Evaluation of Node Position Based on Mutual Interaction in Social Network of Internet Users

Przemysław Kazienko[1], **Katarzyna Musiał**[1], **Aleksander Zgrzywa**[1]

Abstract: A very interesting scientific problem is the assessment of the node position within the directed, weighted graph that represents the social network of Internet users. The weights of graph arcs are extracted from the data about user mutual communication or common activities. The new method of node position analysis, which takes into account both the strength of the connections between network nodes and dynamic of this strength is presented in the paper. The results of experiments on email dataset were described as well.

Keywords: node position assessment, social network of Internet users, social network analysis

1. Introduction

The various kinds of e-commerce and e-business solutions that exist in the market encouraged the users to utilize the Internet and available web-based services more willingly in their everyday life. Many customers look for services and goods that have high quality. Thus, not only the information provided by vendors is important for potential customers but also the opinions of other users who have already bought the goods or used the particular service. It is natural that users, to gather other people opinions, communicate with each other via different communication channels, e.g. by exchanging emails, commenting on forums, using instant messengers, etc. This information flow from one individual to another is the basis for the social network of Internet users (*SNIU*). This network can be represented as a directed graph, in which nodes are the users and the edges describe the information flow from one user to another. One of the most meaningful and useful issue in social network analysis is the evaluation of the node position within the network. Since the social network describes the interactions between people, the problem of assessment the node position becomes very complex because humans with their spontaneous and social behavior are hard predictable. However, the effort should be made to evaluate their status because such analysis would help to find users who are the most influential among community members, possess the highest social statement and probably the highest level of trust (Golbeck, Hendler, 2004), (Rana, Hinze, 2004). These users can be representatives of the entire community. A small group of key persons can initiate new kinds of actions, spread new services or activate other network members (Kazienko, Musiał, 2007). On the other hand,

[1] Intitute of Applied Informatics, Wroclaw University of Technology, Wyb. Wyspianskiego 27, 50-370 Wroclaw, Poland
e-mail: {kazienko,katarzyna.musial,aleksander.zgrzywa}@pwr.wroc.pl

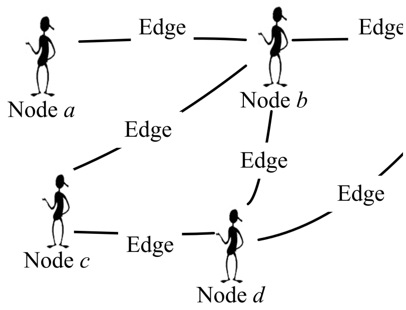

Figure 1. A regular social network

users with the lowest position should be stimulated for greater activity or be treated as the mass, target receivers for the prior prepared services that do not require the high level of involvement. In order to calculate the position of the Internet user, the new measure called node position is introduced in the further sections. It enables to estimate how valuable the particular node within the *SNIU* is. In contrary to the PageRank algorithm that is designed to assess the importance of the web pages, the presented node position measure take into account not only the significance of the direct connections of a node but also the quality of the connection.

2. Related Work

The main concept of a regular social network (Figure 1) appears to be simple as it can be described as a finite set of nodes that are linked with one or more edges (Garton, Haythorntwaite, Wellman, 1997), (Hanneman, Riddle, 2005), (Wasserman, Faust, 1994). A node of the network is usually defined as an actor, an individual, corporate, collective social unit (Wasserman, Faust, 1994), or customer (Yang, Dia, Cheng, Lin, 2006) whereas an edge named also a tie or relationship, as a linkage between a pair of nodes (Wasserman, Faust, 1994). The range and type of the edge can be extensive (Hanneman, Riddle, 2005), (Wasserman, Faust, 1994) and different depending on the type and character of the analyzed actors. The notation that is widely used to represent a social network is the graph theory. The nodes of a graph are actors while edges correspond to the relations in the social network (Wasserman, Faust, 1994). The social networks of Internet users somewhat differ from the regular ones and because of that they yield for new approaches to their definition and analysis. *SNIU* is also called an online social network (Garton, Haythorntwaite, Wellman, 1997), computer-supported social network (Wellman, Salaff, 1996), web community (Gibson, Kleinberg, Raghavan, 1998), (Flake, Lawrence, Lee Giles, 2000), or web-based social network (Golbeck, 2005). Note that there is no one coherent definition of *SNIU*. Some researchers claim that a web community can also be a set of web pages relevant to the same, common topic (Gibson, Kleinberg, Raghavan, 1998), (Flake, Lawrence, Lee Giles, 2000). Adamic and Adar argue that a web page must be related to the physical individual in order to be treated as a node in the online social network. Thus, they analyze the links between users' homepages and form a virtual community based on this data. Additionally, the equivalent social network can also be created from an email communication system (Adamic, Adar, 2003). Others declare that computer-supported social network appears when a computer network connects people or organizations (Garton, Haythorntwaite, Wellman, 1997), (Wellman, Salaff, 1996). On the other hand, Golbeck asserts the view that a web-based social network must fulfil the following criteria: users must explicitly establish their relationships with others, the system must have explicit support for making connections, relationships must be visible and browsable (Golbeck, 2005).

