaring problem exists; the negative attitudes some people have toward online social networking and the subsequent banning of many popular websites in schools. The reality is that students will engage with these websites in out-of-school contexts. We are doing a disservice to students to ignore this globally popular form of communication in formal education. Until this issue is resolved, teachers can capture screen/video shots of online profiles to be analysed by students offline – similar to the way they would accumulate newspaper cuttings in the past.

The process of applying the multimodal analysis framework presented in this Chapter enables educators to identify the skills required, both technically and socially, for online social networking. There is a need for educators at all levels to be aware of, and educating toward, these skills and behaviours. This is particularly important because, while the future impact of online social networking is unknown, it certainly has the potential to affect child and adolescent users well into adulthood. Improper presentation of oneself online can have negative consequence on relationships and in future education and employment endeavours. The instructional approach suggested here teaches students how to apply social semiotic multimodal analysis to interpret online profiles so they, in turn, appreciate how to effectively construct their own online identity.

REFERENCES

Alexa, (2009), *Alexa Top 500 Global Sites*. Available: http://alexa.com/topsites. Accessed 16 April 2009.

Alexa (2010). *About Myspace (myspace.com): Social Networking Site.* Available: http://www.alexa.com/siteinfo/myspace.com. Accessed 8 February 2010.

Bateman, J. and Delin J. (2001). *From Genre to Text Critiquing in Multimodal Documents.* Presented at Workshop on Multidisciplinary Approaches to Discourse: Improving Text: From text structure to text type. Yttre, Belgium, 5-8 August 2001.

Boyd, D. (2006). *Identity production in a networked culture: why youth heart MySpace.* In American Association for the Advancement of Science. St Louis MO.

Boyd, D. and J. Heer. (2006). *Profiles as conversations: Networked identity performance on Friendster.* In Proceedings of the Hawaii International Conference on System Sciences.

Chan, A. (2006). *Social Interaction Design Case Study: MySpace*. Gravity7. Available: www.gravity7.com.

Considine, D., J. Horton and G. Moorman. (2009). Teaching and Reading the Millennial Generation Through Media Literacy. *Journal of Adolescent & Adult Literacy* 52(6), 471-481.

Crystal, D. (2001). *Language and the Internet*. Cambridge: Cambridge University Press.

Crystal, D. (2005). *The scope of Internet Linguistics*. American Association for the Advancement of Science.

Daley, E. (2003). Expanding the Concept of Literacy. *Educause Quarterly* March/April.

Dowdall, C. (2009). Impressions, improvisations and compositions: reframing children's text production in social networking sites. *Literacy* 43(2), 91-99.

Hansford, D. and R. Adlington. (2009). Digital spaces and young people's online authoring: Challenges for teachers. *Australian Journal of Language and Literacy* 32(1), 55-68.

Jewitt, C. (2005). Multimodality, "Reading", and "Writing" for the 21st Century." *Discourse: studies in the cultural politics of education*. 26(3), 315-331.

Kress, G. and van Leeuwen, T. (2006). *Reading images: the grammar of visual design.* New York: Routledge.

Levine, P. and Scollon, R. (Eds). (2004). *Discourse & Technology: multimodal discourse analysis*. Washington: Georgetown University Press.

Machin, D. (2007) *Introduction to Multimodal Analysis.* New York. Oxford University Press.

Miller, H. and Arnold, J. (2003). Self in Web Home Pages: Gender, Identity and Power in Cyberspace. In *Towards CyberPsychology: Mind, Cognitions and Society in the Internet Age*, Giuseppe Riva and Carlo Galimberti (Eds) Amsterdam: IOS Press

O'Halloran, K. L. (Ed). (2004). Multimodal Discourse Analysis. London: Continuum.

Perkel, D. (2006). *Copy and Paste Literacy: Literacy practices in the production of a MySpace profile*. Available: http://people.ischool/berkeley.edu/~dperket/media/dperkel_literacymyspace.pdf Accessed: 24 July 2007.

Read/Write Web. 2007. *Latest SNS Numbers - MySpace Streaks Ahead*. Available: www.readwriteweb.com/archivess/latest_sns_numb.php Accessed: 24 July 2007.

Rodriguez-Illera, J. L. (2004). Digital Literacies. *Interactive Educational Multimedia* 9, 48-66.

Rosen, D. and Nelson, C. (2008). Web 2.0: A New Generation of Learners and Education. *Computers in the Schools*. 25(3), 211-225.

Royce, T. D. and Bowcher W. L. (Eds). (2007). *New Directions in the Analysis of Multimodal Discourse*. New Jersey: Lawrence Erlbaum Associates.

Simpson, B. (2005). Identity manipulation in cyberspace as a leisure option: Play and the exploration of self. *Information & Communication Technology Law* 14(2), 115-131.

Snyder, J., Carpenter, D. and Slauson, G. J. (2006). *MySpace.com - A social networking site and social contract theory*. ISECON Conference. Normandy.

Techcrunch. (2008). *Facebook no longer the second largest social network.* Available: http://www.techcrunch.com/2008/06/12/facebook-no-longer-the-second-largest-social-network/. Accessed: 13 August 2008.

Valkenburg, P. M., Schouten, A. P. and Peter, J. (2005). Adolescents' identity experiments on the Internet. In *New Media & Society.* London: SAGE Publications. 383-402.

van Dikj, J. (1999). The Network Society: Social Aspects of New Media. London: SAGE.

In: Advances in Communications, Volume 8
Editor: Anthony V. Stavros
ISBN: 978-1-61324-794-5

Chapter 9

CRM IMPLEMENTATION: THE MANAGEMENT OF A TECHNOCHANGE IN A FRENCH TELECOMMUNICATIONS COMPANY

Adel Beldi
IESEG School of Management in Lille and Paris
France

INTRODUCTION

Since the 90 major changes have occurred in the telecommunications sector in Europe and particularly in France. Until this period, a governmental administration was in charge of managing post and telecommunications services in France. An owned corporation was created to manage telecommunications services and a separation between the two activities was occurred. The opening of the telecom market to competition in France has accelerated changes in strategic, organizational and human patterns of the company. The French Telecommunication Corporation (FTC) aimed to increase profits in a context characterized with a high competitors' aggressiveness. These factors have led the company to change its strategy, organization and the customer relationship management approach. In the 2000s, the board of directors decided to implement a major customer relationship management (CRM) software in commercial divisions of the company. In "business" branch, the project started in 2001 was the occasion to change internal processes, management methods and organization of the local agencies.

The new customer-focused approach aims to put the customer at the center of corporate strategy at the local level. This tool would reduce the operating costs of doing business, customer service and sales administration while improving the quality of service. In its first phase of the CRM project was to target functions: sales, customer service and marketing. It is centrally managed at the branch "business". The project team includes business and technical specialists. Two levels of project management are observed: a central level and another at local agencies. The implementation of the CRM tools was primary considered as more technical project. But few months after the beginning of the project, the board of managers recognized that the project should cover non-technical issues. It is a real techno change having an impact also on organization.

This case reports on the implementation of CRM approach in a “business” division in the French Telecommunication Corporation. This particular programme set out to develop a new strategy focused on customers. It was defined and launched in 2001 and features a complex web of activities, roles, and IT tools, affecting thousands of employees.

The case considers a CRM approach as a major new infrastructure, consisting of processes, people, and technology, and analyses its inner dynamics of launch, deployment, implementation, use, and ongoing change.

The remainder of the chapter proceeds as follows. First we present briefly the main events related to the launching of the CRM approach in the company and describe its information system. Second, we will present the CRM project: team organization and implementation process. Third, the main changes observed will be described. Finally, concluding results will be discussed.

BRIEF HISTORY OF THE COMPANY

Since the 90s, the French government which recognizes the importance of technological change and openness to competition in the telecom market in Europe has launched a program to modernize this sector. The most important axis of this program is the gradual privatization of the public company to enable it to have sufficient financial resources for its development.

Before the end of the year 90, the telecom operator in France turned into a corporation which is totally owned by the State. With this change in status, the French government also pledged to respect the status of “civil servant” for employees working in the company.

The telecommunication Sector opened fully to competition in Europe on 1 January 1998. Between years 98 and 2000, the opening to public company's capital has allowed its development and acquisition of other European operators, paving the way for the formation of a cluster containing the entire fixed, mobile and Internet in the telecom sector.

THE INFORMATION SYSTEM OF THE COMPANY

The development of FTC activities towards new products and the integration of many branches in the world require a highly responsive and powerful communication tools. In this context, the use of ITs that are built on a web interface based in "internet" protocols and standards can realize this ambition.

At the beginning of the 90s, the company was a part of the administration of (Post, telegraph and telecoms) PTT, which is responsible for operating the telephone system monopoly in France. At its end, the company is struggling in its territory with numerous competitors and has no monopoly. It was also established in many countries in Europe and in the World.

The ambition of the company as part of the evolution of its information system is to give each employee a secured access using an internet browser to all business applications in the office or outside. The new IS has to provide also access to customers and other partners in order to share and diffuse information.

The new technical infrastructure is built towards a human-machine interface (the workstation with a browser and Web servers involved), new business applications, data warehouse and legacy system.

The technical architecture is characterized by ease of integration software market (CRM, ERP) while coupling them with existing applications. This is made possible thanks to its flexibility and scalability.

The main purpose of the information system is to facilitate the achievement of firms' objectives through improved customer service and increased responsiveness in the offer of new services and products. The company seeks to develop an IS to develop more efficient business, retain customers and increase process efficiency

The company has built its new IT architecture using Internet technologies. Users routinely use a web browser in order to access to all business applications. There is nothing installed on the workstation, except the browser. The IS is based on a business approach. Each task performed by an employee may be linked to a business process (taking in account a new command, implementation a phone line, billing etc.).

Each business process is based on software tools. It is implemented through a human-machine interaction. The tool facilitates the automatization of tasks and activities.

The company has evolved from logic of customization to the implementation of commercial software packages. A standard software-based solution begins with an understanding of what it is, what it does well, what it can do and what we should not ask it to do.

PROGRAM MANAGEMENT

The success of such a transformation required to catalyze the energy and manage with precision the profound changes that will be observed. Each project is organized independently from the others. It has its own organization, its own team and resources allocated.

Project teams have been selected according to a certain balance between managers, end users and technicians.

Each project has a high degree of autonomy: the orientation is determined by its own steering committee. It is managed by a director. Reporting is done regularly by the director to the executive committee of the branch.

THE CRM PROJECT

The company is familiar with the Internet and intranet technologies. The implementation of a CRM package is also based on the obligation to make applications available from an enterprise portal using Internet technologies.

The CRM was implemented on FTC, particularly on the division responsible for "corporate clients" in 29 branches that drives business. FTC is a corporation providing telecoms and mobile phones, Internet access and content publishing services. The group is structured into divisions, we are interested in the distribution division (or business) responsible for selling all products and services for the SME market and large corporations. The field itself is the agency. The functions concerned are: the commercial, marketing and the customer service support. There are about 10,000 end users of CRM The purpose of the project deals of the replacement in 2001 of a set of IT applications ("legacy system") used by actors related to customers (sellers, customer service, marketers) in the commercial agencies, locally by a software package developed by an external editor. By launching this project, management sought to address the significant loss of market share and declining profitability in telephone and Internet. The levers are focused on improving the quality of customer service, reducing operating costs and replacement of old computer applications by independent local and integrated software. Our presence in this project was done through participant observation for nine months. Our knowledge of the company has been of great use to quickly grasp the vocabulary used and soak up the culture.

CRM: An Overview

Plakoyiannaki and Tzokas (2002) define CRM as an approach based on technology that identifies, develops and integrates all information related to the client. It relies on different skills of the business to create value for the customer. Thus, this approach has three components: a client-oriented strategy, the use of information technology and communication and integration.

CRM (customer relationship management) can be defined as the set of tools and techniques designed to capture, process and analyze information about customers and prospects, in order to retain them by offering the best service. These software packages deal directly with the customer, whether in terms of sales, marketing or service. It is often grouped under the term "front office”, as opposed to “back office”. The techniques used are based on computer telephone integration (CTI), interactive voice response (IVR) and automatic call distributors (ACD).

The goal of CRM was to identify the sequence of activities needed to complete the whole process which have the aim to propose an offer to the customers. This approach based on an IT system required the participation of various actors in the back and front office.

The roles of units and people who perform the activities were supported with a board set of IT applications. Some of the IT applications were implemented in the past years; others are introduced in the same time as the project. The full layout of CRM is very complex.

CRM Project Organization

The main purposes of the CRM implementation project were: (1) to have a global view of the customer (2) to be able to offer a customized products and services to each client 3) to increase internal effectiveness and enhance value for the customer (Official Presentations and Memos, Steering Committee of 16/05/2001). By launching this project, top managers were attempting to find a solution to the considerable reduction both in their market share and in the profitability of their telephone and Internet business.

The central project team is composed of the project director, the change management director, senior users representing the CRM modules to be introduced (i.e., sales, support and marketing) and the human resources change director. This team was responsible for defining the specific requirements for each module, and the schedule of implementation process in local agencies (the Figure 1 shows this organization). At the beginning, the project team hired a consulting company, to help schedule the project and to prepare change management. Another unit was charged with the technical customisation of the CRM software, and its adaptation to the needs of the end users, as defined by the project team. This unit was affiliated to the central information system department.

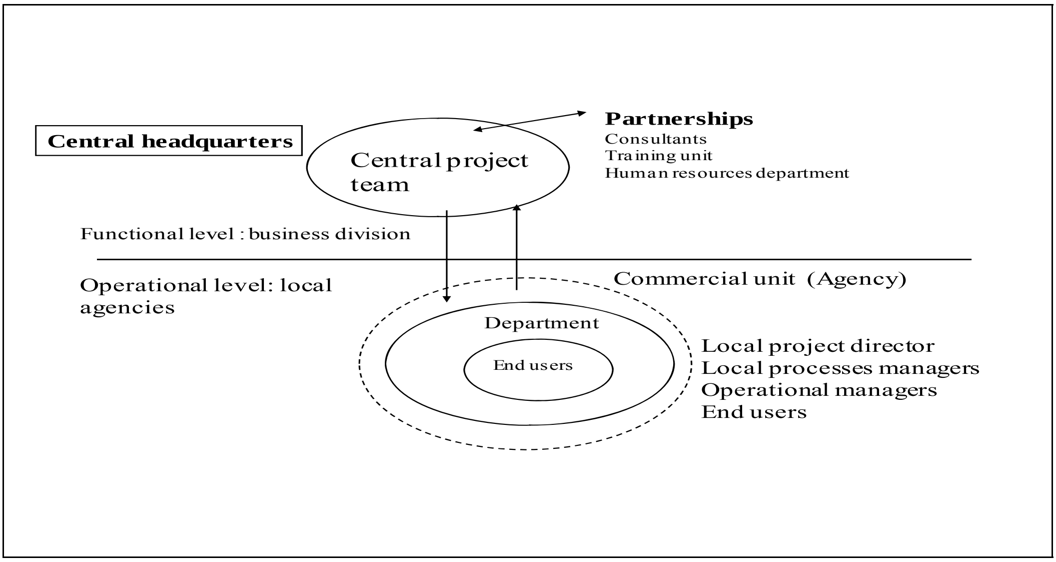


Figure 1. CRM project organization.

THE IMPLEMENTATION PROCESS

The top managers were conscious that implementing CRM in the firm would imply vast, radical changes that could take place a long time. The design of the process was influenced by the technical experience and education by those who decide. The implementation of new technical applications and technologies were considered as usual in the company. A classical and standard process in managing technical project was designed to implement the CRM approach. We can divide this process into three main phases: launching, implementation and post-implementation.

Phase 1- launching (February 2001-March 2002)

In this phase, the project team defined the functional architecture of the IT infrastructure along with its periphery. It was decided that only operational CRM (front-office) modules would be implemented. The analytical CRM was to be based on the legacy systems.

The incremental adaption was to be performed step by step, with continual improvements, and it was not expected to involve a radical change in the company's organization and culture. The choice of the editor PeopleSoft to provide the CRM applications was decided by the top managers. In the beginning, the firm decided to implement: Sales solution, Service solution and Marketing solution.

Phase 2- Implementation (April 2002-February 2003)

This phase started in April 2002, by implementing the modules as planned in phase 1. During this phase, communication between the project team, the designers and the users was emphasised in order to implement a user-friendly system. In this stage, the steps and tasks, and the actors implicated in the deployment of the CRM application, were identified. Other topics related to the local strategy in each agency, and its links with customer relationships, the structural re-organisation, the evolution of employees' behaviors in relation to the tool and how to use the experiences of pilot sites in order to get success in the future phases.

During this phase, senior managers were given presentations to explain the deployment process, essential step to make the whole approach understood.

Training was organized top down and by groups. First some key users who were volunteers were trained by specialists. Then they were asked to train other users, in each agency. A kit of presentations, videos, manuals and guidebooks were distributed. Once key users had been exposed to the new tool and had used it, they would start autonomous training courses in their businesses.

The project team identified the following activities, which were considered necessary at this stage, to facilitate the implementation of the CRM software: (i) Training programs (including via Intranet) for salespeople, to acquaint them with the use of the technical functionalities; (ii) Training programs (including via Intranet) for sales and service staff, to enable them to deal with business issues regarding the implementation of the CRM system. The training sessions were conducted by "expert users", who were volunteers. They were users from local agencies, and they worked in the commercial, customer service and marketing departments.

Phase 3 – Post-implementation (March 2003-December 2003)

This phase took place over the period March 2003-December 2003. This stage concerned the evaluation of the end users' enactment of the technology. The relaying trainers replicated the knowledge they had acquired in their local agencies. A newsletter was distributed to explain the evolution of the project. A questionnaire was designed to evaluate the end users' satisfaction with the pilot agencies. .

In this phase, an upgrading of the CRM application was, the connection time was improved and the ergonomic patterns of the tools were enhanced. The modules and functionalities to be implemented were available on time. The project team, along with the unit in charge of customising the application, had undertaken a technical enhancement of the tool during the previous phase, to meet the end users' needs.

The Table 1 shows the details of the implementation process.

Table 1. Process implementation

Phases	*launching*	*Implementation*	*Post-implementation*
Period	Feb 2001 –Jan 2002	Jan 2002-October 2002	March 2003-october 2003
Actors involved	- Board of directors - managing directors of the project- - IT architects and technicians	- CRM team - IT architects - end users - local managers	- CRM team - end users - local Managers -other partners (training)
CRM team activities	- Definition of project characteristics Internal communication	- implementation Training of some end users » communication end users support	- improving Human machine interface - Training of all end users - support change
Local project team activities	Those teams are not constituted.	- training by local expert users - managing the project in the agency Communication Involvement of the hierarchy	- check data accuracy and reliability - larger communication - leadership
Major decisions	Implementing the project per lot Choice of modules to implement	- select pilot agencies Implement modules	Each agency adapt the CRM project to its local context
Difficulties	- technical, IT complexity - human : resistance to change, fear from change.	- technical, - Human and cultural related to commercial cultures described as individualistic - difficulties linked to insufficient training	- organizational - human linked to support change

CHANGES OBSERVED

Van de Ven and Poole (1995, p.512) define change as a type of event in which the form, quality or state of an entity differs over a period of time.

We focus on the analysis of the change dimensions, and on the process by which CRM systems are implemented. Bygstad (2003) argues that the research on Critical Success Factors does not provide much guidance on how CRM systems should be implemented. Lucas (1981) called IS implementation an ongoing process which includes the entire development of the system from the original suggestion through the feasibility study, systems analysis and design, programming, training, conversion, implementation and evaluation of the system”.

Several studies point out that an increasing number of companies are discontented with CRM implementations, because the majority of them fall far short of reaching the expectations that preceded them (Zablah et al., 2004; CSO Insights, 2006). Among the reasons for this, the under-estimation of complexity and the consequent changes are emphasised (Gartner Group, 2003).

The success of IT project management depends on the recognition of the diversity of interests (Ward et al., 2005), and of the necessity for strong reactivity, in order to solve problems as soon as they appear (Markus, 1983).

