Is the Small World Effect the result of Trust Based Information Diffussion?

Adeela Huma
Department of Computer Science
Virginia Tech
ahuma@vt.edu

Abstract—In social networks trust plays an important role in information diffusion. Sharing information with right people can result in rapid spreading as well as help in maintaining the integrity of the message. People share information with others based on trustworthiness. Forwarding information to other (nodes) involves many factors such as popularity of a node, connectedness of a node in the network, its perception in the network etc. Finding such nodes is very important as they can help spread or limit the spread of information. The studies performed on small world networks for trust based information diffusion shows very good results as compared to other networks [13]. Is there a relationship between small world networks and trust based information diffusion? If there is then it means we can infer that when information is diffused based on trust the nodes that receive that information are part of small world network within the social network. Finding that sub-network can help find the key nodes in the network and those nodes can be used to spread or limit the information diffusion within the network. This study presents an algorithm to compute the trust among peers and do message forwarding based on that trust and uses path length and number of nodes who received message to find small world effect.

Keywords—Trust networks; information diffusion; small world effect; socio sensors

I. INTRODUCTION

There are two billion people connected on the internet [1] and content created and shared is also huge. Which information to trust and which one to ignore is a critical question in social networks. If information is critical or is good it should be spread quickly and fast so that others can benefit from it. If there is lack of trust among peers in networks then good information can be lost. Also the integrity of the information can be maintained when a message is passed through trustworthy networks.

Trust-based message forwarding can help maximize the spread of good information diffusion (e.g., evacuation for emergency) or minimize the spread of bad information diffusion (e.g., a false rumor). In particular, if a node or an agent or an entity interchangeably chooses another node in its neighbor to forward the information, how fast good or correct information can be propagated over a network or how fast bad information can be stopped being propagated over a network. How to choose a node (a trustee) to forward a message can be based on the trustee's trust estimated by a node itself, called trustor.

How to estimate the peer-to-peer trust is a key. In social network two persons may be connected expressing trust relationship but they still may not trust each other enough to forward the information shared by one node to other nodes. Trust depends on many factors such as: (1) a node's individual trust in terms of its capability and integrity (lying or not); and (2) the node's relational trust in a network which explains how influential it is in a network.

Small world networks have an interesting property of small average path lengths between nodes. It means if information is diffused in small world networks, there is a probability that it will spread quickly in the network [13]. Is there a relationship between small world networks and trust based information diffusion?

If there is a relationship then it means we can infer that when information is diffused based on trust the nodes that receive information are part of small world network within the social network. Finding that sub-network can help find the key nodes in the network and those nodes can be used to spread or limit the information diffusion within the network.

In this paper we perform experiment on friendship network of students [5]. We will use the network structure properties to calculate the peer-to-peer trust and do message forwarding based on that trust and measure small world effect. The metrics used are path length between nodes and number of nodes that received the message.

II. RELATED WORK

There has been a lot of related work done in the domain of trust, information diffusion and small world networks.

A. Concept and Dimension of Trust

A lot of research has been done to identify what constitutes trust in networks and how to find the peer-to-peer trust.

The approach used in [3] calculates matrix of trust T where tij represents trust between i and j and then computes other matrices such as if ‘i’ trust ‘j’ and ‘j’ trusts ‘k’ then it implies that ‘i’ also trust ‘k’. Theoretically it makes sense as this is comparable to transitivity concept but in real social networks if ‘i’ trust ‘j’ and j’ trusts ‘k’; it is not necessary that ‘i’ trusts ‘k’. There might be no relationship between them. ‘i’ and ‘j’ may be friends and ‘j’ and ‘k’ maybe family members. So practically ‘i’ and ‘k’ are not related.

In [4] uses nodes in-degree to measures its influence in the network. In degree represents how many nodes are connected to a given node but this measure only does not guarantee the node's influence in network. If for example a node has higher in degree but no or very small out degree that means it's almost disconnected to network in a sense that information will flow towards this node but useful information will not go out because it is not well connected in the network.

In [6] Sibel Adali et.al propose calculating propagative trust based on conversation and propagation. They suggest that if A and B exchange information and talk often then it means A and B trust each other. And also if B propagates the information that A said to B then its implied that B trust A. Conversation is an important factor that reflects trust between two individuals. We will also be using this property in estimating trust.

