Improving The Quality of The Community Relations Knowledge Using Implicit Connections

Tiago Lopes Ferreira
Instituto Superior Técnico
Lisbon, Portugal
tiagohenrique@ist.utl.pt

Alberto Rodrigues da Silva
Instituto Superior Técnico
Lisbon, Portugal
alberto.silva@acm.org

***Abstract*—This paper presents a research that looks for discovering implicit community-based relations by identifying connection degrees and shared patterns between individuals. It introduces the Newsletter Tracking System (NTS) that gathers individuals' implicit information and supports a Connection Degree Model (CDM) that identifies the grain of connection between individuals. Also, based on that, it can foster the quality of community relations. Still, this paper describes the Nano-Tera.ch case study where this research was developed and evaluated with.**

Keywords-online community; implicit connection; connection degree; newsletter tracking system

I. Introduction

The study of social interaction and its value as information collector started in 2005 when Tim O'Reilly introduced the concept of Web 2.0 [12]. For the first time in History, the power of producing and sharing information was available to all who have access to the Internet. Internet became the main chain providing infrastructure for people to collaborate, coordinate, and cooperate to each other [5]. Physical barriers were overcome and communities had the chance to go virtually by transforming physical interactions into virtual connections. Whereas conventional interactions had physical constraints, non-conventional interactions were able to eradicate those barriers and bring people together.

With communities growing virtually, the concept of "virtual communities" on "online communities" came to life in 1993 [9] as well as its ability to trace interactions. While tracking individuals' connections was extremely difficult in the physical world, it became an easy target in the virtual sphere. People got close with technology as well as their own interests, curiosities, hobbies, and professions. Individuals become increasingly exposed to the virtual world, which easily captures their information and moves it into online profiles. Any interaction can be stored as a connection between an individual and something or someone. Moreover, a connection can be described as an explicit connection when it is fully and clearly expressed by the individual (e.g. his address) or an implicit connection, every time it is implied and not expressed (e.g. individuals' interests).

Both connections are useful to understand a community structure and also to improve the quality of the knowledge established or implied on the relations. Although explicit connections are limited to what individuals' express (e.g. a direct invitation for a friend relationship, as in Facebook), the implicit ones are based on individuals' interactions within their virtual exploitation (e.g. clicking a link). The problem arises when it comes to capture and analyze individuals' implicit connections at the rate they grow: data management becomes difficult to handle and implicit patterns harder to find. As a consequence, "the connection between individuals, groups, and information becomes lost, or forgotten, and individuals and groups become more isolated" [10].

In the presented case-study of Nano-Tera.ch [14], a scientific community at Switzerland, the goal of exploring implicit connections was difficult to achieve due to the increase number of interactions and the higher complexity on capturing researchers' interactions. Nano-Tera.ch was interested in understanding how its own community was implicitly organized but had no way to capture and promote interactions between the researches. Because the community was large and the number of implicit interactions was difficult to manage, Nano-Tera.ch needed an automatic tool that was able to track those connections among researchers and organize them into results. This way, it would be easier to understand the community's relations and improve the quality of the information targeted at the community.

On the other hand, it is important to classify the connections among individuals in order to understand how connected individuals are and how important a connection is regarding the universe of discovered connections. At Nano-Tera.ch, the governing bodies had the goal of highlighting connections by a common degree in order to measure their importance in the network.

The research is based on the NTS to capture implicit connections between individuals and on the CDM to measure the connections' strength. The paper starts by presenting the NTS main features and the CDM as a tool to measure connections. At the end, it is presented the evaluation at Nano-Tera.ch based on the statistical results and the discussion on how discovered connections can improve the quality of a community knowledge, on individuals' connections.

II. The Newsletter Tracking System

Online communities enable people to stay connected without geographical restrictions. People can communicate, collaborate, and share content with no need of face-to-face interaction. The facilities brought by technology, such as messaging, web portals, and electronic mail are very

978-0-7695-4777-0/12 $26.00 © 2012 Crown Copyright
DOI 10.1109/QUATIC.2012.32

powerful tools to allow individuals to communicate between them. As there use become part of people's social life, the number of interactions increases and thereby its usefulness becomes clear and easily to exploit. In fact, the knowledge about individuals' connections can be improved based on their virtual interactions.