IT promises to be a powerful tool for reducing the costs of coordination, provided that information technology is seen in terms of how it supports new or redesigned processes, rather than existing business functions (Davenport and Short 1990).

Technical change

According to Markus (2004), a change of IT has the objective of improving the reliability of information, and reducing running costs and the cost of maintenance of the application tools, without causing substantial changes in the way of working and of organising tasks.

The role of the IT specialists is a central one, and other managers have the role of validating the project, providing financial support and, if necessary, expressing their needs in terms of technical specifications.

Organisational change

It is documented that most problems in CRM implementation are not technical. Instead, common problems include organisational change and fluctuation (Bygstad, 2003).

Markus and Robey (1988) distinguished three types of relations (technological imperatives, organisational imperatives and emerging perspectives) between ITs and organisational changes. The emerging perspectives acknowledge that there is great complexity in the causality phenomena and the organisational change associated with IT. By refusing to recognise the causal perspective of the changes, the emergent perspectives differ, in their logic, on the causal arguments concerning the technological and organisational imperatives.

Technochange

According to Markus (2004), 'technochange' refers to the use of new IT applications along with organisational changes. The objective is to improve companies' global performance. *"Using IT strategically to bring about organisational performance improvements is fundamentally different from both IT projects and organisational change programs. Unlike IT projects, which focus on improving technical performance, 'technochange' involves great potential impacts on the users, processes, and organisational performance"* (Markus, 2004, p. 5).

For Lehtonen and Martinsuo (2008), change programs have become a common vehicle for change and innovation in organisations. Cicmil et al. (2006) focused on project actuality research. They pointed out the complexity embedded in the project, the sense-making in project management and the failure of complex projects.

The critical success factors for these projects are technical, organisational and managerial effectiveness (Markus, 2004).

Despite the availability of successful and more reliable technologies, and companies' use of external skills that are recognised on a technical level, managing CRM implementation projects remains a risky undertaking (Corner and Hinton, 2002; Bull, 2003; CSO Insights, 2006). In fact, the introduction of large-scale integrated IS leads to more significant changes in processes, tasks and people than traditional computing projects (Winter et al., 2006).

THE OPPOSITE END USERS REACTIONS TO THE IMPLEMENTATION OF CRM

We analyze the behavior of end users (divided into two groups-the commercial-business and customer services support) following the implementation of CRM within two modules.

The use of CRM by Sales People

The modes of action and work have not changed following the introduction of the new module “handling of business affairs”, "a rejection or underutilization of the tool are observed. The business has retained the use of their old tools (Excel with paper diaries, while the tool offers automatic features to track appointments and the calculation of commercial offerings). Sellers have noted "why bring information that will be used by all", the retention of customer's information is considered important to achieve the business objectives of each seller.

The use of CRM, by supplying it with all necessary data at the commercial level (customer contacts, visits, results of RDV etc.), will facilitate the control of the superior (sales manager) on the quality the work done. In addition, salespeople are unwilling to enter data on commercial CRM application. They believe that these tasks are the tasks of the marketing support.

There is little use of the functionality of the software by the commercial agencies (a connection rate of 10% of potential users). The sellers have not changed their methods of work or their business approach, using the CRM tool is only to record investments at the end of the month (for calculation of the variable incentives). Behaviors and attitudes of commercial individualists are significant psychological barriers to the use of the application and sharing of information. “We need a change of mentality, a seller is essentially individualistic. The tool is completely transverse everyone can connect to the application and what happens to the client (tracking customer visits etc..)” (Note the project manager). Thus, the seller sets everyone's business; it must support this change of mentality.

The use of CRM by Customer Services Support People:

The CRM Provides a 360-degree customer view, guaranteeing everyone in the service has the information he needs to deliver superior service every time.

The applications of the CRM facilitate access to real-time customer information. Staff can respond quickly and accurately to customer needs every time.

A favorable and extensive use of these tools has been recognized since the implementation of the CRM. The change in legacy applications by the new tool was well accepted. New ways of working are implemented with the process approach using the systems' workflows built into the tool.

The CRM application enables automatic tracking of tasks. It is a tool for sharing real time information to facilitate the collaborative work of users. There has been a change in the way of working actors in connection with the client. An actor does not need to seek information from other employees; everything is available on the tool.

The communication between individuals involved in the process has changed. There is less waiting for a response from a colleague on a particular problem, it will get more information directly available on the tool. This new approach to work was quickly assimilated by the actors involved in the process. Little resistance was observed, making them more effective users. Thus, expressions of opposition and inertia from the newly implemented artifact were minimal.

The current use of CRM application by actors of customer service support is dominated by an assessment and projection of future benefits that can be used effectively on work effectiveness and consequently the customer satisfaction.

Unlike commercial team members, customer service support people adhere more strongly to the CRM approach and use the tool in their daily work. The culture of collective work to resolve complaints or provide solutions to customer facilitates understanding of the benefits of the tool in terms of sharing information, real-time streaming data and full visibility.

CONCLUSION

Throughout the 1990s, the literature on management information systems (MIS) found that companies are implementing new IT and linking them to BPR projects (Davenport and Short 1990; Broadbent and Weill 1997. The results mentioned a set of shared assumptions. IT became an essential tool to enact and support transformation in firms. The distinction between technical change and organizational change don’t encourage firms to transform radically the process and the culture. The implementation of a new approach as CRM is considered as a technochange. It requires more holistic approach that takes into account technical, organizational and cultural issues. Technochange is far more complex, larger and more strategic in its nature (Lehtonen and Martinsuo, 2008).

Managing techno change can’t be done by only technicians or business managers. Project teams have to be composed members from various disciplines and having experiences in techniques and management. A mix of skilled people in different areas is required to achieve the results of implementing CRM approach in large firms.

REFERENCES

Broadbent, M., Weill, P. (1997). Management by Maxim: How Business and IT Managers can Create IT Infrastructures, *Sloan Management Review, Spring*, 77-92.

Cicmil, S., Williams, T., Thomas, J., Hodgson, D. (2006). Rethinking project management: Researching the actuality of projects. *Int. J. Project Man.*, 24, 675-686.

Corner I., Hinton M. (2002). Customer relationship management systems: Implementation risks and relationship dynamics. *Qualitative Marketing Research*, 5(4), 239-251.

CSO Insights (2006). *Sales Performance Optimization - 2006 Survey Results and Analysis*. Boulder.

Davenport, T. H., Short, J. E. (1990). The new industrial engineering: Information technology and business process design. *Sloan. Manage. Rev.*, 31 (4), 11-27.

Lehtonen, P., Martinsuo, M. (2008). Change program initiation: Defining and managing the program-organization boundary. *Int. J. Project Manage.*, 26, 21-29.

Markus, M. L. (2004). Technochange management: Using IT to drive organizational change. *J. Infor. Technol.*, 19, 4-20.

Van de Ven, A. H., Poole, M. S. (1995). Explaining development and change in organizations. *Academy Management Review*., 20 (3), 510-540.

In: Advances in Communications, Volume 8
Editor: Anthony V. Stavros
ISBN: 978-1-61324-794-5

Chapter 10

INTERNET TRAFFIC CLASSIFICATION USING MACHINE LEARNING WITH PERFORMANCE GUARANTEES

Zhiyuan Luo*
Department of Computer Science
Royal Holloway, University of London
Egham, Surrey TW20 0EX, UK

Abstract

The Internet is a globally distributed network of networks supporting many applications and services. In addition to traditional applications (e.g. Email and FTP), new Internet applications such as multimedia streaming, Internet telephony, games and peer-to-peer file sharing become popular. Classification of network applications responsible for the generation of the traffic flows is important for network management tasks such as monitoring trends of the applications in operational networks and effective network planning and design. In particular, accurate and fast classification of network traffic based on statistical flow characteristics can offer substantial benefits to a number of key areas in Quality of Service support and service differentiation, enforcement of security policies, traffic engineering and support for Service Level Agreements for many network operators and service providers. The requirements for such classification of Internet traffic are increasing rapidly due to availability of high capacity communications links, heterogeneity of Internet traffic types and continued evolution of applications such as tunnelling and end-to-end encryption.

Accurate and fast traffic classification requires algorithmic capability, in particular, machine learning algorithms. The idea of using machine learning techniques for Internet traffic classification is not new. Many machine learning algorithms have been successfully used to classify network traffic flows with good performance, but without information about the reliability in classifications. However, not knowing the confidence of predictions makes it difficult to measure and control the risk of error using a decision rule. Modern network resource management systems are becoming increasingly complex and as such require high quality, reliable predictions to optimise

*E-mail address: zhiyuan@cs.rhul.ac.uk

performance. Therefore, introducing more reliable algorithms for network resource management will result in higher quality of service for network users.

In this paper, we consider the problem of reliable Internet traffic classification. We present two recently developed machine learning algorithms, namely the Conformal Predictor and Venn Probability Machine, for making reliable decisions under uncertainty and achieving performance guarantees. Experiments on publicly available real Internet traffic datasets show these two algorithms work well. Performance comparison is also made between these two algorithms. Some guidelines on how to choose between these algorithms are provided and discussed.

Keywords: Performance Guarantees, Machine Learning, Venn Probability Machine, Conformal Predictors, Probabilistic Prediction.

1. Introduction

The Internet is a globally distributed network of networks supporting many applications and services. In addition to traditional applications (e.g. Email and FTP), new Internet applications such as multimedia streaming, Internet telephony, games and peer-to-peer file sharing become popular. Classification of network applications responsible for the generation of the traffic flows is important for network management tasks such as monitoring trends of the applications in operational networks and effective network planning and design. In particular, accurate and fast classification of network traffic based on statistical flow characteristics can offer substantial benefits to a number of key areas in Quality of Service support and service differentiation, enforcement of security policies, traffic engineering and support for Service Level Agreements for many network operators and service providers [7]. The requirements for such classification of Internet traffic are increasing rapidly due to availability of high capacity communications links, heterogeneity of Internet traffic types and continued evolution of applications such as tunnelling and end-to-end encryption.

Accurate and fast traffic classification requires algorithmic capability, in particular, machine learning algorithms. The idea of using machine learning techniques for Internet traffic classification is not new. Many machine learning algorithms have been successfully applied to the problem of network traffic classification [6, 8, 10, 14, 15, 19, 24, 25], but usually these algorithms provide bare predictions, i.e. predictions without any measure of how reliable they are. However, not knowing the confidence of predictions makes it difficult to measure and control the risk of error using a decision rule. Modern network resource management systems are becoming increasingly complex and as such require high quality, reliable predictions to optimise performance. Therefore, introducing more reliable algorithms for network resource management will result in higher quality of service for network users. Algorithms that do provide some measure of reliability with their outputs usually make strong assumptions on the data. Bayesian predictors, for example, require the a priori distribution on the data. It was shown in [23] that if this a priori distribution is wrong, the algorithm provides unreliable confidence numbers (measures of reliability). In this paper we consider the application of reliable predictors which make weak assumptions on the data to the problem of network traffic classification.

In many practical applications of machine learning, it is crucial to know how reliable predictions are rather than have predictions without any estimation of their accuracy. In addition, it would be useful to obtain information regarding how strongly we believe in each individual prediction rather than a whole group of predictions for all examples. When a prediction is made in a classification or regression problem, it is useful to have additional information on how reliable this individual prediction is. Such predictions complemented with the additional information are also expected to be valid, i.e., to have a guarantee on the outcome. Recently developed frameworks of confidence predictors and Venn probability machines [23] allow us to address these problems: confidence measures complement each prediction with its confidence and output region predictions with the guaranteed asymptotical error rate; Venn probability machines output multiprobability predictions which are valid in respect of observed frequencies. Another advantage of these frameworks is the fact that they are based on the i.i.d. (independent and identically distributed) assumption and do not depend on the probability distribution of examples. The remainder of the paper is structured as follows. Conformal predictors and Venn probability machines are presented in Section 2. Experimental results of both conformal predictors and Venn probability machines on publicly available real traffic datasets are described and discussed in Section 3. Finally, some conclusions and discussions are given in Section 4.

2. Machine Learning Algorithms with Performance Guarantees

In machine learning, we are given a training set of examples $z_1, z_2, \ldots, z_n$. Each example z_i consists of an object x_i and its label y_i, $z_i = (x_i, y_i)$. The objects are elements of an object space X and the labels are elements of a finite label space Y. For example, in traffic classification, x_i are multi-dimensional traffic flow statistics vectors and y_i are application groups taking only finite number of values such as {WWW, FTP, EMAIL, P2P}.

Suppose that the label space Y is enumerated for all possible classification labels $1, 2, \ldots, |Y|$. The learner Γ is a function on a finite sample of n training examples $z_1, z_2, \ldots, z_n \in \mathbf{Z}_n$ that makes a prediction for a new object $x_{n+1} \in X$ If in a problem of classification or regression we simply attempt to predict a label for a new object, we look for a function of the type

$$\Gamma:\ \mathbf{Z}^n \times X \to Y$$

which we call a *simple predictor*. Such predictor for any finite sequence of labelled objects $z_1, z_2, \ldots, z_n$ and a new object x_{n+1} without a label outputs $F(z_1, z_2, \ldots, z_n, x_{n+1})$ as its prediction for a new label y_{n+1}. There are many examples of single predictors [16, 20, 22].

Alternatively, a probability forecasting can be made

$$\Gamma:\ \mathbf{Z}^n \times X \to [0,1]^{|Y|}$$

Probability forecasting estimates the conditional probability of a possible label given an observed object. For each new object x_{n+1} (with true label y_{n+1} withheld from the learner), a set of predicted conditional probabilities for each possible labels are produced.

To construct a reliable algorithm, we need to make some assumptions on the data generating mechanism. Our standard assumption used in the paper is the i.i.d. assumption. The examples z_i are assumed to be generated independently by the same probability distribution P on $\mathbf{Z}$, i.e., the infinite sequence of examples $z_1, z_2, \ldots$ is drawn from the power probability distribution P^∞ on $\mathbf{Z}^\infty$ ($\mathbf{Z}^\infty$ is the set of all infinite sequences of elements of $\mathbf{Z}$.

Traditionally, most learning algorithms give predictions in the form of a single class label with no measure of how confident the algorithm is in that prediction. In many situations, it is crucial to know how reliable predictions are rather than have predictions without any estimation of their accuracy. For example, in medical diagnosis, this would give practitioners a reliable assessment of risk error; in drug discovery, such control of the error rate in experimental screening is also desirable since the testing is expensive and time consuming.

This paper focuses on machine learning frameworks of conformal predictors and Venn probability machines, which were introduced in [23] and represent a new generation of hedged prediction algorithms. These newly developed methods have several advantages. Firstly, they hedge every prediction individually rather than estimate an error on all future examples as a whole. As a result, the supplementary information which is assigned to predictions and reflects their reliability is tailored not only to the previously seen examples (a training set) but also to the new object. Secondly, both frameworks of confidence and Venn machines produce valid results. Validity is an important property of algorithms and in this case has a form of a guarantee that the error rate of region predictions converges to or is bounded by a pre-defined level when the number of observed examples tends to infinity or that a set of output probability distributions on the possible outcomes agrees with observed frequencies. The property of validity is based on a simple i.i.d. assumption (that examples are independent and identically distributed) or a weaker exchangeability assumption and does not depend on the probability distribution of examples. The latter assumption can be often satisfied when data sets are randomly permuted. The property of validity is theoretically proved [23] in the online mode, when examples are given one by one and the prediction is made on the basis of the preceding examples.

2.1. Conformal Predictors

Conformal Predictors is a Machine Learning technique of making predictions (classifications) according to how similar the current example is to the representatives of different objects' classes. The algorithm is based on the idea of *Nonconformity Measure*, a function which gives some measure of dissimilarity an example to other examples. The higher the value of this function on an example, the more unlikely that this example belongs to the selected group of examples. To exploit this idea, we can assign different labels to the current example, calculate the dissimilarity of this example to other examples with the same label, calculate the dissimilarity of this example to the examples with other labels and using this value classify the example. In order to describe Conformal Predictors more precisely we need to introduce some formal definitions.

First we define a Nonconformity Measure. For each $n = 1, 2, \ldots$ we define the function

$$A_n : \mathbf{Z}^{n-1} \times \mathbf{Z} \to \overline{\mathbb{R}}$$

as the restriction of A to $\mathbf{Z}^{n-1} \times \mathbf{Z}$. The sequence $(A_n : n \in \mathbb{N})$ is called a Nonconformity Measure.

To give an idea on what a Nonconformity Measure is, we need to use the term *bag*. A bag (a multiset) is a set of elements, where each element has a multiplicity, i.e. a natural number indicating how many memberships it has in the multiset. Given a Nonconformity

Measure (A_n) and a bag of examples $z_1, \ldots, z_n$ we can calculate the nonconformity score:

$$\alpha_i = A_n\left(z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_n, z_i\right) \tag{1}$$

for each example in the bag. The nonconformity score is a measure of how dissimilar an example is from a group of examples. For example, in the case of using 1-NN algorithm as the underlying algorithm in Conformal Predictors (the algorithm used to calculate the nonconformity score), the nonconformity score can be calculated as

$$A_n\left(z_1, \ldots, z_n, z\right) = \frac{min_{i \in \{1, \ldots, n\}, y_i = y} d(x_i, x)}{min_{i \in \{1, \ldots, n\}, y_i \neq y} d(x_i, x)}$$

it means that we consider an example as nonconforming if it lies much further from the selected group of objects than to all other examples (not form the selected group).

Now, to use the nonconformity scores we want to compare the current example (which we are trying to classify) and the examples from the training set. In this way, we can see how the current example fits into the "whole picture" of the dataset. Now we introduce the term p-value. The number

$$\frac{|\{j = 1, \ldots, n : \alpha_j \geq \alpha_n\}|}{n}$$

is called *p-value* of the object $z_n = (x_n, y_n)$. The Conformal Predictor receives a significance level (a small positive real number reflecting the confidence we want to achieve, it equals to one minus Confidence Level) as a parameter, to output a label, the Conformal Predictor calculates p-values of the pairs $(x_n, \text{possible label})$ and outputs all such pairs that their p-values are greater than the confidence level.

Algorithm 1 Conformal Predictor Algorithm for Classification

Input: training example sequence $\{z_1 = (x_1, y_1), z_2 = (x_2, y_2), \ldots, z_l = (x_l, y_l)\}$
Input: new example x_{l+1}
Input: non-conformity measure A
Input(optional): confidence level γ
for $y \in Y$ **do**
  $z_{l+1} = (x_{l+1}, y)$
  for j in $1, 2, \ldots, l+1$ such that $y_j = y$ **do**
    $\alpha_j = A(z_j, \{z_1, \ldots, z_l, z_{l+1}\})$
  end for
  $p(y) = \frac{\#\{j=1,\ldots,l+1: y_j = y\}, \alpha_j \geq \alpha_{l+1}\}}{\#\{j=1,\ldots,l+1: y_j = y\}}$
end for
Output(optional): predictive set $R^{\gamma}_{l+1} = \{y : p(y) \geq 1 - \gamma\}$
Output: single prediction $\hat{y}_{l+1} = \arg\max_y \{p(y)\}$
Output: confidence $conf(\hat{y}_{l+1}) = 1 - \max_{y \neq \hat{y}_{l+1}} \{p(y)\}$
Output: credibility $cred(\hat{y}_{l+1}) = \arg\max_y \{p(y)\}$

The conformal predictor calculates $p(y)$ which is a *p-value* associated with a hypothetical completion $y_{l+1} = y$ for the test example (see Algorithm 1). Using these *p-values*, predictions made by conformal predictors can be presented in following two ways:

- Either the user inputs a required degree (level) of certainty and the algorithm outputs a predictive set: a list of classifications that meet this confidence requirement.
- Or the algorithm outputs so called *single prediction*: an individual predictions together with its confidence, which is the minimal level of certainty at which the output is certain (predictive set consists of only one label) and credibility, which indicates whether the test object is representative of the training set, in order to compare with the other conventional learning methods.