In [7] Thomas et al. predict trust and distrust in social networks by computing path probabilities with spring embedding. In [8] Sibel Adali & et.al define a trust based agent framework for network decision-making by using competence and willingness attributes of agents.

B. *Information Diffussion*

A lot of work has been done in information diffusion about identifying models to do inter and intra community diffusion. In [9] Azadeh et.al discuss that there exists an optimal network structure where global (inter and intra community) diffusion of information is possible. In [10] author discuss about effect of external effects in information diffusion in network. Other studies have been studied about the importance of week ties in information diffusion etc.

C. *Small World Effect*

Stanley Milgram [11] in his behavioral study of obedience performed an experiment and suggests that human society is a small world type network. In his experiment participants were asked to get a message to a specific target individual by passing it from acquaintance to acquaintance through the social network. Milgram famously found that messages that arrived at the destination passed through only about six people on their way, which is the origin of the popular concept of the "six degrees of separation."

Small world experiment has been repeated by Dodds et al. [12] using the modern medium of email. In this version of the experiment participants forwarded email messages to acquaintances in an effort to get them ultimately to a specified target person about whom they were told a few basic facts. The experiment found the average path length was somewhere between five and seven steps—very similar to Milgram's result.

Yan et al. [13] conducted experiment to see the effects of trust on information diffusion in online social networks and used networks like regular networks, small world networks, and random networks. They observe that the enhancing of trust relationship has obvious influence on information diffusion in small world networks. The study concludes that "as the increasing of initial transmitting probability, small world networks perform best, and the information diffusion is better than random, regular and scale-free networks."

This study proposes that there is a relationship between faster trusts based information diffusion and small world effect. Human beings tend to pass along information from peers that they trust more [6]. In Stanley Milgram's experiment, people probably passed along packets based on trust level. In social networks random information diffusion is not usually seen.

III. PROPOSED APPROACH

The proposed algorithm uses network structure (in-degree, out-degree) and explicit trust declaration between two nodes.

In this study we propose an algorithm to do trust based information diffusion and will measure if it shows small world effect. To do so, we need to

A. Estimate the node to node trust

B. Information diffusion strategy

A. *Trust Estimation*

To estimate node-to-node trust, the proposed algorithm will use following information.

- In degree of a node represents how popular a node is in the network [4] and gives an indication that how influential the user is in the network. Such nodes in the network represent kind of hub. For our algorithm we represent it by 'I'. Hub nodes are the ones that show funneling effect as described by Stanley Milgram [11]
- Out-degree of a node represents how well connected a node is in the network with other nodes. If a node is connected to many other nodes in the network; it means this node has access to other nodes and can transfer information to other nodes depending on its out degree. For our algorithm we represent it by 'C'.
- Even if there is an edge between two nodes in the network that does not mean that they would transfer information to each other. Sharing information among other factors also depends on the trustworthiness between two nodes. For our algorithm we represent it by 'W'. The algorithm assumes that 'W' is explicitly expressed by connected nodes (weighted directed networks).

This algorithm will be applied to directed trust networks. Trust networks has a weighted edge between nodes; positive edge representing a trust relation and negative or zero edge represents a distrust relationship. For each node in the network we will calculate the (I, C) values.

B. *Difussion mechanism*

After calculating node-to-node trust, the algorithm uses breadth first search approach to diffuse information from source node to destination node.

Following table shows how 'I' and 'C' will be used to calculate Trust level of a node.

Total Trust Level/Probability of Fast information diffusion	I (In-degree)	C (out-degree)
Medium	High	Low
Medium	Low	High
Low	Low	Low
High	High	High

Total trust of node can be defined as sum of in-degree and out-degree:

Total Trust (T) = I + C

The algorithm introduces two more parameters for information diffusion that control the rate of information diffusion.

1. Diffusion Threshold (α) - If T > α then information can be passed to the node as it seems to be very connected and popular in the network
2. Trustworthiness Threshold (γ) - If W > γ then information can be passed to the node.

C. Algorithm

The algorithm (Fig-1) used to diffuse information from source to destination node is breadth first search with trust. It uses trust metric (I, C, W) defined above to diffuse information. Inputs to algorithm are:

G- directed network

s – source node

d - destination node

Information is diffused to neighbors only if they have not received it before from some of the neighbor and total trust is greater than threshold-trust value and trustworthiness between nodes is greater then threshold-trustworthiness. Another thing to mention here is that information is not diffused to all neighbors, since it is a directed network so it is only diffused to out neighbors of the nodes.

