Brandmap: an Information Visualization Platform for Brand Association in Blogosphere

Amadeu Sa de Campos Filho, Fred Freitas, Alex Sandro Gomes
Informatic Center
Federal University of Permanbuco, UFPE
Recife, Brasil
e-mail: ascf2@cin.ufpe.br, fred@cin.ufpe.br, asg@cin.ufpe.br

Jairson Vitorino
IT Department
Elife Company
Mannheim, Germany
e-mail: jairson@elife.com.br

Abstract - Word-of-mouth in the Internet is rapidly increasing, but the dissemination of this information is sometime difficult to interpret. This paper presents a visualisation platform (BrandMap). It was used to structure data of brand association distribution in the blogosphere as a case study that analyses the word-of-mouth of two smartphone brands in Brazil. BrandMap is a prototype that was developed based on extensive research about word-of-mouth in the blogosphere and how to visualize brands monitoring. BrandMap uses a novel approach for information visualization of complex data. The result was the evaluation of BrandMap analyses that compared the user's perception about the advantages and disadvantages of two brands in social networks.

Keywords— Information Visualization, Blogosphere, Social Network, Semantic Web.

I. INTRODUCTION

Since the introduction of the World Wide Web in the mid-90’s, the volume of unstructured data has increased exponentially. Visualizing attributes semantically structured from data collections like person names, places, zip codes, dates among others, has become progressively more difficult. The challenge is to provide good clues to right questions, which is the key to actual scientific discoveries in many areas. A variety of visualization techniques exist to tackle this problem. The development of techniques for supporting appropriate visualization methods in browsing large amounts of data is relevant in order to help end users to navigate through the knowledge repositories created by processing raw information networks.

Furthermore information visualization and the construction of Graphical User Interfaces (GUIs) is not a trivial task particularly when there are large amounts of dynamic data collections being represented. It involves the following issues:

- The nature of the data: numerical, ordinal or categorical
- Number of data dimensions: univariate, bivariate, trivariate or multivariate
- Structure of the data: linear, temporal, spatial or geographic, hierarchical or network
- Type of interaction: static, transformable or manipulable
- Represented features: instances, concepts, hierarchical relations, discover patterns, or detect important links in information, particularly when the data set is complex and uncertain, according to the relevant users needs. [1]

There are two kinds of information visualization: graphical visualization and textual visualization. Each one use specific domain, but some applications may be understood better using a graphical approach than a textual one. For example, we can choice between the possibilities of an alternative route or the existence of historical buildings in the neighbourhood of the route. The graphical approach, if well applied, might improve the human learning process [2].

Furthermore the potential use of 3D graphics for information visualisation is nowadays a significant unexplored area. The use of 3D graphics not only enhances the design of an information visualization interface but it provides additional information to that of 2D. The challenge of visualizing abstract information in 3D consists in determining effective visual representation and interaction schemes for human analysis and exploration.

However, the most important task for successful visualization is the retrieval of necessary data and the availability of a suitable visualization through metaphors that show the association of data representation to an experience outside the object's environment. Metaphors are very important when it shows real worlds and establishes social interactions, especially in virtual reality [3]. So, the choice of a metaphor is most of the time essential to enhance the usability of a visualization interface. One main problem with many graphical designs is that they have no intuitive interpretation, and the user must learn how to understand them. Metaphors found in nature or in the real world avoid this problem by providing a graphical design that the user already understands.

This paper proposes a visualization platform for the structured data on the web. In particular our focus is on a specific knowledge domain, which we explain in the technological background section of this article, namely brand association distribution in the blogosphere. To validate this proposal we conducted a case study to track and measure the distribution of brands in the blogosphere through visualization maps (BrandMap) for these brands.
in terms of its dissemination throughout the Portuguese blogosphere.

In the following sections, the paper presents a brief description of the blogosphere brands monitoring, the technological background and the main concepts of the BrandMap platform following a case study. The paper then concludes with a discussion of the main contributions and future work.

II. BRAND MONITORING IN THE BLOGOSPHERE

Blogs appeared circa the beginning of the last decade in 2000. It was quickly adopted by millions of people everywhere first as a way to express themselves to friends and family and then evolving to professional writing projects, which actually start providing financial income to the producers. Ten years later social media and social networking is everywhere and although blogs are not hot news anymore, it keeps its steady growth in the web. Technorati [4] accounted for more than 133 millions blogs in 2007.