The common way of presenting the prediction results is to choose a significance level ("degree of certainty") $\gamma < 1$ and output the (γ)-*predictive set (region)* containing all labels with p-value equal or greater than $1 - \gamma$:

$$R^{\gamma}_{l+1} = \{y :\ p(y) \geq 1 - \gamma\}. \tag{2}$$

Prediction errors can occur when the prediction set fails to contain the true label. If the size of the predictive set R^{γ}_{l+1} is 1 or 0, then this prediction is *certain*, otherwise it is uncertain and we have multiple predictions. In the context of conformal predictors, uncertain predictions are not errors, but these are indications that the amount of information is not sufficient to make a certain decision at the selected level. Naturally, the higher the confidence level is, the more multiple predictions will appear. If non-conformity measure is specified adequately for the problem at hand, the percentage of uncertain predictions should not be too high.

The main advantage of conformal predictors is their property of validity: the asymptotic number of errors, that is, erroneous region predictions, can be controlled by the significance level the error rate we are ready to tolerate which is predefined by the user (the prediction is considered to be erroneous if it does not contain the true label). All precise definitions will be given later. However, the property of validity is achieved at the cost of producing region predictions: instead of outputting a single label as a prediction, we may produce several of them any of which may be correct. Predictions that contain no labels are called empty predictions, those that contain one label are called certain predictions, and those comprising more than one label multiple predictions. Such multiple predictions are not mistakes: they are output when the conformal predictor is not provided with sufficient information for producing valid predictions at a certain error rate. Informativeness, or in other words efficiency, of a conformal predictor can be translated as its ability to produce as small region predictions as possible. Thus, we have to balance validity (the error rate) and efficiency (the number of labels in each prediction): lower error rates will result in larger region predictions, and vice versa. This feature makes conformal predictors a very flexible tool.

2.2. Venn Probability Machine

Venn Probability Machines are based on the idea of dividing examples into groups and, when a new object arrives, somehow assigning it to one of the groups. Then we can use the frequencies of labels in the group containing the current object as probabilities for the new object's label.

More formally, let N be a predefined number of examples used as the input by the Venn Probability Machine. *Venn taxonomy* is a sequence A_n, $n = 1, \ldots, N$, where each A_n is a finite partition of the space $\mathbf{Z}^{(n-1)} \times \mathbf{Z}$. The expression $\mathbf{Z}^{(n-1)} \times \mathbf{Z}$ here means that A_n does not depend on the order of the first $n-1$ arguments. Let $A_n(\omega)$ be the element of the partition (A_n) that contains $\omega \in \mathbf{Z}^{(n-1)} \times \mathbf{Z}$. With every taxonomy (A_n) we associate a Venn Probability Machine. After choosing a taxonomy, we can start making predictions. Each prediction is a class associated with the current example by the Venn Probability Machine.

Consider the following protocol. Let us call the examples' possible class (P_n). On each step reality makes a move x_i, the classifying system has to assign the new example to one of the classes P_n. Firstly, we consider a class $y \in Y$ and partition $\{(x_1, y_1), \ldots, (x_n, y)\}$ into categories, assigning two points in $\mathbf{Z}$ to the same category if and only if

$$A_n\left(\{z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_n\}, z_i\right) = A_n\left(\{z_1, \ldots, z_{j-1}, z_{j+1}, \ldots, z_n\}, z_j\right),$$

where $z_i = (x_i, y_i)$, $i = 1, \ldots, n-1$ and $z_n = (x_n, y)$. In this case we assume that $\{z_1, \ldots, z_k\}$ means a multiset (a bag), i.e. a set of elements, where each element has a multiplicity, i.e. a natural number indicating how many memberships it has in the multiset.

The category T containing $z_n = (x_n, y)$ is nonempty. Let p_y be the empirical probability distribution of the labels in this category T:

$$p_y\{y'\} := \frac{|\{(x^*, y^*) \in T : \; y^* = y'\}|}{|T|}, \quad y' \in Y.$$

This is a probability distribution on Y. The Venn Predictor determined by the taxonomy is the multiprobability predictor $P_n := \{p_y : \; y \in Y\}$. The set P_n consists of between one and $|Y|$ distinct probability distributions on Y. For each probability distribution, for each possible label we get the probability(frequency) that this label is the correct classification of the current example. from this matrix (with the number of rows equal to the number of columns equal to the power of label space) we can get a probability interval for each label that this label is true (as the interval with the minimum and the maximum probabilities of this event. All this information is the output of the Venn Probability Machine. The prediction here is the label with the highest probability to be correct (e.g. with the largest mean of the probability interval), see Algorithm 2. More information on Venn Probability Machines and Conformal Predictors can be found in [23].

Venn probability machines output multiprobability predictions a set of probability distributions of a label. This output can be also interpreted in a different way: as a prediction with the assigned interval of probability that this prediction is correct. Venn machine outputs are always valid (precise definitions will be given later). The property of validity is based only on the i.i.d. assumption, that the data items are generated independently from the same probability distribution. This assumption is much weaker than any probabilistic assumption, which allows Venn machines to produce valid predictions without knowing a real distribution of examples. Venn machines represent a framework that can generate a range of different algorithms. Similarly to conformal predictors, practically any known machine learning algorithm can be used as an underlying algorithm in this framework and thus result in a new Venn machine. However, regardless of the underlying algorithm, Venn machines output valid results.

Algorithm 2 Venn Probability Machine Algorithm

Require: training examples $\{(x_1, y_1), (x_2, y_2), \ldots, (x_{n-1}, y_{n-1})\}$, label space Y,
Require: new example x_n,
Require: taxonomy A_n.
for $y = 1$ to $|Y|$ **do**
combine (x_n, y) with the training examples (We now have N examples $\{(x_1, y_1), \ldots, (x_{n-1}, y_{n-1}), (x_n, y)\ \}$)
partition the N examples into categories T, assigning (x_i, y_i) and (x_j, y_j) to the same category if and only if $A_n(\{z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, (x_n, y)\}, z_i)$ = $A_n(\{z_1, \ldots, z_{j-1}, z_{j+1}, \ldots, (x_n, y)\}, z_j))$
for $i = 1$ to $|Y|$ **do**
$F_{y,i} = \frac{|\{(x*, y*) \in T : y* = y\}|}{|T|}$ (The category T containing (x_n, y) is nonempty, because it has at least this one element)
end for
end for
$Q = \arg\min_y F_{y,i}$
$i_{best} = \arg\max_i Q_i$
predict $\hat{y} = \arg\max_y F_{y,i_{best}}$
Output: predicted probability interval for $\hat{y}$ as $[\min_y F_{y,\hat{y}}, \max_y F_{y,\hat{y}}]$
Output: error probability interval [1-$\max_y F_{y,\hat{y}}$,1-$\min_y F_{y,\hat{y}}$]
Output: predicted mean probability for $\hat{y}^t$ as $\text{mean}_y F_{y,\hat{y}^t}$

2.2.1. Region prediction for Venn Probability Machine

At each step Venn Probability Machines provide one prediction corresponding to the highest probability. On the other hand, Conformal Predictors can output region predictions. To compare the two algorithms we want to convert probabilistic predictions of Venn Probability Machine into confidence region predictions. Intuitively, we use the fact that the probability predictions are estimates of conditional probabilities and also assume that labels are mutually exclusive [3]. Therefore summing these predictions becomes a conditional probability of a conjunction of labels, which can be used to choose the labels to include in the confidence region predictions at the desired confidence level. Consider the case when a Venn Probability Machine outputs a sequence

$$p_1, p_2, \ldots, p_l,$$

where p_i is the probability that label i is correct. Let us consider the case where $p_1 > p_2 > \cdots > p_l$. If it is not true we can always rename the labels. Now to make a region prediction we carry out the following procedure: we output all labels i such that $\sum_{j=1}^{i} p_j \le 1 - \epsilon$, where ϵ is the significance level, i.e. the level of mistakes we can tolerate, as a possible label. For the label n such that $\sum_{j=1}^{n-1} p_j \le 1 - \epsilon$, but $\sum_{j=1}^{n} p_j > 1 - \epsilon$ we need a more sophisticated approach: we will output label n as a possible label if and only if $\big(\text{rand}() >$

$\left(1-\epsilon-\sum_{j=1}^{n-1} p_j\right)/p_n\big)$, where $\mathrm{rand}()$ is a random number between 0 and 1. The reason for this condition is that we want the probability for a label from the output to be correct with a predefined probability $1-\epsilon$. A more rigorous description of these steps is given in Algorithm 3.

Algorithm 3 Region Predictions for Venn Probability Machines

Require: significance level $\epsilon > 0$
Require: probability forecasting $(p_1, p_2, \ldots, p_l) : p_1 > p_2 > \ldots > p_l$

```
P = 0
i = 1
EXIT_FLAG = 0
while (P < 1 − ϵ) AND (EXIT_FLAG == 0) do
  if P + p_i > 1 − ϵ then
    EXIT_FLAG = 1
    if (rand() > (1 − ϵ − P)/p_i) then
      add label i to the set of predictions
    end if
  else
    P = P + p_i
    add label i to the set of predictions
  end if
  i = i + 1
end while
```

3. Experiments

In this section we describe the underlying algorithm which we used to build Venn taxonomies and nonconformity measures, the publicly available Internet traffic datasets and experimental results of the implemented conformal predictors and Venn probability machines.

3.1. Nearest Neighbours

Both conformal predictors and Venn probability machines are general framework and can be implemented on on top of a so-called underlying algorithm, i.e. an algorithm which gives predictions (or classifications) without any measure of reliability [2–5]. These predictions are then processed and the decision making system receives a number reflecting each prediction's reliability. The performance of the underlying algorithm has an impact on the performance of Venn Probability Machines and Conformal Predictors: if the underlying algorithm performs badly, the reliable predictors will output predictions of poor quality with low reliability, i.e. they will specify that the output is unreliable. There are different ways to define the nonconformity measure, the core element of any conformal predictors. Almost any machine learning algorithm can be used to construct it. There are known implementations based on such algorithms as Support Vector Machines [**?**, 18], k-nearest neighbours [1, 17], nearest centroid [1], linear discriminant [21], naive Bayes [21], kernel perceptron [13]. The most successful and the most widely used ones have been nonconformity measures derived from k-nearest neighbours.

The Nearest Neighbours (NN) algorithm is one of the simplest methods in Machine Learning for making predictions. For example, when the number of NN is 1 (the simplest case), this algorithm consists in finding the current label's nearest neighbour using a predefined distance function and assigning its label to the current example's label. This algorithm can be expanded on the case with several NN. In this case, the algorithm finds n NN, where n is the number of nearest neighbours, and according to its labels make a decision on the current example's label, e.g. the algorithm can choose the most common label among the found label set. More information on the NN algorithm can be found, for example in [23].

In our experiments with Venn Probability Machines we used two types of taxonomies based on the NN algorithm [2]. One type, called "voting", says that two objects belong to the same taxonomy if and only if for each object the most frequent label among its N (parameter of the NN algorithm) is the same. The second type of taxonomies (called "all info") says that two objects belong to the same taxonomy if and only if for each of them the ordered sequence of its N nearest neighbours' labels is the same.

For Conformal Predictors we want to scale the distance to the nearest neighbour and compare relative values, but not distances. The nonconformity measure α_i can be defined as follows [3]:

$$\alpha_i = \alpha_i(x_i, y_i) = \frac{\min_{y_j = y_i} d(x_j, x_i)}{\min_{y_j \neq y_i} d(x_j, x_i)}.$$

This implies that an object is considered to be nonconforming if it is far from objects with the same label and close to objects labelled in a different way. In the case of several nearest neighbours we take their mean value instead of x_j. Again we look for nearest neighbours with the same label as (x_i, y_i) to calculate the upper term of fraction and with different labels to calculate the lower term of fraction. In this paper we use the Euclidean distance measure:

$$d(a, b) = \sqrt{\sum_{i=1}^{n} (a_i - b_i)}.$$

3.2. Dataset

The Network Traffic Classification Datasets (NTCD) is a collection of thousands of examples of hand-classified TCP flows (we used the datasets after feature selection, performed by other researchers ([10])). The collection is comprised of 11 datasets and is publicly available at *http://www.cl.cam.ac.uk/research/srg/netos/nprobe/data/papers/sigmetrics/index.html*. Flows are labelled according to the type of traffic they represent, for example WWW, EMAIL, FTP or P2P traffic. For each flow, there are 248 features describing simple statistics about packet length and inter-packet timings, and information derived from the TCP protocol such as SYN and ACK count, for details see [12].

Table 1 shows the number of TCP flows in each dataset. Eight types of network applications were considered: WWW (1), EMAIL (2), FTP (3), ATTACK (4), P2P (5),

Table 1. Traffic types by dataset

Dataset	1	2	3	4	5	6	7	8
NTCD-1	18211	4146	1511	122	339	238	87	206
NTCD-2	18559	2726	1701	19	94	329	150	220
NTCD-3	18065	1448	2736	41	100	206	136	200
NTCD-4	19641	1429	600	324	114	8	54	113
NTCD-5	18618	1651	928	122	75	0	38	216
NTCD-6	16892	1618	521	134	94	0	42	82
NTCD-7	51982	2771	484	89	116	36	36	293
NTCD-8	51695	2508	551	129	289	43	33	220
NTCD-9	59993	3678	1577	367	249	15	0	337
NTCD-10	54436	6592	930	446	624	1773	0	212
NTCD-12	15597	1799	1513	0	297	295	0	121

DATABASE (6), MULTIMEDIA (7) and SERVICES (8). As it is pointed out in [11] the simplest type of network traffic classification, which is by port number, is not as good as Machine Learning techniques and here we try to contribute to solving this problem.

3.3. Results

In this section we present and discuss the experimental results on the NTCD datasets. All experiments were carried out in the online setting, i.e. we ran the algorithms on the training sets first and then at each step after making a prediction we were adding the current object with its true label to the training set. This construction allows the prediction system to improve its performance over time.

Both conformal predictors and Venn probability machines were implemented on top of the Nearest Neighbours (NN). Single and multiple (region) predictions were considered. For multiple predictions, significance levels of 1%, 5% and 10% were used. As discussed in Section 2, the performance of these predictors is measured by validity (measured in error rates) and efficiency (measured in average number of labels observed). In multiple prediction, if the true label is a member of the prediction set, then the prediction is correct; otherwise, we have a prediction error. Here we are interested in two factors: validity and efficiency. Validity means that the error rate is closed to the expected value, which is equal to a predefined constant (significance level). It is important to note that while conformal predictors provide guaranteed validity, Venn probability machines do not provide such a guarantee. Efficiency reflects how useful the predictions output by a predictor are. Ideally we would like our prediction system to output only one label at each step for multiple predictions.

For illustration purpose, Tables 2 and 3 provide a summary of the experimental results of conformal predictors (CP) and Venn probability machines (VPM) on dataset NTCD-1 and NTCD-2 with different number of nearest neighbours. In these tables , VPM-V stands for Venn Probability Machine based on “voting” taxonomy and VPM-AI stands for Venn Probability Machine based on “all info” taxonomy. Some of the main conclusions which can be drawn from the results are as follows:

Table 2. Performance comparison on dataset NTCD-1

Prediction algorithm	Underlying algorithm	Error rate (%)				Average number of labels		
		1%	5%	10%	Single	1%	5%	10%
CP	1NN	1.13	4.88	6.25	0.66	1.38	1.34	1.33
	2NN	1.10	4.89	6.67	0.62	1.38	1.34	1.33
	3NN	1.12	4.79	6.83	0.68	1.38	1.34	1.32
	4NN	1.16	4.88	6.94	0.69	1.38	1.34	1.32
	5NN	1.19	4.85	7.02	0.71	1.38	1.34	1.32
	6NN	1.17	4.94	7.11	0.72	1.38	1.34	1.32
	7NN	1.17	4.84	7.23	0.78	1.38	1.34	1.32
	8NN	1.16	4.92	7.35	0.81	1.36	1.33	1.30
	9NN	1.16	4.90	7.47	0.84	1.36	1.32	1.30
	10NN	1.17	4.92	7.54	0.86	1.36	1.32	1.29
VPM-V	1NN	1.19	4.96	10.20	0.58	1.06	0.96	0.90
	2NN	0.93	4.91	10.05	0.59	1.06	0.96	0.90
	3NN	1.09	4.81	10.11	0.63	1.07	0.96	0.90
	4NN	0.95	5.30	9.83	0.66	1.07	0.96	0.91
	5NN	1.00	4.89	9.27	0.68	1.09	0.96	0.91
	6NN	0.82	5.12	9.16	0.65	1.09	0.97	0.91
	7NN	0.94	4.81	9.91	0.66	1.09	0.97	0.91
	8NN	1.19	5.09	9.77	0.71	1.09	0.97	0.91
	9NN	1.13	5.47	10.32	0.76	1.09	0.97	0.90
	10NN	1.01	5.62	9.75	0.76	1.09	0.97	0.91
VPM-AI	1NN	1.19	4.96	10.20	0.58	1.06	0.96	0.90
	2NN	0.99	4.94	10.09	0.65	1.05	0.96	0.91
	3NN	1.09	4.78	10.00	0.62	1.04	0.96	0.91
	4NN	0.95	5.32	9.81	0.58	1.04	0.96	0.91
	5NN	0.75	3.15	5.74	0.66	1.38	1.31	1.25
	6NN	0.75	3.27	6.06	0.63	1.38	1.31	1.25
	7NN	0.74	3.25	6.40	0.68	1.38	1.31	1.24
	8NN	0.83	3.37	6.08	0.63	1.37	1.27	1.17
	9NN	0.88	3.52	6.61	0.64	1.38	1.31	1.24
	10NN	0.75	3.67	6.27	0.62	1.38	1.27	1.17

- Venn Probability Machines and Conformal Predictors have similar performance, both from a validity and an efficiency point of view.

- Venn Probability Machines in the multiple prediction mode provide empirical validity, although not as well as Conformal Predictors.

- The number of nearest neighbours does not significantly affect the performance of the Nearest Neighbours algorithm.

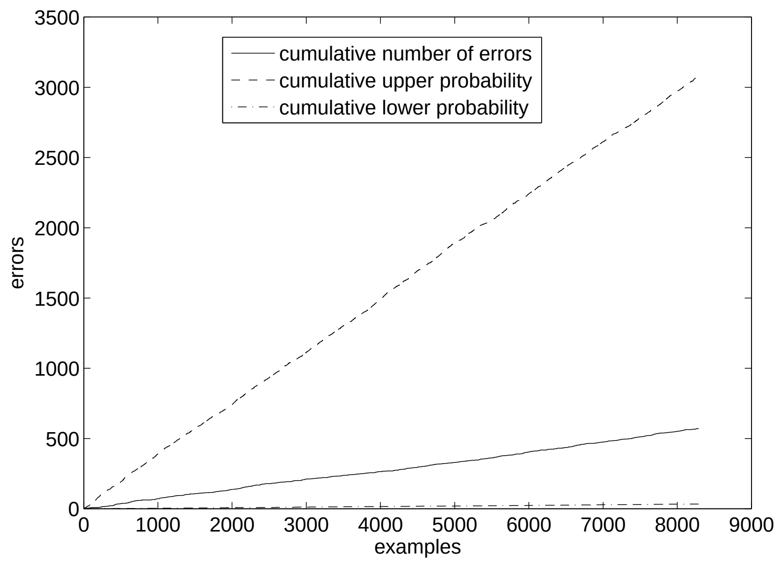


Figure 1 Cumulative number of errors and upper and lower bounds (NTCD-1, 5 Nearest Neighbours, "all info" taxonomy, single predictions).

Figures 1 shows examples of experimental results for Venn Probability Machines on NTCD-1 data in the online setting. Venn Probability Machines here are run in the single prediction mode. It can be seen that in the long run the total number of errors lies between the cumulative lower and upper bounds and this property of Venn Probability Machines guarantees the prediction quality. A decision making system which employs a Venn Probability Machine can use this property to improve its performance, for example by ignoring a prediction if such probability is too low.