IV. Dataset

In this study we have used the konect "Adolescent health" [5] network.

This directed network was created from a survey that took place in 1994/1995. Each student was asked to list his 5 best female and his 5 male friends. A node represents a student and an edge between two students shows that the left student chose the right student as a friend. Higher edge weights indicate more interactions and low edge weight shows that there is no common activity at all.

The network has over 2,539 nodes and 12969 directed edges with weights indication the intensity of interaction.

This dataset is chosen because the edge weight indicates the number of interactions nodes has among themselves. This is equivalent to 'W' in our algorithm. The more interactions are implied as more trust.

```
BFS_WITH_TRUST(G, s, d)

## Initialize each vertex as not visited

for each vertex u in V[G]
    color[u] := WHITE
    d[u] := infinity
end for

## set Initial node as visited

color[s] := GRAY
d[s] := 0
ENQUEUE(Q, s)
nodeFound = FALSE

while (Q != Ø OR !nodeFound)

u := DEQUEUE(Q)

for each vertex v in Adj[u]

##edgeWeight[uv]   == W
##totalTrust[v]    == T
## Information is diffused to neighbors if
##      1) it has not got the information before
##      2) W > γ
##      3) T > α
   if (edgeWeight[uv] > γ &&
       totalTrust[v] >  α &&
       color[v] = WHITE )
        color[v] := GRAY
        d[v] := d[u] + 1
        ENQUEUE(Q, v)

        if (d == v)
           nodeFound = TRUE
        end if
    end if
end for
end while
```

Figure 1- bfs_with_trust algorithm

V. Results

To find out if small world effect is the result of trust based message diffusion we will calculate the metrics that are found in small world networks such as

1. Path length between source and destination node
2. Number of nodes who received the information

The experiment chooses 100 unique pair of nodes (source, destination) randomly and then diffused information between nodes using bfs_with_trust. Diffusion Threshold 'α' was set to

4. The value of 'γ = 3' was chosen based on edge weight (W) of the network. Edge weight is equivalent to the number of communications between two nodes.

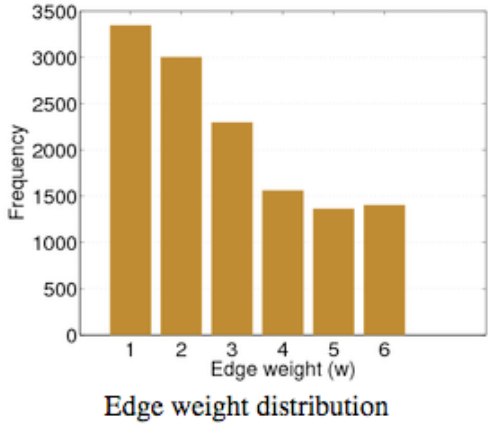

Figure 2 - Edge Weight Distribution

For each pair, the experiment recorded the path length. 20 out of 100 nodes received the message and 80 did not receive the message. This could be because there is no path between nodes or the trust chain was broken and message was not delivered.

The 20 nodes that received the message the average path length was 6.85. Below is the frequency distribution plot of path lengths:

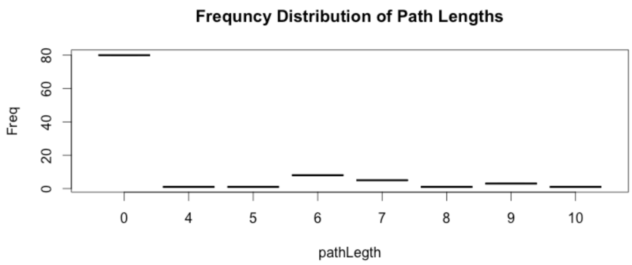

Figure 3- Path Length Frequency Distribution

It can bee seen that mostly nodes who received this information has path length less than 7.

VI. Future Work Extension

There is a lot of room of improvement in refining the experiment performed for this study.

