From Subjective Trust to Objective Trustworthiness in On-line Social Networks: Overview and Challenges

David Zejda

University of Hradec Králové

Faculty of Informatics and Management

david@zejda.net

Abstract: Nowadays dozens of people share their content in the current Web 2.0 space, talk with friends in social networking sites such as Facebook and live on the Net in many other ways. They do all this quite naturally, forgetting the healthy cautiousness sometimes. In real life we rely on trusted people. Do we know how to reflect real-world trust mechanisms into on-line social software? In the article we focused to bring overview on state of the art in main ideas behind a trust processing in online social networking systems. What are common sources of subjective trust, how the trust emerges and how can be captured into the systems? How can be explicit trust processed to infer indirect trust, the trust between users who do not know each other? And what are the ways to to infer trustworthiness, the objective metric of trust? Finally, we point out selected challenges related to the trust in current highly dynamic social networks.

Key words: trust, trustworthiness, reputation, social networks, social networking, inferred trust

1. The trust

The conception of trust has a key role in social exchange theory [1]. Both dynamics of our social relations and also individual social interactions are highly influenced, if not even governed by the trust. Trust may be defined as "the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party." [2]

In the real life the trust emerges from our experiences with others and also from recommendation or guarantee from those, who we trust already. We deal differently with trusted people than with strangers. The level of trust which we feel toward someone helps us to decide how thoroughly to check his proclamations or promises. The trust helps us take the right decisions, such as whom to entrust certain information, task or other person to care for.

2. Trust in virtual milieu

We all belong to a global-world village. As expressed in a small world phenomenon, everyone is connected with anyone else through only several steps of relations [3]. New social strategies are necessary to cope with the social and information overload [4]. Web becomes not only bigger every year, but also semantically richer and more driven by a community. Besides milieu for implicit socialization [5], the web provides dating sites, community portals and social networking sites, such as Facebook. Actually, reputation of social networking sites has been affected by many incidents. Is it possible to join a site with millions of users and trust all of them? Of course not. Though there are risks and we may say well known risks, many people are still being attracted to not only join, but also communicate carelessly, and even reveal quite personal and exploitable information. [6]

With the recent incidents on mind, importance of better trust solutions in social software is increasingly apparent. Besides the ability to react quickly on malicious attempts to attack the site or it's users, we are in need of solutions to foster convenient, natural and safe trust formation among users and their shared content.

What are key characteristics of trust in context of virtual communities? Meo at al. [7] define three main aspects. We may refer to them as multidimensionality, contextuality and scope of relevance. Finally, we added one more, not discussed in the referenced article – a lack of soft indices.

1. *Multidimensionality*. There are many factors to be considered to evaluate the trust. Usually we have to take many traits of the party into account, such as honesty, experience, precision, efficiency or cooperativeness. The broader social space may bring further dimensions.
2. *Contextuality*. Not only the social context does matter, but also the purpose of trust evaluation or the 'contextual theme'. We may talk about theme-contextual trust, such as when you are in search for reliable advices and you trust more to experts on certain domain.
3. *Scope of relevance*. Trustor in our scenario performs his trust evaluation within a virtual community. The result reflects his subjective view. Besides this, we may talk about community-wide or system-wide metrics, refered to as reputation or trustworthiness.
4. *Lack of soft indices*: In a virtual space we actually miss many relevant non-verbal indices which usually help us in the process of trust formation. We do not see the person in real, sometimes we even do not see him at all. It is also more likely

that there are no other trustful people around who could share their opinions based on their direct personal experiences.

Further in the article we describe how explicit trust emerges and further evolves, followed by paragraphs about how could be the trust used to infer subjective indirect trust between users, who do not know each other.

