A New B2B Platform based on Cloud Computing

Gang Xiong[1,2], Jiachen Hou[2,1]
[1]State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, CHINA
Email: gang.xiong@ia.ac.cn, jiachen.hou@gmail.com

[2]Dongguan Research Institute of CASIA, Cloud Computing Center, Chinese Academy of Sciences, Songshan Lake, Dongguan, 523808, CHINA
Email: dongxs@casc.ac.cn

Xisong Dong[3,2]
[3]Jiang Su China Sciences Intelligent Engineering Co., Ltd Enterprise Academician Workstation, 215000,CHINA

Timo R. Nyberg[4]
[4]BIT Research Centre, Aalto University, Aalto, FINLAND
Email: timo.nyberg@aalto.fi

***Abstract*—with the fast development of Internet and its data scale, B2B (Business to Business), whose speed and high availability advantage is based on Internet, is eroding more and more market share of traditional business. In recent years, the new data processing technologies, such as cloud computing, assure the enhancement of computer's computing capacity; it has become possible for researchers to process and analyze the massive business data, including large scale customer data, and large scale commodity information. A new B2B platform frame, based on cloud computing with less running time and better response efficiency, is proposed here to promote transaction handle efficiency. By adoption of complex network theory and cloud computing technology, the new platform has been evaluated to decrease transaction handle time, and make more benefits.**

Keywords-B2B; cloud computing; efficient promotion; massive data; complex network

I. INTRODUCTION

EC (Electronic Commerce) refers merchants using electronic method to handle business over remote or distance transaction.The definition of Electronic Commerce from Wikipedia [1] is "consists of the buying and selling of products or services over electronic systems such as the Internet and other computer networks. The amount of trade conducted electronically has grown extraordinarily with widespread Internet usage. The use of commerce is conducted in this way, spurring and drawing on innovations in electronic funds transfer, supply chain management, Internet marketing, online transaction processing, electronic data interchange (EDI), inventory management systems, and automated data collection systems. Modern electronic commerce typically uses the World Wide Web at least at some point in the transaction's lifecycle, although it can encompass a wider range of technologies such as e-mail, mobile devices and telephones as well."

From the definition above, we can find that EC is a brand new business operation mode, which was created and developed with comuputer technonogy and Internet technology. And in some fact, we can say that the computer technology can stimulate the development of EC, and the latest comuputer technonogy will decide the deepth and breadth of EC.

On the one hand, the development of EC relies on the development and usage of network computing platform; On the other hand, the development of EC has acturally provided a good application platform, and thrown some new light and key R&D topics on the development direction of network software. We can say, network software and EC stimulate their development together.

In this paper, to research the relationship of the existing EC online software package, which has strict coupling modules, and require much maintainance, we propose a new method and frame to analyze the features of network B2B software of EC, which is based on the new technologies, such as complex network [2] and cloud computing [3]. The new software frame has been proved to promote software maintaining efficiency and descrease business handling time.

The rest of this paper is organized as follows: section II introduces the categories of EC, Influence of cloud computing to EC, and Statistical characteristics of networked software. Section III presents the new B2B e-commerce software platform, which is based on complex network technology and cloud computing technology. Section IV calculates the existing statistical features of e-commerce software platform packages, and then gives the experimental results of its inherent statistical characteristics of the new e-commerce software platform package in the real data, and the experimental results of the software efficiency improvement. Section V draws the final conclusions of this paper.

II. RELATED WORK

A. Categories of EC

E-commerce was firstly proposed in the 1960s, and it has gained its rapid development since 1990s, mainly for the reason of the rapid development of Internet. Normally, E-commerce can be divided into the following seven basic modes:

1) B2B: Business-to-Business
2) B2C: Business-to-Customer
3) C2C:Customer-to-Customer
4) B2M:Business-to-Manager or Marketing

5) M2C:Manager-to-Customer

6) B2A and B2G: Business-to-Administration or Business-to-Government

7) C2A and C2G: Customer-to-Administration or Customer-to-Government

B. Influence of cloud computing to EC

According to the definition of NIST, cloud computing can be defined as:

Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.