Figure 2 represents a classical experimental result for Conformal Predictors. The graph of the number of errors over time should be a straight line and the percentage of errors at each point should approximately equal to the significance level. In many applications we can estimate a threshold on the number of errors we can tolerate. Conformal Predictors is a guaranteed method (i.e. there exist a theorem saying that the performance is guaranteed, see [23]) of making predictions with a predefined number of errors we can tolerate. As it can be seen from the figure the graphs of the number of errors are not straight lines but this is due to statistical fluctuations as the result on Conformal Predictors is probabilistic.

Figure 3 shows experimental results of efficiency for conformal predictors for different significance levels. It can be seen that the conformal predictors can ensure validity by giving

Table 3. Performance comparison on NTCD-2

Prediction algorithm	Underlying algorithm	Error rate (%)				Average number of labels		
		1%	5%	10%	Single	1%	5%	10%
CP	1NN	0.96	5.18	7.58	0.34	1.90	1.85	1.83
	2NN	0.98	5.16	7.87	0.38	1.89	1.85	1.82
	3NN	0.96	5.22	8.06	0.38	1.88	1.84	1.81
	4NN	1.01	5.22	8.14	0.40	1.88	1.84	1.81
	5NN	1.05	5.22	8.19	0.42	1.88	1.84	1.81
	6NN	1.03	5.24	8.24	0.42	1.88	1.84	1.81
	7NN	1.02	5.23	8.31	0.43	1.88	1.84	1.81
	8NN	1.02	5.27	8.38	0.44	1.88	1.84	1.81
	9NN	1.02	5.29	8.46	0.45	1.84	1.80	1.77
	10NN	1.00	5.23	8.53	0.47	1.66	1.62	1.59
VPM-V	1NN	0.83	5.00	9.95	0.34	1.00	0.95	0.90
	2NN	1.20	5.11	10.07	0.34	0.99	0.95	0.90
	3NN	1.21	4.85	9.88	0.37	1.00	0.96	0.90
	4NN	1.08	4.99	9.72	0.38	1.00	0.95	0.91
	5NN	1.06	5.18	10.10	0.39	1.01	0.95	0.90
	6NN	1.10	5.03	9.58	0.39	1.00	0.95	0.91
	7NN	1.03	5.45	9.61	0.40	1.01	0.95	0.91
	8NN	0.98	5.00	10.22	0.40	1.01	0.96	0.90
	9NN	0.92	5.07	10.40	0.42	1.01	0.96	0.90
	10NN	0.91	5.08	9.82	0.42	1.01	0.96	0.91
VPM-AI	1NN	0.83	5.00	9.95	0.34	1.00	0.95	0.90
	2NN	1.22	5.11	10.12	0.37	0.99	0.95	0.90
	3NN	1.26	4.88	9.91	0.38	0.99	0.96	0.91
	4NN	1.07	5.05	9.76	0.38	1.00	0.95	0.91
	5NN	1.07	5.26	10.12	0.37	1.00	0.95	0.90
	6NN	1.10	5.02	9.53	0.37	1.00	0.96	0.91
	7NN	1.05	5.43	9.61	0.37	1.00	0.95	0.91
	8NN	0.96	5.03	10.21	0.37	1.00	0.95	0.90
	9NN	0.98	5.08	10.41	0.37	1.00	0.96	0.90
	10NN	0.90	5.11	9.90	0.37	1.00	0.95	0.91

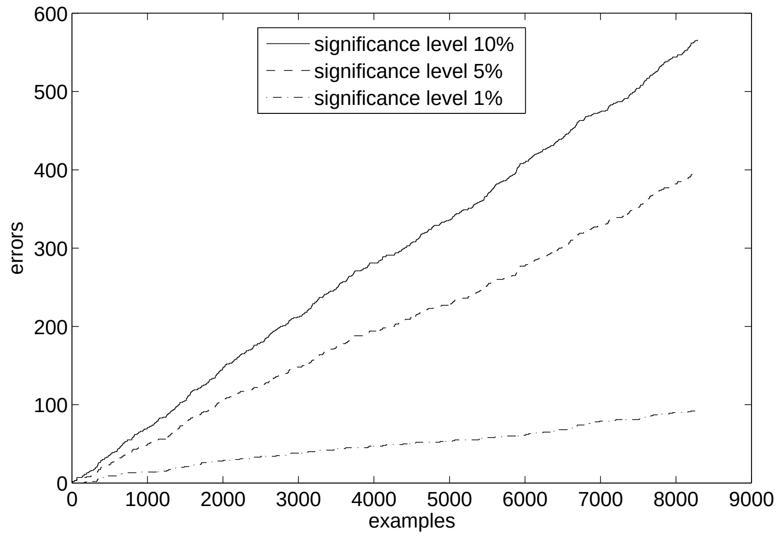


Figure 2 Number of errors for different significance levels (NTCD-1, 3 Nearest Neighbours, multiple predictions).

a region prediction. Instead of predicting a single class label for a new example, it can suggest several, any of which may be correct. Higher confidence levels can be accommodated by giving less efficient region predictions to hedge the classifier's decision.

4. Conclusion

In this paper we discussed and compared two types of machine learning algorithm capable of providing performance guarantees, namely Venn Probability Machines and Conformal Predictors. They represent one type of algorithms which produce predictions complemented with the information on their reliability. Algorithms with online validity represent a flexible framework and can use practically any known machine learning method as an underlying algorithm for designing a nonconformity measure (for conformal predictors) or a Venn taxonomy (for Venn machines).

Since these algorithms have theoretically guaranteed property of validity, their performance can be measured by their efficiency. Good performance in terms of accuracy of the underlying algorithm is usually translated into good efficiency of the corresponding conformal predictor or Venn probability machine. Therefore, efficiency of algorithms with online validity varies across the range of underlying algorithms and depends on the data analysed. For this reason, we are looking for new nonconformity measures and Venn taxonomies that could result in efficient predictions on certain types of data.

In contrast to simple predictors, conformal predictors and Venn probability machines hedge predictions, i.e., express how much a user can rely on them. We described two measures of performance of conformal predictors and Venn probability machines: validity and efficiency. Validity demonstrates how correct predictions are; efficiency is concerned with how informative they are. For conformal predictors, validity implies that the number

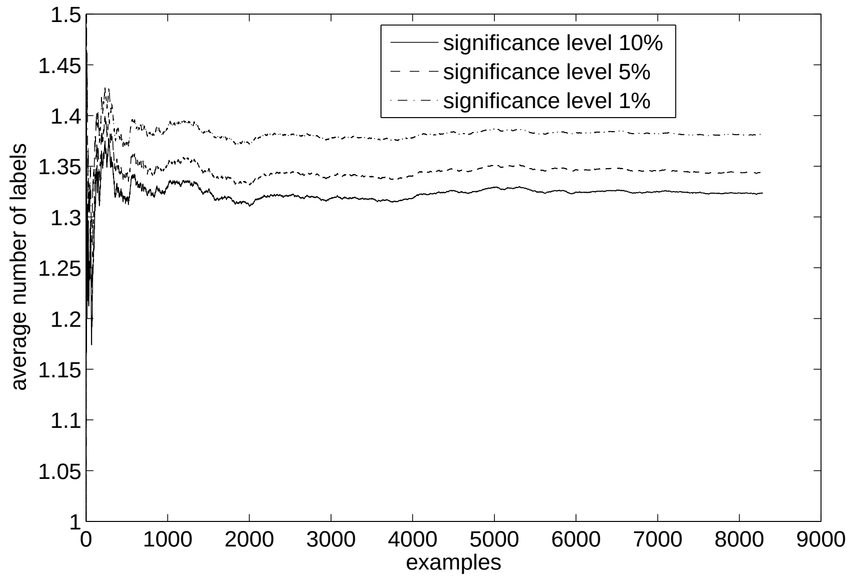


Figure 3 Average number of labels for different significance levels (NTCD-1, 3 Nearest Neighbours, multiple predictions).

of errors is close to the preset significance level, and efficiency means outputting as few as possible multiple predictions. For Venn probability machines, validity results in output probability distributions agreeing with observed frequencies. A probability interval output by Venn probability machine is efficient if it is narrow and close enough to 1.

We demonstrated that these algorithms have distinct advantages over other methods that hedge predictions or estimate overall performance on publicly available Internet traffic datasets. These advantages could be summarised as follows:

1. Property of validity is based on a simple i.i.d. assumption, which can be often satisfied when data sets are randomly permuted. Valid predictions do not depend on the assumed probability distribution of examples.

2. In the case of conformal predictors , the error rate can be controlled rather than computed, which makes these methods a flexible tool.

3. Methods with online validity are not single algorithms, but a flexible framework: each of them depends on a core element (a nonconformity measure for conformal predictors; a nonconformity measure and a Venn taxonomy for Venn probability machines), and practically any machine learning algorithm can be used to define this core element. Thus, the framework can give rise to a set of different algorithms which can perform well on different types of data.

The main conclusion which can be drawn from the experimental results is that a researcher willing to choose between these two algorithms should consider the one which has the theoretical guarantee on its performance, i.e. if a single prediction is required then it is reasonable to choose the Venn Probability Machine and in the case of region predictions it is better to choose Conformal Predictors. However, it is important to choose the underlying algorithm carefully as if it cannot perform well on the initial dataset both Venn Probability Machines and Conformal Predictors will perform poorly.

Acknowledgement

This work was supported by EPSRC through grant EP/E000053/1, “Machine Learning for Resource Management in Next-Generation Optical Networks”.

References

[1] T. Bellotti, Z. Luo, A. Gammerman, F.W. van Delft, and V. Saha. Qualified Predictions for Microarray and Proteomics Pattern Diagnostics with Confidence Machines. *International Journal of Neural Systems*, Vol. 15, No. 4, pp. 1–12, 2005.

[2] M. Dashevskiy and Z. Luo, Reliable Probabilistic Classification of Internet Traffic, *International Journal of Information Acquisition*, Vol. 6, No. 2, pp. 133–146, 2009.

[3] M. Dashevskiy and Z. Luo: Predictions with Confidence in Applications. International Conference on Machine Learning and Data Mining (MLDM 2009), pp. 775–786, 2009.

[4] M. Dashevskiy, and Z. Luo: Network Traffic Demand Prediction with Confidence. GLOBECOM 2008, pp. 1453–1457, 2008.

[5] M. Dashevskiy, and Z. Luo: Reliable Probabilistic Classification and Its Application to Internet Traffic. ICIC (1) 2008: pp. 380–388, 2008.

[6] A. Este, F. Gringoli, and L. Salgarelli. 2009. Support Vector Machines for TCP traffic classification. *Computer Networks*, Vol. 53, No. 14, pp. 2476–2490, 2009.

[7] H. Kim, K.C. Claffy, M. Fomenkov, D. Barman, M. Faloutsos and K. Lee. Internet traffic classification demystified: myths, caveats, and the best practices. In Proceedings of the 2008 ACM CoNEXT Conference (CoNEXT ’08). ACM, New York, NY, USA, , Article 11 , 12 pages, 2008.

[8] Z. Li, R. Yuan, and X. Guan, Accurate Classification of the Internet Traffic Based on the SVM Method, Proceedings of IEEE International Conference on Communications, ICC 2007: 1373–1378, 2007.

[9] A.W. Moore. Discrete content-based classification — a data set. *Technical report, Intel Research*, Cambridge, 2005.

[10] A.W. Moore, and D. Zuev. Internet traffic classification using Bayesian analysis techniques. *SIGMETRICS*, 2005.

[11] A.W. Moore, and D. Papagiannaki. Toward the accurate identification of network applications. *In Proceedings of the Sixth Passive and Active Measurement Workshop (PAM 2005)*, March 2005.

[12] A. Moore, D. Zuev, and M. Crogan. “Discriminators for use in flow-based classification” *Technical Report RR-05-13, Dept. of Computer Science, Queen Mary, University of London*, ISSN 1470-5559.

[13] T. Melluish, C. Saunders, I. Nouretdinov, and V. Vovk. The typicalness framework: a comparison with the Bayesian approach. Technical report 01-05, Royal Holloway, University of London, 2001.

[14] T.T.T. Nguyen and G. Armitage. A Survey of Techniques for Internet Traffic Classification using Machine Learning, *IEEE Communications Surveys and Tutorials*, Vol. 10, No. 4., pp. 56-76, 2008.

[15] J. Park, H.-R. Tyan, and C.-C. Jay Kuo. GA-Based Internet Traffic Classification Technique for QoS Provisioning. In Proceedings of the 2006 International Conference on Intelligent Information Hiding and Multimedia (IIH-MSP ’06). IEEE Computer Society, Washington, DC, USA, 251–254, 2006.

[16] D. Partridge, T.C. Bailey, R.M. Everson, J.E. Fieldsend, A. Hernandez, W.J. Krzanowski, and V. Schetinin. Classification with Confidence for Critical Systems. *Developments in Risk-based Approaches to Safety*, Vol. 6, pp. 231–239, 2007.

[17] K. Proedrou, I. Nouretdinov, V. Vovk, and A. Gammerman. Transductive confidence machines for pattern recognition. Technical report 01-02, Royal Holloway, University of London, 2001.

[18] C. Saunders, A. Gammerman, and V. Vovk. Transduction with confidence and credibility. Proceedings of the 16th International Joint Conference on Artificial Intelligence (IJCAI-1999), pp. 722–726, 1999.

[19] D. Schatzmann, W. Muehlbauer, T. Spyropoulos and X. Dimitropoulos, Digging into HTTPS: Flow-Based Classification of Webmail Traffic, ACM SIGCOMM IMC Internet Measurement Conference, pp. 322-327. Nov. 2010.

[20] Tóth, N. and Pataki, B. On Classification Confidence and Ranking Using Decision Trees. *In Proceedings of 11th International Conference on Intelligent Engineering Systems*, 2007.

[21] S. Vanderlooy, L. van der Maaten, and I. Sprinkhuizen-Kuyper. Off-line learning with transductive confidence machines: an empirical evaluation. Proceedings of the 5th International Conference on Machine Learning and Data Mining in Pattern Recognition, pages 310–323, 2007.

[22] Vapnik, V. N. (1998) *Statistical Learning Theory*, Wiley, New York.

[23] Vovk, V., Gammerman A. and Shafer, G. Algorithmic learning in a random world. *Springer*, 2005.

[24] N. Williams, S. Zander and G. Armitage. 2006. A preliminary performance comparison of five machine learning algorithms for practical IP traffic flow classification. SIGCOMM Comput. Commun. Rev. Vol. 36, No. 5, pp. 5–16, October 2006.

[25] S. Zander, T. Nguyen, and G. Armitage. Automated Traffic Classification and Application Identification using Machine Learning. In *The IEEE Conference on Local Computer Networks*, pages 250–257, 2005.

In: Advances in Communications, Volume 8
Editor: Anthony V. Stavros
ISBN: 978-1-61324-794-5

Chapter 11

ATTAINING MUTUAL AWARENESS OF THE AVAILABILITY STATUS

Agnieszka Matysiak Szóstek[1*], Panos Markopoulos[2] and Berry Eggen[2]
[1] Institute for Information Processing, Interactive Technologies Laboratory
Al. Niepodleglosci 188b, 00-608 Warsaw, Poland
[2] Industrial Design Department, Eindhoven University of Technology
Den Dolech 2, 5600MB Eindhoven, The Netherlands

Abstract

As there is no generally accepted convention regarding which behaviours for negotiating communication initiation are considered as socially acceptable, people tend to develop their own conventions that define the best practices for different situations and communities. Since system users rely on representations of their presence and actions to understand each others' degree of availability, finding ways to share relevant contextual information is necessary to help creating more refined and socially acceptable practices. However, current technologies tend to focus on provision of mechanisms that relate to functional rather than to social requirements and often disrupt the exchange of cues regarding the context communicators are in. For example, phones give the means to contact others anytime anywhere, yet do not provide any indications regarding whether the recipient is able to accept an incoming communication. As a result people are provided neither with sufficient information ab out the context, in which their communicators operate nor with means helping them to manage their communications in a way that is efficient and yet complies with social conventions similar to those used in face-to-face encounters. The goal of the study presented in this article is to investigate ways to ascertain mutual awareness about the recipient's availability status in Instant Messaging applications. We argue that availability indication alone is insufficient to leverage social behaviours in mediated communication and that it is crucial to introduce mechanisms stimulating mutual awareness regarding the communicative needs of communicators. To reach this goal a prototype named Do$^{\text{NT}}$Bother was implemented, which was evaluated in a web design company for a period of three weeks. The contributions of this research include the quantitative and qualitative measurements assessing the proposed solutions and a set of implications

*E-mail addresses: aga.szostek@opi.org.pl, {P.Markopoulos, N.H.Eggen}@tue.nl

that promise to inform the design of future mechanisms supporting the attainment of mutual awareness in Instant Messaging applications.

1. Introduction

We live in times, in which both professional as well as social relationships are created and maintained not only in a physical but also in the digital world. However, if one takes a closer look at many communications conducted through mediated tools, it appears that people are not able to act in social ways with similar grace and cohesion as during face-to-face encounters. Messages tend to arrive at the wrong moments, are often imprecisely formulated, uneasy to act upon and tend to cause information overload. Yet, does it mean that we, who so elegantly interact with each other in the physical world, suddenly turn selfish and insensitive when moving to the digital domain? Or maybe the fault is of technology that insufficiently supports social behaviours?

In working environments people use a variety of communication means to support information sharing and contact with others. A communication channel can be defined as a line of communication that is used to transmit information from a sender to a recipient and it can be rooted in either the physical or the digital domain. The channel selected to communicate plays an integral role in the context and the process of a communicative exchange [14, 22, 20, 8]. Basically, any decision behind channel selection can be seen as an optimization process: people consider which channel helps them best express their intentions and needs [31]. A strong correlation is shown between, for example, the medium selection and the complexity of a desired interaction [1]. Richer media such as face-to-face or phone are more likely to be chosen to deal with potentially complex and ambiguous tasks such as conflict solving or decision taking. Less rich media, such as email or Instant Messaging tools, are more likely to be employed for resolving simple tasks like scheduling or confirming earlier agreements. The choice of the communication channel can be further influenced by the initiator's knowledge of the recipient's context as well as the need for either synchronous or asynchronous communication (which, for example, could relate to the immediacy of the communication subject)[14].

Richer and more synchronous channels better embed social rituals [12] that determine ways interaction unfolds. However, they also impose higher asymmetry in control over the communication between communicators that *'arises because while initiators benefit from rapid feedback, the recipients are forced to respond to the initiator agenda'* [23, 19]. Such an asymmetry can be further gauged by the social and professional relationship between the actors [21, 25] or aspects such as, for example, the recipient's own time-pressure [17, 6, 13]. Less rich channels tend to offer the recipient a possibility to asynchronize communication which, as a consequence, brings them more control over its pace. However, lack of richness can also to lead to misunderstandings regarding each other's communicative needs. As the recipient is largely unaware of the context within which communication was initiated, he or she is more likely to miss signs indicating, for example, relative urgency or importance of this communication. Moreover, in the less rich media, the initiator lacks sufficient indications as to what is happening with the communicative exchange he or she initiated, which may lead to an increase of social tension between the actors and which, as a consequence, could have a negative impact on their future interactions. Based on these observations, we define the goal of this research as follows:

In the study described in this article we aimed to investigate ways to ascertain mutual awareness about the recipient's availability status in Instant Messaging applications in line with the guidelines posed by the Social Translucence framework. Based on the previous results described in [32], we argue that availability indication alone is insufficient to leverage social behaviours in mediated communication and that it is crucial to introduce mechanisms stimulating mutual awareness regarding the communicative needs of communicators. To reach this goal a prototype named Do^{NT}Bother was implemented, which was evaluated in a web design company for a period of three weeks. The contributions of this research include the quantitative and qualitative measurements assessing the proposed solutions and a set of implications that promise to inform the design of future mechanisms supporting the attainment of mutual awareness in Instant Messaging applications.

