A Proposed Trust Model for the Semantic Web

Hoda Waguih

Abstract—A serious problem on the WWW is finding reliable information. Not everything found on the Web is true and the Semantic Web does not change that in any way. The problem will be even more crucial for the Semantic Web, where agents will be integrating and using information from multiple sources. Thus, if an incorrect premise is used due to a single faulty source, then any conclusions drawn may be in error. Thus, statements published on the Semantic Web have to be seen as claims rather than as facts, and there should be a way to decide which among many possibly inconsistent sources is most reliable. In this work, we propose a trust model for the Semantic Web. The proposed model is inspired by the use trust in human society. Trust is a type of social knowledge and encodes evaluations about which agents can be taken as reliable sources of information or services. Our proposed model allows agents to decide which among different sources of information to trust and thus act rationally on the semantic web.

Keywords—Semantic Web, Trust, Web of Trust, WWW.

I. INTRODUCTION & MOTIVATIONS

THE World Wide Web (WWW) is the greatest repository of information ever assembled. It contains documents and multimedia resources concerning almost every imaginable subject, and all of this information is instantaneously available to anyone with an Internet connection. Though, since on the WWW, anyone is allowed to make any statement with no requirements about its accuracy or truthfulness, information from different web sources have to be seen as claims rather than facts. This makes finding reliable information a serious problem on the WWW. The problem will be even more crucial for the Semantic Web, where agents will be integrating information from multiple sources. If an incorrect premise is used due to a single faulty source, then any conclusions drawn may be in error.

On the current Web, human can deal with such problem by relying on their intuition and personal judgment for any page or web site they visit. When reading a web page, humans make many judgments based on the appearance of the page and the source of the information. Although someone could lie about his or her sources, or provide information with intent to mislead, yet it is relatively easy to generate at least some information about the source. However, the Semantic Web being processed mainly by machines cannot depend only on the user for trusting or not trusting the site content. It will have to decide if a particular resource found on the web is trustworthy before using it. This is the main motivation behind our work and what makes trust at the heart of the Semantic Web vision.

The Semantic Web is an open, dynamic network of independent information providers all having different views of the world, different levels of knowledge, and different intentions, interacting with each other. Thus, information retrieved from such system can be of uncertain reliability. This is what makes the so-called web of trust one of the ultimate goals for the Semantic Web. There should be a way to deal with uncertainty in the information sources on the semantic web through deciding which among many possibly inconsistent sources of information is most reliable. Enabling trust on the Semantic Web will ensure more efficient agent interaction. Most of the work concerning trust in computer science has been concentrated in the area of security. They focus on some security technologies to guarantee a certain level of trust for the sake of reliable communication. Among those technologies are cryptographic algorithms for privacy and digital signatures, authentication protocols for proving authenticity and access control methods for managing authorization [2]. However, these methods are not well suited as general models of trust as they cannot manage the more general concept of 'trustworthiness' as conceived in real world social life.

In this paper, we investigated the problem of managing trust for the semantic web. Particularly, how we could exploit real world social characteristics of trust and reputation for managing uncertainty of information sources on the semantic web. We wish to be able to decide which among different sources of information to trust and thus act rationally on the semantic web. As a result of our investigation, we proposed a model for managing trust on the semantic web. The proposed model is inspired by the use trust in human society. Trust is a type of social knowledge and encodes evaluations about which agents can be taken as reliable sources of information or services. Our proposed model allows agents to decide which among different sources of information to trust and thus act rationally on the semantic web.

The presented paper is organized as follows: First, we begin by clarifying our adopted definition and characteristics of trust drawn from our reviews of social studied literature. In section III, we briefly define what reputation is and show how it is closely related to the concept of trust. Section IV presents our proposed trust model for the semantic web. The proposed model is based on the real world social characteristic of trust and the reputation mechanism described in previous sections. Finally, conclusions are presented in section V.

II. TRUST ON THE SEMANTIC WEB

Trust, as discussed before, is at the heart of the Semantic Web vision, which was introduced by the W3C[1] in six main principles, Table I [15]. In one of those principles, it was stated that on the semantic web, there is no need for absolute trust and that trust or more pragmatically trustworthiness will be evaluated by each application that processes the information on the web. And in [4], the authors view the semantic web as a collection of agents interacting with each other in a collaborative environment. Moreover, they pointed out that each agent will have to make subjective trust judgment about other agents with respect to the information they provide or the services they claim to be able to supply.