Cloud computing has some essential characteristics, which can be summarized as:

1) On-demand self-service. A consumer can be automatically provided the computing capabilities, such as server time and network storage, whenever needed, without requiring human interaction with each service's provider.

2) Broad network access. Capabilities are available over the network and the services can be accessed through standard mechanisms, so its usages are promoted by heterogeneous thin or thick client platforms (e.g. mobile phones, laptops, and PDAs).

3) Resource pooling. The provider's computing resources are pooled to serve multiple consumers by using of a multi-tenant model, where different physical and virtual resources are dynamically assigned and reassigned according to consumer demand. There is a sense of location independence, the customer generally has no control or knowledge over the exact location of the provided resources, sometimes can specify the location at a higher level of abstraction (e.g., country, state, or data center). Examples of the resources include storage, processing, memory, network bandwidth, and virtual machines.

4) Rapid elasticity. Capabilities can be rapidly and elastically provisioned, in some cases automatically, to quickly scale out and rapidly released to quickly scale in. For the consumer, the capabilities availability often appears to be unlimited, and can be purchased any quantity at any time.

5) Measured service. Cloud systems automatically control and optimize resource usage by leveraging a metering capability at some abstraction level of appropriate to the service type (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported, which provides the transparency for both the provider and the consumer of the utilized services.

We can find that cloud computing is not a new technology but a new resource application mode, and it can be cited as the promotion of grid computing and transparent computing aiming to make the best usage efficiency by using the coordination of the foundmental devices on lower level and application sofware on higher layer. In cloud computing system, hardware resources such as servers, will be visualized for software applications, and then can dynamically added and removed. The most basic thing of cloud computing system is the computing plus resource visulation.

C. Statistical characteristics of networked software

Networked software [4] is the special software which can run in complex network environment, and whose structure and behavior will evolve accordingly. The typical characteristics of networked software include that:

1) It has a huge number of developers and users for the reason that networked software is always developed and used on Internet by various developers, which results in the unlimited scale and edge of networked software.

2) It has a large number of software modules which changes all the time, for the reason that the user requirements are always changed. So, a large number of fine-grained and configable modules are necessary, which can be combined to meet the customized requirement of a huge number of customers.

3) It has software modules which are loosely coupled, and requires the interoperation easily. The mesh size of networked software should be as small as possible, the interoperation requirement can easily be met by using fast coupling and decoupling of application logics module, where cloud computing system applies software method to tolerate the node failure.

There are some typical networked softwares, such as SAAS (Software as a Service) and iPhone App Store, which is for online query and offline usage.

There are also some research results of kernel modules of networked software.It is reported that the kernel modules of Linux and Tomcat has some characters of scale-free to consists small world network by java classes and reference among them module [5] [6] [7].

In recent years, how to build up EC models and how to construct the new networked EC software by using of complex network analysis technology, have become the research hot points of EC recently.

III. CONSTRUCTION OF NEW B2B PLATFORM BASED ON CLOUD COMPUTING

A. New B2B EC platform based on PAAS

Cloud computing is new type computing mode of distributed computing for shared resources design, development, deployment and operation. Cloud computing model is divided into three layers for different functions such as SaaS (Software as a Service), PaaS (Platform as a Service) and IaaS (Infrastructure as a Service). The detailed description of SaaS, Paas and IaaS are as followed:

1) Cloud Software as a Service (SaaS). The capability, which is provided to the consumer, is to use the provider's applications running on a cloud infrastructure. The applications are accessible from various client devices through a thin client interface such as a web browser (e.g., web-based email). The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible

exception of limited user-specific application configuration settings.

