

Churn Calculation based on Swarm Intelligence Algorithms in Social Networks

Hamidreza Sotoodeh
Electrical and Computer Engineering Department
Azad University of Qazvin
Qazvin, Iran
hr.sotoodeh@qiau.ac.ir

Elahe Daei
Mechatronic Research Laboratory
Azad University of Qazvin
Qazvin, Iran
e.daei@qiau.ac.ir

Farshad Safaei
Faculty of ECE, Shahid Beheshti University G.C
Evin 1983963113
Tehran, Iran
f_safaei@sbu.ac.ir

Parima Mohammadi
Electrical and Computer Engineering Department
Azad University of Qazvin
Qazvin, Iran
p.mohammadi@qiau.ac.ir

Abstract—Today, the churn phenomenon has been considered in many applications as an important outcome. Social networks can be considered as one of the most important applications with the mentioned outcome. Churn in social networks depends on the users' activity in a communication environment and appears if this activity is less than a required extent. Swarm Intelligence algorithms (SI), assumed to be the proper tools to model the communications in a social network. This bunch of algorithms according to the local agents' behavior, try to result the global behavior. This paper aims to measuring the user's churn by the mentioned method and including the communication messages transferred by the users in the network. Considering the measured activity rate, a churn threshold in various areas of communication will be obtained. Simulation results referring to confirmed the presented model of communication. The model validation and other values are obtained by a discrete event simulator. The communications used in this simulation result from mining a data set including real communications for one species of the mentioned networks.

Keywords: Social Networks, Churn, Swarm Intelligence Algorithm (SI)

I. INTRODUCTION

Social network can be defined as a structure that its social entities communicate with each other in different ways. We define these social entities as users [1]. Many previous studies investigated features of offline social networks. For example, experiment of Traverse and milgram[2]. In this experiment, each person has some letter that should deliver them to specific person by its relatives. Considering letter reaching destinations show that each letter after passing some hops (approximate 6 hops) reaches its destination. This result was one of the first experimental proof and effect of small word [3]. Also, it clarified that most of the pair nodes of networks have been connected to each other by a small internal route. However, not only have recently social networks such as Facebook and Myspace prospered, but also they have formed applications of internet that operate under web. In these networks, users can inform each other by writing their comments on social forums. However, how active one user is in the network will play considerable role for us. In this paper, we first calculate this activity. Then, determine percentage of

churned individuals of social communities that have been formed due to motivations of users in these networks.

Before explaining churn concept, we propose a model that describes how components of a social network communicate. In this paper, proposed model is based on swarm intelligence algorithms (SI) [4,5]. SI algorithms are population-based methods that present a unique global pattern by attention to the interaction between internal components. By using this concept and also by considering property of this technique, we can define distributed, self-organized and fault tolerance protocols. In addition, in social networks communication protocol based on this technique will be defined.

SI in social networks is inspired by observing on how ant colonies forage for food [4].

After presenting how the users communicate in social networks, it is turn to calculate and discover the users suffering from the churn phenomenon. When the ratio of the average activity amount of a user to the average activity amount of all users is less than a threshold [6], that user suffers from churn. This definition is comprehensively described in Section III.

The rest of paper will be organized as follow: Section II, presents a review of works about churn field. In section III, we study churn and define it in social networks. Also, section IV proposes a protocol based on Swarm Intelligence (SI) in order to create route between nodes of networks (users). Section V presents simulation model and evaluates it. Also, we calculate churn of sample communities that extracted from AS data set. Finally, in section VI, we conclude the paper.

II. RELATED WORK

The churn phenomenon has been used in a wide range of applications. The telecommunication applications have the highest share of analysis on this phenomenon [7, 9]. Other applications such as banking [9], online games [12], internet service providers [10], and P2P networks [11], have also done analyzes on this phenomenon. The interface among all mentioned applications on churn is to lose users or customers. For instance, a telecommunication industry

user suffers from churn when s/he leaves the current service provider for some reasons and joins another provider [8].

In [14] the influenced effects of other users has been considered and the decisions a user makes influenced by other users are discussed as a parameter affecting the churn phenomenon. Metcalfe's law which is the consequence of these impacts is expressed as follows: A communication network is valued based on the number of connected and active users of the network [14]. Here the presented model for churn is based on observing the users' behavior in a "join- participation- leave" chain and through a specific time period. Since churn in P2P networks is based on the users' independent behavior, it has a different concept from the one in social networks. The outcomes of most of the studies on these types of networks imply that it is less expensive to maintain the current users than attracting new ones [7] and also more beneficial [8].

