Trust Judgment in Knowledge Provenance

Jingwei Huang and Mark S. Fox
Enterprise Integration Laboratory, University Of Toronto
{jingwei, msf}@eil.utoronto.ca

Abstract

Knowledge provenance is an approach to determining the validity and origin of web information by means of modeling and maintaining information sources, information dependencies, and trust structures. This paper explores trust structures in social networks and constructs a trust judgment model for knowledge provenance. Trust judgment includes: trust assessment (to assess trust degree) and trust decision (to make decision of either trusting or distrusting). We reveal a general structure of trust decision, from which (i) the threshold of trust degree to make decision of trusting and (ii) a measure of importance of trust judgment situation are derived. Regarding trust assessment using social network, a major concern is how to aggregate friends' opinions. We propose two new methods: (1) to find most compatible solution to all opinions; (2) to request friends one by one until a set of consistent opinions is obtained. They are close to people's thinking patterns.

1. Introduction

With the widespread use of WWW as a globally accessible information/knowledge repository, comes a problem: anyone is able to produce and distribute information on the web; however, the information may be true or false, current or outdated, or even outright lies. The concerns regarding how to determine the validity of web information are receiving more and more attention. Interest in addressing the issue of trust on the web has appeared under the umbrella of the "web of trust" which is identified as the top layer of the Semantic Web and is still in its infant stage of development (see [2] slides 26&27).

Knowledge Provenance (hereafter, referred to as KP) is proposed in [6] to create an approach to determining the origin and validity of web information by means of modeling and maintaining information sources, information dependencies, as well as trust structures. The major questions KP attempts to answer include: Can this information be believed to be true? Who created it? Can its creator be trusted? What does it depend on? Can the information it depends on be believed to be true? This proposed approach can be used to help people and web software agents to determine the validity of web information.

Four levels of KP have been identified as follows: Level 1 (Static KP) focuses on provenance of static and certain information [6]; Level 2 (Dynamic KP) considers how the validity of information may change over time [10]; Level 3 (Uncertainty-oriented KP) considers uncertain truth value and uncertain trust relationships in determining the validity of information [11]; Level 4 (Judgment-based KP) focuses on social processes to make trust judgment.

In levels 1, 2&3, we assume that trust relationships have been known and mainly focus on the issues of information sources and information dependencies. In this paper, we turn to the issue of trust judgment. The major questions concerned are: when an information user does not know the information creator, how does the user make trust judgment by using social networks? Is trust transitive? How does a person aggregate his/her friends' opinions about trust?

Many trust models using social networks have been proposed. However, only several very simple trust aggregation methods such as "maximum", "minimum", weighted average are applied. In practice, "maximum" and "minimum" are the extreme ends of credulous and skeptical attitudes in judgment, so they are not suitable for many regular situations. Weighted average is widely used, but it may output a result far from all requested friends' opinions. Therefore, trust aggregation methods reflecting people's thinking patterns need to be further investigated. In addition, how to use trust degree to make a trust decision also has not been well studied. Finally, the representation of "trust", "distrust", and "unknown" is fundamental for trust modeling. It is important to explore an appropriate representation that has clear and consistent semantics. This paper attempts to address these problems.

Proceedings of the 16th International Workshop on Database and Expert Systems Applications (DEXA'05)
1529-4188/05 $20.00 © 2005 **IEEE**

The content of this paper is organized as follows. Section 2 discusses related research; section 3 introduces a motivating example; section 4&5 explore the structure of trust and trust judgment and construct a framework of trust judgment; section 6 presents trust assessment using social networks; section 7 discusses implementation; finally section 8 gives a summary.

2. Related Research

Trust is widely concerned in many disciplines. Many researchers contribute to the conceptualization of trust, e.g. Rotter (1967) defines "interpersonal trust" as "an expectancy" that an individual will keep what he promised; Luhmann (1979), Zucker (1986), Lewis & Weigert (1985) identify different types of trust; Deutsch (1973) studies trust with game theory and build up a framework to formalize trust; Gambetta (1988) defines trust as "a particular level of subjective probability". These studies provide important conceptual foundation for trust modeling.