2) Cloud Platform as a Service (PaaS). The capability, which is provided to the consumer, is to deploy onto the cloud infrastructure. The consumer-created or acquired applications are all created by using of programming languages and tools supported by the providers. The consumer does not manage or control the underlying cloud infrastructure, including network, servers, operating systems, or storage, but has control over the deployed applications and possible application hosting environment configurations.

3) Cloud Infrastructure as a Service (IaaS). IaaS provides the consumer such capabilities as the processing, storage, networks, and other fundamental computing resources, where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over operating system, storage, deployed applications, and possibly limited control of select networking components (e.g., host firewalls).

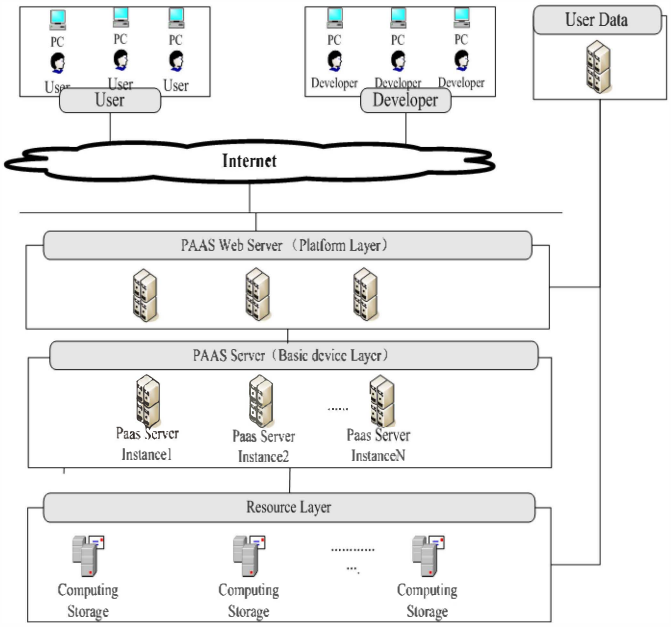

Figure 1. Logical Framework of Target System

In order to promote system efficiency of target Electronic Commerce system, its logical framework can be found in Figure 1. In order to construct real target system, the new B2B platform framework, which is based on cloud computing, can be found in Figure 2.

The new B2B platform framework in Figure 2 can be divided into three layers including Web Service Layer (belonging to SAAS), EC Logical Layer (belonging to PAAS), and Cloud Platform Layer (belonging to IAAS).

The detailed descriptions of the three layers are as followed:

1) Web Service Layer: In this layer, the back end system using work flow engine to handle EC online transaction by sub operations such as EC Authority, B2B large scale data search for customer, B2B Transaction, B2B Account Management and so on. This layer was constructed by SaaS to offer two different service types (e.g., user type and rent type).

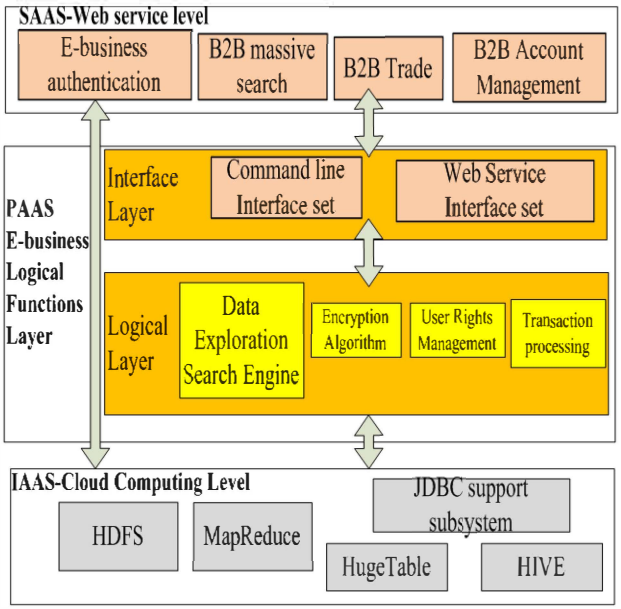

Figure 2. New B2B Platform Framework Based on Cloud Computing

2) EC Logical Layer: In this layer, the specified atomic EC logical function was constructed such as atomic EC logical software components of data search, atomic software components of Encryption Algorithm, atomic software components of User Role Management and atomic software components of Transaction Handle. The component of this web service layer can be called by two ways, Command Line Interface or Web Service Interface.

