Enhancing Network Security By Using Social Contacts

Syed Safi Ali Shah
Aalto University School of Science
Technology
safi.shah@aalto.fi

Abstract

As the popularity of online social networks increases every day, ideas have been suggested where information from such networks is fed into real world communication systems to facilitate the secure and quick exchange of information. One such idea is establishing trusted links on an ad-hoc basis between previously unknown parties based on a certain k-level distance between them on an online social network. This can be understood as a variation of the idea of creating web-of-trust, where communicating parties can effectively develop trust consensus about each other without the need for a secure infrastructure. The applicability of such a scheme finds great potential in mobile ad-hoc networks and more so in delay tolerant networks where the communication with a trusted central certificate authority is not always possible. This paper aims to discuss the various implications such a scheme might have and presents a few ways which can facilitate the use of data from online social networks to create trusted links in real world communication systems.

1 Introduction

Social networking has recently gained immense popularity throughout the world. People have used online social networking portals to either establish new relations or to strengthen the ones that they already had. A very obvious effect that online social networking has had on our lives, is that such portals have lead to an expansion of our social circles. This means that while in the confines of the real world, a user might only interact with say, 30 people on a regular basis, but his online social circle would have at least a hundred direct connections [interaction09], and then through his direct connections, the user is also connected to thousands of other people whom he might not know personally but can establish a trust relation with them, if and when needed, through their mutual friends.

The online social network Facebook claims that it has more than 500 million active users currently, and an average user has 130 friends [facebook stats]. Recent studies have proposed the use of such online communities to develop trust relations in real world communication applications. This would help develop secure communication channels without the need to authenticate with a pre-issued certificate or digital signature. This idea finds great applicability especially when communication with the central certificate authority is not possible or is not feasible.

However putting such a system to wide spread use, and its success greatly depends on a number of uncertainties which need further investigation. For instance how realistically can the statistics and friendship links from online social networks be applied to real world scenarios. Are the online social links generally trust worthy enough to be used for critical information exchange in the real world. Moreover taking into account that a certain proportion of online identities are fake or forged, how would they threaten the applicability of a real world social network. Should such sybils be tolerated or completely removed. How efficiently can the sybils or fake identities be detected, and how would people react if they were forced to link their real world identities to their online social networking accounts. This paper aims to answer these questions and debate over the various possible methods to ensure more authentic social linkages in online social networks. The paper is organized as follows, section 2 aims to discuss the main reasons that motivate the need for using social contacts to establish trust in networks. Section 3 describes how data from an online social network can be gathered and presented in the form of concrete statistics which can be analyzed for their application in real communication systems. Section 4 describes the concept of web of trust, how it is used in the context of Pretty Good Privacy (PGP) and its comparison to the web of trust built on social contacts. Section 5 then describes the concept of interaction graphs and how they prove to be better indicators of trust in online social networks. Section 6 then discusses briefly a few shortfalls in the proposed trust model, and what hinders its widespread deployment. Section 7 then concludes with a few last remarks.

2 Motivation

Talking of peer to peer based networks, it is often seen as an ideal scenario where everyone can talk or exchange information with everyone without the need for any central infrastructure. However one of the greatest challenges to realizing this is that in the absence of a central and managed infrastructure, through which all information must flow, there is a constant security risk. In peer to peer networks, nodes communicate directly with each other, but how should one make sure that the entity trying to communicate is in fact a trustworthy entity and not some malicious attacker sending unwanted data.

A scheme much followed in networks of past and today is that of Public Key Encryption [1]. It basically involved presenting signed certificates before a communication channel can be established for the first time. However such certifi-

cates need to be verified by a certificate authority and/or a trusted third party. Such a scenario may not be practical and applicable in all cases. Such as:

1. The nodes lie in a disconnected ad-hoc network such as a Mobile Adhoc Network or a Delay Tolerant Network where the link with the central certificate authority is not readily available.

2. Asking every user to present a certificate prior to communication causes overhead and runs counter to the open membership policy which is the heart and soul of the success of many peer to peer systems. [9]

Also there are a number of inherent risks associated with the use of central certificate authorities [11], such as:

1. How to select and establish one certificate authority which everyone accepts and trusts globally. This has actually lead to the existence of many certificate authorities in the world wide web today.

2. One single certificate authority can be a single point of failure in the network.

3. The central certificate authority can be the target of denial of service attacks and other such threats.

All of these reasons and others open up grounds for a more de-centralized system by which secure and trustworthy connections can be established on the go.

