Trust evaluation on Facebook using multiple contexts

Tomáš Švec[1], Jan Samek[2]

Brno University of Technology, Faculty of Information Technology,
Božetěchova 2, 612 66 Brno, Czech Republic
[1]xsvect00@stud.fit.vutbr.cz, [2]samejan@fit.vutbr.cz
http://www.fit.vutbr.cz

Abstract. This paper applies the term trust from the point of view of artificial intelligence to social network analysis methods. It evaluates current available interactions for a model of trust considering various social networks. A mathematical model of trust for Facebook is designed. This model is implemented in Python programming language. Experiments are conducted on a sample amount of Facebook users and furthermore analysed from the perspective of both artificial intelligence and social psychology.

Keywords: social network, trust, multi-context trust, Facebook

In a networked world, trust is the most important currency.

Eric Schmidt

1 Introduction

The CEO of Google accurately commented on the current state of human emotions in a networked world in his speech for the University of Pennsylvania. The definition of social interaction has been radically transformed more than once in the past and present century. We reveal more and more of our inner selves on the Internet and there are a growing number of people in our vicinity called *friends* who we have never actually met. Although the artificial intelligence is still miles away from passing the Turing test [8], we still begin to answer the question whether it is possible to use patterns of human behaviour to simulate emotions.

This paper is aimed at creating a model of trust from the point of view of AI which would make use of social psychology in social networks. Basically, it is assumed that as the term *trust* originates in sociology and social psychology, it should be possible to apply this principle in its original field after 20 years and observe the differences. To achieve this, several terms have to be defined both in social psychology and artificial intelligence – similarities are observed and highlighted. Several examples of current social networks will also be briefly analysed and a representative network will be chosen for implementation.

The designed model of trust itself will be mathematically described, keeping in mind the necessity to minimize specific dependencies to be able to implement this

model in a number of other networks. Reasons for correlation between various types of interactions and trust between entities will also be considered. Most importantly, the whole model and its implementation will be validated on real social network users and consequently summarized in the form of an exploratory investigation.

2 Social network analysis, trust and reputation

As the field of trust and reputation lies on the border of two scientific disciplines, sociology and computer science, it is sometimes impossible to adhere to strict technical description and mathematical definitions. A universal apparatus for describing human emotions has not been invented yet, after all. Despite these facts, this thesis leans toward computer science and therefore takes definitions from the field of artificial intelligence.

2.1 Trust

Bruce Schneier, a specialist on computer security and cryptography, considers the ability of building trust between individuals to be the cornerstone of modern society [9]. This mechanism may be tracked back to reciprocal altruism in some species. The following definition comes from one of the most renowned sociologists, *Anthony Giddens* [5]:

> *Trust is related to absence in time and in space. There would be no need to trust anyone whose activities were continually visible and whose thought processes were transparent, or to trust system whose workings were wholly known and understood. It has been said that trust is a device for coping with the freedom of others, but the prime condition of requirements for trust is not lack of power but lack of full information.*

In 1994 computer science was enriched by *Stephen Paul Marsh*, who influenced the field of artificial intelligence in a major way. He introduced trust into multi-agent systems in his doctoral thesis *Formalising Trust as a Computational Concept* [7]. His original understanding of the concept of trust came from the field of Humanities. Despite the precision and technical accuracy of his thesis, the term "trust" has never been fully defined in computer science, or to be more accurate, it has been defined in too many contexts and too many various situations. Marsh himself uses the definition from a famous psychologist, *Dr Deutsch* [2]:

1. The individual is confronted with an ambiguous path, a path that can lead to an event perceived to be beneficial ($Va+$) or to an event perceived to be harmful ($Va-$).
2. He perceives that the occurrence of $Va+$ or $Va-$ is contingent on the behaviour of another person.
3. He perceives the strength of $Va-$ to be greater than the strength of $Va+$.
4. If he chooses to take an ambiguous path with such properties, I shall say he makes a trusting choice; if he chooses not to take the path, he makes a distrustful choice.

2.2 Reputation

For the sake of readability, formal definitions of reputation are omitted in this text. It is, however, worth mentioning that the field of Humanities does not recognize reputation as a valid term. Social acceptance or trust perceived in groups of people or organizations is connected to social prestige instead. In the designed model, reputation could be derived from trust using an algorithm of arithmetic mean or similar techniques. In the context of trust and reputation, it is also important to describe what the *Dunbar number* is. We have seen a rapid rise of human society in the last few thousands of years. Biological evolution could, however, in no way compete against the pace of changes required for the human brain to adapt to modern society. As a result, we still have a fixed number of people we can keep track of in the matter of reputation. It happens to be the exact same number as the average population of a Neolithic settlement and also a rough average of the number of friends on *Facebook*. Today's scientists lean toward the value 150 [3].