3) Cloud Platform Layer: This layer was constructed on HDFS and other open source cloud computing model component to store and manipulate very large scale data in files. And cloud computing technology promotes back end data handle efficiency.

B. Betweenness promotion of software coupling

The network models of EC software discussed in this paper are undirected and unweighted one, which can be expressed as $G=(V,e)$. V is the set of nodes with number $N=|V|$ and e is the set of edges with number $L=|e|$.

The most common classification method of complex network is to divide it into three different types by node degree distribution such as ER random network model [8], small world network model [9], scale free network model [10] and LUHNM network model [11].

1) ER random network model: this model was proposed in 1960 by Erdos and Rényi, and this model can be considered the most basic model of complex networks. In their 1959 paper [8], Erdos and Rényi introduced a model to generate random graphs consisting of N vertices and M

edges. Starting with *N* disconnected vertices, the network is constructed by the addition of L edges at random, avoiding multiple and self connections. Another similar model defines *N* vertices and a probability *p* of connecting each pair of vertices. The latter model is widely known as Erdos and Rényi (ER) model.

2) Small world network model: Many real world networks exhibit what is called the small world property, i.e. most vertices can be reached from the others through a small number of edges. This characteristic is found, for example, in social networks, where everyone in the world can be reached through a short chain of social acquaintances [9]. This concept originated from the famous experiment made by Milgram in 1967 [9], who found that two US citizens chosen at random were connected by an average of six acquaintances. Another property of many networks is the presence of a large number of loops of size, i.e. if vertex *i* is connected to vertices j and *k*, there is a high probability of vertices j and k being connected; for example, in a friendship network, if B and C are friends of A, there is a high probability that B and C are also friends. ER networks have the small world property but a small average clustering coefficient; on the other hand, regular networks with the second property are easy to construct, but they have large average distances. The most popular model of random networks with small world characteristics and an abundance of short loops were developed by Watts and Strogatz [9] and is called the Watts-Strogatz (WS) small-world model. They showed that small-world networks are common in a variety of realms ranging from the C elegance neuronal system to power grids. This model is situated between an ordered finite lattice and a random graph presenting the small world property and high clustering coefficient.

3) Scale free network model: After Watts and Strogatz's model, Barabási and Albert [9] showed that the degree distribution of many real systems is characterized by an uneven distribution. Instead of the vertices of these networks having a random pattern of connections with a characteristic degree, as with the ER and WS models, some vertices are highly connected while others have few connections, with the absence of a characteristic degree. And the model proposed by Barabási and Albert is so called the *scale free network model.*

4) LUHNM network model: this model was proposed by Chinese researcher Prof. J. Q. Fang in 2007 which combine the random connection probability and definite connection probability into one network generate model. And LUHNM network model is the further development of ER random network model [8], small world network model [9], and scale free network model [10].

In research of networked EC software, the class or function is abstracted into the node of complex network, and the inherent of calling relation between different classes, is abstracted into edge of complex network. And it is already proved that the software network of networked EC software belongs to scale free network. In the networked EC software, there exists some kernel java class or C++ function, which occupies about 5%, are in charge of about 95% of the software logics [12].