III. CHURN IN SOCIAL NETWORKS

Users' activities in a network are a part of social capital of that network, so the loss of this capital by a reduction in activities or the users' departures endangers the values of the network. In the networks such as telecommunication network, a factor such as the relocation costs can reduce the churn rate in the network [15]. But this factor is too low and close to zero in social networks, so the user can leave the community and participate again after a while. The main differentiation point in the concept of churn in social networks is to move from a binary mode (active vs. inactive) to threshold based activities. This new definition brings up a question as follows: when does churn reach an acceptable level? The answer to this question is variable considering the type of communities' activities. Also considering that the users are members of the communities and have influences on other users and the whole network, we can analyze the churn phenomenon according to the users, communities, forums, etc. which their relative descriptions are addressed in [16].

Below we describe churn in social networks according to the individual activity [6].

Definition1: Churn (individual based): Previous Activity (PA) window is a time window including time units t_1 to $t_1 + n - 1$ while $\mu_{PA}(v_i)$, $n \in N, n > 0$. shows the average activity of the user v_i in PA window. Churn window (C) is also a time window including m time units $t_2 = t_1 + n$ to $t_2 + m - 1$ while $\mu_C(v_i)$, $m \in N, m \geq 0$ also shows the average activity of the user v_i in C window. According to the above assumptions the user v_i suffers from churn through the C window if:

$$\mu_C(v_i) \leq T(S) \mu_{PA}(v_i) \quad (1)$$

While $T(S)$ is a threshold dependent on the system parameters(S). ($0 \leq T(S) < 1$)

The average users' activity will be separately examined in specific churn windows through (1) and the churn rates will be determined. On the other hand, we can calculate the average activity for all the users at all the time stages having the activity rate of each user at time t , $a(v_i, t)$.

$$\mu_u = \frac{1}{N|V|} \sum_{v_i \in V} \sum_{t=1}^N a(v_i, t) \quad (2)$$

N in (2) refers to the all the observable time stages and V refers to the total number of users in the network. in calculating the churn rate, replacing $\mu_{PA}(v_i)$ with μ_u will leads to a churn calculation at a broader level compared with the user level. This measure is defined as the global churn in [6]. In Section (V-III) we will use this method on a real data set to calculate the churn rate.

IV. PROTOCOL TO BUILD PATH AMONG USERS IN SOCIAL NETWORKS

The protocol presented in this section is inspired by the wireless sensor networks gradient based algorithms [5]. Among existing routing protocols, the gradient based routings allow to build paths toward the sinks while the paths provide some conditions according to the node topologies [17]. It can be also a similar structure in social networks, while the sink nodes of sensor networks will be changed to popular nodes. We have used this idea here to redevelop this algorithm according to social networks.

Comparing with the biological process of pheromone release, we can say that each node sends the communication signaling packets containing its own pheromone level which will be updated considering the pheromone level of its neighbors. The pheromones' gradient will form in the direction of the popular node. As a result, the pheromone level forms the network communication process meaning that the data packets will be directed to the node with the highest pheromone level among its neighbors. Our expectation from the path building protocol is to operate in a way that the connected nodes to the most favorite node have the most impacts on the other nodes.

The important point is that any changes in the network level affect the path building protocol. It is due to the fact that the dependent nodes and path selecting decisions vary according to the changes in the pheromone levels and result in a redirection in the communication gradient path. It is implied that if a node or a link is lost, the network will recover itself through other close nodes and the path gradient will be formed again.

In the following, we will reconstruct the SI based routing method expressed at [5] according to social networks. Our goal is how to build paths among the nodes in social networks. We focus on our goal and avoid going into

details of data exchange; data exchange in the whole implementation process implies the exchange of data packets header containing nodes' pheromone and the hop count to the popular node.

We assume that there are 2 types of nodes at the network level including popular node and regular nodes; the pheromone amounts are divided to P different levels starting from 0 to P-1. The more pheromone the node has, the more valuable and prominent it is. The relationships among the nodes are based on pheromone packets exchange. It is expected that the gradients generated from the paths are directed to the popular node; meaning that if the popular node is connected to an regular node, that node will be included in a lower degree popular nodes category. The popular node is responsible to introduce the connected node to the others and vice versa. The first mode is of course more visible.