2. Related Work

Instant Messaging systems are near-synchronous communication tools that facilitate communication between a person and their 'buddy list' by supporting an exchange of short textual messages (sometimes supported by a video channel [27]). The near-synchronous nature of the tool allows communications to be paced according to the preferences of both communicators. The great success of instant messaging can be attributed to its flexible nature [18] and a relatively low cost of interruptedness [11]. However, as the use of instant messaging is growing, particularly at work, the insufficient support for managing people's availability for communication tends to lead to communication breakdowns, which, in turn, can have negative effects on the social relationships between the system users [4, 29]. The aforementioned problem is not new and a vast body of research was conducted on this subject [24, 3, 4, 15, 2, 34].

Voida et al [34] observed that, in Instant Messaging systems, while it might be convenient for the initiator to start a conversation at a particular moment, it may be undesirable for the recipient to engage in that conversation at that moment. The recipients must then face a trade-off between continuing their current task or engaging in communication. Nardi et al [24] saw that information exchange in Instant Messaging systems can be successful only through subtle negotiations of availability as a way to establish connection by inhabiting and maintaining a shared communication space. The process of negotiating availability binds people tightly together for a specific interaction as they establish a particular *attentional contract* and this is likely to have consequences for their future communications.

In the context of mediated communication, the ability to provide awareness regarding availability status should be seen as one of the most important features of Instant Messaging clients. They typically indicate whether a user is online and whether he or she is currently active or idle by measuring keyboard activity. Most Instant Messaging clients also allow users to enter short status messages that remain visible in the 'buddy list' view until changed or deleted. In [32], it was showen that contextualized yet abstract availability indication forms a base for correct understanding of each other's communicative state. It allows initiators not only to see that their buddies are unavailable but also to understand why they are unavailable. The same study also showed that availability indication alone is insufficient to leverage socially salient behaviours of initiators.

So far, some awareness mechanisms were designed enabling initiators to 'grab' recipient's attention through the use of various audio-visual alerts and alarms. These alerts, however, miss out on indicating the nature of a communicative attempt. Hsieh et al [15] and Tang et al [33] showed that tagging of instant messages might be a valuable way for the initiator to indicate the importance of an incoming communication. However, little has been done to address the asymmetry in control over communicative exchange [23]. This study focused on investigating attainment of mutual awareness regarding the communicative state of the recipient by attempting to answer the following research question:

> How to attain mutual awareness regarding the availability of communicators in mediated communication as means to form a foundation for accountability based on the mutual knowledge of each other's actions?

To answer this question, we investigated the extent to which people were willing to comply with the availability status of others when provided with that status during conversation initiation. I also looked into the possible effect of different status indications on people's decision to initiate communication. Moreover, we were interested to see whether mutual awareness of the recipient's status enabled participants to discard poorly timed communications. Finally, we wanted to elicit design implications regarding the attainment of mutual awareness that could help to design future Instant Messaging systems.

3. Design

To address our research question, a prototype of an Instant Messaging application was designed and deployed. The Do^{NT}Bother system was implemented in Java and based on an open source Jabber client: JBother [35]. JBother was chosen as it allowed for transport registration, so that participants could integrate commercial Instant Messaging clients (such as MSN or GTalk) with the prototype and receive all messages through one unified application. In this way I wanted to lower the acceptance threshold for Do^{NT}Bother and assure that participants had the possibility to integrate all their contacts and to execute all their communications through Do^{NT}Bother rather than using multiple separate Instant Messaging applications.

Firstly, we were interested to see whether having multiple options to describe one's status could render some additional advantages to the way the status is presented and perceived by others. The users were offered to choose among the following means to present their availability (see: Figure 1):

- The study described in [32] showed that an abstract visualization of availability was perceived as a sufficient and privacy-respecting means to describe people's communicative state. Based on that observation, *availability levels* graphically representing an overall state of participants' availability on a 5-point scale ranging from available to unavailable and represented by colours spanning from green to red were proposed. Participants could manually choose a status representation from the given range that best presented their availability state.

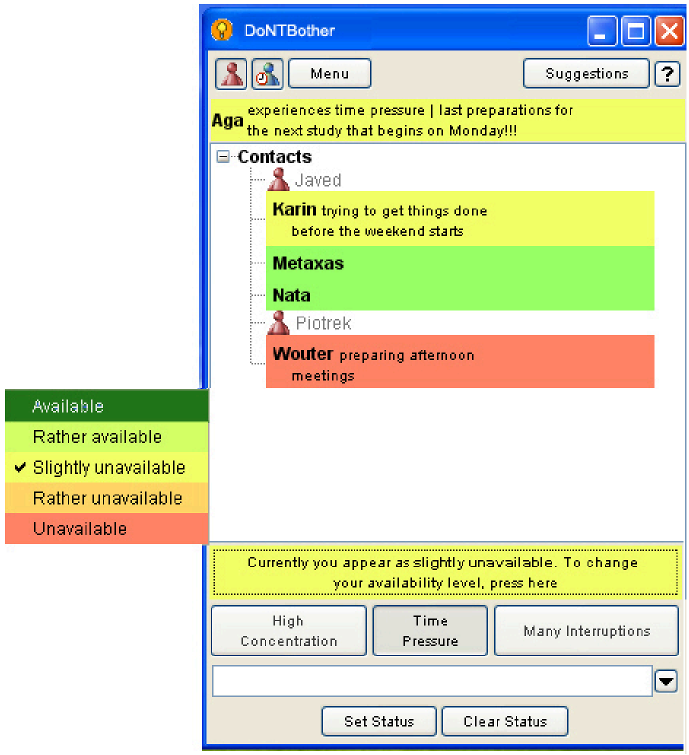


Figure 1. The availability status representations offered by Do$^{\mathrm{NT}}$Bother : an expandable menu for indicating the availability level (on the left), three status buttons and a text box to enter a status message with a status repository (on the bottom of the buddy list). Both the status generated by the buttons and the status manually entered are visible to others next to the name of the user on the top of the buddy list .

- The study described in [32] also showed that a contextualized explanation of one's abstract availability state helps initiators to better judge the meaning of that state. The three descriptors chosen to describe participants' status in the study: 'concentration', 'time pressure' and 'many interruptions', proved to suitably explain their communicative state. Following these results, three manual *status buttons* ('Concentration', 'Time Pressure' and 'Many Interruptions') were designed, each generating a predefined message ('Ann is concentrated', 'Ann experiences high time pressure' and 'Ann experienced many interruptions'). Participants could use these predefined status messages to add context to their status representation by pressing one or more of the buttons appearing at the bottom of the buddy list.

- Erickson and Kellogg [9] argued that text is a powerful means for conveying social information as it allows to specify one's state. Based on that argument we decided to once more add a *status message*: a text box, in which a personalized explanation of one's status could be entered. In Do$^{\mathrm{NT}}$Bother , any status message could also be stored in a *status repository*: a drop-down list accessible after clicking on the button on the right side of the text box. Participants could enter a relevant status description through a status text box appearing at the very bottom of the buddy list view. They could also store selected messages to be later reused in the repository. Any message could be deleted from the repository by clicking the 'Delete' button appearing next to it.

The main difference between Do^{NT}Bother and other Instant Messaging systems was that, besides showing status information in the buddy list, Do ^{NT}Bother also showed that information in each newly opened chat box (that was visible both on the initiator and, after sending also on the recipient's side) in two ways (see: Figure 2):

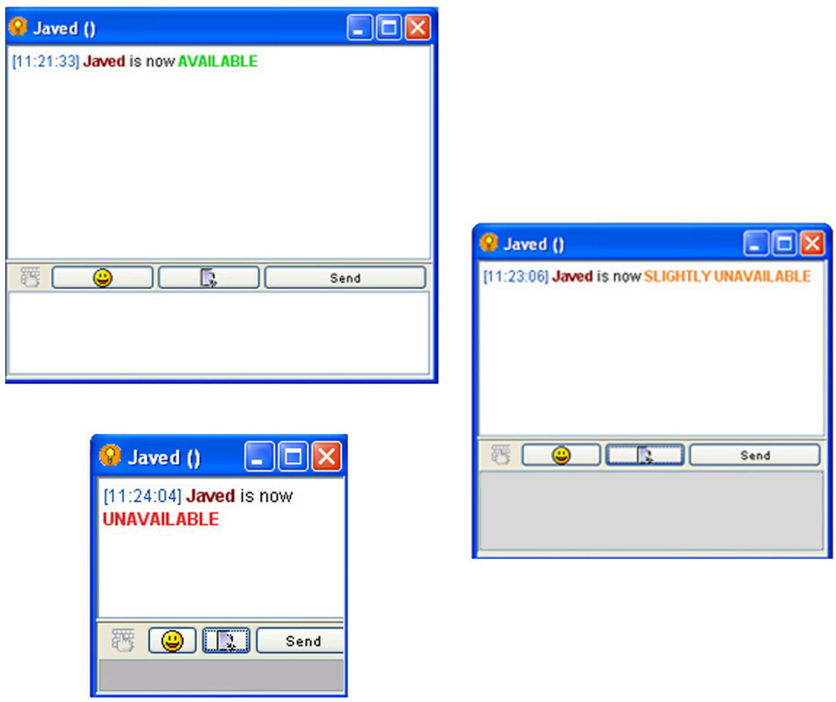


Figure 2. Displaying recipient's status prior to communication initiation: the representation of three different textual and graphical status updates in the chat box.

- *A textual status update*: the recipient's current status, including the availability level and also the status description if it was entered by the recipient, was shown as a line of text appearing on the top of the chat box of a newly initiated Instant Messaging conversation, so that both communicators could see it before engaging in that conversation.
- *A graphical status update*: any newly opened chat box changed its size depending on the availability level indicated by the recipient, so that it opened in full size if one indicated full availability and obtained a gradually smaller size if the status was set to other levels. Also the entry space of the chat box gradually changed colour from white to grey depending on recipient's availability. Change in the physical parameters of the text entry box was intended to indicate to the initiator the possible cost of initiating

communication. The initiator was in no way prohibited from entering the text of any length. If necessary the chat box could be enlarged by dragging its right-bottom corner and the background colour could be changed by using the background-colour palette. As soon as the reply to a message was received (meaning that the recipient showed interest in engaging in a conversation) the chat box returned to its original size and colour.

The rationale behind the textual status update was to ensure that (i) the initiator was aware of the recipient's status before initiating the conversation and (ii) that the recipient was notified that the initiator was aware of his or her status. The rationale behind the graphical status update was to visually represent to the initiator the potential cost the recipient needs to face when interrupted at an inappropriate moment. The additional effort required to resize the chat box and to change the background colour of the text box was aimed at helping the initiator realize these costs.

4. Participants

Erickson and Kellogg [9] said that systems leveraging social processes should be studied in real work settings rather than in laboratory conditions. According to them, only in the real situation a real social dependence between communicators is formed, which, in turn, could begin to serve as basis for creation of new social rules. Having already examined Social Translucence in a laboratory setting, it was necessary to also study communication initiation behaviours in the field. Therefore, the study was conducted in a web-design company employing 35 people divided over five departments who were located on two floors of the same building. Prior to the study, all employees participated in a time-management course [1], during which it became apparent that a substantial amount of time was lost due to untimely communications and that they all should improve their communication practices .

Out of all employees, a total of 10 persons (8 male, 2 female) agreed to participate in the study. Participants were members of three different departments and were distributed across four open-office spaces. Three participants worked on the same project, others were involved in different projects. They all collaborated with each other in the past and were also likely to work together in the future. Half of the participants worked part-time (4 days a week). The data obtained through the demographics questionnaire distributed to them prior to study initiation showed that all participants were acquainted with at least one Instant Messaging application (mostly MSN). Instant Messaging was used only for professional purposes by 5 out of 10 participants. The remaining participants used the Instant Messaging system equally often for both professional and social communications. They reported to use it to communicate with other co-workers (both collocated and distributed), clients, service providers, friends and family. A total of 8 participants reported to use it daily and 2 to use it a few times a week. Only 1 participant frequently updated her status, 6 did it sometimes and 3 persons had never set their status before. The remaining employees were not willing to switch to another Instant Messaging client and did not agree for their interactions to be logged.

[1]The timing of the time-management course and the study itself was coincidental.

4.1. Study setup

The system was presented to the participants during a one-hour presentation before the study initiation. All participants received a software package to install Do ^{NT}Bother and also to integrate it with their other Instant Messaging clients. They were provided with assistance during the installation phase and received a three-page instruction regarding the access to the features available in the system. The study lasted three weeks. After its completion two Focus Group sessions were conducted (with 5 participants present at each session).

The choice of a Focus Group was motivated in the following way. Focus Groups support the collection of a consistent set of qualitative data that is shared by the majority or by all participants. Each opinion is accordingly motivated and made precise through the discourse between participants who favour that opinion and those who do not [10]. Focus Groups also provide quality control of the collected data as participants check and verify each other's statements so that false or extreme ones are either corrected or rejected [26]. During each one-hour session participants were asked to describe the way they used Do^{NT}Bother with respect to: (i) indicating communicators' availability status, and (ii) leveraging mutual awareness of that status. Participants were also encouraged to compare the features of Do^{NT}Bother with other commercial Instant Messaging clients. The sessions were recorded and transcribed.

4.2. Data analysis

In this study, data was collected from two data sources. The summary of participants' interactions (see: Table 1 and Table 2) was intended to illustrate the relationship between their availability status and initiated communications. For that purpose, the following data was logged:
– user ID,
– time of the status update,
– use of the availability level indication,
– use of the status buttons,
– creation of a status message,
– use of the status repository,
– time of communication initiation,
– recipient's availability status at the moment of the initiation,
– whether the chat box was opened and also if communication was initiated or not,
– whether the initiation was responded to or not.

Data about setting the availability status aimed to help us understand how participants would construct their availability indications. Furthermore, we wanted to see the extent to which participants were willing to comply with the indicated availability status of their colleagues after being presented with status indication in the chat box. The data about whether the initiation was executed gave insights into the possible effect of the status indication in the chat box on the decision to initiate communication. The data about whether or not the initiation was responded to showed the extent to which the mutual awareness of the recipient's status provided participants with an opportunity to discard poorly timed communications.

The qualitative analysis of the Focus Group sessions aimed at providing insights into participants' opinions regarding using the representation of their availability status as means to negotiate communication in Instant Messaging systems and to see whether mechanisms supporting mutual awareness can be seen as the means to leverage socially salient behaviours. For that purpose, the transcripts from the two Focus Group sessions were, first, assessed by the author of this thesis. A total of 39 statements that did not provide any motivation for expressed opinions (e.g., 'I feel the same way' or 'It all is just personal') were removed. The remaining 106 statements were analyzed using the Direct Content Analysis [16] according to two predefined categories: (visibility and awareness. They were coded by two independent coders in two iterations (interrater agreement of 82% for the first iteration and of 91% for the second). Coding resulted in 47 statements categorized as expressing participants' observations regarding the differences in setting the availability status between Do^{NT}Bother and other Instant Messaging tools (specifically, MSN). 69 statements reflected participants' opinions regarding attaining mutual awareness of the recipient's availability status.

5. Results

In this section, the results from the two data sources are described. The summary of participants' interactions aims at illustrating the ways they initiated communications for different availability status indications. The analysis of the Focus Group sessions provides qualitative insights into participants' opinions regarding the advantages and disadvantages of the given prototype in comparison to other Instant Messaging applications, especially regarding ways of both presenting the availability status and invoking mutual awareness about that status.

5.1. Logs

It was observed that availability levels were the most frequently chosen representation of participants' status, who adapted it approximately 3 times a day (see: Figure 3). Level 1 (available) was the most often used to represent that status [2] (see: Table 1). This result confirms the findings of Szostek et al [32] that an abstract and graphical status representation is the preferred means to indicate people's communicative state.

The status buttons were the second most used means to represent participants' status ($n = 121$, approximately 1 update per person per day). In 99 cases more than one button was used to depict one's communicative state. The 'Concentration' button was used 116 times, 'Time Pressure' button 67 times, and the 'Many Interruptions' button was used 37 times. This result show that people are willing to provide an explanation of their status, especially if the effort to do so is minimal and the explanation itself does not threaten their privacy.

Status messages were the least often used means to indicate participants' status ($n = 60$). New messages were entered 21 times and existing messages were selected from

[2] The large number of Level 1 indications could be explained by the fact that Level 1 was a default status used whenever a person connected to the system.

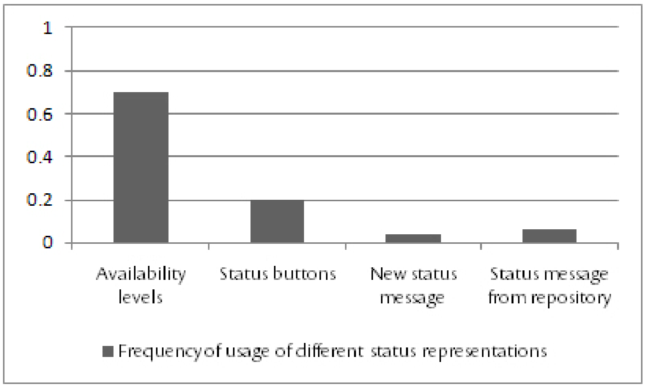


Figure 3. A graphical representation of the frequency of usage of different status indicators by the study participants.

the status repository 39 times. The messages fell into two distinct categories: they described a project or a task participants were working on like: 'administration' or 'Vodafone' ($n = 40$), or aimed at further defining their communicative state like: 'Do not disturb' or 'Drinking coffee' ($n = 20$). In line with the findings reported by Szostek et al [32], the results have shown that the higher the unavailability level was indicated the more additional indicators were used to strengthen the meaning of that status, for example, participants would indicate their unavailability by using the availability level indication, pressing the three status buttons and also adding a status message like, e.g., 'Don't disturb' at the same time.

A total of 173 communication initiations were recorded during the study (approximately 2 conversations per person a day). 73% of these initiations were commenced when participants' availability was indicated as available; 6% as rather available; 10% as slightly unavailable; 10% as rather unavailable and only 1% as unavailable (see: Tab. 2). This result indicates that, in general, communicators were willing to initiate communications at an appropriate moment based on the status information presented in the 'buddy list' view and also in the chat box.

It was also found that 30% of the recorded initiations were never responded to (n = 52). This is not a surprising result; similar findings were reported by Nardi et al [24] and also Avrahami and Hudson [5]. Interestingly, no relationship between the lack of response and the availability status could be found. The majority (n = 43) of unanswered communications were initiated when the recipient's status was indicated as available. From that observation a conclusion was drawn that sending a message when the recipient appeared available did not necessarily guarantee receiving a reply. There might have been two reasons for lack of a response: either the message itself did not require a reply (e.g., because it either got outdated or was resolved though a different channel) or that the status indication visible in the system was not up to date and the recipient was, in fact, not able to answer the message. Alternatively, the recipient might have not considere d an instant message as an initiation of synchronous communication or might have postponed a response based on reasons unrelated to his or her current availability status. A detailed analysis of the conversations is needed to support further analysis of these observations. Such an analysis

Table 1. Frequencies regarding the use of different availability indicators to represent the recipients' availability status.

	Availability levels	Status buttons	New status message	Status message from repository
Level 1: available	314	46	8	11
Level 2: rather available	24	10	5	1
Level 3: slightly unavailable	36	31	4	9
Level 4: rather unavailable	20	14	1	4
Level 5: unavailable	22	20	3	14
Total:	416	121	21	39

Table 2. Frequencies showing the relationship between the availability level presented by the recipients and number of communications started by initiators (i.e., sending of a message).

Availability level	Initiated communications
Level 1: available	126
Level 2: rather available	11
Level 3: slightly unavailable	17
Level 4: rather unavailable	17
Level 5: unavailable	2
Total	173

was, however, not possible in this study due to the a priori agreement with the company pertaining to the privacy of the logged data.

Finally, 25 events were recorded during which a chat box was opened but no communication was initiated. The majority of these events ($n = 17$) occurred when the recipient's availability status was other than available (for Level 2 - 6 cases, for Level 3 - 6 cases, for Level 4 - 2 cases and for Level 5 - 3 cases). This result hints at the positive impact of the status update in the chat box on increasing awareness regarding the recipient's status at the point of communication initiation. Participants seemed to withdraw from initiating communication after being presented with their colleague communicative state in the chat box. This observation seems to verify the design rationale behind the proposed mechanisms. A longer study , however, would be necessary to realize a true impact of that design on socially responsible behaviours of communication initiators.