3. Explicit trust

What are common sources of trust among users within social networking systems? Online transactions are only technically flavoured variants of similar transactions usually performed as off-line in real world, without the technical support. [6] Existing trust may be captured from the real social background by technical means and mapped into the system [4]. So, if you personally invite someone to join a networking site, you probably know him and trust him, at least at certain level. Besides this, new trusted friendships may arise out of vital interactions within the site, usually during some period and based on certain level of interaction. Last to mention, similarity is not equal to trust, of course, but correlation between similarity and trust evolves quite often [4], thus some matching functions may help users to find prospective trustees.

We already mentioned the contextuality of trust. Yan et al. [8] redefine trust as "trustor A trusts trustee B for purpose P under condition C based on root trust R". The main difference is in the element C, condition to trust. Trustor should be informed about any distrustful behaviour of the trustee according to the conditions and a trust is considered as something dependant on the conditions. Level of trust considered sufficient for certain purpose differs. [9] Trustors may have e.g. different personal preferences and requirements on a time and deepness of communication before falling into trust. The preferences may depend on many factors – on type of relationship or transaction, on previous history of the possible-trustee within community, and so.

Of course, trust is inherently dynamic. Caverlee at al. [10] recommend to fold two main sources of information in well-designed trust metric. The first source is the relatively static network topology, as they recommend with quality of relationships taken into account. The other proposed source is users' behaviour. A feedback mechanism capturing influences of behaviour on trust brings the necessary dynamics.

In general, trust grows slowly, but falls sharply [4]. It may take months or years before you trust someone. A single act of betrayal destroys the trust to the roots. Algorithms used by social networking sites should reflect this behaviour. Also, besides positive trust expressions social software users should be granted with means to withdraw the trust and express loss of confidence, the distrust instead. As an example, Moghaddam et al. [11] provide model for rapidly evolving networks, with emphasis on feedback as a source of trust.

Trust may be gained, lowered, or even lost. Goldbeck et al. [12] offer conceptual representations of failures of trust, such as distrust, mistrust, untrust and ignorance. Explicit distrust may be utilized by social networking site maintainers to reveal malicious users, such as scammers or other betrayals. For further examination of the case it may be useful to allow or even require to provide reasons for the distrust expression. The loss of trust is not necessarily terminal – it may be followed by a recovery of trust – when regret followed by forgiveness take place. [12]

4. From explicit to inferred trust

How to establish trusty relationships in on-line social network, where no existing trusted social network is present in the background? [13] How to measure a reliability of advices provided by strangers? We are indeed in need of some metric of indirect trust. On the following lines we do not wish to plunge deeply into certain model. We just point out selected aspects, which are essentially common to computational models of social networks of any kind.

Most common way to represent a social network is by means of discrete mathematics. The network is being viewed as a graph with users as nodes and relationships between them as edges. Though most models use graphs, they differ in computational methods, e.g. Meo et al. [7] distinguish computational graph-based, link-based, and expert-finding trust models.

The trust itself, either explicitly expressed or revealed by inference, once captured may be represented as a binary value (do trust – do not trust), as a discrete scale (levels of trust), or real scale, usually normalized into certain interval. For easier further complementary computations Walter et al. [14] recommend to use the same range 0..1 for all trust-related variables including expressed trust, inferred trust and all variants of objective trustworthiness discussed later, so the values probably need some normalizations.

Assuming the trust is transitive [15], the basic idea of indirect trust inference is to multiply trust values along the path between users. The multiplication effectively discounts the resulting value, thus those whom the user trusts already are being taken more seriously e.g. as a source of recommendations whom else to trust.

Walter et al. [14] present fairly tuned metric. The algorithm does not reduce cycles in a graph before computation as most other algorithms do. In recommender systems the algorithm gives best results when used to find recommendation for a different category of media than in which the user recommended already, such as when user who posted comments on cartoon movies asks for recommendation on drama. Hales at al. [13] use similar algorithm to find cooperative routes among selfish agents acting as players in prisoner's dilemma, in an environment with no central authority.