2.3 Social network

This term is relatively new and dates back into the last century when *Barnes* described his stay in a *Norse village called Bremnes* [1]. Due to family traditions and isolation of this village from the rest of the world, Barnes was able to study some class phenomenon and categorize the inhabitants into groups. These relatively autonomous groups and their relationships were later described as a *social network*. The definition is as follows:

> *A social network is a social structure made up of individuals (or organizations) called "nodes", which are tied (connected) by one or more specific types of interdependency, such as friendship, kinship, common interest, financial exchange, dislike, sexual relationships, or relationships of beliefs, knowledge or prestige.*

3 Social networks: current situation

In spite of the prevailing endeavour to remain disclosed from any details of implementation that would concern a specific network, it was necessary to pick a deputy of social networks to demonstrate the formulated model of trust using real-life data. There were a few requirements concerning this deputy. Three most used social networks in Europe (*Facebook*, *VKontakte* and *MySpace*, according to InSites Consulting [10]) were amended with *Google+* and *Spoluzaci.cz*, two networks with bonds to the Czech environment. These were the desired treats of the deputy:

- more than one form of interaction on this network which can affect trust – these forms of interaction should also be actively used,
- it should be simple to use an API (Application Programming Interface) to access these services, the best option being an alternative from the service provider himself,
- location awareness would help us in the future to consider geographical factors in the analysis conducted using the model,
- the desired social network should be widespread, so that it is easier to collect representative data from real-life users,
- as the data would be collected in collaboration with people from the Czech Republic, it would also be desirable to have a number of Czech speaking users.

The results can be summarized in this table showing *Facebook* as the winner:

	Interactions	API	GPS	Penetration	Czech
Myspace	x		x		
Google+	x	x	x		
Facebook	x	x	x	x	x
Twitter		x	x	x	
Spolužáci	x			x	x

Table 1. Social network properties.

4 Multi-context trust model for Facebook

The mathematical core of this model leans on a theory distributed by *Marsh* in his founding thesis [7]. This theory introduces so-called contexts of trust which represent the fields in which we are capable of trusting the entity. To explain this term in a simplified example, "I trust my brother to drive me safely to the airport, but I would feel very insecure if he were to fly my plane." Dividing trust into contexts is the only reasonable way to comprise a thing as complex as trust while maintaining the possibility of flexible changes and further development. Every context is normalised into the interval from 0 to 1 to facilitate future aggregation.

This model was designed for the possibility of implementation for multiple contemporary social networks. It was, however, necessary to implement and test this theory for a particular social network. Although the described contexts stand on functionality provided by Facebook, chosen interaction types are present in other networks as well. Please note that methods of computation were chosen according to the environment the tests took place in. Several optimizations aimed at robustness, accuracy or speed may be considered, including saturation of values (meaning extreme values shall be restricted not to distort obtained results), omitting larger groups that anyone has a high probability to be a member of, or analysing the content of text, not only its quantitative measures.

4.1 Trust contexts

A short description of the investigated *trust contexts* is described in the following sections.

Interaction time span This context seems to be the most intuitive one. The longer the time between the first and the last interaction, the higher trust we are likely to feel, even though there may be exceptions concerning people we contacted soon after joining the community.

Number of interactions The term *interaction* stands for one-way information channel, in this case – wall posts, comments and "likes". Their overall number should be counted and normalized using the following formula (equation 2):

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

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

I_x stands for number of interactions with person x, A is the average number of interactions and the fraction divisor is a sum of the average number of interactions and absolute deviations of all acquired values. This formula provided the most reasonable results according to the first three respondents and was later empirically confirmed in the experiments.

Exclusion of extreme spikes represented by overly-active users is crucial here. Heuristics for this case include setting a maximum value.

Number of characters Several works in the area of trust study the relation between a number of characters in a message and the credibility of the writer [6]. As these works often belong to another application domain, this context is not given so much impact in the model. Setting a ceiling for the maximum number of characters is very important here, since copy & paste skills would be the easiest way to influence the model for educated users.