5.2. Focus Groups

In the last part of the study, participants shared their opinions regarding setting their availability status and also the extent to which they were aware of their communicators' availability state at the point of communication initiation.

5.2.1. Presenting availability status

Participants appreciated Do^{NT}Bother for allowing them to present their status in three different ways. The threefold way of defining one's status was perceived as providing an acceptable description about one's communicative state and at the same time as a way to disclose limited information about oneself (e.g., by only indicating one's availability level without providing any explanation of that level) or to exaggerate one's status (e.g., by setting the availability level to red and also pressing all status buttons). Different features allowed participants to control the amount of information they wanted to share with their colleagues. In line with findings by Nardi et al [24], also in this study the participants reported to frequently use Do^{NT}Bother to check the status of a person whom they wanted to get in touch with through alternative communication channels (e.g., Face-to-Face or through the phone).

> 'Do^{NT}Bother , unlike MSN, allows you to quickly tell others how busy you are and also on different levels, so not just only "busy" or "not busy".'

> 'Success of Do^{NT}Bother comparing to MSN is the way it allows me to set my status. You want to give others a correct feeling of what's going on with you. In a face-to-face situation if someone comes and I'm busy I give him a dark look and he should turn around and go away. In Do^{NT}Bother I set a very strong status.'

Consistently with the findings described [32], participants most appreciated the possibility to indicate the different availability levels as means to express their communicative state. It allowed them to suggest to their colleagues whether it was a good or a bad moment to start communication but without having to be specific regarding the reasons behind the indicated state. They also repeatedly remarked that the graphical representation and colour coding of availability levels provided them with a quick overview about whom of their colleagues was busy and who was available. Participants thought that the meaning of green and red colours was exceptionally straightforward and easy to interpret: green meant fully available and red completely unavailable; the remaining levels were seen as allowing for initiating communication but the more unavailable one was indicated the more urgent or important the intended communication subject should have been.

> 'The red and green indications are clear and everything in between is negotiable. If someone has a red status then you should have a really good excuse to bother him.'

The status buttons were also perceived as a quick and effortless way to provide additional information about participants' working context, which was indicated by participants to strengthen the meaning of the availability level participants wanted to present. The message generated by the ‘Time Pressure’ button was perceived as the most meaningful information about one’s situation. Participants explained that the notion of time pressure was relatively straightforward to understand especially in a community, where most people were to some extent knowledgeable about the stage of projects and upcoming deadlines of their colleagues. The message generated by the ‘Concentration’ button was seen as more difficult to interpret and only appreciated among participants who were adequately aware of the progress in their colleagues work. The message generated by the ‘Many Interruptions’ button was considered as confusing and therefore the least frequently used. Some participants interpreted the message generated by that button as indicating that one has already experienced many interruptions and therefore did not want to be interrupted anymore. Others thought that it might be better to interrupt right away, because a person was already distracted by others.

> ‘It is good to have them (status buttons) because they require really low effort to say: ‘I am concentrated’ comparing to typing in a message.’

Status messages were seen as an informative means to further explain other status indications: namely the availability level and messages generated by the status buttons. From the perspective of a communication initiator participants frequently mentioned that the status message varied in its notion and had a more significant meaning if they were working on the same project as the colleague they wanted to contact. The message indicating that the recipient was generally unavailable but working on the same project as the initiator was interpreted by team members as ‘potentially more available for my team’, particularly if the issue at hand was related to the project they were busy with. In such a case, even a strong unavailability indication would be easier to break through especially when the communication subject was, indeed, related to the project they both were working on. In a situation when the message indicated that one was working on a different project, participants indicate d to be more conservative regarding initiating communication. Participants also appreciated the possibility to store status messages in the repository as it saved them the time to type them in over and over again. They typically kept messages indicating the name of the projects they worked on and also messages strengthening the notion of the unavailability status.

Similarly to the study of the LILSYS system [7] it was observed that the status indication would not always stop participants from contacting their colleagues in situations, in which they appeared unavailable. Participants based their decision of whether to interrupt on their knowledge regarding the current task of their co-worker. That knowledge was often supported by the status message stating the project a recipient was presently occupied with. Interestingly, it was also found that a decision whether or not to interrupt someone was based on how trustworthy the status appeared. If the status indication seemed plausible, participants were more likely to respect it. The plausibility of the message could be described as a subjective judgment of the communication initiator whether the status possibly is a true representation of recipient’s communicative state. Participants usually assessed the accuracy of the information entered in the system by, e.g., checking whether the status changed over time. They also evaluated recipient’s availability status based on their personal subjective knowledge of recipient’s whereabouts and workload. Whenever they had doubts regarding its reliability, they would easily discard it and initiate communication.

> 'I find it handy when people set the status using messages so that I know that she is working on project X and she is really unavailable. Then if you are in the same project you know how to behave. I know that she is available for me. So any feature that supports setting the status and also letting people know what they are working on now is great.'

> 'Some colleagues set their status to unavailable from 8 a.m. to 5 p.m. If I check it a few times, I think: "I don't believe that you had so many interruptions already at 8 a.m., you just pressed all button". And I would contact him anyway.'

Finally, participants found it sometimes problematic to keep their status up to date. They reported to remember to set their status in the morning and then would forget to update it during the day. They also remembered to indicate in the system whenever they were busy but would frequently overlook to change the status back to available once a particular task has ended.

> 'If you are busy it isn't a problem to remember to tell others that you are busy but when you are not busy anymore you never change it back.'

5.2.2. Invoking awareness of the recipient's status among communicators

Many participants considered the status update in the chat box as a successful reminder prompting them to either postpone communication or keep it short in situations when the recipient indicated limited availability. Participants felt that the status in the buddy list on its own was insufficient to assume that the communicator was aware of their communicative state. The textual status update was seen as a guarantee that one's status was seen and therefore one can use it as an excuse for delaying or deferring poorly timed communication. The graphical status indication was seen to counterbalance the possible negative impact that communication initiation could have on the recipient.

> 'If someone starts a conversation and gets my status, it is like a reminder: You started talking but I am under time pressure, so you should keep it short. And you can always ask this person to look at your status.'

> 'It was very handy that the chat box got smaller if someone was unavailable. Sometimes you don't really check if he is available, you just want to talk to him. And then you see the chat box getting smaller.'

Surprisingly, participants frequently reported misunderstandings regarding the meaning of sending a message at times when the recipient appeared unavailable. Often when participants sent a message to someone who displayed limited availability for communication they did not intend to compromise her need for solitude. Their goal was, instead, to indicate to the recipient that they would like to communicate with him or her in the near future as, for example, proposed by Romero [28]. At the same time, initiators wanted to leave the initiative to the recipient as when to react to that message. However, while messages received at moments when the recipient was available could be easily deferred for later, those received while one's status was set to unavailable, were often seen as urgent and participants felt inclined to at least read and often also act on them immediately.

> 'If I see that someone is busy, I don't expect an answer right away. I already contacted him and it was enough for me to let him know what my problem is and that I would like a response sometime.'

> 'When my status is set to available and something blinks then I don't care. But when my status is set to unavailable and there is something blinking I think: 'It must be really urgent'. I will answer both, of course, but my curiosity is higher when my status is unavailable and still someone tries to reach me.'

Finally, although they appreciated the status update in the chat box, participants felt that Do^{NT}Bother lacked support for communication breakdowns caused by a change in the communicators' situation that occurred in the middle of the conversation and was often caused by a trigger happening outside the application (e.g., receiving a phone call). Participants needed to be able to quickly put an ongoing IM communication on hold if another situation needed their attention but at the same time they wanted to quickly provide an acceptable excuse explaining why they went into an idle state.

> 'It is important to define who should come back to that conversation because it may end up in a situation that someone is waiting for you and you are waiting for him to continue.'

> 'How many times I wrote: 'Wait a minute, phone'. It is an interruption in my conversation with this person and I want to put him on hold but I also want to tell him that I will be back in few minutes. It is better if somebody knows what is going on. But it is also annoying to have to type in the exact same message over and over again.'

6. Discussion

This study investigated ways to achieve mutual awareness of the recipient's availability status in Instant Messaging applications. The results showed that providing the recipient's status in the chat box encouraged participants to respect the communicative state of their colleagues. In line with findings of Begole et al [7], also this study showed that users did not always refrain from initiating communications when their colleagues appeared to be unavailable. Therefore, presenting a textual status update in the chat box guaranteed that to the recipient that his or her communicative state was seen by the initiator. Attaining mutual awareness of the recipient's status provided participants with the opportunity to use it as an excuse to avoid unwanted communication. The graphical representation was seen more as a 'defense mechanism' before the communication was actually initiated as it made the task of typing in the initial message more difficult for the initiator.

The study results further revealed that leveraging mutual awareness of each other's communicative state is not only important during communication initiation but during the entire communicative process. Participants, for example, indicated that they missed a lightweight way to indicate to each other communication breakdowns, especially if the breakdown is caused by external factors (e.g., a visitor or a phone-call). A current work-around was to quickly send an explanatory message like: ‘I am on the phone’ and then leave the conversation. However, such a method was perceived neither efficient (participants kept on typing similar messages every so often) nor lightweight (typing in a message required time and was often inconvenient). Therefore, participants indicated that they would have liked to have lightweight means to indicate the reasons behind pausing communication. One possible solution would be to support users in creation of a list of ‘delay’ messages; a sort of an ‘excuse r epository’, from which typical responses could be easily and efficiently retrieved. Such messages should, however, be personalized rather than system-generated. For example, Romero [30] proposed the *drag&drop* mechanism similar to the status repository concept proposed in the current study. A recipient could create a set of standard reactions and use them to respond to a instant message by a simple drag-and-drop action. However, the evaluation showed that the system users were not likely to use that mechanism. Therefore, a further design effort is necessary to propose a mechanism which is considered as sufficiently lightweight and also purposeful by the users of Instant Messaging applications.

Another way to ascertain mutual awareness communication systems should support communicators in managing their expectations as when a response to a message can be expected. On the one hand, it is desirable that the initiator has the possibility to indicate how urgently the recipient should react to the particular message. On the other hand, recipients should have the possibility to indicate to the initiator when it is feasible to expect a response. Such systems could also support people in indicating who of the communicators should be responsible for communication re-initiation and when the right moment for that re-initiation is. For example, the system could indicate to the initiator that recipient’s status changed to available. An example of such a mechanism was proposed by Romero [30, 29]. She designed a tangible device enabling communicators to indicate a reception of a instant message and also to notify the initiator when the response might be expected. It would be also interesting to expand the proposed functionality by enabling initiators to indicate whether a response is at all expected or not. Such a solution would, however, need to be critically assessed as whether it does not introduce another level of information processing rather than minimize the current level of information overload.

We can see that while the general idea of adding mechanisms to attain mutual awareness may seem to burden the Instant Messaging users with additional effort when initiating communication, it should not be dismissed as these mechanisms bring benefits to the ways mediated communication unfolds. Such mechanisms, when well integrated in systems supporting mediated communication, are likely to provide social benefits to communicators like, for example, helping to formulate rules regarding when communication is desirable and when it is unwanted.

7. Conclusions

The study described in this article empirically investigated ways to attain mutual awareness of recipient’s availability for communication in Instant Messaging systems. For that purpose a prototype of an Instant Messaging system named Do^{NT}Bother was tested. The analysis of the results showed that providing status indication during communication initiation influenced participants to show more respect towards the communicative state of their colleagues comparing to similar situations when the availability status was visible only in the 'buddy list' view. Moreover, this study showed that mutual awareness should be maintained not only in the phase of communication initiation but also during the entire communication. These findings confirm the importance of supporting reciprocal awareness that was already discussed by Erickson and Kellogg [9]. By ascertaining that all systems users know what information is shared among them people are likely to become more sensitive to the communicative needs of others and act in a way that is socially salient.

References

[1] T.J. Allen and O. Hauptman. The influence of communication technologies on organizational structure: A conceptual model for future research. *Information Technology and the Corporation of the 1990s: Research Studies*, pages 475–483, 1994.

[2] D. Avrahami, S.R. Fussell, and S.E. Hudson. IM waiting: timing and responsiveness in semi-synchronous communication. In *CSCW*, pages 285–294. ACM Press, 2008.

[3] D. Avrahami and S. E. Hudson. Communication characteristics of instant messaging: effects and predictions of interpersonal relationships. In *CSCW*, pages 505–514. ACM Press, 2006.

[4] D. Avrahami and S. E. Hudson. Responsiveness in instant messaging: predictive models supporting inter-personal communication. In *CHI*, pages 731–740, 2006.

[5] D. Avrahami and S.E. Hudson. QnA: augmenting an instant messaging client to balance user responsiveness and performance. In *CSCW*, pages 515–518. ACM Press, 2004.

[6] B. P. Bailey, J. A. Konstan, and J. V. Carlis. On the need for attention-aware systems: Effects of interruption on task performance, error rate, and affective state. *Journal of Computers in Human Behavior*, 1(special issue on attention aware systems), 2005.

[7] J. Begole, N. E. Matsakis, and J. C. Tang. Lilsys: Sensing unavailability. In ACM Press, editor, *CSCW*, volume 1, pages 511–514, 2004.

[8] B. Daly. The influence of face-to-face versus computer-mediated communication channels on collective induction. *Accounting, Management, and Information Technologies*, 3(1):1–22, 1993.

[9] T. Erickson and W. A. Kellogg. Social translucence: An approach to designing systems that support social processes. *ACM Transactions on Computer-Human Interaction*, 7(1):59–83, 2000.

[10] U. Flick. *An introduction to qualitative research*. Sage Pubns Ltd, 2009.

[11] R.K. Garrett and J.N. Danziger. IM=Interruption Management? Instant Messaging and disruption in the workplace. *Journal of Computer-Mediated Communication*, 13(1):23–42, 2007.

[12] E. Goffman. *Interaction Ritual: Essays in Face-to-face Behavior*. Random House Inc, 1967.

[13] V. M. Gonzalez and G. Mark. Managing currents of work: Multi-tasking among multiple collaborations. In *CSCW*, volume 1. Springer, 2005.

[14] C. Heyer and M. Brereton. Socialising across channels: group multichannel communication. In *CHISIG*, pages 421–424, 2006.

[15] G. Hsieh, S.E. Hudson, and R. Kraut. Using tags to assist near-synchronous communication. In *CHI*. ACM Press, 2008.

[16] H. F. Hsieh and S. E. Shannon. Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9):1277, 2005.

[17] J.M. Hudson, J. Christensen, and T. Kellogg, W.A. Erickson. I'd be overwhelmed, but it's just one more thing to do. In *CHI*, pages 97–104. ACM Press, 2002.

[18] E. Isaacs, A. Walendowski, S. Whittaker, D.J. Schiano, and C. Kamm. The character, functions, and styles of instant messaging in the workplace. In *CSCW*, pages 11–22. ACM Press, 2002.

[19] M. Kakihara, C. Sorensen, and M. Wiberg. Fluid interaction in mobile work practices, the interaction society: Practice, theories, and supportive technologies. In *The Interaction Society: Practice, Theories and Supportive Technologies*. IDEA-group Inc., 2004.

[20] N. S. A. Karim and R. Heckman. Group communication media choice and the use of information and communication technology to support learning: a case study. *Campus-Wide Information Systems*, 22(1):28–42, 2005.

[21] A. Kendon. *Conducting interaction : patterns of behavior in focused encounters*. Cambridge University Press, 1990.

[22] R. E. Kraut, R. S. Fish, R. W. Root, and B. L. Chalfonte. Informal communication in organizations: Form, function, and technology. *Human reactions to technology: Claremont symposium on applied social psychology*, 1990.

[23] B. Nardi and S. Whittaker. The place of face-to-face communication in distributed work. In P. Hinds and S. Kiesler, editors, *Distributed Work*. MIT Press, p. hinds and s. kiesler edition, 2001.

[24] B. A. Nardi, S. Whittaker, and E. Bradner. Interaction and outeraction: instant messaging in action. In ACM Press, editor, *CSCW*, pages 79–88, 2000.

[25] S. Patil and J. Lai. Who gets to know what when: configuring privacy permissions in an awareness application. In *CHI*, volume 1, pages 101–111. ACM Press, 2005.

[26] M.Q. Patton and M.Q. Patton. Qualitative evaluation and research methods. 1990.

[27] N. Romero, G. McEwan, and S. Greenberg. A field study of community bar: (mis)-matches between theory and practice. In *GROUP*. ACM Press, 2007.

[28] N.A. Romero. *Coordination of Interpersonal Privacy in Mediated Communication* . Technische Universitet Eindhoven, 2008.

[29] N.A. Romero, L. Boer, and P. Markopoulos. Interactive and lightweight mechanisms to coordinate interpersonal privacy in mediated communication. In *INTERACT*. Springer Verlag, 2009.

[30] N.A. Romero and P. Markopoulos. Grounding interpersonal privacy in mediated settings. In *GROUP*, pages 263–272. ACM Press, 2009.

[31] N.M. Su and G. Mark. Communication chains and multitasking. In *CHI*. ACM Press, 2008.

[32] A. Szostek Matysiak, E. Karapanos, B. Eggen, and M. Holenderski. Understanding the implications of social translucence for systems supporting communication at work. In *CSCW*. ACM Press, 2008.

[33] J.C. Tang, E. Wilcox, J.A. Cerruti, H. Badenes, S. Nusser, and J. Schoudt. Tag-it, snag-it, or bag-it: combining tags, threads, and folders in e-mail. In *CHI*. ACM Press, 2008.

[34] A. Voida, W.C. Newstetter, and E.D. Mynatt. When conventions collide: the tensions of instant messaging attributed. In *CHI*, pages 187–194. ACM Press, 2002.

[35] www.jbother.org.

In: Advances in Communications, Volume 8
Editor: Anthony V. Stavros
ISBN 978-1-61324-794-5

Chapter 12

SPACE-TIME CODED COOPERATIVE WIRELESS COMMUNICATION

Hamed Rasouli and Alagan Anpalagan
Department of Electrical and Computer Engineering
Ryerson University, Toronto, Ontario, Canada

Abstract

Wireless relaying can be viewed as a new dimension in resource exploitation in radio communication systems. Relaying is a technique that uses certain terminals, either fixed or mobile, which have good communication links with the destination terminal to act as relay nodes for those who do not have. By applying this technique in a cellular system for example, either the throughput or network coverage can be increased. Multiple antenna schemes provide enormous performance gain in wireless systems by increasing the capacity of the network. Using space-time codes in the relaying scheme provides the diversity benefits of multiple antenna techniques. It has been suggested to use space-time block coding techniques that were proposed for multiple-antenna systems in a wireless relaying system. The idea of making the relay nodes act like multiple antennas is an interesting idea and has been studied extensively in the literature. Space-time coded relaying protocols can provide a higher throughput gain compared to the repetition-based relaying protocols. In this chapter, different relaying techniques are categorized and reviewed. Amplify-and-forward and decode-and-forward are the two main categories based on the relay functioning. Relaying can also be categorized in terms of wireless access protocol that is used in accessing the radio channel. Time-division, frequency-division are the legacy access protocols and space-division is a relative new one. The focus of this chapter is mainly on space-time coded cooperation applied to wireless relays. Following a detailed discussion on distributed space-time coded relaying, some selected results are provided to show the advantages of space-time coded cooperation through wireless relays.

1. Introduction

It is well known that multiple antenna scheme provides enormous performance gain in wireless systems [1]. In a point-to-point multi-input-multi-output (MIMO) channel with n transmit and n receive antenna and i.i.d. flat fading, where the receiver knows the channel state, the capacity of the system at high SNR's is approximately $n\log(\gamma)$ [2, 3], where γ is the SNR between source and destination. Therefore, we say that the capacity is equal to the sum of capacities of n parallel single antenna channels and provides n degrees of freedom in MIMO channel.