Marketing managers promptly realized the potential of such a huge resource of data. Many bloggers mentioned their brands, products and services. They discussed actively with their community a new game release, an advertisement television spot or a global recall of automobile parts. These discussions fuelled even more the on-line buzz, sometimes creating huge successes or crisis stories, depending on which types ideas got spread. Suddenly a group of global influential bloggers could make more than a dent on a product release.

Nielsen, a global company performed a study for over 25,000 Internet consumers world wide, ninety percent of consumers surveyed said that they trust recommendations from people they know, while 70 percent trusted consumer opinions posted online [5]. See graph in the figure below.

![Figure 1. Number of trusted consumer opinions posted online](image)

According to figure 1, consumers trusted recommendations from people they knew 90% of the time. This shows the importance of word-of-mouth marketing.

Many start-ups and established companies have started developing tools to handle this vast collection of information in a way that companies can interpret all this data within the minimal time possible. This proved to be a complex task that involved many computer science sub-areas such as text mining, information extraction and web crawling among others.

The aim of our work was the development of innovative visualization techniques to convey information found in the blogosphere about brands. After crawling a collection of blogs for specific terms (i.e. nike, running shoes, etc) we end up typically with a list of annotated posts and comments. By annotations we mean a set of attributes such as the publication and crawling dates, a term frequency list, author name, location (when given), number of incoming links pointing to that blog (extracted via Google API) among others. From this listing one can imagine how complex it can be to render it visually when it is composed of hundreds or thousands of items.

In the next section, we explain a brief description of the technological background of information visualization platform for brand association in blogosphere.

III. TECHNOLOGICAL BACKGROUND

The general architecture of the new information visualization platform for brand association in blogosphere was based on a Framework Proposal for Developing 3D Graphical User Interfaces on Telehealth Systems. [6] (Figure 2)

![Figure 2. Architecture of framework for information visualization in semantics contents](image)

According to figure 2 the architecture is composed by 3 layers:

- **Data layer**: In this layer, the data is stored from the data provider. In this layer it is not important to know where the data come from (Blogs, Sites, Twitter, etc) and how is their structure (XML, database, RDF, etc.)
- **Business layer**: In this layer all the information in the data layer remains untouched, we have to parse this data to right visualization standard that is based on the metamodel proposed. This metamodel must provide all variables to the visualization. To build this transformation, the user must participate, by selecting parameters to be used for the XML Parser in order to construct the nodes (entities).
- **Graphical user Interface (GUI) layer**: In the GUI layer there will be a VisualModel that is the graphical model of information visualization platform. This VisualModel will be created based only on the MetaModel in business layer no matter what kind of data structure is coming from data layer.

In the next section, we explain a particular useful artefact for displaying the described listings in a way that might lead to insights more quickly than simply browsing...
linearly through the data. We introduce the concept of brand association maps.

IV. BRANDMAP: THE VISUALIZATION PLATFORM

BrandMap is a visualization platform for Consumer-Generated Media analysis that delivers insights into how consumers think and talk about a brand in the Blogosphere through the similarity among terms more frequency in the blog post. In the blogosphere there are typically huge collections of terms that when they are alone, they are only row data and have not meanings. However, this row data when connects to others terms, can deliver information but sometimes it is hard to visualize theirs meanings.

The BrandMap architecture is an instance of framework above and is depicted in Figure 3 In data layer, the system will reuse information retrieval and information extraction technology from XML files to populate Brand Association model. After this, in the business layer the populate Brand association model will fill out the system model to be processed the cluster algorithms in order to group the data with the same similarity. Finally in the GUI layer, this data is then rendered on a interface to represent the brand dissemination.

BrandMap goal is to support the better visualization of huge set of connected terms. It is based on Universe metaphor where the main term, the brand term, is the centre of system (Sun) and the others terms that are correlated with the brand term are displayed around the centre like planets (see Figure 4). Beyond the use of the universe metaphor, the visualization platform used physical characteristic of the object (color, size and shapes) to represent several kinds of key brand dimensions, including product attributes, messaging elements, competitive and category sets, and related concepts and themes.

According to the figure 4 we defined that the distance between the centre and the term is inversely proportional to its frequency of the variable that is studied. In the case of brandmap the variable is the number the times that the term is cited in the blog post.

\[
d = \frac{1}{freq(var)}
\]
The angle between the terms around the centre is calculated by a new technique created by [7] that is based on hierarchy clustering technique that analyses their similarities and dendrogram technique (see Figure 5). Dendrogram is a diagram showing the relationships of items arranged like the branches of a tree. It is common interface to use it to show the process of a hierarchical clustering method.