Interaction regularity Regularity differentiates people engaged in heated, yet scarce discussions which would normally boost a person's computed trust way above appropriate level. It is natural to trust people we communicate with on daily basis more than people that we had contact with in the past. One way to compute this context is using the formula introduced in the thesis [11] (equation 3):

$$x_v^{\Delta T}(A, B) = \prod_{i=1}^{n-1} |t_{i+1} - t_i| \tag{3}$$

There is an implementation issue, however, when we consider the amount of data and the necessity to express time in milliseconds. This formula would bring the most satisfying results at the cost of wide data type range. This context is therefore computed in a simplified manner. A set of perfectly regular intervals for the fixed number of interactions is computed and then compared to the real values.

As this statement may be a little unclear, a simple example shall be provided. Let us say we have four interactions to be considered, all of them occured shortly after the beginning of our friendship with the researched person. Our timespan for analysis is three months. Say we wanted to communicate regularly with this person. That would mean the first interaction occured at the beginning (which is correct and gives a small deviation). The second interaction should have occured after one month (which is still relatively close). The third and fourth interaction, however, should have occured at the end of the second and third month. If we compare these values to the ones close to the beginning, we get a very high deviation.

Based on the previous paragraph, we can see that the more interactions are irregular, the higher the deviation. This fact led to the necessity to invert the value to correspond with the rest of the contexts.

Photo tagging Photo tags have a very important meaning for trust. They usually indicate a link of people in the real world. There are special cases which should be considered (Christmas wishes would be a very good example, their informative value is next to nothing), but generally this context is very important for the resulting model.

Group membership A certain terminological ambiguity should be explained here. Groups and pages were not distinguished in Facebook initial times. Groups in this context represent a set of people who share a common trait, for example people who commute to the same city, people who work on the same project or people from one regional country unit. The more groups two people share, the more likely it is they trust each other. There is an inverse relationship between the size of groups and their importance. A shared smaller group usually means that these two subjects trust each other.

Common interests The only context which does not depend on any interaction and can be computed for any two people around the world. It builds upon the premise that people who share similar interests (like the same page here) are likely to trust each other more. This statement can be found in many papers on the subject, [9] serves as an example. A similar inverse relationship about size can also be applied here. This context, however, is the most time-consuming to compute and requires a lot of bandwidth. In case of time-critical operations, this is the part which should be omitted first, as it serves as more of an experimental feature.

Number of friends Due to the inconsistency in Facebook Graph API [4], this context was not implemented in the final model. It can be related to the previously mentioned "Dunbar's number". The deviation from the standard and widely accepted number of friends could also be considered an interesting factor for computing trust. People who have way more friends than the average number in their country may express similar traits. The same goes for the other extreme. This statement depends on many factors, though, and should be considered in connection with age groups.

4.2 Trust aggregation

These seven (eight) contexts should be aggregated in a way which allows us to establish an order relation. Marsh simply multiplied his contexts and used the resulting values. This approach fails here because of different importance of individual contexts. For this purpose, a priority vector (equation 4) is introduced in this model. It is a vector of numbers where T_x represents the priority for given context.

$$P = (T_S, T_N, T_C, T_F, T_P, T_G, T_L) \tag{4}$$

The final value of trust can be obtained with this formula (equation 5):

$$T_x = \frac{S \cdot T_S + N \cdot T_N + C \cdot T_C + F \cdot T_F + P \cdot T_P + G \cdot T_G + L \cdot T_L}{S + N + C + F + P + G + L} \tag{5}$$

This method of aggregation enables us to attribute each context with its importance. If, for instance, we find a context less contributing to overall trust in our recent

findings, we simply decrease the level of importance in the priority vector. Similarly, a completely new context may be added to the existing set and this expansion is also planned in the nearest future.

As for the particular model used in the experiments, the vector $(1, 3, 2, 2, 1, 2, 3)$ was used. Individual priorities were chosen based on empirical experience of the first 3 experimental users. The values were, however, retrospectively checked in the survey of participating users. Results showed apparent oscillation towards this choice of numbers as well.

5 Implementation and experimental results

As the model was intended to be deprived of any implementing details, the implantation itself shall only be described in a very brief manner. The particular example was implemented in the Python 3 programming language using the application interface supported by Facebook called Graph API [4]. Graph API produces data in JSON format, hence the need of Python's in-built libraries. Authentication is provided by the OAuth 2.0 technology.