TABLE I
THE SEMANTIC WEB PRINCIPLES

Principle 1: Everything can be identified by URI's
Principle 2: Resources and links can have types
Principle 3: Partial information is tolerated
Principle 4: There is no need for absolute truth
Principle 5: Evolution is supported
Principle 6: Minimalist design

A. Defining Trust

Trust, as we all know, is a very important aspect of life. Trust is central to all transactions in human society, where our own actions are dependent on the actions of others. Thus, excluding instances where trust has no influence on our decisions. But, what is the meaning of trust? In [14], the authors address the issue of divergent trust definitions, creating two kinds of conceptual typologies: a *classification system* for types of trust, and a *set of six related types* of trust constructs resulting from the analysis of the classification system. For our purpose, however, we have chosen to use the following definition by Gambetta [9]:

> *"Trust (or, symmetrically, distrust) is a particular level of the subjective probability with which an agent assesses that another agent or group of agents will perform a particular action, both before we can monitor such action, or independently or his capacity to be able to monitor it, and in a context in which it affects his own action."*

In his definition, Gambetta draws attention to four main points. First, trust is a subjective property. Second, trust is affected by some actions that we cannot monitor. Third, trust is context dependent. Fourth, the level of trust depends on how our actions are affected by the other agent's action. Accordingly trust is not an objective property of an agent but a subjective degree of belief about specific agent [14], within a specific context that ranges from complete distrust to complete trust. Though, there is also a situation where an agent does not have an opinion about another agent trustworthiness. Furthermore, as stated in [12], trust is neither a prediction nor some measure of probability. A trusting action is taken despite uncertainty of outcome but of course in

an anticipation of a positive outcome [16]. Thus, when we say we trust someone or that someone is trustworthy, we implicitly mean that the probability that he will perform an action that is beneficial or at least not detrimental to us is high enough for us to consider engaging in some form of cooperation with him. Correspondingly when we say that someone is untrustworthy, we imply that the probability is low enough for us to refrain from doing so.

B. Characterizing Trust

The adopted definition of trust draws some requirements that have to be met when building an environment based on trust. Based on this definition and work by Marsh [13], Grandison and Sloman [11], we can outline the main significant trust features in the following:

- *Trust is context dependent.* Trust has different meaning in different context. An agent a may trust an agent b regarding specific context c_1, but do not trust it for context c_2.

- *Trust is subjective.* Trust is the social knowledge that is derived from personal observations and serves for future personal decision-making.

- *Trust is directed.* A trust relation has a trustor and a trustee. It is asymmetric binary relation such that "*a* trust *b*" doesn't necessarily imply "*b* trust *a*".

- *Trust is a measurable belief.* An agent a may trust an agent b more than a trusts c for the same context. Thus, a trust relation is associated with a value that represents it's strength or degree of truth within a specific context.

- *Trust is conditionally transitive.* Trust in security sense is always intransitive, but in recommendation sense it is partially transitive [7].

- *Trust has a temporal dimension.i.e. It exists and evolves in time.* The fact that agent a trusted agent b in the past does not in itself guarantee that a will trust b in the future. This is because agent b's performance and other relevant information may lead agent a to reevaluate its trust in b. Since trust is learned from past observations, trust values evolve with new observation and experience. Moreover, to account for changes in a trustee's behavior, recent observations carry more weight in deriving trust.

- *Trust between collectives does not necessarily distribute to trust between their members.* On the assumption that an agent a trusts a group of contractors to deliver (as a group) in a collaborative project, one cannot conclude that a trusts each member of the team to deliver independently.

- *Trust is reflexive, yet trust in oneself is measurable.* An agent a may trust an agent b to for a specific context c more than it trusts itself for the same context. Self-assignment underlies the ability of an agent to delegate or offer a task to another agent in order to improve efficiency or reduce risk.