Social network analysis (Wasserman, Faust, 1994) provides some measures useful to assess the node position within the social network. To the most commonly used belong: centrality, prestige, reachability, and connectivity (Hanneman, Riddle, 2005), (Wasserman, Faust, 1994). There exist many approaches to evaluation of person centrality (Freeman, 1979): degree centrality, closeness centrality, and betweeness centrality. Degree centrality takes into account the number of neighbors that are adjacent from the given person (Hanneman, Riddle, 2005). The closeness centrality pinpoints how close an individual is to all the others within the social network (Bavelas, 1950). It tightly depends on the shortest paths from the given user to all other people in the social network. The similar idea was studied for hypertext systems (Botafogo, Rivlin, Shneiderman, 1992). Finally the betweeness centrality of a member specifies to what extend this member is between other members in the social network (Freeman, 1979). Member a is more important (in-between) if there are many people in the social network that must communicate with a in order to make relationships with other network members (Hanneman, Riddle, 2005). The second feature that characterizes an individual in the social network and enables to identify the most powerful members is prestige. Prestige can be also calculated in various ways, e.g. degree prestige, proximity prestige, and rank prestige. The degree prestige takes into account the number of users that are adjacent to a particular user of the community (Wasserman, Faust, 1994). Proximity prestige shows how close are all other users within the social community to the given one (Wasserman, Faust, 1994). The rank prestige (Wasserman, Faust, 1994), is measured based on the status of users in the network and depends not only on geodesic distance and number of relationships, but also on the status of users connected with the user (Katz, 1953).

3. Evaluation of Node Position Based on Mutual Interaction in Social Network of Internet Users

Before the new method for node position measurement will be presented the definition of social network of the Internet users should be established.

3.1. Social Network of Internet Users

The various kinds of definitions of the social network of Internet users (see Section 2) yields for the creation of one consistent approach.

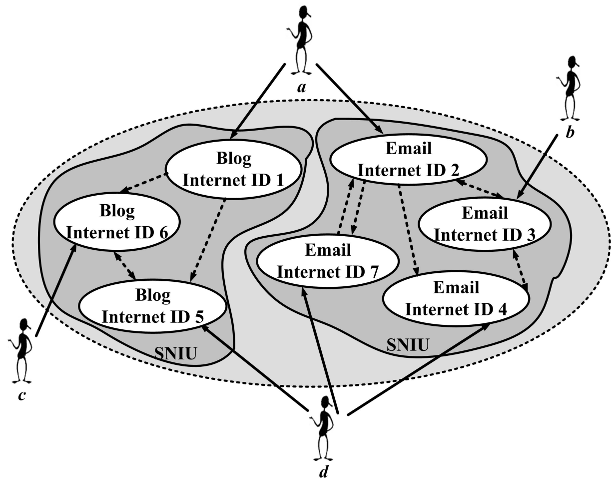

Figure 2. Two social networks of Internet users

Definition 1. *Social network of Internet users is a tuple SNIU=(IID,R), where IID is a finite set of non-anonymous internet identities i.e. the digital representation of a person, organizational unit, group of people, or other social entity, that communicate with one another or participate in common activities, e.g. using email system, blogs, instant messengers.R is a finite set of internet relationships that join pairs of distinct internet identities: R:IID×IID, i.e.* $R = \{(iid_i, iid_j) : iid_i \in IID, iid_j \in IID, i \neq j\}$ *and* $(iid_i, iid_j) \neq (iid_j, iid_i)$*. The set of internet identities IID must not contain isolated members – with no relationships and card(IID)*> 1.