The greatest issue encountered when collecting the data from users was how to get only limited access to their profiles and persuade them that no harm would come to their privacy. For this particular purpose, OpenGraph provides so-called access tokens, which can be generated on the developers' page and can be used to configure privileges for the holder of the token for a limited amount of time.

5.1 Exploratory investigation

As the research could be considered invasive by some users, quantitative research was not a valid option. Users with valid data for experiments consider their internet identities a part of their lives and therefore do not willingly provide access to their profiles. An exploratory investigation was a compromise and provided the possibility to work with a limited number of respondents and to ask relatively simple questions.

The exploratory investigation included 18 respondents randomly chosen in the age interval from 17 to 30 years. Men and women were both equally represented. The analysis was conducted for the time-span from 1.4.2011 till 1.5.2012. Results were verified by the respondents themselves using a questionnaire consisting of closed questions with the utilization of scaling.

Certain criteria had to be met in order for the user to participate in this investigation. The only condition was for the profile to be regularly used. Participating users were sent a short PDF file describing the procedure of generating their access token and also explaining which personal data they were making accessible. While the script was running (around 5 minutes for an average profile), they were given a simple command to record their answers for later use: "Name ten people you trust most on Facebook." Keeping this information to themselves was a key part of the investigation. They would perhaps try to obfuscate the initial guess if they were to show it to another person. This way, they were the only people who knew the answer.

After seeing the results of the scripts, users had to answer these questions with multiple choice answers:

1. How many people you listed actually occurred in the script's results?
   – Possible answers: 0–10.

2. How many people's trust was wildly mismatched?
   – Possible answers: 0–2, 3–5, 6–8, 9–11, 12 or more.
3. What actions among friends do you find most important for trust on Facebook?
   – Possible answers: Values 1–5 for these categories:
     - private messages,
     - comments,
     - "Like" tags,
     - common photographs,
     - common groups,
     - interaction regularity.

5.2 Experiment results

Question number 1 was the key element of this questionnaire. Resulting values form a fairly regular Gaussian curve. Most results converge to the number 5 and the arithmetic mean of all the values is 4.83. The figure 1 shows the number of respondents with each individual answer.

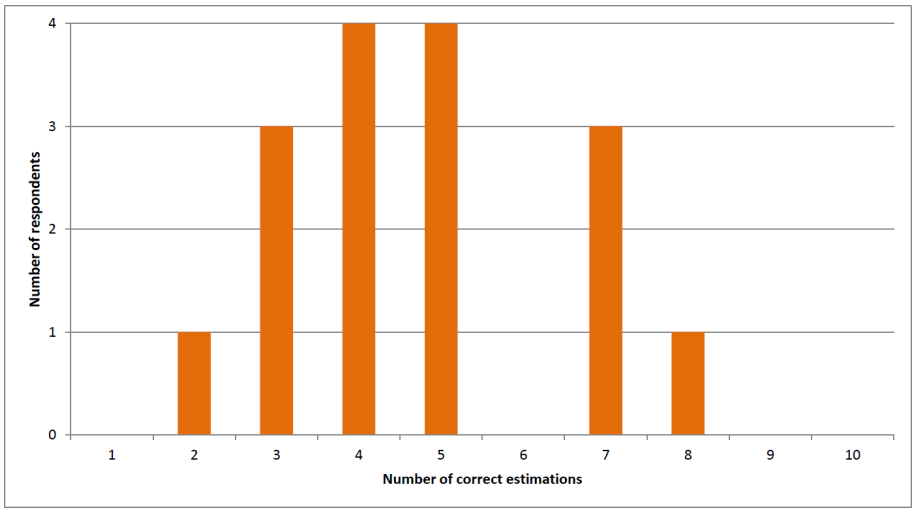

Fig. 1. Answers corresponding to expected results.

Question number 2 was designed to detect the most significant flaws of the model. Respondents were given the possibility to state whether someone's trust was way below/above expected values and this way they also verified the results themselves. Most respondents (11 answers 0–2) stated there were not as many deviations as one would expect. The figure 2 shows the deviation for respondents.

Question number 3 aimed at the credibility of the used priority vector. In this state, there must be a person setting the priority vector according to his/her preferences and acquired statistical data. One of possible expansions, however, relies on the possibility to change this model dynamically according to amount of collected data. So far, users seem to copy the initially set priority vector values.