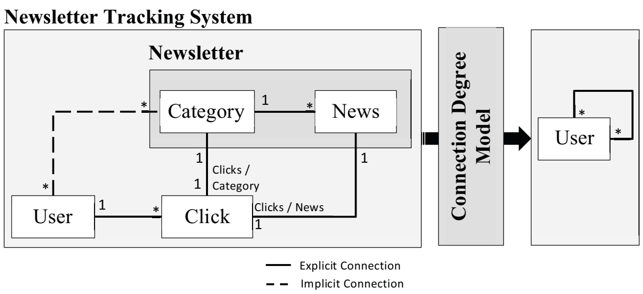

Figure 1. Newsletter Tracking System Main Concepts

The NTS' "Newsletter" (see Figure 1.) consists in a set of "News" linked to "Categories" to which "Users" are related through "Clicks". A "User" is able to interact with a "Newsletter" by clicking on the "News". On each "Click" the "User" is implicit related with the "News" and its "Category". In addition, the "User" can explicit express his preferences on a "Category" and be explicitly linked to it.

Once the relations are gathered, the community can use the CDM to discover implicit relations between "Users" and assign them a connection degree value. This process converts the relations Individual-Link and Individual-Category into relations Individual-Individual with a CD.

A. Defining and Sending Newsletters

The NTS is an online tool that uses email technology and PHP scripting to send and collect individuals' information. The core engine is responsible for preparing the newsletters and gathering the information from individuals' interactions. A community interested on improving their knowledge on individuals' connections can build their own newsletters and use the NTS to send, capture, and process the data. On the other side, individuals can receive the newsletters on their email addresses and share them with online social networks. NTS is also able to capture the interactions that come from the outside, e.g. shared emails, social networks, and online newsletter [11].

The process of building a network of implicit connections between individuals is break down into several stages in which information is treated in different ways. Initially communities are the main player once they are the best to know what kind of news their community would like to receive and to share. Thus, communities are responsible for creating their own newsletters, with the only restriction as HTML format, and to upload them into the system. After that, the NTS will track the links in order to capture individuals' interaction and store the newsletters on the server. At this stage the news are typed - "text" or "image" - and categorized. News can be typed as "text" if it represents a text block or "image" if it symbolizes an image. The symbolism allows communities to understand to what kind of news individuals interact more.

On the other side, categorization is based on the news' categories. Each news should be connected to a category that is previously defined by the community. This categorization is critical to improve a community's knowledge about individuals' interaction once it is used to create a network of implicit connections. A community should define the subjects according to what it wants to study about their community. For example, if a community of researchers wants to study its researchers' interactions it would be a good idea to define the newsletters' categories the same as its research fields. Once a community properly defines its categories and associates them with the newsletters' news, the system is able to bring to light hidden relationships, share patterns, and help to improve the quality of the knowledge on individuals' relationships.

The second stage is related with the process of sending the newsletters to the community. At this level communities are responsible for selecting the list of individuals, with whom they share the newsletter, and gather their email addresses into an input file. The NTS uses electronic mail to spread the newsletter across the contacts and enable communities to include individuals' first and last names to personalize the newsletter.

B. Tracking User Interactions

When a community includes all the input information the system is able to start the process of preparing the newsletter to be sent. Once the translation for the news in the NTS is the URL, every news is uniquely related with a hyperlink. This way an interaction on the news can be gathered by tracking the respective link. For all of this, the links in the newsletter are redefined through the standard URL "`base-url/code/user-id/link-id/newsletter-id/`", where "`base-url/`" is a common prefix to all the links, namely the path to the server where the NTS is working. The "`code/`" defines one of the four possible actions: open the email, click the news, share the newsletter, and see it online. The "`user-id/`" identifies the individual who trigged the action, "`link-id/`" recognizes which link was clicked, and "`newsletter-id/`" the newsletter in what the action was performed. However, due to the variable "`user-id/`", the task of redefining the links is made for each email address. At the end NTS creates a unique newsletter per receiver.