Algorithm1: Popular Nodes

```

1: P ← P-1
2: SetTimer( $T_1$ )
3: loop
4: e ← WaitForEvent()
5: if e = TIMER_EXPIRED then
6:   Upload (p)
7: end if
8: end loop

```

Algorithm2: Regular Nodes

```

1: P ← 0
2: SetTimer( $T_1$ )
3: SetTimer( $T_2$ )
4: loop
5: e ← WaitForEvent()
6: if e = DATA_RECEIVED then
7:  $P_n, h_n \leftarrow GETDATA()$ 
8:  $P \leftarrow Reinforcement(P_n)$ 
9: else if e = TIMER_EXPIRED then
10:  $t \leftarrow getTimer()$ 
11: if  $t = T_1$  then
12:   Upload (p) or
13:   Selective_NeighborPost (p)
14: end if
15: else if { $t = T_2$  and don't least one replied }
16:  $p \leftarrow max(0, evaporation\_function())$ 
17: end if
18: end if
19: end loop

```

The structure of gradient based communication building protocol will be formed by introducing the popular node with the highest pheromone level (P-1). The algorithms 1

and 2 describe the way of path creation among the nodes, separately for both popular and regular nodes.

As expressed before, the popular node/s will be introduced according to the maximum amount of their pheromone (algorithm 1, line 1) and consecutively broadcast their pheromone level to the neighbor nodes through uploading data (algorithm 1, lines 5 to 7). T_1 shows the time period for the popular node to rebroadcast.

The regular nodes are assumed to have an zero initial value of pheromone (algorithm 2, line 1) and this value will be updated through the process of reinforcement – evaporation for each node. When a node receives a data packet, it will send a data packet back to the sender as a response and enter the reinforcement phase and will update its pheromone level according to the received packets (algorithm 2, lines 6 to 8). The regular nodes can also broadcast their packets at the time period of T_1 . It implies that the users can broadcast their posts or comments to their neighbors in this time period (algorithm 2, lines 10 to 13). The nodes enter the evaporation phase if they would not respond to an upload or received post in a specific time period of T_2 (algorithm 2, lines 15 to 16). In the following we will focus on the process of, neighbor selection to sending data, pheromone reinforcement, and pheromone evaporation.

A. ASSUMPTIONS

- Only the regular nodes will be involved in the mentioned triple process while sending or receiving data packets and the popular node/s just in specific time periods send the data to their neighbors.
- Table I shows the symbols used in the mentioned process. The symbols are according to [18].

TABLE.I THE APPLIED SYMBOLS IN NEIGHBOR SELECTION BY EACH USER

Symbols	Description
T_{ij}	Shows the pheromone level of the path from node i to node j.
$\gamma(i, j)$	Shows the reverse of hop counts from node j to the popular node of the group which j is a member of, plus 1.
β	The determining parameter for the effectiveness of the function $\gamma(i, j)$, $\beta > 0$
N_i	Shows the number of neighbors for node i.
R_j	Shows the number of sender nodes pass through node j
$\Sigma(H_{ij})^{-1}$	The reverse of hop counts from node j to the popular node
$(h_j)^{-1}$	Inverse of number of hops from node j to popular node
θ	The path pheromone evaporation parameter
P	The pheromone level for each node

B. NEIGHBOR SELECTION PROBABILITY CALCULATION

A sender selects the neighbors with the most relationships with the popular node as intermediate nodes, in order to direct its data packets towards the popular node [18]. $p_k(i, j)$ in (3) shows the probability of node j as the next selected node among the neighbors. This node will be selected by the ant k starting from node i , ending to the popular node.

$$p_k(i, j) = \frac{T_{ij} \times \gamma(i, j)^\beta}{\sum_{u \in N_i} T(i, u) \times \gamma(i, u)^\beta} \quad (3)$$

C. PHEROMONE REINFORCEMENT

In this step each node according to the (4) to (8) updates its pheromone amount under some conditions (algorithm 2, lines 8 to 9).