In networks, the greater number of paths in which a vertex or edge participates, the higher importance of this vertex or edge is for the network. Thus, assuming that the interactions follow the shortest paths between two vertices, it is possible to quantify the importance of a vertex or an edge in terms of its betweenness centrality [13] defined as:

$$B_u = \sum_{i,j} \frac{\sigma(i,u,j)}{\sigma(i,j)} \tag{1}$$

Where $\sigma(i,u,j)$ is the number of shortest paths between vertices i and j that pass through vertex or edge u , $\sigma(i,j)$ is the total number of shortest paths between i and j ,and the sum is over all pairs *i* and, j of distinct vertices.

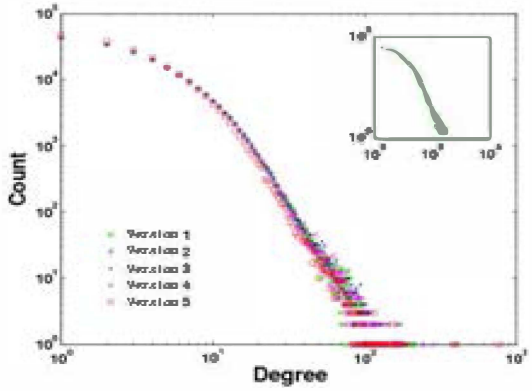

Figure 3. Degree of Old Electronic Commerce Software

In the reconstruction of real networked EC software which was programmed with Java, class is abstracted into node of complex network, the inherent relation and aggregation relation of different classed are abstracted into edge of complex network. The statistic result of node degree and betweenness of the formal five version of this networked EC software can be found in Figure 3 and Figure 4 [14].

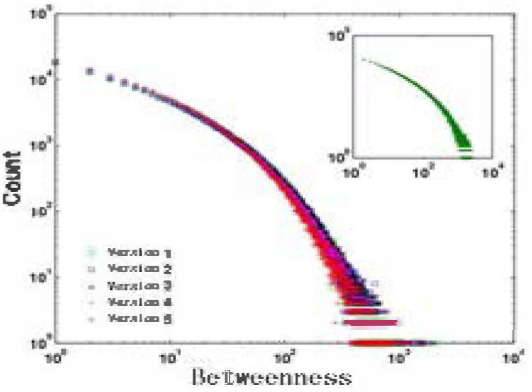

Figure 4. Betweenness of Old Electronic Commerce Software

It can be clearly found in Figure 3 and 4 is that there some 5% nodes' degree and betweenness of the EC software in charge of 80% degree and betweenness values of the

whole EC software package [15]. These nodes are very important for software reconstruction. And in the process of software coupling reconstruction, we decompose these important java classes into smaller java classes for more efficient software component reused to decrease software maintaining in business logic changing.

IV. VALIDATION OF NEW B2B PALTFORM BASED ON CLOUD COMPUTING

A. Result of software transaction response time experiment

In order to validate the performance of the new B2B platform framework based on cloud computing, the software transaction response time was tested on the real EC system from Figure 3 and 4. The web presentation layer was coded in Flex in java and Tomcat6.0 is used to be the web service component container. Cloud computing layer is implemented by APIs from Hadoop-0.19.0.core.jar.The back end database was constructed in the cluster of 20 nodes with 16 GB Memory, 12TB hard disk, 4*2GHz CPU each on RedHat 6.

In the deployment process of our project, the new B2B platform system based on cloud computing consists of five kernel sub systems and four assistant sub system. The promotion rate of Average transaction handle time calculated by formula (2) can be found in Figure 5 and Figure 6 [16].

$$\text{Promotion Rate} = \left(\frac{\text{AVG Pre Response Time-AVG Cur Response Time}}{\text{AVG Pre Response Time}} \right) \quad (2)$$