The last stage is described as the most important once individuals are the main players. After they receive the newsletter on their email address the process of gathering interactions' information only boots when the action of clicking the news takes place. For each click an individual performs he is redirected to the track engine that is hosted in a server and has the sources to convert the tracked link into the real link and forward the individual. During this process the NTS engine stores all the information listed above and also the variables of time and geographical location. Once the individual is forward to the real link the NTS loses the track on the individual and no more information is stored.

C. Data Analysis

One of the main features about the proposed solution is its statistical and data analysis purpose. Once the data on individuals' interactions is stored, the analysis feature allows communities to study individuals' connections and discover implicit relationships between them. Thus NTS enables the data to be organized in several ways and uses JpGraph library [13] to present it into online visuals. The analysis starts by presenting the total number of clicks, of individuals that have clicked, the amount of individuals that opened the newsletter email, as well as the total number of individuals to which the newsletter was sent, and the time passed.

The analysis is based on individuals' interactions with the news from several newsletters. The first analysis describes the way news can be presented in a newsletter - "text" or "image". Each action on the news takes into this representation in order to allow communities to understand the best approach to design their newsletters. If a community interacts more with news based on images maybe it would be a good idea to design an image-based newsletter. The goal is to understand in what way newsletters' quality can be evolved in order to increase individuals' interest.

The following analysis is crucial for the NTS to discover implicit relationships among individuals. Based on news' categorization, the analysis is performed through the relation between the news and categories. In fact, besides the importance of categorization to discover which subjects individuals interact more it is relevant to group individuals by category and understand how related they are. Moreover the relation through categories is also used to calculate the connection degree between individuals.

The last analysis is related with communities interested in social networking. NTS allows individuals to share the newsletters within online social networks, such as Facebook, twitter, and digg, and it is able to capture any interaction that comes from those social networks. When an individual decides to share the newsletter a new one is created in order to identify the social network and the individual that share it. The goal is to allow communities to share their content in order to reach other communities through online sharing.

At this stage communities are able to study their individuals' interactions and improve their knowledge on connections through information collected and data analysis. In addiction NTS allows communities to extract the data on individuals' interactions in order to give them the freedom to work the data the way they want.

III. CONNECTION DEGREE MODEL

Once information on individuals' connections increases it is important to understand how relevant discovered connections are and how they can organize by a value in the network. The presented solution tries to answer this need by proposing a model named "Connection Degree Model". This model uses the two types explicit and implicit connections as input values to calculate the connection degree for every two related individuals. Both connections are gathered to calculate the explicit and implicit degree in the final equation, respectively.

The explicit degree is based on individuals' preferences on the categories that a community has for its newsletters. An individual can express his interests by explicitly checking his preferences in the categories. Thus, a checked category is understood as a positive preference, an unchecked category as a negative preference, and with no value if no explicit action is performed. Then the explicit degree will affect the final connection degree result through a "category importance" (ci) value defined by the community and calculated for each individual. Moreover, the value represents the importance for a community about this explicit behavior in the connection degree calculation.

$$explicit\ degree\ (individual){:}\ ed(I) =$$

$$\begin{cases} ci, & positive\ preference \\ 0, & no\ preference\ shown \\ -ci, & negative\ preference \end{cases} \qquad 0 \leqslant c_i \leqslant 1 \quad \text{(a)}$$

On the other hand implicit degree is calculated based on the individuals' interactions with news and categories. The action of clicking the news is stored as an implicit connection between the individual and the news. Once the news is linked to a category the action also implicitly relates the individual with the news' category. Although the implicit relations on news and categories are worked in differently, both have a common strength in the final connection degree. The implicit relation between individuals and news is organized into a matrix $n{\times}m$ where both rows and columns represent all individuals, and the elements the total number of news that both individuals have clicked in common.