"Web of trust" as the top layer of semantic web is receiving considerable attentions. The concept of "web of trust" perhaps is first developed in PGP as a trust model used for public key validation. For the purposes of secure web access control, Blaze et al (1996) propose "decentralized trust management" to separate trust management from web applications. Khare and Rifkin (1997) further propose several basic principles for trust management. Recently, several projects focusing on trust in social networks have emerged. For examples, FOAF project (http://foaf-project.org/) attempts to create social networks on the web by facilitating people to describe acquaintance relationships in machine-readable web pages; Abdul-Rahman&Hailes (1997) propose a distributed trust model. They discern direct trust and recommender trust and addressed that the latter is transitive; Yu&Singh (2000) propose a model of reputation (trust) propagation and building among agents; Mui et al (2002) formalize interpersonal trust as Bernoulli trials; Golbeck et al. (2002) extend the acquaintance relationships in FOAF model by introducing levels of trust and applied the model for filtering emails; Richardson et al. (2003) propose a formal model of trust propagation and applied it in bibliography recommendation; Guha&Kumar(2004) construct a trust propagation model considering distrust.

3. Motivating Example

In this section, we reveal the concepts of trust judgment in the following motivating example.

John finds a piece of news on the web that says that Google's stock price may reach $300 per share from $85 as it went to public in 2004 summer, which is forecasted by David Garrity, an analyst in Caris & Co. For the purpose of investment, John wants to determine the trustworthiness of the information.

John himself is not an investment expert and he does not know both David Garrity and Caris & Co. John may ask his friends about the trustworthiness of David in the field of stock analysis, and his friends may also ask their friends. For example, Kevin is John's financial advisor, and John highly believes what Kevin's belief in investment, so, if Kevin believes David's forecast, John also tends to believe it. When John obtains his friends' opinions, he needs to combine these opinions to form his own opinion about David's trustworthiness in stock analysis.

John and his friends may not know David Garrity but believe Caris & Co. in stock analysis. Since David is an analyst of Caris & Co., John indirectly believes David in the field.

This example reveals the following important points for trust judgment: (1) the trustworthiness of information creator can be used to represent the trustworthiness of the information created; (2) trust can be placed in what the trusted individual behaves like Kevin's trust in David in stock analysis. This type of trust is intransitive; (3) trust can be placed in what the trusted friend believes to be true in a field, like John's trust in Kevin in investment. This type of trust is transitive and can propagate in social networks; (4) trust in an organization in a field can be transferred to a professional member of the organization. This is another approach of trust propagation.

4. Trust

Trust is the psychological state in which (1) the trustor believes that the trustee behaves as expected in a specific context, based on evidence of the trustee's competence and goodwill; (2) the trustor is willing to be vulnerable to that belief.

Here, trustor and trustee can be any entity, such as an individual, an organization, a software agent, or a web service; the behavior of the trustee that is expected by trustor may be trustworthy information or cooperative action; competence is the capability of the trustee to conduct the expected behavior; goodwill is the good intention of the trustee to conduct the expected behavior; Trust is context-specific, for example, John trusts his financial advisor in investment but doesn't trust the financial advisor in health-care. Context can have very complex structures. For

simplicity, we use topic or knowledge field to represent context.

Types of Trust

According to the approaches of trust, trust can be classified into three types:

(1) Interpersonal trust: the direct trust between two entities based on their experiences of interaction.

(2) Relational trust (or social networks-based trust): the indirect trust via friends in social network, e.g. the use of references / recommendation in social life.

(3) System trust [15]: the trust placed in the function of a system, including characteristics-based trust [22], professional membership-based trust [22], institutional-based trust [22], and as regularity-based trust [17].

According to the types of the expected behaviors, we identify the following types of trust:

(1) trust in behavior, i.e. trust in what trustee makes, more specifically, trust in the information created or the action conducted by the trustee;

(2) trust in belief, i.e. trust in what trustee believes, specifically, trust in the information that the trustee believes to be true or the action that the trustee believes to be done as expected;

Trust in belief is transitive [12], which grounds trust propagation in social networks.

5. Trust Judgment

As defined earlier, trust is comprised of (1) belief and (2) willingness to be vulnerable to that belief. Corresponding to this, trust judgment is comprised of (1) **trust assessment** and (2) **trust decision**. In trust assessment process, the trustor assesses the degree of belief (called **trust degree**) that the trustee behaves as trustor expects; in trust decision process, according to trust degree and the situation which the trustor is facing to make trust judgment, the trustor decides to either trust or distrust. If the trustor chooses to trust even with the risk that the trustee may not behave as trustor expects, the trustor is willing to be vulnerable (i.e. to take risk).