The example of two separate social network of Internet users is presented in Figure 2. Note that an individual human can simultaneously belong to many social networks in the Internet. Moreover, they can also maintain several Internet IDs – see person d in Figure 2. The internet identity is a digital representation of the physical social entity. These are objects that can be unambiguously ascribed to one person (individual identity), to a group of people or an organization (group identity). This representation must explicitly identify the social entity (a user, group of users or an organization). This mapping enables to define the connections between social entities based on the relationships between their internet identities. An individual identity possesses individuals, whereas a group identity corresponds to a group of people, e.g. family that use only one login to the family blog, as well as to an organization, e.g. all employees use one e-mail account to respond customers' requests. Such group identities can by identified by content analysis.

A relationship connects two internet identities based on their common activities. Every social entity that is represented by the internet identity can be conscious of such relationship or not, depending on the profile of activities. Three kinds of social relationships can be distinguished: Direct relationship – it connects two internet identities with a direct connector. The direct connector is an object that is addressed to the specific type of internet identities and related communication, e.g. email addresses (internet identities) are connected with messages exchanged among them. Thus, the direct connector can be email communication, phone calls (or VoIP), etc. Quasi–direct relationship – two internet identities are aware of the fact that they are in the relationship but they do not maintain the relationship, e.g. people who comment on the same blog. Indirect relationship – the internet identity is not aware of the fact that is similar to other internet identity. Two internet identities are connected by indirect relationship when their profiles are similar, e.g. people who examine and similarly rate the same photos published in the Internet. The examples of *SNIU* based on the established definition are: a set of people who date using an online dating system (Boyd, 2004), a group of people who are linked to one another by hyperlinks on their homepages (Adamic, Adar, 2003), the company staff that communicate with one another via email (Culotta, Bekkerman, McCallum, 2004), (Shetty, Adibi, 2005), etc.

3.2. Node Position Evaluation

Based on the data derived from the source system, we can build a graph that represents the connections between users and then analyze the position of each node within such network. Nodes of the graph represent the Internet users who interact, cooperate or share common activities within the web-based systems while edges correspond to the relationships extracted from the data about their common communication or activities. Node position function $NP(a)$ of node a respects both the value of node positions of node's a connections as well as their contribution in activity in relation to a, in the following way:

$$NP(a) = (1 - \varepsilon) + \varepsilon \cdot (NP(b_1) \cdot C(b_1 \rightarrow a) + ... + NP(b_m) \cdot C(b_m \rightarrow a)) \quad (1)$$

where: ε – the constant coefficient from the range $[0, 1]$. The value of ε denotes the openness of node position on external influences: how much a's node position is more static (small ε) or more influenced by others (greater ε); b_1,...,b_m – acquaintances of a, i.e. nodes that are in the direct relation to a; m – the number of a's acquaintances; $C(b_1 \rightarrow a)$,...,$C(b_m \rightarrow a)$ – the function that denotes the contribution in activity of b_1,...,b_m directed to a. In general, the greater node position one possesses the more valuable this member is for the entire community. It is often the case that we only need to extract the highly important persons, i.e. with the greatest node position. Such people surely have the biggest influence on others. As a result, we can focus our activities like advertising or marketing solely on them and we would expect that they would entail their acquaintances. The node position of a node is inherited from others but the level of inheritance depends on the activity of the users directed to this person, i.e. intensity of common interaction, cooperation or communication. Thus, the node position depends also on the number and quality of relationships. To calculate the node position of the person within the social network the convergent, iterative algorithm is used. This means that there have to be a fixed appropriate stop condition τ .

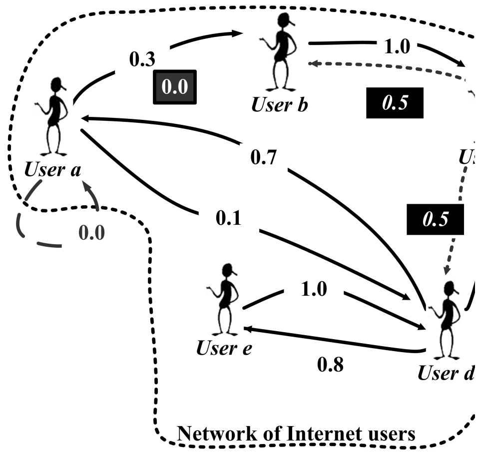

Figure 3. Example of the social network of Internet users with the assigned commitment values

3.3. Commitment Function

The commitment function $C(b \rightarrow a)$ is a very important element in the process of node position assessment, thus it needs to be explained more detailed. $C(b \rightarrow a)$ reflects the strength of the connection from node b to a. In other words, it denotes the part of b's activity that is passed to a.