$$P_i = T_{ij} \times \frac{P_i}{P_i + P_j} + P_i \quad (4)$$

$$P_j = T_{ij} \times \frac{P_j}{P_i + P_j} + P_j \quad (5)$$

Where

$$T_{ij} = (1 - \theta)T_{ij} + \theta \Delta T_{ij} \quad (6)$$

$$\Delta T_{ij} = [1 + (h_i - h_j)] \times \Delta \omega \quad (7)$$

$$\Delta \omega_j = \sum_{i \in R_j} (H_{ij})^{-1} + (h_j)^{-1} \quad (8)$$

In (4) and (5), two head nodes of the communication link update their pheromone amount according to that of the link and the other party. While the pheromone amount of the current link will be shared among the transferring nodes, according to a weight based equation. Node pheromones updating require updates in the communication link's pheromone amount which according to (6) will be updated by the evaporation parameter θ . In (7) a value of more than zero for $h_i - h_j$ means that node j is closer to the favorite node, so the algorithm allocate a more amount of pheromone to the path including j . Zero value for $h_i - h_j$, implies that both nodes i and j are in the same distance from the favorite node, so the algorithm adds the amount of $\Delta \omega_j$ to this path; and if $h_i - h_j$ was less than zero, the pheromone amount will stays unchanged.

Equation (8) calculates the amount of $\Delta \omega_j$ which shows the total hop counts starting from the sender, passing through node j , and ending at the popular node. The less the hop counts are, the more pheromone will be allocated to i - j path (as showed in (7)). This excites the ants to join this path.

D. PHEROMONE EVAPORATION

Those nodes will be subjected to evaporation which haven't respond to at least one sender in a time period of T_2 (algorithm 2, lines 16 to 17). This action will be according to (9) to (11):

$$P_i = P_i - T_{ij} \times \frac{P_i}{P_i + P_j} \quad (9)$$

$$P_j = P_j - T_{ij} \times \frac{P_j}{P_i + P_j} \quad (10)$$

Where

$$T_{ij} = (1 - \theta)T_{ij} \quad (11)$$

The equations used for pheromone evaporation are similar to the equations for reinforcement, however here the updated amount will be subtracted from the pheromone amounts of nodes based on the weight ratios (10 and 11).

V. SIMULATIONS

In order to have a better understanding and also evaluating the parameters of analysis model, we have presented a discrete event simulator. The existing software tools were not able to provide us with our goal to model evaluation and threshold determination for evaluating users' activities. Among the open source tools we chose SNAP [19] for clustering the network graph.

The presented simulator is developed by C# programming language. The simulator receives a graph as input and shows the evaluation results as the output of applied path building algorithms (algorithms 1 and 2). The input we have considered here is a graph extracted from the data set of an automated systems network [20], which it will be described in section (V-A).

To ease working with the mentioned graph, we have organized the existing nodes in a 10×100 grid while each cell contains one node (i.e. $\xi^n(v) = 1.0 \frac{\text{node}}{\text{cell}} \forall v \in V$). The real

relations existing in the data set have been considered as the relations among the graph nodes in order to obtain a relationship between the nodes' positions and numbers. Social networks users tend to join the communities, so we have clustered the graphs by the help of SNAP, each calculated in a different time period (totally 35 clusters). In each cluster we have assumed the node with the highest degree as the popular node. Comparisons between the graphs showed that the nodes and the edges were being

added and removed; this lets us to model the users' dynamic behavior such as joining and leaving the network while the program is running. We can also obtain different consecutive graphs through the program operation and the algorithms (algorithm 1 and 2) running and consider them as the program's input and then calculate the evaluation parameters. In order to model the users' behavior in social networks, we have activated each user with an average delay of 1.0 s following the exponential distribution.

Through the simulation there are some parameters listed in table I with default values and relative symbols according to [5, 18].

TABLE.II THE DEFAULT VALUES USED THROUGH SIMULATION

notation	Description	value
P	The pheromone level	[0-10]
β	Determining parameter for effectiveness of the function $\gamma(i, j)$	20
θ	The parameter of pheromone evaporation on the path	0.3
\bar{N}	Shows the average number of neighbors for the user in a group	8
T_1	Sending time period for each user (algorithms 1 and 2)	8
T_2	The evaporation time period for each user (algorithms 1 and 2)	1,2,4

A. DATA SET DETAILS

The graph of routers forming the internet can be organized as sub graphs called autonomous systems (AS). Each AS exchanges its traffic flows with its neighbors. Having the reports from BGP protocol, we can form the communication graph. In this paper we have used an AS data set [20] which has been collected from 2004 to 2007. From 2004 to 2006 there exist a specific file for each week, and after that the files have been collected monthly. Each file contains a complete AS graph obtained by BGP tables.

