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

David Zejda
University of Hradec Kralove
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 what are the sources of trust dynamics? How can be trust 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 infer objective metrics of trust, the reputation or trustworthiness? 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 should we 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 need solutions to foster convenient, natural and safe trust formation among users.

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 and the contextual setting. When you search for reliable advices on accounting, you trust to experts on the accounting domain, whereas their opinions on machinery are probably not so much relevant.
3. *Scope of relevance*. Trustor in our scenario performs his trust evaluation within a virtual community. The result reflects his subjective view, useful for certain purposes. Besides this, we may talk about community-wide or system-wide metrics, refered to as reputation or trustworthiness, suitable for other purposes.
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. Sources of 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] So, trust existing already may be captured from the real social background by technical means and mapped into the system [4]. For example, if you personally invite someone to join a networking site, you probably know him and trust him, at least at certain level. The trust has been established already, based on personal experiences, such as on a willingness of the party to help you in difficulties, on promises which has been kept, and so. In this case, you as the user do not ask the system to help you to establish the trust, but rather you are the one providing explicit trust indices to the system. So, the pre-existing trust based on personal experience may be one of the sources of explicit trust.

Besides this, new trusted friendships may arise out of vital interactions within the site, usually during a sufficient period of time and based on a sufficient level of activity. Similarity is not equal to trust, but correlation between similarity and trust evolves quite often [4], thus some matching functions may help users to find prospective trustees. Once initial touch established, users may begin to trust each other if they exchange several messages or some other content. Of course, dangers of various types of frauds affect the emergence of trust. If users have the dangers on mind, it increases their requirements on the proofs which the other party has to provide. Deeper, longer, more factual communication may be needed. We may identify two complementing types of errors which affect the trust formation:

1. *Excessive prudence.* User is too suspicious that he indeed does trust virtually nobody, even though there are enough positive indices.
2. *Trust to deceiver.* User is either careless and does not check proclamations enough and believes quickly, or he is prone to fraud attempts.

While longer trust formation brings certain losses, such as lowered convenience, harm of the latter type of error may be much higher. Abuser who enjoys the confidence has wide opportunities to attack the party. Actually it was the strategy of many deceivers to gain confidence first, which has led to the most severe abuse incidents on social networking sites. The disproportion of possible harms should be taken into account when designing trust-related systems. Users should be supported to be able to distinguish who deserves their confidence and who does not.

4. Dynamics of trust

Of course, trust is inherently dynamic. Caverlee at al. [8] 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 further users' behaviour. A feedback mechanism capturing influences of behaviour on trust brings the necessary dynamics. But the behaviour of users is not the only source of trust dynamics. We may identify other, both general and more particular, subjective, and subtle influences:

- *Changes in social topography.* The graph of social relations evolves as new users are joining the network and others are leaving.
- *Other users' behaviour.* Users live in the sites, posting content, sending messages, changing their profiles.

- *Evolution of preferences.* Peoples' ideas, views and preferences evolve in time. Every interaction, every experience change us somehow. Even aggregated population or community preferences may be described in terms of trends or evolving inclinations.
- *Context.* Certain user may utilize the site in various ways, in different contexts and for various purposes. The immediate need of trust varies according to the context.
- *Mood.* There are even really soft and subtle influences, such as current fettle or vein of certain user.

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. [9] provide model for rapidly evolving networks, with emphasis on feedback as a source of trust.

We already mentioned the contextuality of trust. It brings further dynamics to the possible model. Yan et al. [10] 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. [11] 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.

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 takes place. [12]

5. 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 trust in 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. We may find some models based on undirected graphs, assuming that trust is something inherently mutual, but majority of models use graphs directed, where levels of 'trusting' and 'being trusted' may vary.

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 may 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. This main idea has been extended by authors to overcome some of it's weaknesses. For example Walter et al. [14] present fairly tuned metric. The algorithm does not reduce cycles in a graph before computation as most other algorithms do. Applied 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.

6. Objective trust

Besides the individual trust mentioned earlier, either expressed or inferred, for certain purposes we may be in need of more general and more objective trust metric. Meo et al. [7] offer a model of trust with metrics highly parametrizable, e.g. with preferences on so-called correctness/novelty ratio, but that's not essential for us now. We rather wish to mention their classification of trust-related metrics. When they distinguish 'trust', 'reputation' and 'reliability', they are taking scope of relevance into account. Besides trust and reputation, we see two more levels of trust characteristics:

1. *Subjective trust.* The personalized trust metric, either explicit or inferred. The one we were talking about in the previous paragraphs.
2. *Community-wide reputation.* The credit of the user in the community within the social networking site, such as within certain formally established and named group or within informal interrelated cluster of users. The reputation may be based on community-related activity of the users and evaluated in the context of the community.
3. *System-wide trustworthiness.* Even wider metric, abstracting from community-related characteristics. Someone 'trustworthy' may be no way similar to the user asking for the value, but the user is probably not a scammer.
4. *World-wide trust identity.* Measure of confidence of certain user, exceeding borders of particular systems and shared among them. Such as if certain user trustworthy on Facebook would be considered trustworthy on Twitter and LinkedIn too. We are still lacking solutions for this level of trust.