$$n{\times}m = \begin{bmatrix} TC_{11} & \cdots & TC_{1m} \\ \vdots & \ddots & \vdots \\ TC_{n1} & \cdots & TC_{nm} \end{bmatrix} \text{(b)}$$

As regards the implicit connections with categories, individuals are firstly organized into a matrix where rows represent all combinations of two individuals and columns the categories. For the pair (I_i, I_j) and the category c_l the implicit value is represented by the following equation:

$$(I_i, I_j){\times}c_l = min(a_{il}, a_{jl}) * (ed_i + ed_j) \text{ (c)}$$

Where a_{il} represents the total number of clicks that the individual i gave in the category l and ed_i the explicit degree for the individual i. Moreover, the value for the explicit degree is included in this equation once the equation measures the level of connection between individuals and categories. The calculation ends by also organizing individuals into a matrix $n{\times}m$ where rows and columns represent individuals and values the sum of all (c) equations for all categories l in the newsletter.

$$n{\times}m = \begin{bmatrix} \sum_{i=0}^{l} (I_1, I_1){\times}c_i & \cdots & \sum_{i=0}^{l} (I_1, I_m){\times}c_i \\ \vdots & \ddots & \vdots \\ \sum_{i=0}^{l} (I_n, I_1){\times}c_i & \cdots & \sum_{i=0}^{l} (I_n, I_m){\times}c_i \end{bmatrix} \text{(d)}$$

The final value for the connection degree between individuals is calculated by performing a syntax sum of both matrix (c) and (d) but using the multiplication operator to calculate the final elements' value. At the end only one of the sides in the matrix is taken into account, called as lower or upper triangular matrix, and excluded the diagonal.

$$connection\ degree : n \times m =$$

$$\begin{bmatrix} (b)_{11} \times (d)_{11} & \cdots & (b)_{1m} \times (d)_{1m} \\ \vdots & \ddots & \vdots \\ (b)_{n1} \times (d)_{n1} & \cdots & (b)_{nm} \times (d)_{nm} \end{bmatrix} \quad \text{(e)}$$

For the last operation it was used the multiplication operator so a click has a significant importance. Once the system is based on newsletters and the number of interactions is not high, each click should have a significant value so it can influence the connection degree between two individuals. The proposed model on connection degree tries to assign to each connection a value that enables communities to measure their universe of discover connections and understand which are more important.

IV. Results

The evaluation was done inside Nano-Tera community [14], a scientific Swiss federal program with more than 40 projects. Nano-Tera projects' diversity spans from Health to Security and Environment where take part hundreds of researchers around the world. Moreover, Nano-Tera had a goal of improving their knowledge on researchers' interactions by bringing to light hidden relationships.

During 6 months the system was evolved and 4 newsletters were sent. However, the presented results are relative to the last 2 newsletters, during 2-months of evaluation, where the newsletters were sent to the Nano-Tera community as well as a set of outside communities, such as universities and media contacts. In the end, there were collected almost 300 clicks from around 200 researchers and interested people of the 280 that opened the newsletters.

A. News' Type

Nano-Tera community showed that its community is strongly text oriented with more than 96% of interactions on text news and 4% on news linked to images. The results can be related with the fact that Nano-Tera is a strong scientific-base community and with the reduced number of images news regarding the total number of news attached to text.

B. News' Categories

The results showed that Nano-Tera community is strongly interested in the three main categories of "Health" (38%), "Environment" (27%), and "Security" (22%). The results showed that the community's interest is well balanced between the projects. Moreover, it appears a category on "Nano-Tera" (8%) with almost all interactions coming from outside Nano-Tera community.

C. Clusters Detection

The results on community detection were performed using Vizster [6] that applies the Girvan-Newman algorithm on graph data. In this case, NTS is responsible for creating a graph where individuals represent the nodes and implicit connections the edges between them.

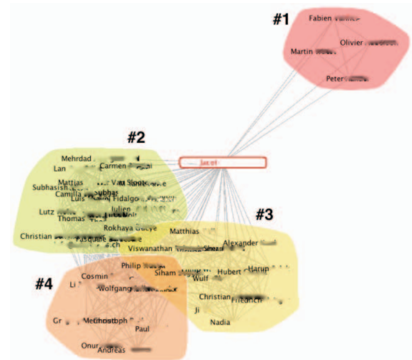

Figure 2. Groups detection with Vizster.