Alamouti proposed a simple transmit diversity scheme [4] which was further extended by Tarokh *et.al*. [5, 6] to a general class of block codes named "space-time block code (STBC)". The simplicity of STBC attracted many researchers to invest on the idea of space-time coding in different areas of wireless communications. One of the areas which has gained interest recently is cooperative wireless networks [7]. The conventional antenna diversity schemes [5] have some limitations. The spacing between antenna elements needs to be larger than half a wavelength to avoid fading correlation and antenna coupling. In many practical wireless applications, wireless devices are so miniaturized that such a spacing between multiple antenna can not be employed. One of the ways to achieve spatial diversity is through the relay terminals, where each relay retransmits the information received from a remote source to the destination [8]. Cooperative relay terminal is an important physical layer concept that can help the wireless network to achieve higher throughput, lower energy consumption and a longer lifetime of wireless routes.

A cooperative relaying scheme has been proposed and studied in [9] and [10]. In this scheme, wireless relays do not exchange information with one another; but, rearrange the received signal according to the Hurwitz-Radon matrices which are the mathematical base of space-time block codes [6]. Hurwitz-Radon matrices can be used to construct space-time codes for parallel wireless relays which are located between source and destination. Based on Alamouti space-time codes, another distributed space-time coding technique was proposed for a system with one or two assisting relays [11, 12]. Approximate and asymptotical formulas for BER of distributed space time block coding (DSTBC) system with BPSK modulation are obtained in [12] for cases with one and two relays. It is shown that this design achieves a diversity gain over a point-to-point system with the same bandwidth. This method requires two independent frequency channels to achieve diversity gains over the non-cooperative case. In some cases, where two frequency channels are correlated, the diversity gain would be less than what is proposed.

In this chapter, a brief review of space-time block codes is presented. Then, application of Alamouti codes in a cooperative scheme is reviewed. A new concept in cooperative wireless communications named "on-channel relaying" is introduced and some results are presented next. Finally, the chapter ends with the possible future work in the related areas of cooperative wireless communications.

2. An Overview of Space-Time Block Codes

Due to severe time-variant attenuation in wireless environment, it is extremely difficult for the receiver to determine the transmitted signal. Therefore, some form of diversity is needed to compensate for this weak point. Time or frequency diversity methods overuse the bandwidth - the most precious resource in the wireless network. Receiver diversity which benefits from multiple-antenna schemes at the receiver is an efficient method to achieve diversity gain in the network. In receiver diversity, diversity is achieved by maximal-ratio combining (MRC) of different version of the received signal by the antenna elements at the receiver. Alamouti proposed his technique to use multiple-antenna capability of the receivers in transmit mode to form a transmit diversity [4]. Alamouti technique is only applicable for a transmitter with two antenna elements. The generalization of Alamouti method was done by Tarokh *et.al.* [5, 6] in which transmit diversity can be achieved for large number of antenna elements. In the following, we will review the space-time block coding which was proposed by Alamouti and will end the section by an overview of higher order space-time block codes.

2.1. Alamouti Space-Time Codes

It is assumed that the transmitter has two antennas where two signals are transmitted from them simultaneously as shown in Figure 1. Transmission and reception of signal are performed for a block of two symbols. Table 1 shows the encoding of a signal transmitted by an antenna with two elements in two consecutive time intervals; however, signal could be transmitted in two adjacent (correlated) carriers.

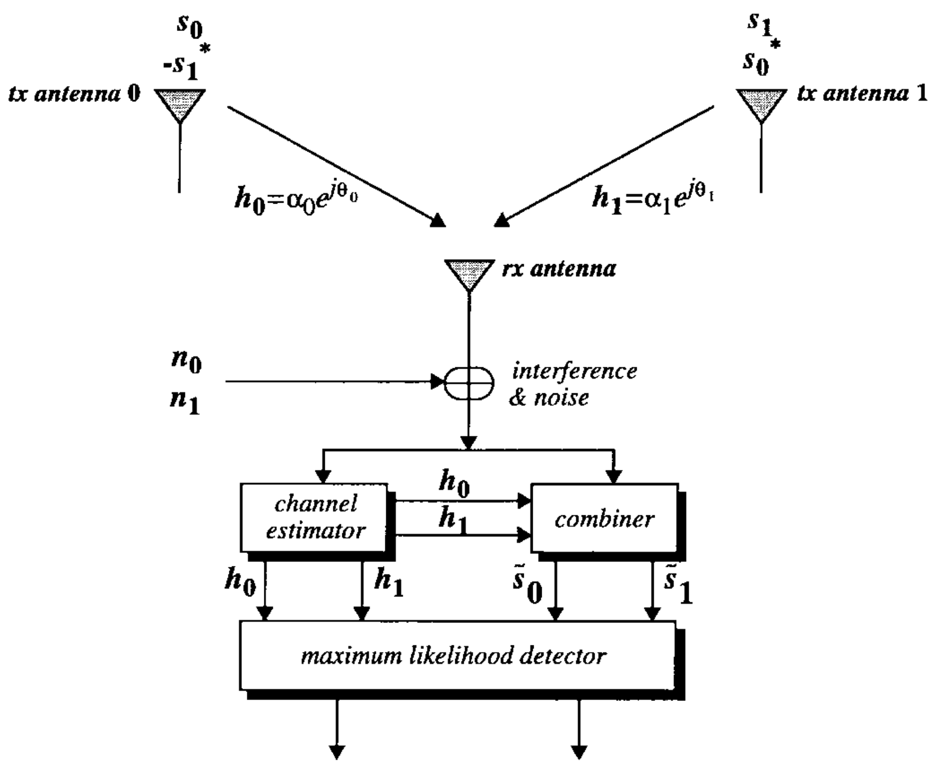


Figure 1. Distributed STBC system model with one assisting relay [12].

It is assumed that the channel gain remains constant during the transmission of one block of signal and is modeled by a normal complex multiplicative process h_0 for transmit

Table 1. Transmission sequence for two-branch diversity

	antenna 0	antenna 1
time t	s_0	s_1
time t+T	$-s_1^*$	s_0^*

antenna 0 and h_1 for transmit antenna 1 as:

$$\begin{aligned} h_0 &= \alpha_0 e^{j\theta_0}, \\ h_1 &= \alpha_1 e^{j\theta_1}, \end{aligned} \tag{1}$$

where α_i and θ_i are the amplitude and phase of h_i, respectively, $i = 1,2$.

The received signal at destination can be expressed as:

$$\begin{aligned} r_0 &= h_0 s_0 + h_1 s_1 + n_0, \\ r_1 &= -h_0 s_1^* + h_1 s_0^* + n_1, \end{aligned} \tag{2}$$

where r_0 and r_1 are the received signal at time t and $t+T$ and n_0 and n_1 are complex Gaussian random variables representing the receiver noise. Assuming that the receiver knows the channel gains, it uses a maximal ratio combiner to form the decision variable as follows:

$$\begin{aligned} \tilde{s_0} &= h_0^* r_0 + h_1 r_1^* = (\alpha_0^2 + \alpha_1^2) s_0 + h_0^* n_0 + h_1 n_1^*, \\ \tilde{s_1} &= -h_1^* r_0 - h_0 r_1^* = (\alpha_0^2 + \alpha_1^2) s_1 - h_0 n_1^* + h_1^* n_0 \end{aligned} \tag{3}$$

These combined signals are sent to a maximum likelihood detector where, for each of the transmitted signals s_0 and s_1, it makes its decision based on the minimum Euclidean distance.

As it is seen in (3), the combined signal in the receiver has a diversity gain of two. It is shown in [4] that by using two transmit and m receive antenna, a diversity gain of $2m$ can be achieved. It should be noted that transmit diversity scheme needs twice as much as power compared to receiver diversity. Also, the spacing between antenna elements should be large enough to guarantee diversity gain.

2.2. Improved Alamouti Space-Time Codes

Alamouti code discussed in the previous section is not suitable for the cooperative relaying scheme. Another version of Alamouti space-time coding which was chosen for universal mobile telecommunication system (UMTS) standard [13] is overviewed here. In this scheme, one antenna sends the signal stream without performing any modification (conjugating) and the other antenna transmits the modified symbols. Table 2 shows the encoding of signal transmitted by two antenna in two orthogonal channels (time or frequency).

The received signal at destination can be expressed as:

$$\begin{aligned} r_0 &= h_0 s_0 - h_1 s_1^* + n_0, \\ r_1 &= h_0 s_1 + h_1 s_0^* + n_1, \end{aligned} \tag{4}$$

Table 2. Improved transmission sequence for two-branch diversity

	antenna 0	antenna 1
time t	s_0	$-s_1^*$
time t+T	s_1	s_0^*

where r_0 and r_1 are the received signal at time t and $t+T$ and n_0 and n_1 are the added noise at the receiver noise. Assuming that the receiver knows the channel gain, it uses a maximal ratio combiner to form the decision variable as follows:

$$\begin{aligned}\tilde{s_0} &= h_0^* r_0 - h_1 r_1^* = (\alpha_0^2 + \alpha_1^2) s_0 + h_0^* n_0 + h_1 n_1^*,\\ \tilde{s_1} &= -h_1 r_0^* + h_0^* r_1 = (\alpha_0^2 + \alpha_1^2) s_1 - h_1 n_0^* + h_0^* n_1.\end{aligned} \tag{5}$$

The combined signals are sent to a maximum likelihood detector where for each of the signals s_0 and s_1, it makes its decision based on the minimum Eudlidean distance.

Higher order of STBC are discussed in [5, 6]. It was shown that full-rate codes exist for real constellations and half-rate codes exist for complex constellations. Alamouti coding is a special case of complex STBC that achieves full-rate transmission. There are also 3/4-rate codes for special case of multiple antenna of three and four elements.

3. Distributed Space-Time Coding for Relay

Using nearby mobile users for total throughput enhancement was proposed in [14]. The main idea is that after selecting a partner from the in-cell mobile users, each user detects a faded and noisy version of the partner's transmitted signal and combines this information with its own data to construct its transmitted signal. It is shown that in a flat fading environment the code division multiple access (CDMA) cooperative system of [15, 16] achieves a higher throughput than the one-hop CDMA system.

Distributed space-time coding was first introduced by Laneman [7] where the author divided cooperative transmission in two phases; noncooperative phase (phase I) and cooperative phase (phase II). In phase I, the source broadcasts the information signal to the destination where relays are also listening to the wireless medium and receiving the signal. In phase II, relays cooperate with the source based on the pre-defined protocol which can be repetition-based or space-time-coded. In repetition-based cooperation, relays either amplify and forward, or decode and repeat the source signal in orthogonal channels in the second phase; while in space-time coded cooperation, relays utilize a suitable space-time code and therefore can simultaneously transmit on the same channel. Outage capacity analysis in [7] shows that for high-SNR case, full spatial diversity can be achieved in the number of cooperating relays for both cases.

The proposed cooperative protocol by Laneman has also been investigated by Anghel [12] for special case of Alamouti space-time coding. In this section, we review the proposed method by Anghel which provides diversity gain using improved Alamouti space-time codes with one and two assisting relays. Unlike the approach in [7] that focuses on asymptotic bounds of outage capacity, the average BER is used in [12] to analyze the distributed space-time block coding (DSTBC) system.

3.1. Distributed Space-Time Coding with One Relay

A wireless communication system with three terminals of source, relay and destination is considered as shown in Figure 2. It is assumed that every terminal has one antenna and relay terminal is fully synchronized with the source. Relay is non-regenerative and can not simultaneously transmit and receive on the same channel (time or frequency). Therefore, two independent frequency bands of A and B are allocated for reception and transmission of signals at relay, respectively.

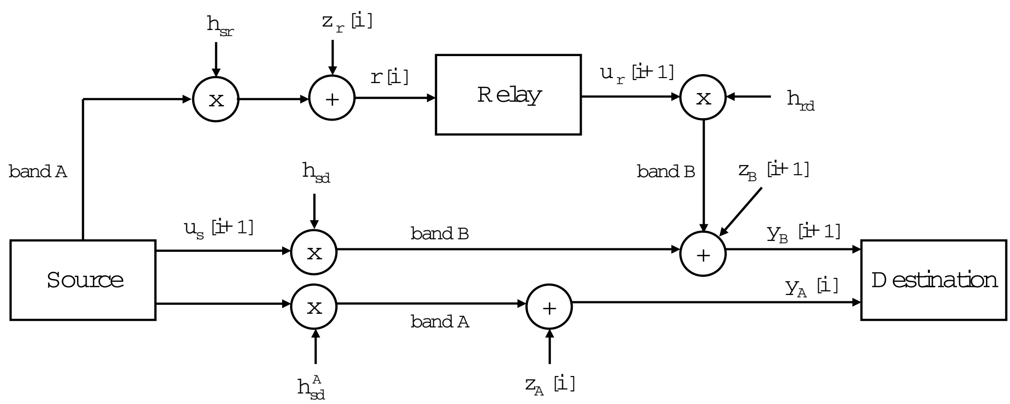


Figure 2. Distributed STBC system model with one assisting relay [12].

Using frequency band A, source broadcasts information block of $\sqrt{\alpha_A}\mathbf{s}[i] = \sqrt{\alpha_A}\left[s[2i], s[2i+1]\right]^T$. The received signal block at relay R, $\mathbf{r}[i]$ is given as:

$$\mathbf{r}[i] = h_{sr}[i]\sqrt{\alpha_A}\mathbf{s}[i] + \mathbf{z}_r[i] \ , \tag{6}$$

where h_{sr} is the slow varying fading channel between source and relay R, α_A is the transmit energy in channel A, and $\mathbf{z}_r[i]$ is the noise vector with each entry being a complex circular Gaussian random variable with variance $N_0/2$ per dimension. Provided that the relay can perfectly estimate the channel gain h_{sr}, it amplifies the received signal to construct a soft estimate of the signal. During the next time slot, $i+1$, source and the relay transmit in band B using improved Alamouti space-time codes [4]. Source transmits $\mathbf{u}_s[i+1] = \sqrt{\alpha_B}\mathbf{s}[i]$; which is the same signal block transmitted in band A during the previous block except for the transmit energy. Source transmits continuously in band A as well as in band B. Relay R transmits

$$\mathbf{u}_r[i+1] = \sqrt{\alpha_R}\, c_{sr}[i]\left[r^*[2i+1], -r^*[2i]\right]^T \ , \tag{7}$$

where $\sqrt{\alpha_R}$ is the fixed amplification at relay, and c_{sr} depends on the channel between source and relay. Destination receiver is designed to use the orthogonal structure of the space-time coding pairs $\mathbf{u}_s[i+1], \mathbf{u}_r[i+1]$ to separate the signals $s[2i]$ and $s[2i+1]$.

As it is shown in Figure 2, the signal at destination is received in channels A and B. The equivalent baseband model of these two signals can be expressed at time slot i, as:

$$\mathbf{y}_A[i] = \left[y_A[2i], y_A[2i+1]\right]^T = h_{sd}^A[i]\sqrt{\alpha_A}\mathbf{s}[i] + \mathbf{z}_A[i], \tag{8}$$

$$\begin{aligned}\mathbf{y}_B[i+1] &= \left[y_B[2i+2], y_B[2i+3]\right]^T \\ &= h_{rd}[i+1]\mathbf{u}_r[i+1] + h_{sd}[i+1]\mathbf{u}_s[i+1] + \mathbf{z}_B[i+1]. \end{aligned}\tag{9}$$

The effect of flat fading is modeled by h_{sd}^A, h_{sd}, h_{sr} and h_{rd}. These channel gains are assumed to be constant over one signal block, but could vary from block to block. The first subscript specifies the transmitter and the second subscript identifies the receiver. For example, h_{sd} denotes the channel gain between source and destination. The superscript A means that the gain belongs to channel A.

After acquiring $h_{sr}[i]$, the relay selects $c_{sr}[i] = h_{sr}^*[i]/|h_{sr}[i]|$ to compensate for the phase of the channel between source and the relay. Space-time decoding of received signal at channel B is performed by processing $\mathbf{y}_B[i+1]$ with matrix $\mathbf{H}_1$ as follows:

$$\mathbf{x}[i+1] = \mathbf{H}_1\, \mathbf{y}_B[i+1], \tag{10}$$

where $\mathbf{x}[i+1] = \left[x[2i+2], x[2i+3]\right]^T$ is the decision variable in channel B, and according to improved Alamouti decoding, matrix $\mathbf{H}_1$ is defined as follows:

$$\mathbf{H}_1 = \begin{bmatrix} \sqrt{\alpha_B}h_{sd}^*[i+1] & -\sqrt{\alpha_A\alpha_B}|h_{sr}[i]|h_{rd}[i+1] \\ \sqrt{\alpha_A\alpha_B}|h_{sr}[i]|h_{rd}^*[i+1] & \sqrt{\alpha_B}h_{sd}[i+1] \end{bmatrix}. \tag{11}$$

A maximal ratio combiner combines two orthogonal signals of $\mathbf{y}_A[i]$ and $\mathbf{x}[i+1]$ with proper weights. Then, receiver detects the signal by using a maximum likelihood decision rule.

3.2. Distributed Space-Time Coding with Two Relays

There are some cases where the two channel gains between source and destination, i.e., $h_{sd}^A[i]$ and $h_{sd}[i+1]$, are not independent. In case of dependency of these two channels, diversity gain of scheme would be lost due to the channel correlation. To improve the performance of the system, another relay would be required [11, 12]. As it is depicted in Figure 3, we assume that source cooperates with two relays R_1 and R_2 in order to perform improved Alamouti space-time coding.

Relays R_1 and R_2 cooperate after receiving the signal block in channel A transmitted by the source. Source does not transmit $\mathbf{u}_s[i+1]$ and this role is transferred to relay R_2. The received signal at destination from source would be the same as in (8); replacing $\sqrt{\alpha_A}$ with $\sqrt{\varepsilon}$, where ε is the transmit energy of source in channel A. The transmitted signal of two relays in channel B can be given as follows:

$$\mathbf{u}_{r_1}[i+1] = \sqrt{\alpha_{R_1}}\frac{h_{sr_1}^*[i]}{|h_{sr_1}[i]|}\left[r_1^*[2i+1], -r_1^*[2i]\right]^T, \tag{12}$$

$$\mathbf{u}_{r_2}[i+1] = \sqrt{\alpha_{R_2}}\frac{h_{sr_2^*}[i]}{|h_{sr_2}[i]|}\left[r_2[2i], r_2[2i+1]\right]^T \tag{13}$$

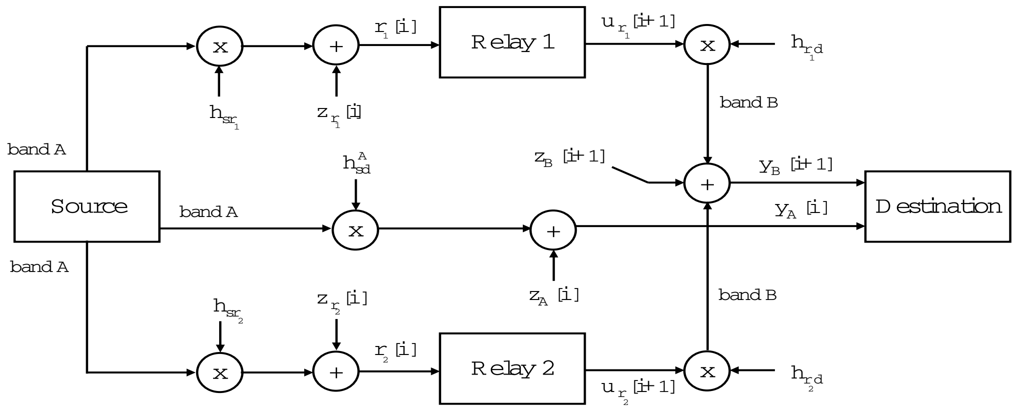


Figure 3. Distributed STBC system model with two assisting relays [12].

where $\sqrt{\alpha_{R_j}}$ is the amplification factor at relay R_j, and $\mathbf{r}_j[i]$ is the received signal block at relay R_j which is expressed as:

$$\mathbf{r}_j[i] = \left[r_j[2i], r_j[2i+1]\right]^T = h_{sr_j}[i]\sqrt{\varepsilon}\mathbf{s}[i] + \mathbf{z}_{r_j}[i], \quad j = 1, 2. \tag{14}$$

The received signal at destination in channel B is given by:

$$\mathbf{y}_B[i+1] = h_{r_1d}[i+1]\mathbf{u}_{r_1}[i+1] + h_{r_2d}[i+1]\mathbf{u}_{r_2}[i+1] + \mathbf{z}_B[i+1]. \tag{15}$$

The receiver structure is the same as the system with one assisting relay, except that, instead of $\mathbf{H}_1$, the receiver uses matrix $\mathbf{H}_2$ as given by:

$$\mathbf{H}_2 = \begin{bmatrix} \sqrt{\alpha_{R_2}}|h_{sr_2}[i]|h^*_{r_2d}[i+1] & -\sqrt{\alpha_{R_1}}|h_{sr_1}[i]|h_{r_1d}[i+1] \\ \sqrt{\alpha_{R_1}}|h_{sr_1}[i]|h^*_{r_1d}[i+1] & \sqrt{\alpha_{R_2}}|h_{sr_2}[i]|h_{r_2d}[i+1] \end{bmatrix}. \tag{16}$$

Then the received signal in channel A and the processed signal of channel B are combined using a maximal ratio combiner to form the decision variable for the maximum likelihood detector.