Trust Assessment

How does a trustor assess the degree of trust in a trustee? For interpersonal trust, trust assessment focuses on evaluating the behavior of the trustee and reinforcing the trust relationship; for relational trust, trust assessment focuses on trust propagation in social networks, especially, the aggregation of trust degrees given by different friends; for system trust, trust assessment focuses on identifying the characteristics of a trustee. We will discuss trust assessment in section 6.

Trust Decision

In real world, there are various trust decision situations. By applying decision tree and utility theory, we propose a general structure of trust decision situation as shown in figure 1.

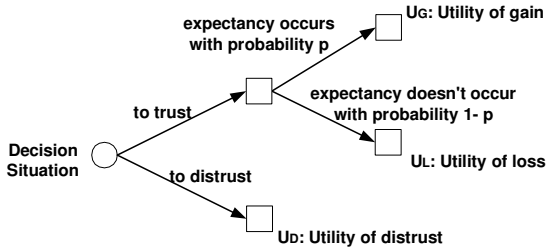

Fig. 1 Structure of trust decision situation

The condition for a rational decision of trusting is that the expected utility of trusting is greater than the expected utility of distrusting, i.e.

$$p \cdot U_G + (1-p) \cdot U_L > U_D$$

For decision analysis, we only need to consider the case $U_G > U_D > U_L$. In this case, the condition to make trusting decision can be derived as:

$$p > (U_D - U_L) / (U_G - U_L)$$

In our model, trust degree has the semantics of probability p, so that $(U_D - U_L) / (U_G - U_L)$ is the threshold to make trusting decision. Furthermore, an index of the importance of trust judgment situation can be defined as:

$$I = ||U_D|| \cdot (U_D - U_L) / (U_G - U_L)$$

$||U_D||$ is normalized utility ranging from 0 to 1. This index can be used in trust assessment to represent how important a trust judgment situation is.

Representation of trust, distrust and unknown

The representation of "trust", "distrust", and "unknown" is a basic issue in trust modeling. Basically, "unknown" is different to "distrust". "Unknown" is the case in which the trustor does not know the trustee in the concerned context; "distrust" is the case that the trustor knows the trustee but does not trust in the trustee. Some researchers (e.g. [9] [21]) use "1.0" to represent trust and "-1" to represent distrust, which causes the difficulties to interpret and to process; some others (e.g. [19]) do not discern the differences between "distrust" and "unknown".

From our point of view, "known" is a precondition of "trust". That is to say, they are related but different variables, so that status "trust", "distrust", and "unknown" cannot be represented with one variable. Therefore, a Boolean variable K needs to be introduced, K=1 represents "known" and K=0 represents "unknown".

Recall that "trust degree" is defined as the degree of belief (subjective probability) that the trustor believes the trustee behaves as the trustor expects in a specific context. In this way, trust degree (probability) of 1.0 is corresponding to "completely trust"; 0 is corresponding to "completely distrust", i.e. the expectancy happens with probability of 0. In this paper, the degree of "trust in belief" is denoted as $td_b(x,y,k)$, where x is trustor, y is trustee, and k is context; similarly, the degree of "trust in behavior" is denoted as $td_m(x,y,k)$. When the type of trust and the context are not concerned or they are clear in the context of discussion, for simplicity, trust degree is denoted as $td(x,y)$.

Under certain conditions, the status of "unknown" can be mapped into a value of trust degree. The degree of trust in an "unknown" trustee can be assigned as the trust degree with which the trustor trusts average people; when there is completely no any information, according to entropy theory, 0.5 can be assigned as the degree of trust in this unknown trustee (see [11]). However, this specific value of trust degree is not necessarily corresponding to "unknown", because the trustor may just trust with this value (see [16]).

6. Trust Assessment Using Social Networks

This section studies trust assessment by using social networks. Each entity in social networks has a certain amount of trust relationships. The collection of trust relationships of all entities in social networks forms a huge and complex graph. Since "trust in belief" is transitive, trust can propagate in such a huge network. Fortunately, according to the principle of "six degrees of separation", it is acceptable to search only limited length of trust paths.