AS relations are of 3 different types i.e. customer-provider, provider-customer, and peer-to-peer. The collected data between the dates Jan 2004 to Jan 2005 has been used for this paper which includes 16301 nodes and 65910 edges. The dynamism of nodes and edges in being added or removed helps us to calculate churn in the network which we will focus on (V-B).

B. CHURN CALCULATION

According to the definition 1, a user will suffer from churn if the ratio of its average activity rate to the average activity rate of total users is less than a threshold. Considering this definition and the proposed equations of Section III we can calculate this measure for the autonomous systems' data. The goal is to achieve an average pheromone rate for the nodes suffering from churn. Having this rate, we can propose some methods to keep the pheromone's threshold balanced. To calculate the churn rate in a graph we use the (1) described in Section III. (0.2,

0.5, 0.7) are thresholds for T(S) which respectively shows the average of high, medium, and low users' activity [6].

The pheromone density is also dependent on the pheromone emission rate λ and evaporation rate μ . In other words, the reinforcement-evaporation process depends on the number of neighbors of an agent (\bar{N}). Considering this description we can define a new measure such as r , expressing the ratio of the broadcast rate to the evaporation rate:

$$r = \frac{\lambda \bar{N}}{\mu} \quad (12)$$

In table III churn rate has been calculated and expressed for a sample community and the average pheromone rate to achieve this threshold has been proposed. In these calculations PA window and CW are assumed to have the variable values of m and n respectively and $r=8$. According to the results, as the churn threshold is lower 0.2, the number of nodes suffering from churn is also lower and vice versa. In the following we can calculate the average pheromone rate of the nodes with churn. As the threshold goes is higher, the pheromone rate of mentioned nodes goes upper. It is to be considered that r is assumed to be a constant value, and the sample nodes are the nodes with churn.

TABLE.III PHEROMONE MASS THRESHOLD FOR THE NODES WITH CHURN IN GROUP 240, WITH DIFFERENT VARIABLE VALUES OF T(S), N, AND M

M	n	T(S)	Churn rate percentage, considering the (1)	Average pheromone density for the nodes with churn
13	5	0.2	10.7	0.2
13	5	0.5	20.0	0.7
13	5	0.7	25.3	0.8
5	13	0.2	17.5	0.26
5	13	0.5	24.0	0.41
5	13	0.7	30.3	0.68
5	5	0.2	18.5	0.38
5	5	0.5	28.5	0.94
5	5	0.7	33.3	1.04

For instance, according to Fig. 1, the churn threshold to determine the pheromone density in the stable mode belongs to the community 24 including 39 nodes and 520 edges. The parameters for creating this diagram extract from last line in table III. This diagram shows if the pheromone density threshold becomes 1.04, then churn rate of that community that in churn will be nearly 33%. This value is specified with a frame in Fig. 1.

C. THE EVALUATION OF THE GENERATED GRADIENT

Fig. 2 shows the distribution of pheromone density on the surface v for different values of r , in a stable condition. In order to have a better illustration we have chosen a community with large number of nodes i.e. community 10 with 400 nodes and 5630 edges.

Considering this community, our work space would be an area of 20×20 with the maximum pheromone rate of 10 for the popular nodes. The effectiveness of r 's value in the formation of pheromone density rate is visible. Considering $r=1$, Fig. 2(a), implies that the nodes mostly suffer from evaporation and a few number of the nodes participate in message sending process. A higher value of r ends to more sending process, shown in Fig. 2(b) and 2(c). Since nodes activation follows an

exponential distribution, for all 3 different values of r there are some nodes suffering from churn. Some of them can even have a zero value of pheromone rate. Zero pheromone rates are caused when a node did not have any sending either when it has experienced consecutive evaporations. Any popular node can be considered as the popular node for different communities.