Approximate and asymptotical formulas for BER of DSTBC system with BPSK modulation are obtained in [12] for both cases of one and two relays. Based on BER analysis, it is concluded that the DSTBC system with two relays is more versatile than the one relay system.

4. On-Channel Distributed Space-Time Coded Relaying with Feedback

In this section, a new DSTBC technique which does not require an extra frequency channel for cooperation is presented. It is called *on-channel DSTBC*. Similar to the model in the previous section, there are three terminals: source, relay and destination. The improved

Alamouti space-time block code discussed earlier is used for this cooperative space-time coding technique. It is assumed that all the receivers are regenerative and channel gains for relay-destination and source-destination are independent. Also, a perfect error-free link between the source and the relay is assumed. The proposed method requires a feedback from the destination to the source which can be done via the reverse channel between the destination and the source. The proposed scheme is further improved by introducing the recursive on-channel DSTBC technique.

4.1. On-Channel DSTBC

Transmission and cooperation phases are mixed over time domain here. Sequence of transmission by source and relay is shown in Table 3. Every transmission block is composed of three time slots. In the first time slot, source broadcasts its signal s_0 to the relay and the destination. In the next time slot, source transmits s_1 while relay after decoding s_0 transmits conjugate of decoded s_0, $\hat{s}_0$. This constructs the second row of improved Alamouti STBC. In the third time slot, source repeats conjugate of s_1 with a different amplitude which is set by the feedback from the destination. This will help the destination receiver to build up the first row of improved Alamouti STBC. Table 3 shows the signal transmission sequence of source and relay in one signal block.

Table 3. Transmission sequence for on-channel DSTBC

	source	relay
time t	s_0	-
time t+T	s_1	$\hat{s}_0^*$
time t+2T	$\left(\frac{h_{rd_0}}{h_{sd_0}}\right) s_1^*$	-

In Table 3, h_{sd_0} and h_{rd_0} are the channel gains between source and destination, and relay and destination, respectively. The received signal at destination is then given by:

$$\begin{aligned} r_0 &= h_{sd_0}s_0 + n_0, r_1 \quad = h_{sd_0}s_1 + h_{rd_0}s_0^* + n_1, \\ r_2 &= h_{rd_0}s_1^* + n_2 \end{aligned} \tag{17}$$

The receiver decodes the received signal and builds up the decision variable as follows:

$$\begin{aligned} \tilde{s}_0 &= h_{sd_0}^*(r_0 - r_2) + h_{rd_0}r_1^* \\ &= (\alpha_{sd_0}^2 + \alpha_{rd_0}^2)s_0 + h_{sd_0}^*(n_0 - n_2) + h_{rd_0}n_1^*, \\ \tilde{s}_1 &= -h_{rd_0}(r_0 - r_2)^* + h_{sd_0}^* r_1 \\ &= (\alpha_{sd_0}^2 + \alpha_{rd_0}^2)s_1 - h_{rd_0}(n_0^* - n_2^*) + h_{sd_0}^* n_1. \end{aligned} \tag{18}$$

As it can be seen in (18), the new on-channel coding scheme achieves theoretically a diversity gain of two using one more time slot (1.5-fold bandwidth). In the next part, this coding scheme is modified by a simple recursive method.

4.2. Recursive On-Channel DSTBC

By a slight modification of the previous method, it is possible to increase the diversity gain of on-channel DSTBC. The last symbol sent by the source in the last block of three time slots, can be used as the first symbol in the next block. The transmission sequence is shown in Table 4. In this table, h_{sd_i} and h_{rd_i} are the channel gains in the i^{th} block, between source and destination and relay and destination respectively.

Table 4. Transmission sequence for recursive on-channel DSTBC

	source	relay
time t	s_0	-
time t+T	s_1	$\hat{s}_0^*$
time t+2T	$\left(\frac{h_{rd_0}}{h_{sd_0}}\right) s_1^*$	-
time t+3T	$\left(-\frac{h_{rd_0}}{h_{sd_1}}\right) s_2^*$	$\hat{s}_1$
time t+4T	$\left(\frac{h_{rd_1}}{h_{sd_1}}\right) s_2$	-
time t+5T	$\left(\frac{h_{rd_1}}{h_{sd_1}}\right) s_3$	$\hat{s}_2^*$
time t+6T	$\left(\frac{h_{rd_2}}{h_{sd_2}}\right) s_3^*$	-
$\vdots$	$\vdots$	$\vdots$

The received signal at destination can be formulated as follows:

$$\begin{aligned}
r_0 &= h_{sd_0}s_0 + n_0, \\
r_1 &= h_{sd_0}s_1 + h_{rd_0}s_0^* + n_1, \\
r_2 &= h_{rd_0}s_1^* + n_2, r_3 &= -h_{rd_0}s_2^* + h_{rd_1}s_1 + n_3, r_4 = h_{rd_1}s_2 + n_4, \\
r_5 &= h_{rd_1}s_3 + h_{rd_2}s_2^* + n_5, \\
r_6 &= h_{rd_2}s_3^* + n_6, \\
&\vdots
\end{aligned} \tag{19}$$

The receiver decodes the received signal by the proposed method presented earlier. The block of r_0, r_1 and r_2 is used to build the decision variable as in (18) to detect s_0 and s_1. Then block r_2, r_3 and r_4 is used to construct the decision variable for s_1 and s_2. Then r_4, r_5 and r_6 are used for detecting s_2 and s_3. Since decoding of each block is done by borrowing a symbol from the previous block, we call this method *recursive on-channel DSTBC relaying*. Decision variables for the first three blocks are as follows:

$$\tilde{s}_0 = h_{sd_0}^*(r_0 - r_2) + h_{rd_0}r_1^* = (\alpha_{sd_0}^2 + \alpha_{rd_0}^2)s_0 + z_0, \tag{20}$$

$$\tilde{s}_1 = -h_{rd_0}(r_0 - r_2)^* + h_{sd_0}^* r_1 = (\alpha_{sd_0}^2 + \alpha_{rd_0}^2)s_1 + z_1, \tag{21}$$

$$\tilde{s}_1 = h^*_{rd_1} r_3 + h_{rd_0}(r_2 + r_4)^* = (\alpha^2_{rd_0} + \alpha^2_{rd_1}) s_1 + z_2, \quad (22)$$

$$\tilde{s}_2 = -h_{rd_0} r^*_3 + h^*_{rd_1}(r_2 + r_4) = (\alpha^2_{rd_0} + \alpha^2_{rd_1}) s_2 + z_3, \quad (23)$$

$$\tilde{s}_2 = h^*_{rd_1}(r_4 - r_6) + h_{rd_2} r^*_5 = (\alpha^2_{rd_1} + \alpha^3_{rd_2}) s_2 + z_4, \quad (24)$$

$$\tilde{s}_3 = -h_{rd_1}(r_4 - r_6)^* + h^*_{rd_2} r_5 = (\alpha^2_{rd_1} + \alpha^2_{rd_2}) s_3 + z_5, \quad (25)$$

where $z_i's$ are the equivalent additive noise for initial decision variable of $\tilde{s}_i$. Two consecutive decoded blocks are combined to form a higher order of diversity at the receiver. For example, to detect s_2, (23) and (24) are combined to form the decision variable for s_2 as follows:

$$\tilde{s}_{2_t} = (\alpha^2_{rd_0} + 2\alpha^2_{rd_1} + \alpha^2_{rd_2}) s_1 + z_3 + z_4. \quad (26)$$

According to (26), it is concluded that by using two time slots for one symbol (two-fold bandwidth) a diversity gain of three is achieved.

4.3. Performance of On-Channel DSTBC

It was shown theoretically that the proposed on-channel relaying scheme increases the diversity gain in the wireless network. In this section, the BER performance of the on-channel relaying system is evaluated by Monte-Carlo simulation.

Six different schemes are simulated and they are: no diversity, time diversity (order two and three), Alamouti scheme, on-channel DSTBC and recursive on-channel DSTBC. To have a better estimate of the performance of the proposed technique, two different signal constellation of BPSK and QPSK are considered.

In the simulation of on-channel DSTBC, it is assumed that a block of two symbols, which is randomly generated with an equiprobable signal generator, is transmitted by the source. Only the reception of signal is simulated at the receiver side. This is accomplished using the derived formulas (26) and (5). The average SNR received at destination is set by the variance of AWGN; i.e., after generating a unit variance noise, the generated noise is divided by SNR value. The unit variance signal is then added to the noise and processed by the receiver in the simulator. The resulting BER graphs after simulation are depicted in Figures 4 and 5.

As it is seen in these two figures, the average BER performance of the proposed on-channel DSTBC relaying scheme behaves the same as the time-diversity system with order of two. It means that by using 1.5 times more bandwidth a diversity gain of two is achieved. Compared with the system without diversity, on-channel DSTBC achieves 8dB SNR gain for BER of 10^{-3}. Also, for the recursive on-channel DSTBC it is seen that a 3dB SNR gain is achieved for BER of 10^{-3} compared to time diversity of order two. Recursive on-channel relaying behaves like a time-diversity system with order of three. This complies with what was shown theoretically in the previous section. Therefore, it can be said that the recursive on-channel DSTBC relaying scheme achieves diversity order of three.

If a time division duplex scheme is used to divide the forward and reverse channels, then it is possible to use the received signal in the reverse channel to estimate the channel state in the forward channel and there would be no bandwidth loss due to the feedback of h_{sd_i} to the source; however, the feedback is still needed for the source to know h_{rd_i}, in order to adjust its transmit power.

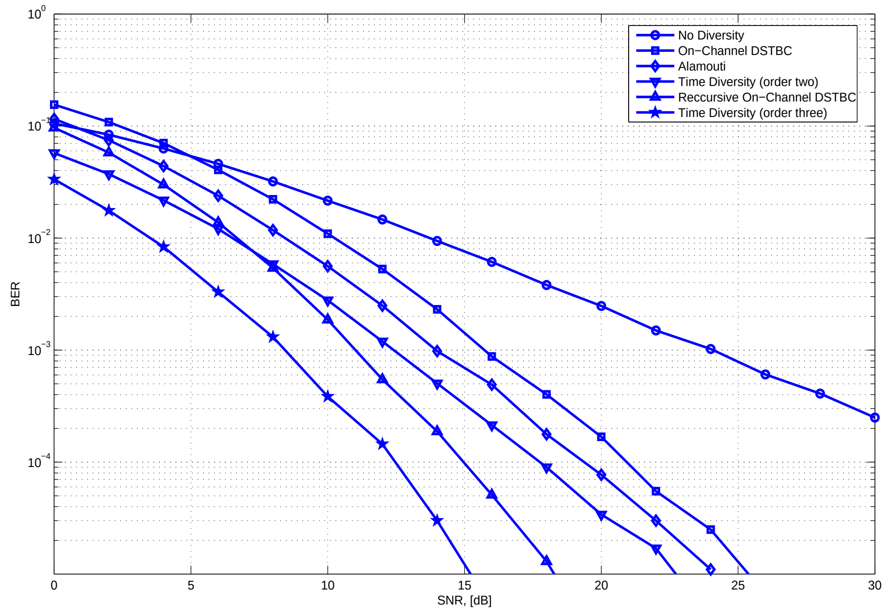


Figure 4. BPSK BER performance of different space-time techniques.

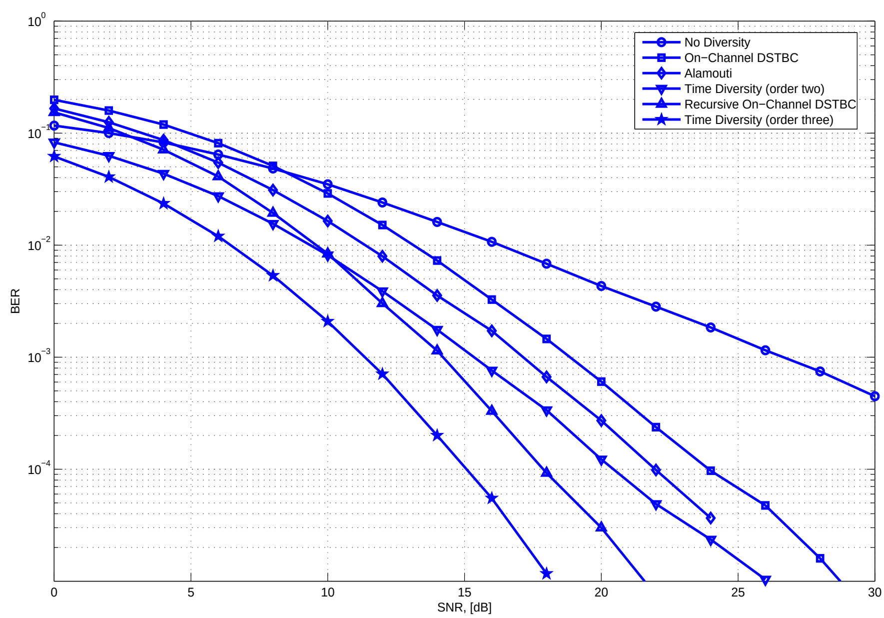


Figure 5. QPSK BER performance of different space-time techniques.

Another question that should be answered is when would the cooperation result in performance improvement compared with no diversity scheme. As it is seen in Figures 4 and 5, the cooperation results in BER improvement in average to high SNR's; thus, it is possible to design an adaptive system where source cooperates with the proper relay above a certain SNR threshold. What signifies the proposed method from [7] is that the transmission and cooperation phases are not divided in orthogonal channels (time or frequency), but mixed in time domain. It is possible to perform cooperation in an adaptive way according to the average received SNR at destination. If the average received SNR is below a certain threshold, cooperation would be stopped. For example, in recursive on-channel DSTBC, diversity order would be decreased from three to two.

The proposed scheme does not require any infra-structure in the network; therefore, it can also be implemented in an ad hoc network. In an infra-structure based network such as a cellular system, source terminal is the mobile station, while the cooperating relay can be either an individual fixed element in network (relay station) or a nearby mobile station and the destination terminal is the base station. Fixed relays can be located in some certain positions in order to have a better communication link with base station. This can itself improve the performance of the proposed method.

According to (18), the received signal at destination in on-channel DSTBC with feedback depends on α_{sd_i} and α_{rd_i}. This means that if the link between relay station and base station is in good condition, signal would rarely go into deep fades. As can be seen in (26), the received signal at base station for the recursive scheme is only vulnerable to the link between relay and destination which can further increase the reliability of the scheme.

Another important issue that should be considered here is the level of amplification in the source which helps to construct the DSTBC at destination. The amplification factor at source terminal is determined by two factors: h_{sd_i} and h_{rd_i}. In other words, the channel gain between destination and cooperating terminals (source and relay) determines the amplification factor at source terminal. If relay allocation procedure can be performed in a way that the link between relay and destination is always in good condition then amplification factor at source would only be dependent on the link between source and destination. A threshold can be set for the amplification level at source terminal to prevent saturation at the source transmitter. At this point, cooperation would be stopped and diversity order would be decreased to two.

5. Research Challenges and Directions

The new trend of using space-time block codes to increase the spectral efficiency of the relaying protocols by constructing a virtual antenna array can help the users in the poor channel conditions benefit from spatial diversity. A DSTBC protocol was proposed in [12] for cooperation of relays in the system which needs two frequency channels for cooperation. An on-channel relaying DSTBC was proposed here and it was shown to have diversity improvement in the system. This method uses only one relay which has two modes of reception and transmission. Compared to the similar techniques [7, 12], this method does not require the relay to transmit at all the time slots. Space-time code construction is done by sending a feedback from destination to source in the reverse channel.

The main fact that can lead to the construction of higher order on-channel distributed space-time block codes is that in a wireless network with a sufficient number of potential relays around the source terminal, always, there is one relay that knows the symbol sequence transmitted cooperatively on the channel by the user and all the other relays. This problem can be viewed from another interesting perspective. All the proposed methods for constructing a virtual antenna array by the cooperative relays use the space-time codes designed for a multiple-antenna user. By the unique properties of the on-channel relaying system, perhaps, a different class of space-time codes can be designed with respect to the cooperative system properties which has a less complexity and a higher bandwidth efficiency.

References

[1] D. Tse and P.viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[2] I. E. Telatar, "Capacity of multi-antenna gaussian channels," *European Transactions on Telecommunications*, vol. 10, pp. 585–595, Nov./Dec. 1999.

[3] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in rayleigh flat fading," *IEEE Transactions on Information Theory*, vol. 1, pp. 139–157, Jan. 1999.

[4] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 1451–1458, Oct. 1998.

[5] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block codes from orthogonal design," *IEEE Transactions on Information Theory*, vol. 45, pp. 1456–1467, Jul. 1999.

[6] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block coding for wireless communications: performance results," *IEEE Journal on Selected Areas in Communications*, vol. 17, pp. 451–460, Mar. 1999.

[7] J. Laneman and G. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Transactions on Information Theory*, vol. 49, pp. 2415– 2425, Oct. 2003.

[8] J. Boyer, D. Falconer, and H. Yanikomeroglu, "Multihop diversity in wireless relaying channels," *IEEE Transactions on Communications*, vol. 52, pp. 1820–1830, Oct. 2004.

[9] Y. Hua, Y. Mei, and Y. Chang, "Parallel wireless mobile relays with space-time modulations," *IEEE Workshop on Statistical Signal Processing*, pp. 375–378, Sept./Oct. 2003.

[10] Y. Hua, Y. Mei, and Y. Chang, "Wireless antennas - making wireless communications perform like wireline communications," *IEEE Topical Conference on Wireless Communication Technology*, pp. 47–73, Oct. 2003.

[11] P. Anghel and M. Kaveh, “Multi-user space-time coding in cooperative network,” *IEEE ICASSP*, vol. 4, pp. 73–76, Apr. 2003.

[12] P. Anghel and M. Kaveh, “On the performance of distributed space-time coding systems with one and two non-regenerative relays,” *IEEE Transactions on Wireless Communications*, vol. 5, pp. 682–692, Mar. 2006.

[13] M. Dohler, E. Lefranc, and H. Aghvami, “Space-time block codes for virtual antenna arrays,” *IEEE PIMRC*, vol. 1, pp. 414–417, Sep. 2002.

[14] A. Sendonaris, E. Erkip, and B. Aazhang, “Increasing uplink capacity via user cooperation diversity,” *IEEE Symposium on Information Theory*, p. 156, Aug. 1998.

[15] A. Sendonaris, E. Erkip, and B. Aazhang, “User cooperation diversity - part i: System description,” *IEEE Transactions on Communications*, vol. 51, pp. 1927– 1938, Nov. 2003.

[16] A. Sendonaris, E. Erkip, and B. Aazhang, “User cooperation diversity - part ii: Implementation aspects and performance analysis,” *IEEE Transactions on Communications*, vol. 51, pp. 1939– 1948, Nov. 2003.

INDEX

A

B

C

D

E

F

G

H

M

N

O

P

Q

R

S

T

U

V

W

Y