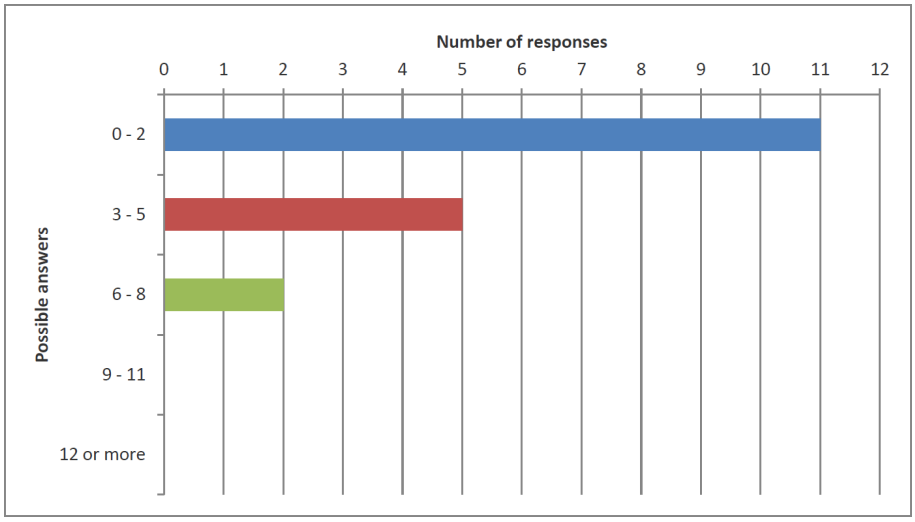

Fig. 2. Significant deviation of the model.

6 Conclusion

The goal of this paper was to analyse current situation in social networks from the point of interactions, design a model of trust for social networks, implement it and test its correspondence to the real world. The best evaluation of fulfilling these tasks is the experimental result:

Based on the respondents' answers, the model can evaluate correct trust with 48.3% probability. This number may seem like an unsatisfactory result. On the other hand, the model was given only information that (in most cases) is freely available on the web to anyone administering any Facebook account. Considering the best safety available, this information can still be seen by our friends, whose numbers, as we have learned, vary around the number 150. Would it be disturbing to the users that these 150 people can guess half the people they trust most on this network and use them for social engineering?

There are multiple paths this model could take in development. Since the very beginning, new contexts were intended to be added to this model, for example the similar number of friends or private messages analysis. Another possibility is to dynamically adjust the priority vector according to the amount of collected data. Users would also welcome an HTML interface for conducting the research themselves. Almost all respondents who participated in the exploratory investigation expressed this wish.

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) and by the project CEZ MSM0021630528.

References

1. J. A. Barnes. Class and committees in a norwegian island parish. *Human Relations*, 7, 1954.
2. Morton Deutsch. *Cooperation and Trust: Some Theoretical Notes*, pages 275–320. Nebraska University Press, 1962.
3. Robin Dunbar. *Grooming, Gossip, and the Evolution of Language*. Harvard University Press, 1998. ISBN 0674363361.
4. Inc. Facebook. Graph API [online]. `https://developers.facebook.com/docs/reference/api/`, 2011-04-02 [cit. 2013-04-16].
5. Anthony Giddens. *The Consequences of Modernity*. Polity Press, Oxford, UK, 1990. ISBN 0745607934.
6. Audun Josang, Roslan Ismail, and Colin Boyd. A survey of trust and reputation systems for online service provision. *Decision Support Systems*, 43(2), March 2007.
7. Stephen Paul Marsh. *Formalising Trust as a Computational Concept*. PhD thesis, University of Stirling, April 1994.
8. Graham Oppy and David Dowe. The turing test [online]. `http://plato.stanford.edu/archives/spr2011/entries/turing-test/`, 2011-01-26 [cit. 2013-04-16].
9. Bruce Schneier. *Liars and Outliers*. Wiley, 2012. ISBN 978-1118143308.
10. S. van Belleghem. Social media around the world 2011 [online]. `http://www.slideshare.net/stevenvanbelleghem/social-media-around-the-world-2011`, 2011-09-13 [cit. 2011-11-12].
11. Lizi Zhang, Cheun Pin Tan, Siyi Li, Hui Fang, Pramodh Rai Yao Chen, Rohit Luthra, Wee Keong Ng, and Jie Zhang. The influence of interaction attributes on trust in virtual communities. In *International Conference on User Modeling, Adaptation and Personalization (UMAP)*, 2011.