We follow several basic principles: (1) trust yourself first: when trustor a has interpersonal trust relationships with trustee c, a uses its own trust relationships with c to make trust assessment; (2) listen to your friends: for an unknown trustee c, trustor a requests its friends first, prior to applying system trust, because friends' opinions about c are more specific than system trust; (3) apply system trust: If both the trustor and its friends do not know the trustee but know the trustee "belongs to" a trusted "system", the trustee's behavior may be assumed to follow the "system".

Now we consider when entity a does not know entity c, and a requests a set of friends $b_1, b_2, \ldots, b_n$ regarding the trustworthiness of c, how should a aggregates these opinions to calculate the degree of trust in c? If only one friend's opinion is considered, the aggregation is sequence aggregation. The most common aggregation method is multiplication, i.e. $td(a,c) = td(a,b) \cdot td(b,c)$. We also used the operator because it is consistent with the semantics of trust degree. The most interesting question here is: when several friends' opinions need to be considered, how entity a aggregates these opinions and the degrees of trust in these friends? The answer to this question depends on how people make trust assessment in real world. Unfortunately, there seems no one single general trust assessment pattern. Different people in different situations tend to use different ways.

As discussed in section 1, the most frequently used methods "minimum", "maximum", "weighted average" have some drawbacks. Here, we propose two new aggregation methods.

(1) Most compatible aggregation

In real world, we often find that an opinion owner usually accepts the opinions close to his/her opinion. In other words, one opinion is compatible with the other opinions close to it. The smaller the distance between two opinions is, the more they are compatible with each other. In this paper, the degree of compatibility is defined as follows,

$$c(x_1, x_2) = \begin{cases} 1 - d(x_1, x_2)/r & d(x_1, x_2) \le r \\ 0 & otherwise \end{cases}$$

Here, $d(x_1, x_2) = |x_1 - x_2|$, and $r = 0.5$

This proposed method attempts to find the solution that is most compatible to all friends' opinions by a weighted voting process. In this process, each friend, say b_k, has voting weight $td(a, b_k)$, votes a solution x with $c(td(b_k,c),x)$. The most compatible solution x^* is the solution of the following optimization problem,

$$x^* = \underset{x}{opt} \max_{0 \le x \le 1} \sum_{k=1,\ldots,n} td(a,b_k) \cdot c(td(b_k,c),x)$$

(2) Consistency aggregation

The thinking pattern underlying this aggregation method is that the trustor requests friends one by one until a set of consistent opinions is found.

In this aggregation process, trustor a requests friends $b_1, b_2, \ldots, b_n$ one by one till consistency rate τ is not less than a predefined threshold τ_0 (dependent on the importance index). Consistency rate is defined as:

$$\tau = (\sum_{bi \in S1} td(a,b_i)) / (\sum_{bj \in S} td(a,b_j))$$

where S is the set of requested friends who have answered their degrees of trust in trustee c, and S_1 is the largest subset of friends whose opinions are consistent with each other, i.e. $S_1 = \{ b_j \mid for\ any\ b_k \in S_1, |td(b_j,c) - td(b_k,c)| < \lambda \}$, λ is predefined constant.

The trust degree from a to c will be:

$$td(a,c) = td(a,b_k) \cdot td(b_k,c)$$

where, $b_k = \underset{x}{opt} \max_{x \in S_1} td(a,x)$.

If all friends have been requested but the consistency rate is smaller than the threshold, the result

should be “unknown”. However, the trustor can activate other trust aggregation methods.

7. Implementation

There are two computing paradigms to implement trust assessment using social networks: (1) centralized computing; (2) distributed computing. The first requires collecting other people’s trust relationships. This approach is difficult to use in practice for the reason of privacy protection; the second paradigm can be realized by P2P technology. In this approach, each entity in social networks only uses its own trust relationships and requests its friends to calculate their degree of trust in a specific trustee.

The proposed trust judgment model has been implemented in Prolog. We are in progress to implement the model on the semantic web.

In order to facilitate defining trust relationships, we have defined a trust description language in RDFS. The following is a piece of example containing a trust relationship.

```
<tr:TrustRel rdf:ID=
    “http://www.example.com/people/~john/tr-kevin”>
    <tr:trustor rdf:about=
        “http://www.example.com/people/~john/”/>
    <tr:trustee rdf:about=
        “http://www.example.com/people/~kevin/”/>
    <tr:trustDegree> 0.9 </tr:trustDegree>
    <tr:expectancyType>believed</tr:expectancyType>
    <tr:inContext> investment analysis </tr:inContext>
</tr:TrustRel>
```

Because different people in different situations may prefer different ways to make trust judgment, our trust description language allows people to use trust policies to express their preferences in trust judgment.