The value of commitment function $C(b \rightarrow a)$ in *SNIU*(*IID*,*R*) must satisfy the following set of criteria:

1. The value of commitment is from the range $[0; 1]$: $\forall(a, b \in IID)\ C(b \rightarrow a) \in [0; 1]$.
2. Commitment function to itself equals 0: $\forall(a \in IID)\ C(a \rightarrow a) = 0$.
3. The sum of all commitments has to equal 1, separately for each node of the network:

$$\forall(a \in IID) \sum_{a \in IID} = 1 \tag{2}$$

4. If there is no relationship from b to a then $C(b \rightarrow a) = 0$.
5. If a member b is not active to anybody and other n members a_i, $i = 1,...,n$ are active to b, then in order to satisfy criterion 3, the sum 1 is distributed equally among all the b's acquaintances a_i, i.e. $\forall(a \in IID)\ C(b \rightarrow a_i) = 1/n$. The example of network of Internet users with values of commitment function assigned to every edge is presented in Figure 3. According to the above criteria all values of commitment are from the range $[0; 1]$ (criterion 1) as well as the sum of all commitments equals 1, separately for each user of the network (criterion 3). Moreover, the value of commitment function $C(a \rightarrow a)$ equals 0 (criterion 2) and because there is no relationship b to a so $C(b \rightarrow a) = 0$ (criterion 4). Note also that according to condition 5, node c is not active to anybody but two others b and d are active to c, thus commitment of c is distributed equally among all c's connections $C(c \rightarrow b) = C(c \rightarrow d) = 1/2$.

The commitment function $C(a \to b)$ of member a within activity of their acquaintance b can be evaluated as the normalized sum of all contacts, cooperation, and communications from a to b in relation to all activities of a:

$$C(a \to b) = \frac{A(a \to b)}{\sum_{j=1}^{m} A(a \to b_j)} \tag{3}$$

where: $A(a \to b)$ – the function that denotes the activity of node a directed to node b, e.g. number of emails sent by a to b; m – the number of all nodes within the *SNIU*. In the above formula the time is not considered. The similar approach is utilized by Valverde et al. to calculate the strength of relationships. It is established as the number of emails sent by one person to another person (Valverde, Theraulaz, Gautrais, Fourcassie, Sole, 2006). However, the authors do not respect the general activity of the given individual. In the proposed approach, this general, local activity exists in the form of denominator in formula 3. In another version of commitment function $C(a \to b)$ all member's activities are considered with respect to their time. The entire time from the first to the last activity of any member is divided into k periods. For instance, a single period can be a month. Activities in each period are considered separately for each individual:

$$C(a \to b) = \frac{\sum_{i=0}^{k-1} (\lambda)^i A_i(a \to b)}{\sum_{j=1}^{m} \sum_{i=0}^{k-1} (\lambda)^i A_i(a \to b_j)} \tag{4}$$

where: i – the index of the period: for the most recent period $i = 0$, for the previous one: $i = 1$,..., for the earliest $i = k - 1$; $A_i(a \to b)$ – the function that denotes the activity level of node a directed to node b in the ith time period, e.g. number of emails sent by a to b in the ith period; $(\lambda)^i$ – the exponential function that denotes the weight of the ith time period, $\lambda \in (0; 1]$; k – the number of time periods. The activity of node a is calculated in every time period and after that the appropriate weights are assigned to the particular time periods, using $(\lambda)^i$ factor. The most recent period $(\lambda)^i = (\lambda)^0 = 1$, for the previous one $(\lambda)^i = (\lambda)^1 = (\lambda)$ is not greater than 1, and for the earliest period $(\lambda)^i = (\lambda)^{k-1}$ receives the smallest value. The in a sense similar idea was used in the personalized systems to weaken older activities of recent users (Kazienko, Adamski, 2007).