In the further text we are referring to the wider trust metrics, the reputation and the trustworthiness in particular, as to *objective trust*. As a main source of objective trust we may take the subjective trust, both explicit and inferred. More incoming subjective trust logically brings higher objective trust of certain user. But instead of simply summing and normalizing individual subjective trusts, better algorithms are available. Particularly eigenvector-type[1] algorithms [12] are being used widely to weigh the individual values according to trustor's own objective trust. The trustor's objective trust may be viewed as a confidence of his own trust expressions [16]. In result, the objective trust of certain user is dependent on objective trust of his neighbours in the graph of trust [14].

Besides subjective trust, supplementary sources may taken into account when inferring the objective trust. For example, 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 a source too, if it has not been calculated into the trust already. [3]

Not-yet-much-trustworthy users may be allowed to express their trust in others, though these expressions are not treated as much relevant, until the trustors themselves gain enough objective trust. Eigenvector algorithms also take count of outgoing trust expressions into account. If user with certain objective trust expresses his confidence to a single user, the single expression is being considered as of a greater value than if he trusts dozens of other users. Pavlovic [3] further recommends to focus on user's attitude toward trust. The attitude may be used to normalize user's trust expressions.

We may summarize the two possible sources of objective trust mentioned and add one more:

- *Aggregated weighted subjective trust.* Total of subjective incoming confidences, weighted by trustors' own objective trusts, counts of their trust expressions and their trust attitudes.
- *Supplementary sources.* Other indices not calculated into subjective trust, such as those related to user's behaviour within system.
- *Parametrized by contextual preferences.* Parameters of objective trust inference set per-community or per-site, with possible contextually-dependent dangers on mind.

Not only subjective trust but also objective trust has dynamic flavour. It 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. On the other end of it's life-cycle, the objective trust may be 'disputed' when the user loses confidence. The transitions may be fully driven by peers or, alternately, in systems with central authority an approval or check-up by site maintainers may be required for the major transitions. In the evaluation of case, impeacher's own trustworthiness in comparison with trustworthiness of the user whose reliability has been disputed may be used as a weight of the withdrawal.

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

7. Utilization of trust

Once captured, both subjective and objective trust may be utilized in many ways and for many purposes. For example, 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 objective trust, it automatically affects objective trust of those, who received his trust. Further, the trust relations of the disputed user may be examined by maintainers. They may try to find either other victims of deceiver's malicious behaviour in order to warn them or provide other assistance, or try to find his complices on the other hand.

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. [17] 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. [18] provide 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.

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 [19] and for online transactions [20].

8. 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, such as:

- *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? Will you be able to distinguish, whether is it not an attempt to scam you? Yes, scamming is related to other threat widespread nowadays:
- *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.

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 earlier, having broader aspects of the task on mind. One of them is uncertainty and steadily increasing dynamics, where millions of users are joining, performing their activities and leaving. Sometimes besides the humans, there are multiple software agents operating in the same environment [23]. Defective peers [13] and agents with random, selfish or even malicious behaviour [4] have to be taken into account. Caverlee et al. [8] mention two interesting properties of current social networks, which complicates our efforts. 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 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. Sometimes the requirement of safety may be in contradiction with other important requirements, such as ease of use or users' privacy.

9. Conclusions

In the article we have brought overview on trust matters in on-line social networks, starting with discussion about the trust itself, sources of trust and emergence of explicit trust. Further we showed, that the virtual milieu of social networks makes the matters of trust even more complicated, bringing new complexity, while lacking of relevant indices available in a real world. Trust was discussed as something dynamic, continuously influenced by various factors. Further we outlined basic ideas of trust processing and inference of indirect trust and explained that subjective trust, either explicit or inferred, may be used as a source of objective trust metrics, such as community-wide reputation or system-wide trustworthiness. Trust systems may be used in many ways, e.g. to foster reliable interactions among users, to augment utility of shared content providing a property of reliability, as a major source of information for access control systems and for recommender systems.

The battle with deceivers from far is not at the end. Probably never-ending fight with deceivers' strikes and our counter-strikes is ahead. New forms of scamming, profile hijacking, cyberstalking and other malicious efforts in the current highly dynamic milieu of social networks bring new challenges to cope with. It will be furthermore a long path from system-wide trustworthiness to a global trust identity, shared among systems, envisioned by Jordan et al. [21]. We moved further already, meanwhile ontologies seem to provide a good glue to facilitate the interoperability, but we will see.

10. References

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

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

[3] Pavlovic, D.: *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] Walter, F., Battiston, S., Schweitzer, F.*: 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] Wennerberg, P.O., Oellinger, T.: *Ontology Based Modelling and Visualization of Social Networks for the Web*: Discovering Security Related Information from Online News Sites, 2006.

[6] Dwyer, C., Hiltz, S., Passerini, K.: *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] Meo P.D., Nocera, A., Quattrone, G., Rosaci, D., Ursino, D.: *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] Caverlee,J., Liu, L., Webb, S.: *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.

[9] Moghaddam, S., Jamali, M., Ester, M., Habibi, J.: *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.

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

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

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

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

[14] Walter, F.E., Battiston, S., Schweitzer, F.: *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] Huang,J., Fox, M.S.: *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] Yan, Z., Holtmanns, S.: *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] Moturu, S.T., Yang, J., Liu, H.: *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.

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

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

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

[21] Jordan, K., Hauser, J., Foster, S.: *The Augmented Social Network: Building Identity and Trust into the Next Generation Internet*, 2003.

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

[23] Hang C., Wang Y., Singh, M.P.: *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.