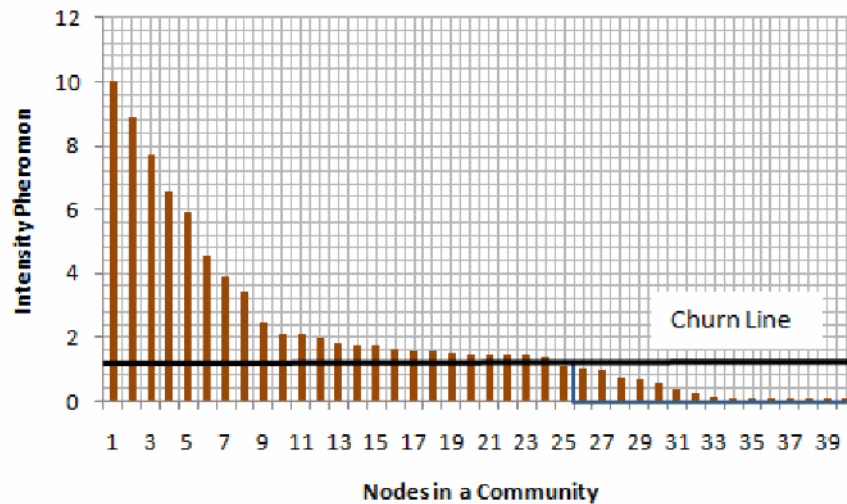


Figure 1. Threshold of pheromone intensity for churned nodes of community 24.

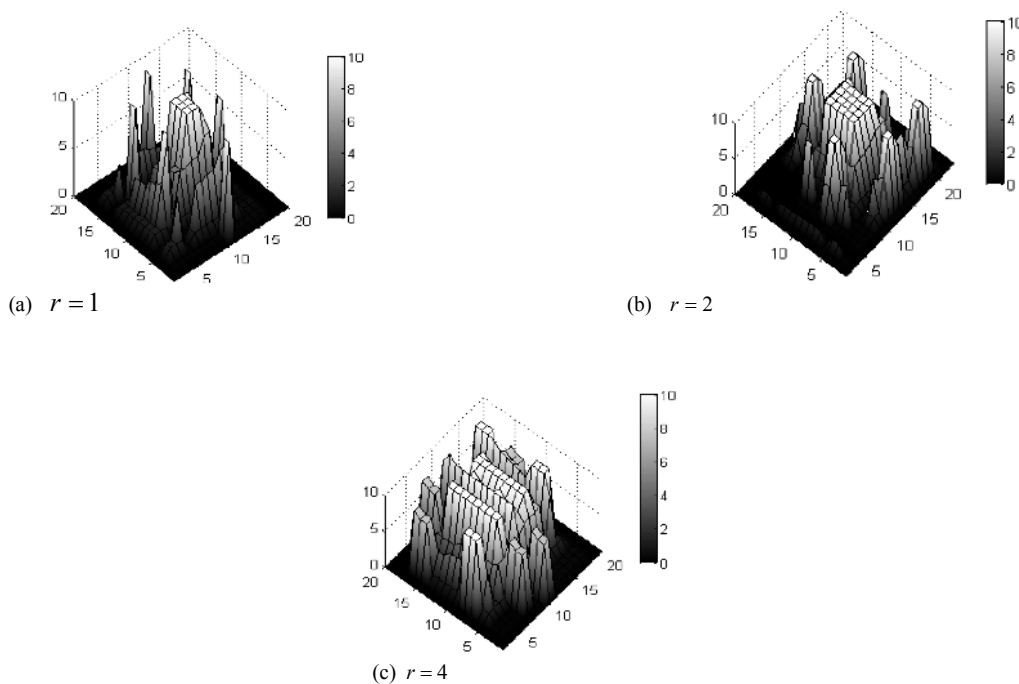


Figure 2. Distribution of pheromone intensity in stable condition for different value of r .

VI. CONCLUSIONS

The analysis based on social networks modeling which has been done in this paper used the interactions among nodes.

in a network. By interaction, we meant the message passing process. This model creates a picture of communication model for the users of a real social network. Since the origin of Swarm Intelligence algorithms is based on pheromone updates and exchanges which do not relates to the network.

topology, scalability is achievable while using these algorithms. Pheromone packets exchanges which are considered as data transfer and users' activity rate let us to the calculate churn rate in a social network. The presented paper uses this specification and introduces churn thresholds for the communities. We have used the existing connections in an autonomous system and simulated the environment using these connections. Finally the simulated results verify communication model that described.