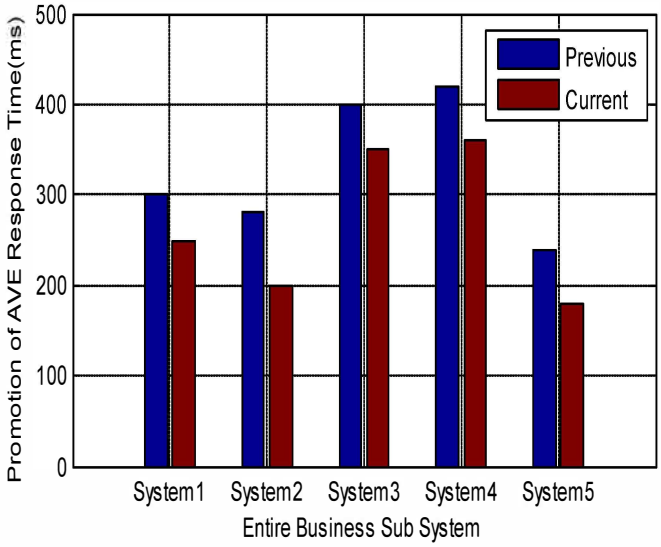

Figure 5. Promotion of Five Core Sub System

It can be calculated from the Figures 5 and 6 that the promotion rate of Average Transaction Handle Time in the five kernel sub system ranges from 12.31% to 29.18%. And the promotion rate of Average Transaction Handle Time in the four assistant sub system ranges from 4.44% to 10.53% less than that in the five kernel sub system for the reason that the collaborative efficiency in assistant sub systems is always lower than the efficiency in kernel sub systems [17].

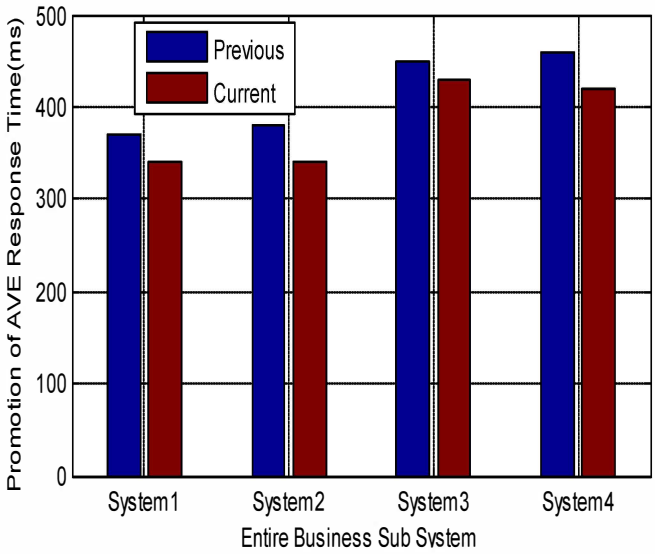

Figure 6. Promotion of Four Assistant Sub Systems

B. Result of software maintaining experiment

In the process of software coupling reconstruction, we decompose these important java classes into smaller java classes for more efficient reusing software component to decrease software maintaining in business logic changing. We tested the software code maintaining rows in time of business logic changing in Figure 7.

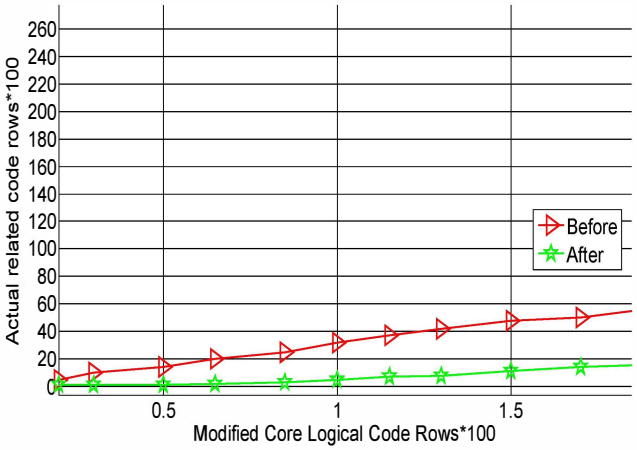

Figure 7. Compare of Logical Code Maintain of Target System

It can be clearly found in Figure 7 that, after class coupling betweenness promotion in target software, the modified code rows decreased. When the core code rows (X axis in Figure 7) changed, the related modified code rows (Y axis in Figure 7) decrease from 25% to 20% after reconstruction. This method is proved really useful in EC software promotion [18].