The results reveal that Nano-Tera community can be break into 4 different clusters: Cluster #1 is based on individuals with interest on "Nano-Tera" category, where outside people are linked; Cluster #2 is strongly connected on "Environment"; and the Cluster #3 and #4 are highly related with "Security" and "Health" respectively.

D. Connection Degree

To better illustrate the results on the CDM it was selected a sample of 3 individuals and 5 news, linked through 3 categories. The NTS allows individuals to explicit expressed their preferences on categories and assigns them a value for their explicit degree as a pair (as shown bellow).

TABLE I. Individuals' Explicit Interests

Individuals' Preferences	Category		
	C_A	C_B	C_C
I_1	1[a]	0	0
I_2	0	0	0
I_3	1	0	-1

×

Category Importance	Category
	All
CI[b]	.4

a. $\begin{cases} -1, & \text{negative preference} \\ 0, & \text{no preference is shown} \\ 1, & \text{positive preference} \end{cases}$

b. Defined by the community

TABLE II. Explicit Degrees Between Individuals

Explicit Degree	Category		
	C_A	C_B	C_C
I_1-I_2	.4[c]	0	0
I_1-I_3	.8	0	-.4
I_2-I_3	.4	0	-.4

c. I_1 preference on C_A + I_2 preference on C_A

On the other hand, individuals are also implicit connected through their interactions with news and categories. In the first relation individuals are connected by the total number of news that they have clicked in common, and in the second thought the total number of clicks per category.

TABLE III. Clicks per Individual per News

Number of Clicks	News					
	N_1	N_2	N_3	N_4	N_5	N_6
I_1	1	1			1	
I_2		1		1	1	1
I_3	1	1	1			1

TABLE IV. Relation Between Individuals through News

News in Common	Individual		
	I_1	I_2	I_3
I_1		2	2
I_2	2		2
I_3	2	2	

The results are then related with the outcomes of the relation between individuals thought categories. Initially individuals are individually related to categories and then related as pairs where their explicit degree is taken into account. At the end individuals are related to each other by their degree of connection with categories. With the following categories per news – C_A for N_1; C_B for N_2, N_3, and N_4; C_C for N_5 and N_6 – it got the following results.

TABLE V. CLICKS PER INDIVIDUAL PER CATEGORY

Number of Clicks	Category		
	C_A	C_B	C_C
I_1	1	1	1
I_2		2	2
I_3	1	2	1

TABLE VI. RELATION BETWEEN INDIVIDUALS AND CATEGORIES

Individuals × Categories	Category		
	C_A	C_B	C_C
I_1-I_2		1 [e]	1
I_1-I_2	1.8	1	.6
I_2-I_3		2	.6

e. Min[TableVI(I_1 , C_A) , Table VI(I_2 , C_A)] + TableII(I_1-I_2 , C_A)

TABLE VII. RELATION BETWEEN INDIVIDUALS THROUGH CATEGORIES

Relation Between Individuals	Individuals		
	I_1	I_2	I_3
I_1		2 [f]	3.4
I_2	2		2.6
I_3	3.4	2.6	

f. $\sum_{i=A}^{C}$ TableVII($I_1 - I_2$, C_i)

With both implicit degree values on news and category calculated, the final connection degree can be obtained by their multiplication.

TABLE VIII. CONNECTION DEGREE

Individuals' Relations	Connection Degree
I_1-I_2	4 [g]
I_1-I_2	7
I_2-I_3	5

g. Table IV(I_1 , I_2) × Table VIII(I_1 , I_2)

V. DISCUSSION

The discussion explores the results on Nano-Tera case study and proposes some of the ways to use the information for the benefit of any other community. On news' types, the results show that Nano-Tera community is text based even when the newsletter is well divided on text and images news. Seen as an expected result, Nano-Tera proved that its community is very scientific focus and its researchers prefer receive their newsletters based on text news than image. In a general application the results can be applied to understand how to better fit newsletters' design to increase individuals' interest and therefore the number of interactions.