3 Social Graphs

Online social networks have been shown to be a promising domain which can be utilized to enhance many communication scenarios. However before they can be put to reasonable use, an analysis of their social characteristics needs to be made. Social graphing is the technique employed to study social networks. A social graph essentially can be defined as Şthe global mapping of everyone and how they are relatedŤ [2]. However as more research has flourished in this field, different variants of social graphs, showing different relationships between users have been introduced. For instance, social graphs can be used to show the social degree of users, which means that how many friends is a node directly connected to. They can also be used to show the clustering coefficient, which represents how closely the nodes are bonded within their own communities or localized cliques. Social graphs might also represent the path length distribution between random users on an online social network. Also, a social graph might represent the interaction levels between users. Such graphs form the foundation of any application relying on connection made through online social networks. To name a few such applications and ideas we see that routing of packets in adhoc networks, especially delay tolerant networks (DTNs) has been shown to improve when social graphs are used to forward packets based on opportunistic connections between mobile nodes [7]. SybilGuard uses social networks to detect sybils or fake identities in a peer to peer network [11]. Reliable Email (RE) uses social connections to filter out unwanted email and at the same time providing a better guarantee that useful emails will not be filtered out, than the current email spam filters [6]. Moreover, content distribution systems such as bit torrent can also benefit from online social networks by preventing selfish behavior of non-cooperating nodes. This is shown in detail in [5].

Quite recently many elaborate studies have been conducted on the most widely used online social networks, and the statistics collected from such studies show even more promise for using social networks to enhance other communication applications. For instance [10] lists detailed statistics of users registered on facebook organized into regional segments representing some of the most populous urban areas in the world. According to this study, facebook users living in big cities like London, Toronto, New York, Manchester etc have on the average an eccentricity of only 5. Eccentricity means the distance between a node and any other node in a social graph. This shows that users are typically more closely knit than we generally would imagine.

4 Web of Trust

The need to eradicate the requirement for a trusted third party or a certificate authority has been seen from as long as the public key infrastructure has been developed. The downside to trusted third parties which can issue and verify public keys or certificates on the go, lies in their central architecture. A central trusted third party means a single point of failure. A single entity which should have 24/7 availability and can cater to an ever increasing population of internet users, sounds a bit of a stretch.

Phil Zimmerman came up with a solution as early as 1991 when he proposed the encryption mechanism called as ŞPretty Good PrivacyŤ. PGP uses the concept of web of trust where users sign each others keys on the basis of earlier contact and personal relationships. Thus instead of having a central certificate authority, the key distribution and verification is done in a distributed fashion by the users themselves. Each userŠs public key can contain a number of digital signatures of the so-called ŞintroducersŤ. This is to say that the introducer can vouch on behalf of the party whose public key it has signed. Thus if A and C donŠt know each other earlier, but AŠs public key has been signed by B, and B is known previously to C, then C can develop a trusted connection with A [3].

Recently however there has been a lot of debate whether social network connections can be used as a good basis for establishing trust linkages between communicating peers. For instance if we maintain a metric k, such that if party A is connected to a party B with not more than k hops on an online social network, then we say that A and B can trust each other, and can thus proceed with regular communication. Here the value of k defines the strictness of the trust establishment. It can have values from 0 to N, where 0 would mean that you only trust your directly connected friends, 1 would mean that you also trust your friends of friends and so on.

[7] also addresses the problem of storing and maintaining the social linkages information on the small sized memory of mobile nodes. There are basically two considerations that

need to be addressed. First is the small memory of mobile nodes, which means that the social graph information stored on such devices has to have certain bounds on its size. The second problem is that of privacy. For sure all information regarding ones social links can not be stored on the mobile device in plain text, lest it might fall in the hands of an adversary. [7] solve these problem by applying first applying community detection on the data derived from online social network and then computing a digest of this information. This digest is called a community digest, and [7] claims that such a digest is small enough to fit in the memory of small mobile devices and also the digest hides the possible privacy sensitive social information from unknown third parties.

Once two devices meet and want to communicate securely, they can exchange their community digests. An intersection of the friendship links can then be performed on the exchanged community digest and a partyŠs own digest. Once a k-level link (meaning that the two parties are connected by k friends in-between them) is found, the communication can proceed under trust.