[1] W3C (World Wide Web Consortium)

III. REPUTATION

Trust, as discussed before, is a social phenomenon. Reputation is more social notion of trust. In our lives we each maintain a set of reputations for people we know. When we need to work with a new, unknown person, we can ask people with whom we already have relationships for information about such person. We obtain information from these other sources by means of word of mouth i.e. a mechanism for propagating reputation. Based on the information we gather, we form an opinion about the reputation of the new person. Positive reputation leads to confidence or trust in that person along with a higher level of social status and power. Negative reputation, on the other hand, leads to a loss of esteem held in society along with social status and power [5]. This system, for propagating reputation, works well, even though there are a lot of people in the society, because communities tend to be highly interconnected, and the number of steps between any two people tends to be rather small. This is known as the Small World effect, and it has been shown to be true for a variety of social and web based systems [3].

Therefore, reputation information is of great important when dealing with sources of information in making effective and informed trust decisions. Reputation is not a single notion but one with multiple parts depending on the context in which it is used. In [17], the author reviews the basic notions of reputation as used in several disciplines, including economics, computer science, evolution biology, anthropology, and sociology. In [16], the author stated "Reputation helps us to manage the complexity of social life by singling out trustworthy people, in whose interest it is to meet promises". Based on his statement, Misztal defines reputation as follows:

> *A reputation is an expectation about an agent's behavior based on information about or observations of its past behavior.*

Accordingly, the reputation information we obtain for a specific agent is based on either our own experience or the experience of others in dealing with this agent. Thus we maintain a set of reputation information by combining our personal opinion and the opinion of others for the same reputation information.

IV. THE PROPOSED TRUST MODEL

This work primarily aims to establish a new model of trust with a primary goal of clarification of the concept of trust. Trust, as discussed before, is mainly a social phenomenon. Thus any artificial model of trust should be based on how trust works in society. In this section, we present our proposed trust model. The proposed model is based on the real world social characteristic of trust and the reputation mechanism described in the previous sections.

A. Model Assumptions

In this section, we begin by drawing attention to some of the assumptions we have made in designing our trust model. These assumptions explain how trust and reputation are defined in our model.

1. Trust Actors

Trust is explained in terms of a relationship either between two agents or an agent and a static entity. That is a trustor and a trustee. A trustor, is an agent who holds a belief for trusting another agent or a static entity, and a trustee, the trusted agent or entity. Based on its trust in the trustee, a trustor can decide whether it should believe the information/services that the trustee claims to provide or not.

2. Trust Relationship

A trust relationship exists between an agent a_1 and an agent a_2 when a_1 holds a belief about a_2 trustworthiness. But, this relationship is directed, that is the same belief in the reverse direction need not exist at the same time i.e. "a_1 trust a_2" doesn't necessarily imply "a_2 trust a_1". However, if it do exists it is represented as a separate trust relationship. Furthermore, although a common assumption of most authentication protocols infer the transitivity of trust. That is (a_1 trusts a_2) & (a_2 trusts a_3) implies that (a_1 trusts a_3). This is not true for our model; trust may be transitive but conditionally. In our model, the transitivity of trust apply only for recommended trust in which the recomendors are trustworthiness agents. Therefore, the properties of a trust relationship in our model are:

- − Trust is always between two agents,
- − It is asymmetric binary relation, and
- − It is conditionally transitive

3. Trust Context

Trust, in our model, is context dependent. Trust, as discussed before, has different meaning in different context. That is an agent a_1 may trust an agent a_2 in specific context c_1, but do not trust it for context c_2. Thus trust in our model refers to an agent's beliefs about the trustworthiness (or usefulness) of other agents' knowledge in a certain domain.

4. Trust Values

Trust, as discussed before, is a measurable belief. Thus any trust relation is associated with a value that represents its strength or degree of truth. That is the level of trust an agent may have in another agent regarding a specific contex although the range of trust, as pointed in [8] can be either Boolean or numeric value. Yet, numeric trust value can be more accurate in capturing the uncertainty of trust. Naturally there is no agreement [7] on the exact numerical trust representation, especially when distrust is considered. In our model, we propose to use a numeric trust system, which has nine grades ranging from absolute distrust to absolute trust i.e. [-1,1]. The proposed trust system was inspired by the trust schema in [Golbeck et al., 2003] that specifies trust values on a scale of 1-9. The values, meaning and description for our proposed system are given below, Table II.