Results on news' categories show that the community's interest on Nano-Tera categories is well balanced. Although "Health" category has higher preference, the difference is not significant and can be explained by the larger number of projects. As an explanation for this surprisingly result can be the equal division of the categories in the newsletter or the proper organization of the human resources by projects at Nano-Tera. On the other hand, results may be influenced by the fact that Nano-Tera was presenting new projects in all categories to the community and updating them with hot news. This attitude can motivate individuals to read about every new project and thus categories. Regarding "Nano-Tera" category it is interesting to notice that most of the interactions came from outside people, revealing people's interest about Nano-Tera's projects. As a general use, the data can be used to understand how a community is organized and what categories individuals prefer.

On clustering detection it is clear that Nano-Tera community is highly divided into 4 main clusters where individuals are implicit related through news and categories. Its organization by size shows that Cluster #2 is the biggest, followed by Cluster #3, #4, and #1, with their main categories on "Environment", "Security", "Health", and "Nano-Tera". It is interesting to see that even with "Health" being the main category in clicks, it is on "Environment" where people are more connect-ed. This shows that people that have clicked "Health" category have also clicked "Environment". The results reveal that people interested in "Health" have also a strong interest on "Environment".

It is also possible to discover what individuals have a high interest on Nano-Tera community by looking at the nodes that link the different groups and are placed on their intersection. In the middle Nano-Tera can see what is the most influence individual in the community. For a general purpose, the data on group detection can be used to understand how a community is implicit organized and what individuals are key connections. An example at Nano-Tera was the use of the implicit groups to predict the number of attendances in conferences or workshops in a category.

The connection degree is crucial to measure how strong individuals are connected. Although clustering detection allows the visualization of individuals' connections, it is important to realize the grain of connection in order to identify the most important ones and have the chance to work the connections based on a common degree. The results show that the relation between the individuals I_1 and I_3 is the strongest with a value of 7. Once interactions can grow and therefore the connection degree, there is no maximum value to compare with. The higher the value, the stronger is the relationship.

Although the number of clicks on news is the same for both relations I_1-I_2 and I_1-I_3, the last one has a stronger connection degree thanks to individuals' explicit interest on C_A and to their implicit connection on C_C. Also the connection degree for I_2-I_3 is higher than the connection for I_1-I_2 due to individuals' implicit connection on C_B. Even with the same number of clicks, individuals I_2 and I_3 have a strong connection on the C_B due to the total number of clicks on news on C_B. In fact, the connection degree could have been higher if I_3 had not expressed a negative preference on C_C.

The values on the connection degrees show that despite all individuals are related their relations' strength can be different as well their value in the network. For the presented sample it can be concluded that the most important connection is the relation between I_1 and I_3. In a general use, the connection degree model can be used to discover which connections are crucial in a community.

VI. Related Work

In 2002, Ducheneaut and Bellotti commented that "even people having offices next to each other, still use email as a principal communication medium" [15]. Additionally, previous work have also used email interaction has a source to discover implicit relations. In 1993, Schwatz and Wood, used email headers to extract shared interests between people using graph theory [8]. However, this approach is very limited to the emails' headers and does not take into account the message's body and the subject, which can be the richest source on individuals' interests.

PeCo collected users' relationship through email headers (From, To, and Subject), but like previous solution does not focus on discovering and explore implicit connections as a network [4]. On the other hand, McArthur and Bruza discover implicit connections by mining semantic associations from people's communications [10]. They proposed a model called HALe that automatically creates a dimensional representation of words based on the email corpus and uses it to discover a network of people implicitly connected. However, that solution does not have any measure to place connections at different levels and compare them.

Together with the referred works there are several track engines used for marking purposes, which are also able to send newsletters and track users' interactions. An example is the system developed by Foulger et al [7]. This system generates an email template and uses it to create a custom email for each email target, who receives it with a custom URL in order to identify the source of interaction. However, besides their strong focus on business, the detection of key connections through a connection degree model can be harder to achieve or difficult to understand once the model works in a black box.