One of the activity types is the communication via email or instant messenger. In this case, $A_i(a \to b)$ is the number of emails that are sent from a to b in the particular period i; and $\sum_{j=1}^{m} A_i(a \to b_j)$ is the number of all emails sent by a in the ith period. If node a sent many emails to b in comparison to the number of all a's sent emails, then b has greater commitment within activities of a, i.e. $C(a \to b)$ will have greater value and in consequence node position of node b will grow. However, not all of the elements can be calculated in such a simple way. Other types of activities are much more complex, e.g. comments on forums or blogs. Each forum consists of many threads where people can submit their comments. In this case, $A_i(a \to b)$ is the number of user a's comments in the threads in which b has also commented, in period i, whereas the expression $\sum_{j=1}^{m} A_i(a \to b_j)$ is the number of comments that have been made by all others on threads where a also commented, in period i.

Table 1. The statistical information for the Enron dataset

No of emails before cleansing	517,431
Period (after cleansing)	01.1999 - 07.2002
No. of removed distinct, bad email addresses	3,769
No. of emails after cleansing	411,869
No. of internal emails (sender and recipient from the Enron domain)	311,438
No. of external emails (sender or recipient outside the Enron domain)	120,180
No. of distinct, cleansed email addresses	74,878
No. of isolated users	9,390
No. of distinct, cleansed email addresses from the Enron domain (social network users) without isolated members the set *IID* in *SNIU*=(*IID*, *R*)	20,750
No. of network users within IID with no activity	15,690 (76%)
Percentage of all possible relationships	5.83%

4. Case Study

The experiments that illustrate the idea of node position assessment were carried out on the Enron dataset, which consists of the employees' mail boxes. Enron Corporation was the biggest energy company in the USA. It employed around 21,000 people before its bankruptcy at the end of 2001. A number of other researches have been conducted on the Enron email dataset (Priebey, Conroy, Marchette, Park, 2005), (Shetty, Adibi, 2005). First, the data has to be cleansed by removal of bad and unification of duplicated email addresses. Additionally, only emails from within the Enron domain were left. Every email with more than one recipient was treated as $1/n$ of a regular email, where n is the number of its recipients. The general statistics related to the processed dataset are presented in Table 1. After data preparation the commitment function is calculated for each pair of members. To evaluate relationship commitment function $C(a \to b)$ both of the presented formulas – 3 and 4 - were used. Formula 3 was utilized to calculate node position without respecting time (*NP*) whereas formula 4 serves to evaluate node position with time factor (*NPwTF*). The initial node positions for all members were established to 1 and the stop condition was as follows $\tau = 0.00001$. The node positions without and with time coefficient were calculated for six, different values of the ε coefficient, i.e. $\varepsilon = 0.01$, $\varepsilon = 0.1$, $\varepsilon = 0.3$, $\varepsilon = 0.5$, $\varepsilon = 0.7$, $\varepsilon = 0.9$. The conducted case study revealed that the time necessary to calculate the node positions for all users tightly depends on the ε value, i.e. the greater ε is the greater processing time is (Figure 4). The similar influence has the value of ε coefficient on the number of iterations required to fulfill the stop condition. Some additional information about the values of node positions provides the average node position within the *SNIU*

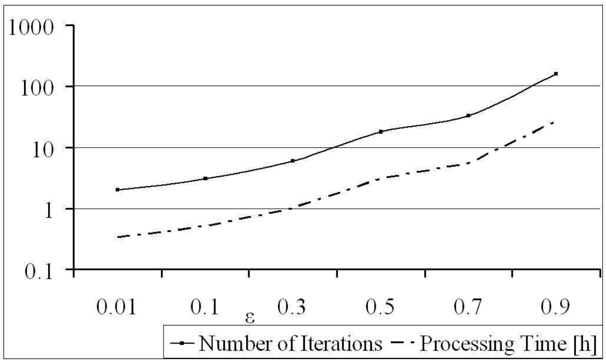

Figure 4. The number of necessary iterations and processing time in relation to ε