5. Trust Types

Trust, as discussed before, is an important aspect of social interactions in human society and potentially in agent societies

as well. There are two classes of trust commonly used in agent society [8]: personal trust and public trust.

TABLE II
THE SEMANTIC WEB PRINCIPLES

Value	Meaning	Description
-1	Absolute Distrust	Completely untrustworthy.
-0.75	High Distrust	Highly distrust
-0.5	Moderate Distrust	Average distrust
-0.25	Slight Distrust	Lowest distrust.
0	Ignorance	Cannot make trust judgments.
0.25	Slight Trust	Lowest possible trust.
0.5	Moderate Trust	Average Trust
0.75	High Trust	A high trustworthy value
1	Absolute Trust	Completely trustworthy

Personal trust, which is usually subjective, is derived from an agent's own social experiences and serves as the basis for the agent's future trust related decisions. Public trust, on the other hand, is based on reputation information. That is, the reported social experiences throughout agent society, and reflects the general opinion about individuals, which is used as initial trust about unfamiliar agents.

Based on our study of trust in human society, we differentiate, in our model, between two types of trust, figure 1, which we prefer to name; direct trust, and recommended trust. Direct trust, refers to an agent's own beliefs about the trustworthiness (or usefulness) of other agents' knowledge in a certain domain. Recommended Trust, on the other hand, which is based on reputation information for a specific agent, results from an agent belief or collected agents' beliefs about the trustworthiness (or usefulness) of other agents' knowledge in a certain domain.

Fig. 1 Model Trust Types

B. Model Description

The proposed model deals mainly with beliefs about the trustworthiness of agents based on domain trust and/or social trust information. Domain trust information refers to an agent's own experience regarding the trustworthiness of another agent for a specific context. While social trust information covers an agent's beliefs of another agent for specific context based on collected reputation information. That is, what we choose to refer in our model by direct trust and recommended trust. Fig. 2 depicts the proposed trust data model. The model takes the context and time frame as input and computes the estimated trust belief value, see table II, based on a) evaluating the trustor personal beliefs in the trustee, or b) evaluating recommended trust belief value. The estimated trust belief value is then rounded to the expected trust belief which is represented in one of three values {trust,

distrust, ignore}. In the rest of this section, we elaborate the main boxed data in the model; 'context and time frame', 'direct trust belief', 'recommended trust belief', 'estimated trust belief', and 'expected trust belief' respectively.

1. Contexts and Time Frame

We agree that any artificial model for trust should be based on how trust works in society. Thus our presented trust model, Fig. 2, is drawn from the real world social characteristic of trust. In this sense, trust, or the trustworthiness an agent a_1 set for another agent a_2 is not absolute value but it depends on a specific context. That is, an agent a_1 may trust an agent a_2 regarding specific context c_1, but do not trust it for context c_2. Moreover, trust has a temporal dimension. That is, the fact that agent a_1 trusted agent a_2 in the past does not in itself guarantee that a_1 will trust a_2 in the future. This is because agent a_2's performance and other relevant information may lead agent a_1 to reevaluate its trust in a_2. Thus we should take into consideration the time frame or the duration during which the agents trust belief value is computed. Accordingly, before computing any trust belief, the context and time frame should be given to the model as inputs from the environment
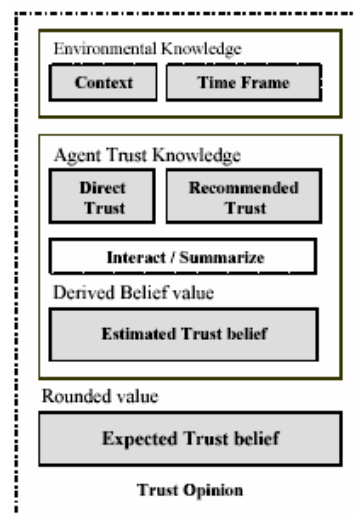