1. Experiment should be performed with varying values of 'α'and 'γ' to see what effect these parameters play.
2. Instead of always randomly choosing new target (destination) node, we pick some *specific* destination nodes and choose many source nodes and then do diffusion to mimic Stanley Milgram's experiment [11] or in other words don't change the destination node.
3. This study used only one dataset [5] to perform experiment. The study can be extended to perform experiment on other networks and compare results.
4. The experiment can be performed with some other diffusion algorithm and see if choice of diffusion algorithm plays any important role.
5. This study can be enhanced to include other trust metrics such as perception of a node about its neighbors that is built over time (P). In social networks [14], even if there is some level of trust between nodes based on individual's perception they might not transfer information. The quality of message delivered (Q) is also important. It can be used to measure how many nodes received the message in its integrity. To study Q we will also need to consider distrust in the network. The message most likely will be changed if there is distrust among nodes.
6. Clustering co-efficient is one of the metric used in small world networks. This study should be extended to create a sub network from actual network and finding clustering co-efficient on sub-network and if that network shows high clustering co-efficient, it can be inferred that those who received the message were part of small world network within the main network. The vertices in the sub-network will be random nodes (source, destination) chosen in experiment along with all the nodes that were traversed while finding the destination node. The edges will also be maintained between nodes.
7. Once we get the sub-network from the main network, we can also find out that why some nodes did not receive the message? Was it because there was lack of trust or was it because there were disconnected in the network (no path between source and destination).
8. Another thing that can be identified in the network is the hub nodes and see if the nodes that received the message; did they get it from a specific node i.e. hub (Stanley Milgram's funneling effect).

This extension study can help solidify the findings of experiment.

VII. Conclusion

Main contribution of this paper is finding the relationship between trust based message diffusion and small world effect. Though not a lot number of nodes received the message but ones who received show an average path length of 6.85. The ones who did not receive it probably were not the part of network.

Trust based information diffusion is very important. Finding trusted nodes in the network is very important as they can help spread or limit the spread of information. And finding a sub-network of trustees within the social network means those nodes in the network as socio-sensors and anytime we want to spread the information we diffuse it to them and anytime we want to limit its spread we focus on these nodes.

VIII. REFERENCES

[1] "Internet usage statistics - the internet big picture: World internet users and population stats," http://www.internetworldstats.com/stats.htm, 2010.

[2] M. Richardson, R. Agrawal, and P. Domingos, "Trust management for the semantic web," The SemanticWeb-ISWC 2003, pp. 351–368, 2003.

[3] R. Guha, Ravi Kumar, Prabhakar Raghavan, and Andrew Tomkins. 2004. Propagation of trust and distrust. In Proceedings of the 13th international conference on World Wide Web (WWW '04). ACM, New York, NY, USA, 403-412. DOI=http://dx.doi.org/10.1145/988672.988727

[4] Adrien Guille, Hakim Hacid, Cecile Favre, and Djamel A. Zighed. 2013. Information diffusion in online social networks: a survey. SIGMOD Rec. 42, 2 (July 2013), 17-28. DOI=10.1145/2503792.2503797 http://doi.acm.org/10.1145/2503792.250379

[5] Adolescent health network dataset -- KONECT, July 2016.

[6] S. Adali et al., "Measuring behavioral trust in social networks," Intelligence and Security Informatics (ISI), 2010 IEEE International Conference on, Vancouver, BC, Canada, 2010, pp. 150-152. doi: 10.1109/ISI.2010.5484757

[7] T. DuBois, J. Golbeck and A. Srinivasan, "Predicting Trust and Distrust in Social Networks," Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on, Boston, MA, 2011, pp. 418-424. doi: 10.1109/PASSAT/SocialCom.2011.56

[8] Adalı, S., Chan, K. & Cho, J. Comput Math Organ Theory (2015) 21: 461. doi:10.1007/s10588-015-9193-x

[9] Nematzadeh, Azadeh and Ferrara, Emilio and Flammini, Alessandro and Ahn, Yong-Yeol, "Optimal Network Modularity for Information Diffusion"

[10] Radas, S. (2005). Diffusion Models in Marketing: How to Incorporate the Effect of External Influence?. Privredna kretanja i ekonomska politika, 15(105), 30-51. Retrieved from http://hrcak.srce.hr/18398

[11] http://www.columbia.edu/cu/psychology/terrace/w1001/readings/milgram.pdf

[12] Dodds, P. S., Muhamad, R., and Watts, D. J., An experimental study of search in global social networks, Science 301, 827-829 (2003).

[13] Yan, Zhijun, et al. "The effect of Trust on Information Diffusion in Online Social Networks." PACIS. 2014.

[14] Barrio, Rafael A., et al. "Dynamics of deceptive interactions in social networks." Journal of The Royal Society Interface 12.112 (2015): 20150798.