5. From the individual trust to a system-wide trustworthiness

Besides the individual trust either expressed or inferred, the subjective personalized trust metric, we may be in need of more general, either per-community related [3] or system-wide trustworthiness. Whereas trust is something between two people, with trustworthiness we mean overall reliability of the user, his credit, site-wide reputation.

As a source of trustworthiness we may take subjective trust, both explicit and inferred. Activity of user within the system in the past,

such as how many times he failed to deliver goods ordered in auction or how often has been his wiki contribution re-edited may serve as the source too even if not calculated into the trust.

Back to subjective trust as a main source of trustworthiness, logically, more incoming trust increase the overall trustworthiness of certain user. But instead of simply summing individual trusts, it is better to apply certain, usually eigenvector-type[1] algorithm [12] to weigh the individual values according to trustor's own trustworthiness. The trustor's trustworthiness may be viewed as a confidence of his own trust expressions [16]. In result, the trustworthiness of certain user is dependent on trustworthiness of his neighbours in the graph of trust [14].

Not-yet-much-trustworthy users may be allowed to express their trust towards others, though these expressions are not treated as much relevant, until the trustors themselves gain enough trustworthiness. Eigenvector algorithms also take count of outgoing trust expressions into account. If user with certain trustworthiness expresses his trust toward a single user, the single expression is being considered as of greater value than if he trusts dozens of other users. Meo et al. [7] offer a model of more trust-related metrics – the trust, the trustworthiness (referred to as reputation in the work) and a reliability. The metrics in their model are parametrizable, e.g. with preferences on so-called correctness / novelty ratio. Pavlovic [3] further recommends to focus on user's attitude toward trust. The attitude may be used to normalize user's trust expressions.

Not only trust but also trustworthiness may undergo transitions, such as from 'unknown' or 'not-yet-trustworthy' to 'trustworthy' when user reaches a threshold, defined either per-community or for the whole site. The whole process may be fully driven by peers. Alternately in systems with central authority an approval by site maintainers may be required for major trustworthiness transitions.

On the other end of it's life-cycle, the trustworthiness may be 'disputed' when the user loses confidence and, maybe, the case is under consideration by site maintainers. Impeacher's trustworthiness in comparison with trustworthiness of the user whose reliability has been disputed may be used as a weight of the withdrawal. Whereas trustworthy users may be granted with more privileges, users who are losing the trustworthiness and becoming 'untrustworthy' lose the privileges in parallel. Further, if someone loses the trustworthiness, it automatically affects trustworthiness of those, who received his trust. The trust relations of the disputed user may be examined by maintainers. They may try to find either other victims of malicious behaviour in order to warn them or provide other assistance, or try to find complices on the other hand.

6. Utilization of trust and trustworthiness

Once captured, trust may be utilized in many ways and for many purposes. Recommender systems, website access control systems or e.g. message filtering may be built on top of trust metrics or augmented using them. [12] The trust is pivotal in social relationships [17] and for online transactions [18].

We may infer quality and credibility of certain content based on trust or trustworthiness of the author or content provider. E.g. Moturu et al. [19] focused on health as the negative impacts are high for this domain. They developed a vital model to quantify utility and trustworthiness of content to guide users toward both relevant and credible information.

Carminati et al. [20] provide another application – rule-based access control mechanism specifying access policies on the resources owned by web social network participants. In their model access control enforcement is carried out client-side and access to a resource is granted when the claimer is able to provide a proof of being authorized.

7. Current challenges

There had been impressive visions of trustworthy Internet, such as Augmented Social Network [21], where internet-wide persistent online identity across systems would facilitate reliable interactions of so called 'citizens of the Net'. A lot of work has been done to make Internet more trusty space already. We have security and trust authorities, security certificates, great algorithms, trust-related ontologies, whole area of trust management, some great systems. But seven years passed since ASN vision and Internet in general is not more reliable than before. The advancement is being effectively outweighed by more sophisticated efforts of deceivers.