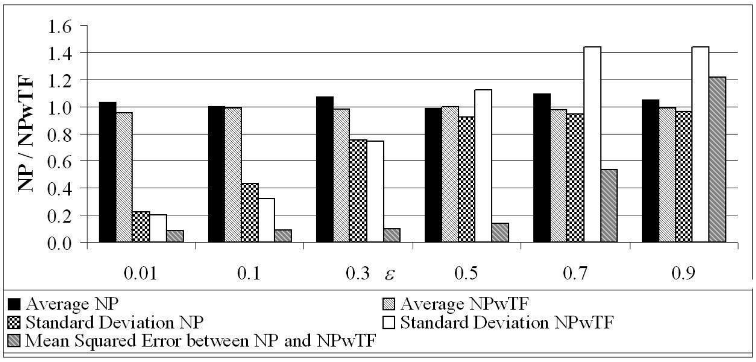

Figure 5. Average *NP* and *NPwTF*, standard deviation of *NP* and *NPwTF*, mean squared error between *NP* and *NPwTF* calculated for different values of ε

and the standard deviation of both node position values *NP* and *NPwTF* (Figure 5). The average node position does not depend on the value of ε. In all cases, it equals around 1 (Figure 5). Its convergence to 1 is formally proved. However, the standard deviation differs depending on the coefficient ε value. The greater ε is, the bigger standard deviation is. It shows that for greater ε the value of the distance between the members' node positions increases, and this can be noticed for both *NP* and *NPwTF*. It can be noticed that the value of node position *NP* for over 93% (see also Table 2) and *NPwTF* for over 95% (see also Table 3) of the community is less than 1 (see also Table 2). It means that only few members exceed the average value that equals 1. This confirms that node position can be the good measure to extract the key users in *SNIU* (Kazienko, Musiał, 2007). The comparison of the values of *NP* and *NPwTF* (Figure 6) reveals that more users obtain higher *NPwTF* position than *NP*. It means that people who have greater *NPwTF* were more active in the latest periods. Node position *NP* denotes the general position of a node regardless of time. Hence, *NP* will be the same for a person a that received n emails

Table 2. The percentage contribution of members in the Enron social network with *NP* ≥ 1 and *NP* where time factor is not included in relation to ε

ε	0.01	0.1	0.3	0.5	0.7	0.9
$NP \geq 1$	6.973	6.973	6.188	5.494	2.251	0.906
$NP < 1$	93.027	93.027	93.812	94.506	97.749	99.094

Table 3. The percentage contribution of members in the Enron social network with *NPwTF* ≥ 1 and *NPwTF* where time factor is not included in relation to ε

ε	0.01	0.1	0.3	0.5	0.7	0.9
$NPwTF \geq 1$	5.865	4.723	4.443	4.371	4.173	0.906
$NPwTF < 1$	95.135	95.277	95.557	95.629	95.827	99.094

from b three years ago and for a user c that also received n emails from b but all in the latest month. Such situation will not appear during calculation of *NPwTF*. In such case the position of node a will be lower then of the node c, because the weight assigned to the earlier period will be lower than the weight assigned to the latest period.

5. Conclusions

Node position is a measure for the importance of a user in *SNIU* that reflects the characteristic of the user's neighbourhood. Its value for a given individual respects both node positions of the nearest acquaintances as well as their attention directed to the considered user. Thus, the *NP* measure provides the opportunity to analyze *SNIU* with respect to social behaviours of individuals.

Node position is crucial for extraction of key network users and can be successfully used to establish project teams (Kazienko, Musiał, 2007), find new potential employees, search the potential consumers for advertising campaigns (Kazienko, Adamski, 2007) or in recommender systems (Kazienko, Musiał, 2006). It can also be utilized in target marketing to search for the appropriate target group of customers (Yang, Dia, Cheng, Lin, 2006). As a result, some specific products or services can be offered to the carefully selected representatives of the network, who are the most important in the population as well as those who potentially have the greatest influence on others.

Acknowledgements

This work was partly supported by The Polish Ministry of Science and Higher Education, grant no. N516 037 31/3708.

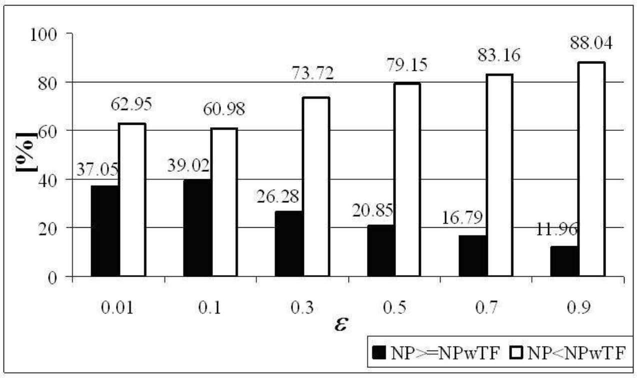

Figure 6. The percentage contribution of members with $NP \geq NPwTF$ and $NP < NPwTF$ within the Enron social network in relation to ε

References

Adamic, L.A. and Adar, E. (2003) Friends and Neighbors on the Web. *Social Networks* **25**, 3, 211–230.

