EVALUATING TRUST OR REPUTATION BASED ON SOCIAL NETWORK ANALYSIS

Tomáš Švec
Bachelor Degree Programme (3), FIT BUT
E-mail: xsvect00@stud.fit.vutbr.cz

Supervised by: Jan Samek
E-mail: samejan@fit.vutbr.cz

Abstract: This thesis takes into consideration the rapid growth of social networking in today's society and tries to apply computational concepts of trust onto it. The set goal is to find out whether any form of relevance may be found between the relationship models of our communication on social networks and the model of trust presented at the end of the last century. An example set of rules is derived for trust between people and applied on Facebook.

Keywords: trust, reputation, social network, Facebook, Open Graph API, Python

1 INTRODUCTION

Social networks have been penetrating peoples' lives with remarkable pace in the past few years. As we expose our inner selves on this growing infrastructure, the risk of misuse is growing at a similar rate. There are people who we scarcely or never meet in person and instead communicate with them on a daily basis using only means of electronic communication. Artificial intelligence has a long way ahead from passing the Turing test [1] to actually being able to impersonate living people. Is there any way to use mathematical formulae, recognize patterns of human behavior? Utilize concepts of trust and use them to determine who one's best friend on any social network is?

This thesis' ambition is to create a model of several contexts of trust and determine the relevance to real-life data. The question remains whether we can actually express human emotion (such as trust) using a computational concept. However likely to malfunction, this idea has already been used for targeted advertising at the very least.

2 ANALYSIS OF SOCIAL NETWORKS

The matter of great importance is the choice of the network to base the model on. A process was necessary to determine which social network would provide sufficient data for our analysis. Basic criteria had been set for the process: diverse kinds of interactions, a useful API, integration of GPS services and penetration on both the Czech and world markets. The only social network being able to fulfill all the set requirements was Facebook.

3 CONTEXTS OF TRUST

This thesis utilizes the concept presented by Marsh in [2]. According to his basic principles, entities we observe may be assigned different kinds of trust (called *contexts*) based on the area in which the trust is applied. The created model contains six different contexts so far. They shall be examined in the extent of their current state of research in the following subchapters.

3.1 Trust based on timespans

This context of trust is the easiest to compute. It is assumed that the longer history of communication people have, the stronger their trust in each other is. The application makes it possible to compute trust for various timespans (hence the possibility to count with reputation as well). We take such a timespan, locate the first and the last interaction in the units of days, subtract them and express trust as a quotient. This quotient is produced by the timespan of active interactions divided by the whole timespan, resulting in a number between 0 and 1 (percentage).

3.2 Trust based on the number of interactions

Trust based on the number of interactions takes into account solely the total number of interactions exchanged between individuals. An *interaction* is a one-sided transfer of information. Speaking in the Facebook terminology, we have posts, comments and so-called *likes*. After a number of empirical experiments it became clear that a threshold should be set to determine when the trust in this context becomes 1 and does not change no matter how many interactions we append. This is a result of chronical Facebook addicts who may publish a large number of interactions and therefore tamper other quotients. These are the final formulae for computing this context of trust:

$$A = \frac{1}{n} \cdot \sum_{x=1}^{n} I_x \tag{1}$$

$$T_N(x) = \frac{I_x}{A + \frac{1}{n} \cdot \sum_{x=1}^{n} |A - I_x|} \tag{2}$$

where I_x is the number of interactions for a particular person x, A is an average number of interactions and the divider of the second equation can be computed as a sum of average and absolute deviation of our set of data.

3.3 Trust based on photo tags

Photo tags are also very simple to incorporate into the model. When a person is tagged in a photo with one or more of his/her friends, it almost always means a real-life interaction. One particular exception are the group photos where multiple users are selected (e.g. Christmas postcards). These exceptions are better to be eliminated. Trust is computed as a portion of the maximum amount of tags in a set of users.

3.4 Trust based on dispersed interactions

This particular mathematical model is derived from [3]. The assumption states that given the same number of interactions with two people, higher trust should be assigned to the person whose interactions are more evenly distributed on the timeline. The formula for this distribution is very intuitive:

$$T_D = \prod_{i=1}^{n-1} |t_{i+1} - t_i| \tag{3}$$

where t is a fixed time of interaction and n is the number of all interactions.

3.5 Trust based on group membership

The concept of group membership has not been fully incorporated into the application as of present day. The author assumes that membership in certain groups may be closely related to real-life connections. As a paradox, an inverse relationship exists between the size of common groups and trust between two entities. The less members this group containts, the more likely it is that two entities know each other in person or at least share a fairly unique treat.

3.6 Trust based on the length of comments

The length of posted comments and its relevance to trust is a matter of ongoing debates. Even though there appears to be a relationship between the quality of comments and their length, this model is designed to pay little importance to this factor by setting a very small coefficient in the resulting vector.

4 COMBINING INDIVIDUAL CONTEXTS TOGETHER

These six contexts appear to have greatest influence over real-life trust. For the purposes of paying a certain ammount of importance to each of these six contexts a vector had been set to determine the influence.

$$T(S,N,P,D,C,M) = \frac{S \cdot T_S + N \cdot T_N + P \cdot T_P + D \cdot T_D + C \cdot T_C + M \cdot T_M}{S+N+P+D+C+M} \quad (4)$$

where S, N, P, D, C and M represent coefficients of importance for their corresponding context of trust and belong to the set of natural numbers.

5 CONCLUSION

Despite the absence of some features' implementation, this model shows promising results when compared to real-life data and judged by several external testers who confirmed relevance to their subjective trust on social networks. Reconnaissance on a larger scale is already designed to confirm results of this drafted model, eventually to modify some properties for more precise reflection of reality.

6 ACKNOWLEDGMENT

This work was partially supported by the European Regional Development Fund in the IT4Innovations Centre of Excellence project (CZ.1.05/1.1.00/02.0070), grant GACR 102/09/H042 and the research plan MSM0021630528.

REFERENCES

[1] Oppy, Graham and Dowe, David. The Turing Test [online]. Latest modification: January 26^{th} 2011. [cit. 2012-03-02]. Available on URL: <http://bit.ly/GUXsup>.

[2] Marsh, S. P.: Formalising Trust as a Computational Concept. PhD thesis, University of Stirling, Stirling, UK, 1994.

[3] Lizi, Z., Cheun, P. T., Siyi, L., et al.: The Influence of Interaction Attributes on Trust in Virtual Communities. In Proceedings of the International Conference on User Modeling, Adaptation and Personalization (UMAP). Nanyang Technological University, Singapore, 2011.