Fig. 2 Trust Data Model

2. Direct Trust and Recommended Trust

In order to calculate the estimated trust belief value, an agent holds for another agent, the model takes input from both environmental knowledge and agent's trust knowledge. In this perspective, our model differentiates between two types of trust: direct trust and recommended trust. Direct Trust, as discussed before, refers to an agent's own beliefs about the trustworthiness (or usefulness) of other agents' knowledge in a certain domain at a certain time. In our model, we represent direct trust (dt), i.e. an agent's belief (a_1) in another agent's (a_2) trustworthiness within a certain context (c_1) at a specific time (t_1) by the formula:

$$dt_{c1}^{t1}(a_1, a_2)$$

Where the value of the trust belief is an element of the set {-1, -0.75, -0.5, -0.25, 0, 0.25, 0.5, 0.75, 1}. These values ranges from absolute distrust, noted by the value –1, to absolute trust, noted by 1. The other semantics for those values are shown in table II. In our model, we use the term ignorance to refer to an agent being unable to make a trust related judgment for another agent regarding a specific context. We agree that, modeling distrust and ignorance can be as useful as trust. As stated in [19], suppose that an agent discovers a document that no one explicitly trusts, but that no one explicitly distrusts either, that is, the state of ignorance in our model. Most likely, this agent will trust this document more than it will trust one that has been explicitly labeled as untrustworthy

Recommended Trust, on the other hand, is based on reputation information about a specific agent. This information usually results from an agent or collected agents' beliefs about the trustworthiness (or usefulness) of this agents' knowledge in a certain domain. In a related work [1], the author used the term recommender trust and refers to it by the belief value an agent assign regarding the trustworthy of another agent for giving recommendations with respect to a context. In our model, though, we differentiate between three classes of reputation information as shown in figure 3. Moreover, we assume that 1) we are dealing with rational agents. That is, when an agent depends on reputation information in talking trust decisions, it usually ask trusted recomender agent. 2) recommended trust are made honestly and correctly, that is if a_1 says that it trusts a_2 absolutely, then it really does and it is not providing a false trust certification. Therefore, when an agent depends on reputation information in taking trust decision it deals with trusted ones. In our model, we represent recommended trust (rt), i.e. an agent's (a_1) belief in another agent's (a_2) trustworthiness, based on collected reputation information, within a certain context (c_1) at a specific time (t_1) by the formula:

$$rt_{c1}^{t1}(a_1, a_2)$$

Where the value of the trust belief, as discussed before, is element of the set {-1, -0.75, -0.5, -0.25, 0, 0.25, 0.5, 0.75, 1}. These values ranges from absolute distrust, noted by the value –1, to absolute trust, noted by 1.

3. Estimated Trust and Expected Trust

The estimated trust belief is derived from the evaluation of the two types of trust in our model, while the expected trust belief is the target for our trust model upon which an agent will build its trust opinion for different agents in different contexts. Our proposed model, as discussed before, is based on two main types for trust; direct trust and recommended trust. Since trust is mainly a subjective belief, our model begins with evaluating direct trust information. That is, the agent's own experience for determining the trustworthiness of target agent, which will be evaluated to one of the values in table II. The agent will not ask for a recommendation unless it

fails to make a trust related judgment for another agent in a specific context i.e. the trust belief value is zero. In such case, according to our model, the agent searches for recommendation from other agents. We assume here, as discussed before, that the agents acts rationally and thus it only ask trusted recomenders for their opinion. Furthermore, in case there exist a number of trusted recomenders for the target agent in a given context, the model should apply an aggregation function that combine agents beliefs into one belief. The estimated belief is, therefore, evaluated based on either types of trust information to a belief value as shown in Table I.
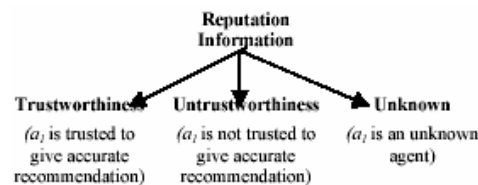

Fig. 3 Types of Reputation Information

The expected belief, on the other hand, is the target of our trust model upon which an agent will build its trust opinion for target agent in different context. The final interpreted trust belief should be interpreted to an element of the set {trust, distrust, ignore}. The aim of our model here is to consider the problem of deriving the final expected belief for the target agent. That is rounding the values of the given estimated belief into some sort of discrete values

V. CONCLUSIONS

A serious problem on the WWW is finding reliable information, since anyone is allowed to make any statement with no requirements about its accuracy or truthfulness. Although on the current Web, human can sometimes handle such problem by relying on their intuition and personal judgment in evaluating the content of any page or web site they visit. The problem is more crucial for the Semantic Web, as machines cannot depend only on the user for trusting or not trusting the site content.