Barão and Silva proposed an holistic and complex model to define the Relational Capital Value (RCV) of organizations as well as online communities [1][2]. Explicit but also implicit relational connections (such as these discovered by the NTS system) are important for the RCV model application, hence for the determination of online communities relational value.

VII. Conclusion

The NTS allows communities to improve the quality of their knowledge related individuals' relationships gathering from individuals' interactions on a common newsletter. Based on implicit connections NTS is able to bring to light hidden relationships and to measure the degree of these connections by proposing a "Connection Degree Model".

Information on individuals' explicit connections is gathered through their preferences on newsletters' categories and used in the connection degree model together with an importance value defined by the community. On the other hand, implicit connections are collected via interactions with newsletters' news. When a user clicks the news, the action is stored as an implicit connection to the news and its category.

The data on individuals' relations and the application of the connection degree model allow communities to better understand the way individuals are related according to their actions and preferences. Using Vizster software, a community can be break down into groups and individuals with a strong interest can be highlighted. Moreover, the connection degree model supports the results by assigning to the connections a value that allows connections to be compared to each other.

We believe that NTS and its connection degree model can help communities to explore and study individuals' connections and thus improve their knowledge on interactions.

Acknowledgment

This work was supported by the Strategic Executive Committee of Nano-Tera.ch which is a program financed by the Swiss Government. Special thanks to all Community Knowledge Development team Dr. Peter Bradley, Dr. Nitesh Khilwani, and Madhur Agrawal. Also to Prof. Chris Tucci from CSI/EPFL who supported and trigged the project and to Mariana Araújo who helped on the newsletters' design study.

References

[1] A. Barão, A. Silva, "How to value and monitor the relational capital of knowledge-intensive organizations?", Research on Enterprise 2.0: Technological, Social, and Organizational Dimensions, IGI Global, 2012.

[2] A. Barão, A. Silva, "A Model to Evaluate the Relational Capital of Organizations (SNARE-RCO)", Conference of Enterprise Information Systems (Centeris'2011), Springer, 2011.

[3] C. Aggarwal, "An introduction to social network data analytics", Springer Science And Business Media, LLC 2011.

[4] H. Ogata, Y. Yano, "Collecting oganisational memory based on social networks in collaborative learning".WebNet, pp. 822-827, 1998

[5] J. Harley, N. Blismas, "An Anatomy of Collaboration Within the Online Environment", Springer-Verlag Berlin Heidelberg 2010, 14-34.

[6] J. Heer, D. Boyd, "Vizster: Visualizing Online Social Networks", 2005 IEEE Symposium on Information Visualization, 2005.

[7] M. Foulger, T. Chipperfield, J. Cooper, A. Storms, "System and method related to generating and tracking an email campaign", IC Planet, 2006.

[8] M. Schwartz, D. Wood, "Discovering shared interests among people using graph analysis of global electronic mail traffic", Communication of the ACM, 1993

[9] P. Zaphiris, C. Ang, Social Computing and Virtual Communities, Chapman and Hall/CRC, 1 edition, 2009.

[10] R. McArthur, P. Bruza, "Discovery of implicit and explicit connections between people using email utterance", Kluwer Academic Publishers, 2003, 21-40.

[11] T. Ferreira, A. Silva, P. Bradley, K. Nitesh, M. Agrawal. "Discover Implicit Relationships Between Researchers Using Email Tracking", CAPSI 2011.

[12] T. O'Reilly, "What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software", O'Reilly Media, Sebastopol (CA) USA, 17-37.

[13] URL: http://jpgraph.net/

[14] URL: http://www.nano-tera.ch/topdownbottomup/index.html

[15] V. Bellotti, N. Ducheneaut, Howard, M., Smith, I., E. Grinter, "Quality Versus Quantity: E-Mail-Centric Task Management and Its Relation With Overload. Human-Computer Interaction", vol. 20, pp. 89-138. , Lawrence Erlbaum Associates, Inc., 2005.