REFERENCES

- [1] M. Gjoka, Measurement of Online Social Networks DISSERTATION, DOCTOR OF PHILOSOPHY Thesis, UNIVERSITY OF CALIFORNIA, 2010.
- [2] S. Milgram, "The Small World Problem". *Psychology Today*, 1(1), May 1967. pp 60 – 67.
- [3] D. Eppstein, M.T. Goodrich, M. Loffler, D. Strash, L. Trott, "Category-based routing in social networks: Membership dimension and the small-world phenomenon", *Computational Aspects of Social Networks (CASoN), International Conference on Digital Object Identifier: 10.1109/CASON*, pp. 102 – 107, 2011.
- [4] M. Hinchey, R. Sterritt, C. Rouff, "Swarms and swarm intelligence", *IEEE Computer* (2007) 111–113.
- [5] D. Bruneo, M. Scarpa, A. Bobbio, D. Cerotti, M. Gribaudo, "Markovian agent modeling swarm intelligence algorithms in wireless sensor networks," *Elsevier Performance Evaluation*, Volume 69, Issues 3–4, March–April 2012, Pages 135–149.
- [6] Karnstedt, M. Hennessy, T. Chan, J. Hayes, "Social Churn in Social Networks: A Discussion Boards Case Study Computing (SocialCom)", 2010 IEEE Second International Conference on Digital Object Identifier: 10.1109/SocialCom.2010.40 Publication Year: 2010, Page(s): 233 - 240.
- [7] M. Mozer, R. H. Wolniewicz, D. B. Grimes, E. Johnson, and H. Kaushansky, "Churn reduction in the wireless industry," in *NIPS*, 1999, pp. 935–941.
- [8] H. Hwang, T. Jung, and E. Suh, "An ltv model and customer segmentation based on customer value: a case study on the wireless telecommunication industry," *Expert Syst. Appl.*, vol. 26, no. 2, pp. 181–188, 2004.
- [9] Y. Xie, X. Li, E. W. T. Ngai, and W. Ying, "Customer churn prediction using improved balanced random forests." *Expert Syst. Appl.*, vol. 36, no. 3, pp. 5445–5449, 2009.
- [10] B. Q. Huang, M. T. Kechadi, and B. Buckley, "Customer churn prediction for broadband internet services." in *DaWaK*, ser. Lecture Notes in Computer Science, vol. 5691, 2009, pp. 229–243.
- [11] O. Herrera and T. Znati, "Modeling churn in P2P networks," in *Annual Simulation Symposium*. IEEE Computer Society, 2007, pp. 33–40.
- [12] J. Kawale, A. Pal, and J. Srivastava, "Churn Prediction in MMORPGs: A Social Influence Based Approach," in *CSE '09*, 2009, pp. 423–428.
- [13] S. M. Keaveney, "Customer switching behavior in service industries: An exploratory study," *The Journal of Marketing*, vol. 59, no. 2, pp. 71–82, 1995. [Online]. Available: <http://www.jstor.org/stable/1252074>.
- [14] B. Metcalfe, "Metcalfe's law: A network becomes more valuable as it reaches more users," 1995.
- [15] J. P. Maicas, Y. Polo, and F. J. Sese, "Reducing the level of switching costs in mobile communications: The case of mobile number portability," *Telecommunications Policy*, vol. 33, no. 9, pp. 544 – 554, 2009. [Online]. Available: <http://www.sciencedirect.com/science/article/B6VCC-4WCT02M-1/2/737ee54afa660372dbd56e74d6fec8c4>.
- [16] M. Karnstedt, M. Rowe, J. Chan, H. Alani, C. Hayes, "The Effect of User Features on Churn in Social Networks," In *Proceedings of the ACM International Conference on Web Science*, June 2011, Koblenz, Germany.
- [17] K. Akkaya, M. Younis, "A survey on routing protocols for wireless sensor networks", *Elsevier Ad Hoc Networks* 3 (2005) 325–349.
- [18] W. Liao, Y. Kao, C. Fan, "Data aggregation in wireless sensor networks using ant colony algorithm," *Elsevier Journal of Network and Computer Applications*, Volume 31, Issue 4, November 2008, Pages 387–401.
- [19] <http://snap.stanford.edu/snap/>.
- [20] <http://snap.stanford.edu/data/as-caida.html>.