V. CONCLUSION

In this paper, we proposed a new B2B platform framework software based on cloud computing, which is proved by experiment to promote the transaction processing throughput the rate of business operating system with decreasing average business transaction handle time. Furthermore, the new B2B platform framework software, which is based on cloud computing, can decrease the related modified code rows in business logic changing.

To the best of our knowledge, this is a new study of real EC system construction, which is based on cloud computing and complex network concept. We believe that this research will pave the way for further study of information system construction of EC industry.

ACKNOWLEDGMENT

This work is supported in part by NSFC 61174172, 90924302, 60921061, 70890084, 90920305; CAS 2F11D03, 2F12E03, 2F11N01. Finnish TEKES Cloud Software Program and Finnish TEKLA. R&D Fund.

REFERENCES

[1] http://en.wikipedia.org/wiki/Electronic_Commerce

[2] J. Leskovec, J. L. Kevin , Michael Mahoney.Empirical comparison of algorithms for network community detection. 19th international conference on World Wide Web, pp. 631-640. 2010.

[3] J. Dean, S. Ghemawat. Mapreduce: Simplified Data Processing on Large Clusters , pp.137-150. 2004.

[4] J. Zhou. Modeling and Control of Complex Network. Doctor'sThesis Paper of Wuhan University.2009.

[5] K. Q. He, Y.T.Ma. Software Network.Science Press.Beijing, 2008.

[6] J. Liu.Software architecture analysis and optimation based on complex network. Doctor'sThesis Paper of Wuhan University.2007.

[7] D. Y. Li, H. S. Zhang . Internet Computing Beyond Turing Machine.Journal of CCF，12(5):8-16. 2009

[8] P. Erdös, A. Rényi. On the Evolution of Random Graphs, Publ. Math. Inst. Hung. Acad. Sci., Ser. A 5, pp.17-61. 1960.

[9] D. J.Watts and S. H.Strogatz. Collective Dynamics of "Small World" Networks. Nature, 393 (4): 440-442 , 1998.

[10] [A. L.Bababasi and Albertr. Emergence of Scaling in Random Networks. Science, 286: 5092512 , 1999.

[11] J. Q. Fang,X. F.Wang. DynamicsComplexity research on Non-linear network. progress in natural science,29(1):1-74. 2009.

[12] H. C. Wang. Multi-Granularity Dynamic Characteristic Analysis of Networked Software. Doctor'sThesis Paper of Wuhan University.2010.

[13] C. Kiss, A. Scholz, M. Bichler. Evaluating Centrality Measures in Large Call Graphs. IEEE Joint Conference on E-Commerce Technologyand Enterprise Computing, E-Commerce and E-Services, 2006.

[14] K. Chen, W. M...Zheng,. Cloud Computing:System Instances and Current Research. Journal of Software.vol 5. pp:1337-1348,2009.

[15] Z. W. Xu, H. M. Liao, H. Y. Yu L. Zha. Notes on Classifying Network Computing Systems. Chinese Journal of Computers.vol 9.pp:1509-1515 . 2008.

[16] D.G. Feng, D. Guo, M. Zhang, Y. Zhang, Z. Xu. Study on Cloud Computing Security.Journal of Software. vol. 1.pp:1-132011.

[17] J. W. Han, M. Kamber. Data Mining: Concepts and Technique. The Morgan Kaufmann Series in Data Management Systems 2nd Edition, Morgan Kaufmann Publishers, 2006 .

[18] B.Bollobas. Modern Graph Theory. New York: Springer, 1998.