8. Summary

This paper explored the trust structures in social networks and constructed a trust judgment model for knowledge provenance. (1) we revealed a general structure of trust judgment, from which (i) the threshold of trust degree to make decision of trusting and (ii) a measure of the importance of trust judgment situation have been derived; (2) we gave a representation of “trust”, “distrust” and “unknown” that has clear and consistent semantics; (3) we proposed two new aggregation methods: “most compatible aggregation” and “consistency aggregation”. They are close to people’s thinking patterns in judgment.

This research was supported, in part, by Novator Systems, Ltd.

9. References

[1] A. Abdul-Rahman, S. Hailes, “A Distributed Trust Model”, Proceedings of the 1997 workshop on New security paradigms, 1997, pp.48-60.

[2] T. Berners-Lee, “Semantic Web Status and Direction”, Int. Semantic Web Conf. 2003, keynote. http://www.w3.org/2003/Talks/1023-iswc-tbl/

[3] M. Blaze, J. Feigenbaum, and J. Lacy, “Decentralized Trust Management”, Proceedings of IEEE Conference on Security and Privacy, May, 1996.

[4] K. Blomqvist, “The Many Faces of Trust”, Scand. J. Mgmt, Vol. 13, No. 3, 1997, pp. 271-286.

[5] M. Deutsch, “Trust and Suspicion: Theoretical Notes”, (1957), in Deutsch eds. The Resolution of Conflict, Yale University Press, 1973, pp. 143-176.

[6] M.S. Fox and J.Huang, “Knowledge Provenance: An Approach to Modeling and Maintaining the Evolution and Validity of Knowledge”, EIL Technical Report, University of Toronto. http://www.eil.utoronto.ca/km/papers/fox-kp1.pdf

[7] D. Gambetta, “Can We Trust Trust?” in D. Gambetta eds., Trust, B. Blackwell, 1988, pp.

[8] J. Golbeck, J. Hendler, and B. Parsia, “Trust Networks on the Semantic Web”, University of Maryland, 2002.

[9] R. Guha and R. Kumar, “Propagation of Trust and Distrust”, WWW2004, New York, 2004.

[10] J. Huang and M.S. Fox, “Dynamic Knowledge Provenance”, in Proceedings of Business Agents and Semantic Web Workshop, London, Canada, 2004, PP.11-20.

[11] J. Huang and M.S. Fox, “Uncertainty in Knowledge Provenance”, In Bussler, C. Davies J., Fensel D., Studer R. (eds.) The Semantic Web: Research and Applications, Lecture Notes in Computer Science 3053, 2004, pp.372-387.

[12] J. Huang and M.S. Fox, “A logic structure of trust”, EIL Research Report, University of Toronto, 2005.

[13] R. Khare and A. Rifkin, "Weaving and Web of Trust", World Wide Web Journal, Vol. 2, No. 3, pp. 77-112.

[14] J. D. Lewis and A. Weigert, “Trust as a social reality”, Social Forces, V. 63 N.4, pp967-985, 1985.

[15] N.Luhmann, Trust and Power, John Wiley&Sons,1979.

[16] S.P. Marsh, Formalising Trust as a Computational Concept, Ph.D. Thesis, University of Stirling, 1994.

[17] N. Minsky, “Regularity-Based Trust in Cyberspace”, iTrust2003, 2003.

[18] L. Mui and A. Halberstadt, “A computational model of trust and reputation”, Proceedings of the 35th Hawaii International Conference on System Sciences, 2002.

[19] M. Richardson, R. Agrawal, and P. Domingos, “Trust Management for the Semantic Web”, Proc. of Int. Semantic Web Conf. 2003, PP.351-368.

[20] J. Rotter, "A new scale for the measurement of interpersonal trust", J. Personality. 35(1967) 651-665.

[21] B.Yu and M.P. Singh, “A Social Mechanism of Reputation Management in Electronic Communities”, Proceedings of 4th Int. Workshop on Cooperative Information Agents, 2000, pp.154-165.

[22] L.G. Zucker, “Production of Trust”, Research in Organizational Behavior, Vol. 8, 1986, pp. 53-111.