The selection of the terms that will be displayed in the map is correlated to the frequency that these terms appear in relation of the central brand in blogs. The selection technique is based on information retrieval techniques. The terms are grouped in category according to data model of the search engine system and they are stored in an XML file.

A. A Case Study

In September 2010, electronics companies W launched in Brazil a smartphone X to compete with a smartphone Y of a competitor Z. Due to the similarity of features of both devices, the company W made use of various advertising campaigns attacking the weaknesses of its competitor Z. To observe whether the word-of-mouth of the Internet against the advantages and disadvantages of the devices reflected the market's behavior in relation to smartphone X of the company W, the BrandMap platform was used to compare the perception of smartphone X and Y in social networks.

For producing the BrandMaps of brands, we use BuzzMonitor search engine [8] to search for most frequently terms mentioned in the blogs in the Portuguese language related to smartphone X and smartphone Y, from January to September 2010. In a top 50 list of terms, step words of Portuguese language was excluded and selected about 10 to 20 most interesting terms associated with the brands mentioned.

After the terms are selected, they are categorized according to the characteristics that they have, for example, "Jailbreak" and "Operator" related to unlock the smartphone Y, and "Bada" and "Android" related to the comparison of Operating Systems from smartphone X. The analysis of word-of-mouth on blogs is made in the BrandMap by the proximity of terms associated with the central term. The terms closer to the center are those that appear most frequently in postings that have mention to the central brand. Thus the BrandMap platform allows one to develop insights into the opinions of Internet users and their perception of the brands in social media by quickly examining how categories, terms and brands relate to each other on the map.

The BrandMap of smartphone Y is depicted in Figure 6 and proved to be very related to reception problems in its new model, which interfered with if handled by the body of the device. It was necessary the use of a case to avoid the problem. The term "Reception" at the center of the brandMap shows that this was the main point of criticism of the new device on the blogs. The terms "antenna", "failure", "problem" and "case" also deal with the same problem with signal reception.
The exclusive contracts with certain providers brings several testimonies on how to unlock the devices, with tips on how to make the "Jailbreak" in the smartphone Y.

On the other hand in the BrandMap of smartphone X (Figure 7), "Bluetooth" 3.0, which allows a high rate of files transfer including multimedia files, was the functionality most often mentioned. In addition to this feature, the Android operating system, the camera for "videos" in HD, the touchscreen, GPS and memory card 2GB were also mentioned. There were also comparisons with others competing smartphones, especially the iPhone (3G version), Nokia and Motorola. The launch of the iPad tablet also generated high expectations. There are many comparisons especially with the tablet of the American company.

By analysing the map the user find three fundamental elements: terms, categories (discriminate by colors) and their location on their plan. For example, by examining the location of color green in figure 5, the user can for example note that many terms of the category features are aligned together and in particular somewhat separated from other categories. The only exception is Firefox that is very close to touchscreen (maybe people were relating both terms together?). By looking those relations between terms and categories on the 2D space, the user is led to questions which in turn lead to insights. This is the main goal of such a tool: support a market analyst in the search of questions and correspondingly insights.

V. CONCLUSION AND FUTURE WORK

This Paper showed an innovative software platform for information visualization of brand dissemination over the Internet. The results of this platform may provide the means to monitor, analyze and advise clients regarding the brand perception and services, behavior and consumer habits, trends and relationship with the consumer 2.0, among others, by communicating word-of-mouth and public dialogues and spontaneous consumer online line.

The platform may also give to the users a deep understanding how your brand is perceived, discussed and understood among consumers online.

The contributions of this platform are:

- An original platform to visualize large spatio-temporal datasets representing Internet Brand dissemination, including 3D visualization techniques, metaphors and the incorporation of virtual world features when applicable.
- The proposed platform may be used by other researchers interested in investigating communicating word-of-mouth in several disciplines (psychology, sociology, communication sciences, anthropology, etc).

For future work we will evaluate these visualization platforms applying usability questionnaires in order to improve the user interface techniques and to define, which one is the best to apply for the Social Network domain case. Furthermore we will include interaction design methodologies to construct the new interfaces in different contexts as in learning management systems where data represent interaction and collaboration between users.

Finally we will develop a new version of the system including a variety of 3D information visualisation techniques. We expect to provide users with better tools for the visualization of systems that manipulate the so called big data. The incorporation of 3D techniques will contribute to the advance of the state of the art in the field of 3D graphical user interfaces.

ACKNOWLEDGMENT

We thank CAPES/DAAD PROBRAL and the Brazilian National Council for Scientific and Technological Development (CNPq) for financial support. This research was supported by an FACEPE grant (ACE-0158-1.03/12).

REFERENCES