4.1 Comparison to PGP Web of Trust

The idea proposed in this paper in some ways in similar to the idea of the web of trust being used extensively by PGP. A loose logical relation can be established between the PGP trust model and the social network trust model. In a social network, each user has a profile which contains information visible to other users on the social network. Thus a profile can be understood as a public key certificate belonging to the owner of the profile. Each profile then has a number of ŞfriendsŤ mentioned on it. These friends can be understood as the digital signatures of people who trust the owner of the profile and have thus chosen to be his friends (and have thus digitally signed his profile).

However there is more promise in using social networks than PGP web of trust model. In PGP trust model, a public key certificate can either be signed by another PGP user or not signed at all. This is a two level trust establishment, meaning that A is either trusted by B or not trusted at all by B. However, when we come to social graphs, they contain much more information than this. Community detection shows which users share how much of their interests and are how closely bonded. Social Interaction graphs [10] show how often two users communicate and thus can be a direct indicator of the level of trust they both have in each other.

5 Trust Indicators in Online Social Networks

The discussion from the last section brings us to the question: Are social linkages online, really a representation of trust two parties have in each other.

We see that on facebook, as well as on many other online social networks, a common user is allowed to add logically unlimited other users as his friends. Users of such social networks typically have a tendency to add online profiles of celebrities or such other famous profiles, which then become as hubs in online social networks [8]. Such relationships although are bidirectional but, are not representative of bidirectional trust in general. The presence of these hubs in the social networks, distort the community structure such that a typical social graph can not be used to derive trust relationships straight away. In general all social links are not equally useful when analyzing the social networks, since only a considerably smaller percentage of users are the ones which are actively engaged in the network [10]. In other words, not all social links represent active social interaction. To study the interactions between users, instead of just merely basing our analysis on social links, [10] introduces what is called as interaction graph.

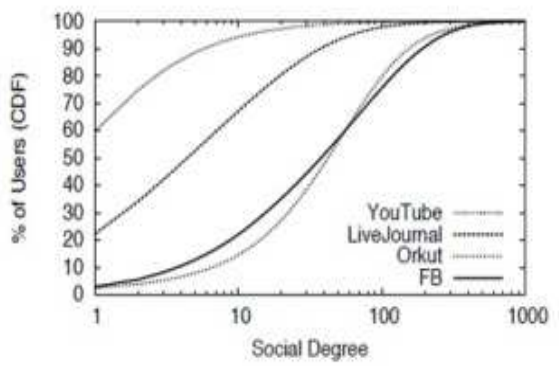

Figure 1: A social graph showing the social degree of users on different online social networks [10]

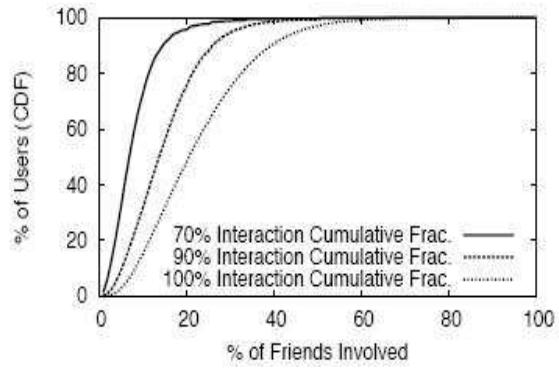

Figure 2: An interaction graph showing that majority of the interactions occur with only a small percentage of friends [10]

Interaction graph distinguishes between a userŠs active relationships and those which are merely associated for the sake of sheer name. It measures the number of interaction events between two users within a finite window of time. This makes up an interaction rate. An interaction graph thus clearly shows the social links which have an interaction rate greater than a certain lower threshold. Such links between two users, which fall above a lower bound of interaction rate, can be considered as trusted social links.