We may mention some of the most proliferous:

- *Scamming and phishing:* Scammers are increasingly more proficient, with both technical and social skills. You would probably never give money based on poorly written scam e-mail. But what if you are being contacted by your friend or relative, who has been coincidentally trapped somewhere without a coin?
- *Impersonating and profile hijacking*: One of trends is creating false profiles or hijacking profiles for scamming or similar fraudulent purposes. Ironically, the illusion of security on sites which take safety seriously may lower cautiousness of users, leading to even higher dangers if the fraud occurs. [22]
- *Cyberstalking:* Social networks give vital ground for cyberstalking or cyberbullying, varying from false accusations to gathering information for further harassment.
- *Trust authority compromising:* Institutional trust authorities are targeted often by attackers and they are vulnerable. Besides this, power-law distribution where rich becomes even richer works in trust systems similarly – trusted nodes tend to receive even more trust. It leads to constitution of so called 'trust hubs' [3], informal trust authorities in a space of the social network. Importance of institutional or informal trust authority intensifies the impact when the authority is being compromised.

Caverlee et al. [10] noted two aspects of current social networks, which makes the dangers even worse. First, the small world phenomenon causes, that there is a short distance in the network between any two participants. Even if user is able to control his

1 One of well-known eigenvector-type algorithms is PageRank by Google.

direct friends, malicious users may be only few hops further. Second, the user has limited network view, so even if he controls his friends and maybe friends of friends, he has no idea about credibility of other participants.

More work is needed in areas such as:

- *Complicated settings:* Algorithms and metrics to manage trust information shared e.g. in multi-agent systems [23] and under conditions of uncertainty.
- *Continuous fight with deceivers:* Continuously search for ways how to keep networks useful and sufficiently safe and trusty, despite increasing activities of defective peers [13] or agents with random, selfish or even malicious behaviour [4]. It's a never-ending fight with strikes and counter-strikes.
- *Privacy:* Keep privacy questions on mind while dealing with the security. Sometimes the requirements may be contradictory.
- *Trust identity:* It will be furthermore a long path from system-wide trustworthiness to a global trust identity, shared among systems. Ontologies seems to be a good glue to facilitate the interoperability, but we will see.

We have to secure the social software itself, foster growth of confidence among users and their content and deal with all those matters of trust mentioned in the systems with steadily increasing dynamics, where millions of users are joining, performing their activities and leaving.

Conclusions

The article has brought overview on trust matters in on-line social networks. We mentioned some basic ideas about the trust itself, about sources of trust, emergence of explicit trust, mechanisms for trust processing and inference of indirect trust. Further we explained that subjective trust, either explicit or inferred, may be used as a source of objective, either community-wide or system-wide metric, the trustworthiness. Well designed trust and trustworthiness models and algorithms could be used in trust systems to foster reliable interactions among users, to augment utility of shared content providing a property of reliability. Trust may be used also as a major source for access control and recommender systems. Fraudulent scamming, cyberstalking and other malicious efforts in the current highly dynamic milieu of social networks brings new challenges to cope with.

Bibliography

[1] D.M.E. Roloff, *Interpersonal Communication: The Social Exchange Approach*, Sage Publications, Inc, 1981.

[2] R. Mayer, J. Davis, a D. Schoorman, "An Integrative Model of Organizational Trust," *The Academy of Management Review*, vol. 20, 1995, s. 734, 709.

[3] D. Pavlovic, "Dynamics, Robustness and Fragility of Trust," *Formal Aspects in Security and Trust: 5th International Workshop, FAST 2008 Malaga, Spain, October 9-10, 2008 Revised Selected Papers*, Springer-Verlag, 2009, s. 97-113.

[4] F. Walter, S. Battiston, a F. Schweitzer, "A model of a trust-based recommendation system on a social network," *Autonomous Agents and Multi-Agent Systems*, vol. 16, nor. 2008, s. 57-74.