Bavelas, A. (1950) Communication patterns in task – oriented groups. *Journal of the Acoustical Society of America* **22**, 271–282.

Botafogo, R.A., Rivlin, E. and Shneiderman, B. (1992) Structural analysis of hypertexts: identifying hierarchies and useful metrics. *ACM Transaction on Information Systems* **10**, 2, 142–180.

Boyd, D.M. (2004) Friendster and Publicly Articulated Social Networking. *CHI 2004, ACM Press* **22**, 1279–1282.

Culotta, A., Bekkerman, R. and McCallum, A. (2004) Extracting social networks and contact information from email and the Web. *CEAS 2004, First Conference on Email and Anti-Spam.*

Flake, G., Lawrence, S. and Lee Giles, C. (2000) Efficient identification of web communities. *In: Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 150–160.

Freeman, L.C. (1979) Centrality in social networks: Conceptual clarification *Social Networks* **1**, 3, 215–239.

Garton, L., Haythorntwaite, C. and Wellman, B. (1997) Studying Online Social Networks. *Journal of Computer-Mediated Communication* **3**, 1.

Gibson, D., Kleinberg, J. and Raghavan, P. (1998) Inferring Web communities from link topology. *In: Proceedings of the Ninth ACM Conference on Hypertext and Hypermedia.*

Golbeck, J. (2005) Computing and Applying Trust in Web-Based Social Networks. *Dissertation Submitted to the Faculty of the Graduate School of th Universtity of Maryland.*

Golbeck, J. and Hendler, J.A. (2004) Accuracy of Metrics for Inferring Trust and Reputation in Semantic Web-Based Social Networks. *EKAW 2004, LNCS 3257, Springer Verlag*, 116–131.

Hanneman, R. and Riddle, M. (2006) Introduction to social network methods. Online textbook, available from:
`http://faculty.ucr.edu/~hanneman/nettext/`.

Katz, L. (1953) A new status derived from sociometrics analysis. *Psychometrica* **18**, 39–43.

Kazienko, P. and Adamski, M. (2007) AdROSA - Adaptive Personalization of Web Advertising. *Information Sciences* **177**, 11, 2269–2295.

Kazienko, P. and Musiał, K. (2006) Recommendation Framework for Online Social Networks. *The 4th Atlantic Web Intelligence Conference, Studies in Computational Intelligence, Springer Verlag*, **23**, 111–120.

Kazienko, P. and Musiał, K. (2007) On Utilizing Social Networks to Discover Representatives of Human Communities. *International Journal of Intelligent Information and Database Systems*, to appear.

Priebey, C.E., Conroy, J.M., Marchette, D.J. and Park, Y. (2005) Scan Statistics on Enron Graphs. *Computational & Mathematical Organization Theory* **11**, 3, 229–247.

Rana, O.F. and Hinze, A. (2004) Trust and reputation in dynamic scientific communities. *IEEE Distributed Systems Online* **5**, 1.

Shetty, J., and Adibi, J. (2005) Discovering Important Nodes through Graph Entropy The Case of Enron Email Databases. *3rd International Workshop on Link Discovery, ACM Press*, 74–81.

Valverde, S., Theraulaz, G., Gautrais, J., Fourcassie, V., Sole, R.V. (2006) Self-organization patterns in wasp and open source communities. *IEEE Intelligent Systems* **21**, 2, 36–40.

Wasserman, S., and Faust, K. (1994) Social network analysis: Methods and applications. *Cambridge University Press, New York.*

Wellman, B. and Salaff, J. (1996) Computer Networks as Social Networks: Collaborative Work, Telework, and Virtual Community. *Annual Review Sociol* **22**, 213–238.

Yang, W.S., Dia, J.B., Cheng, H.C. and Lin, H.T. (2006) Mining Social Networks for Targeted Advertising. *39th Hawaii International Conference on Systems Science, IEEE Computer Society.*