Another problem with interactions is that, by nature, an interaction event, such as sending a message to a friend, is a uni-directional event and it does not necessarily means that an interaction in one direction will also trigger an interaction

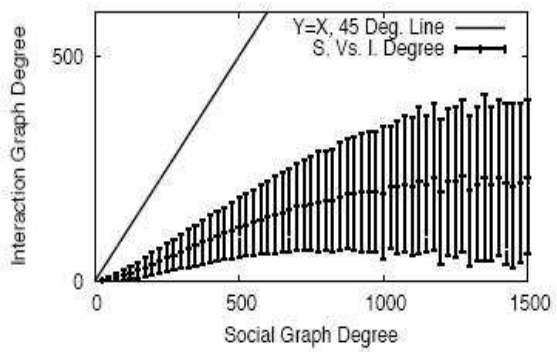

Figure 3: Comparison of social graph degree with interaction graph degree [10]

in the reverse direction. The worst case scenario can be understood as spamming, where one user keeps sending messages to everyone but seldom gets replied. Such a unidirectional link can not be considered as a trusted link. According to studies in [10], for 65

To make the point very clear, the figure below shows the disparity between social degree and interaction degree. That is to say that there is a marked difference between the friends a user has and the active social relationships he maintains out of his friends.

6 Challenges

When it comes to applying data from online social networks, to real world communication systems, one has to be careful if the data is consistent to be mapped to the real world.

One of the major concerns for using online social network links in real world communication is the authenticity of identities online. The existence of sybils in online social networks has been debated very strongly over the past few years. Sybils are nodes, which effectively forge identities, and thus try to gain trust of people by masquerading as someone else [4]. This concept can also be extended to take into account multiple identities that people tend to create in OSNs, where each identity is used to communicate with a different set of people online. Sybils have remained a great challenge for all peer to peer distributed networks and while many schemes have been proposed to defend against such attacks, their applicability and success is still under question.

Presence of sybil nodes greatly challenges the social trust relationship data that is collected from such online networks. Many schemes have been devised to avoid sybils in online social networks. The simplest of which is to somehow force each user to register his account or user profile with his social security or bank account number. However as discussed in the beginning of this paper, such a scheme runs counter to the open membership policy that is the heart and soul of such online systems. Plus a central trust worthy authority has to be installed which can take care of such confidential data. Security of such a central entity is also a challenge. [9] discusses that most of the more elaborate sybil detection algorithms work by detecting community structure in social networks. That is to say that nodes which are more sparsely connected and have few links to other communities, are mostly labeled as sybils. This method has roots in the general idea that it is particularly difficult for a sybil node to form and maintain a large number of links to other real nodes. Also if this is the case, using interactivity levels will limit the effect of sybils on our trust model, since a sybil node will generally lie below the interactivity threshold we set for demarcating trust. Research is underway to make such schemes more efficient.

7 Conclusion

Social networks are still in the phase of evolution, and day by day their footprint on our lives in getting bigger. Today they host huge amounts of valuable information about users, and now their domain is expanding to also cover the mobile media. GPS might soon get incorporated in social networks to maintain detailed location specific data about users.

This paper aims to strengthen the idea that such vast data from online social networks should be fed into real world peer to peer applications to make better use of it. One such idea that is proposed in this paper is the use of social links for better security considerations in peer to peer applications. We try to develop the idea that other communication applications can benefit from the trust relationships established in online social networks, so that a central certificate authority is no longer required. Instead its job is spread out in a distributed fashion where each node verifies its peers on the basis of social trust linkages. A lot of work has already been put into the field of analyzing social networks, however there is a need to better focus the efforts to put the results of such analysis to real life use.

References

[1] Standard specifications for public-key cryptography. Technical report, IEEE, 2009. `http://grouper.ieee.org/groups/1363/`.

[2] Facebook: One social graph to rule them all? Technical report, CBS News, SanFrancisco, April 21, 2010.

[3] A. Abdul-Rahman. The pgp trust model. 1997.

[4] J. douceur. The sybil attack. 2002.

[5] W. Galuba, K. Aberer, Z. Despotovic, and W. Kellerer. Leveraging social networks for increased bittorrent robustness. 2009.

[6] S. Garriss, M. Kaminsky, M. J. Freedman, B. Karp, D. Mazières, and H. Yu. Re: Reliable email. 2006.

[7] P. Hui and N. Sastry. Real world routing using virtual world information. 2009.

[8] B. E. Ur and V. Ganapathy. Evaluating attack amplification in online social networks. 2009.

[9] B. Viswanath, A. Post, K. P. Gummadi, , and A. Mislove. An analysis of social network-based sybil defenses. 2010.

[10] C. Wilson, B. Boe, A. Sala, K. P. N. Puttaswamy, and B. Y. Zhao. User interactions in social networks and their implications. 2009.

[11] H. Yu, M. Kaminsky, P. B. Gibbons, and A. Flaxman. Sybilguard: Defending against sybil attacks via social networks. 2006.