[5] P.O. Wennerberg a T. Oellinger, "Ontology Based Modelling and Visualization of Social Networks for the Web: Discovering Security Related Information from Online News Sites," 2006.

[6] C. Dwyer, S. Hiltz, a K. Passerini, "Trust and privacy concern within social networking sites: A comparison of Facebook and MySpace," *Proceedings of the Thirteenth Americas Conference on Information Systems*, Keystone, CO, USA: 2007.

[7] P.D. Meo, A. Nocera, G. Quattrone, D. Rosaci, a D. Ursino, "Finding reliable users and social networks in a social internetworking system," *Proceedings of the 2009 International Database Engineering & Applications Symposium*, Cetraro - Calabria, Italy: ACM, 2009, s. 173-181.

[8] Z. Yan a P. Cofta, "A Mechanism for Trust Sustainability Among Trusted Computing Platforms," *Trust and Privacy in Digital Business*, 2004, s. 11-19.

[9] Z. Chunying a C. Huajun, "Social network mashup: Ontology-based social network integration for statistic learning," 2008.

[10] J. Caverlee, L. Liu, a S. Webb, "Towards robust trust establishment in web-based social networks with socialtrust," *Proceeding of the 17th international conference on World Wide Web*, Beijing, China: ACM, 2008, s. 1163-1164.

[11] S. Moghaddam, M. Jamali, M. Ester, a J. Habibi, "FeedbackTrust: using feedback effects in trust-based recommendation systems," *Proceedings of the third ACM conference on Recommender systems*, New York, New York, USA: ACM, 2009, s. 269-272.

[12] J. Golbeck, *Computing with Social Trust*, Springer, 2008.

[13] D. Hales a S. Arteconi, "Friends for Free: Self-Organizing Artificial Social Networks for Trust and Cooperation," 2005.

[14] F.E. Walter, S. Battiston, a F. Schweitzer, "Personalised and dynamic trust in social networks," *Proceedings of the third ACM conference on Recommender systems*, New York, New York, USA: ACM, 2009, s. 197-204.

[15] J. Huang a M.S. Fox, "An ontology of trust: formal semantics and transitivity," *Proceedings of the 8th international conference on Electronic commerce: The new e-commerce: innovations for conquering current barriers, obstacles and limitations to conducting successful business on the internet*, Fredericton, New Brunswick, Canada: ACM, 2006, s. 259-270.

[16] Z. Yan a S. Holtmanns, “Trust Modeling and Management: from Social Trust to Digital Trust,” *book chapter of Computer Security, Privacy and Politics: Current Issues, Challenges and Solutions*, 2007.

[17] F. Fukuyama, *Trust: The Social Virtues and The Creation of Prosperity*, Free Press, 1996.

[18] N.W. Coppola, S.R. Hiltz, a N.G. Rotter, “Building trust in virtual teams,” *IEEE Transactions on Professional Communication*, vol. 47, 2004, s. 95–104.

[19] S.T. Moturu, J. Yang, a H. Liu, “Quantifying Utility and Trustworthiness for Advice Shared on Online Social Media,” *Computational Science and Engineering, IEEE International Conference on*, Los Alamitos, CA, USA: IEEE Computer Society, 2009, s. 489-494.

[20] B. Carminati, E. Ferrari, a A. Perego, “Enforcing access control in Web-based social networks,” *ACM Trans. Inf. Syst. Secur.*, vol. 13, 2009, s. 1-38.

[21] K. Jordan, J. Hauser, a S. Foster, “The Augmented Social Network: Building Identity and Trust into the Next Generation Internet.”

[22] S. Barnes, “A privacy paradox: Social networking in the United States,” *First Monday*, vol. 11, 2006.

[23] C. Hang, Y. Wang, a M.P. Singh, “Operators for propagating trust and their evaluation in social networks,” *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems - Volume 2*, Budapest, Hungary: International Foundation for Autonomous Agents and Multiagent Systems, 2009, s. 1025-1032.