In this paper, we have investigated the importance of trust for finding reliable information on the semantic web. As part of our investigation, we proposed a trust model based on the real world social characteristics of trust. Although, much of the current work in the literature is related to sociological concepts that is trust and reputation. However, we are not dealing with those concepts in a real society where the advantages of face-to-face interactions, personal trust and reputation, and physical cues, among others apply [17]. Instead we are dealing with maintaining trust and reputation in a virtual community such as on the Semantic Web. We believe that such social variables have significant roles for enhancing the user experiences as well as agents interaction online. Therefore, we've studied how trust works in society and

outlined a model for supporting trust on the Semantic Web, which is based on agent's own experience and collected reputation information.

In our model, we have set some assumptions based on social characteristics of trust and reputation. Naturally, as discussed before, there is no one universal system for specifying the trust belief value an agent hold for another agent regarding a specific context. In our model, we choose to use a range of nine values inspired by a trust schema which we adopt from our reviews. Future work will include further investigation into more concrete basis for these values so that we can accurately capture the uncertainty of trust, and on the issues of propagation of trust belief values between different agents as well as for the different techniques used for rounding values of the estimated trust belief so as to calculate a final trust opinion, that is what we call in our model expected trust belief. Finally, it will be interesting to test the behavior of our trust model with real experimental data.

REFERENCES

[1] Abdul-Rahman, A. and Hailes, S., "Supporting Trust in Virtual Communities", Paper presented at Hawaii International Conference on System Sciences (HICSS), 2000.

[2] Abdul-Rahman, A. and Hailes, S., "A Distributed Trust Model", Proceedings of New Security Paradigms 97 Workshop, 1997.

[3] Adamic, L., "The Small World Web", Proceedings of ECDL, pp. 443-452, 1999.

[4] Berners-Lee, T. Hendler, J. and Lassila, O., "The Semantic Web", Scientific American, May 2001.

[5] Carter, J. Bitting, E. and Ghorbani, A., "Reputation Formalization within Information Sharing Multi agent Architectures", Computational Intelligence, 2(5):45--64, 2002.

[6] Christianson, B. and Harbison, W. S., "Why isn't trust transitive?" Proceedings of the Security Protocols Workshop, pp. 171-176, 1996.

[7] Ding, L. Kolari, P. Ganjugunte, S. Finin, T. and Joshi, A., "Modeling and Evaluating Trust Network Inference", Proceedings Seventh International Workshop on Trust in Agent Societies at AAMAS 2004, July 2004.

[8] Ding, L. Zhou, L. and Finin, T., "Trust based Knowledge Outsourcing for the Semantic Web Agent" Proceeding of IEEE/WIC International Conference on Web Intelligence, 2003.

[9] Gambetta, D. "Can We Trust Trust?", In Trust: Making and Breaking Cooperative Relations, Gambetta, D (ed.), Basil Blackwell, Oxford, pp. 213-237, 1990.

[10] Golbeck, J., Parsia, B., and Hendler, J., "Trust Networks on the Semantic Web", Proceedings of Cooperative Intelligent Agents, Helsinki, Finland, 2003.

[11] Grandison, T. and Sloman, M., "A Survey of Trust in Internet Application", IEEE Communications Surveys & Tutorials, 3(4), 2000.

[12] Luhmann, N., "Trust and Power", New York, John Wiley, 1979.

[13] Marsh, S. P., "Formalizing Trust as a Computational Concept" Ph.D. dissertation, University of Stirling, 1994.

[14] McKnight, D. H., Chervany, N. L. "The Meanings of Trust", Technical Report 94-04, Carlson School of Management, University of Minnesota, 1996.

[15] Miller, E. Swick, R. Brickley, D. and McBride, B., "Semantic Web Activity" at the W3C, 2001. Available from: http://www.w3.org/2001/sw/

[16] Misztal, B., "Trust in Modern Societies", Cambridge MA, Polity Press, 1996.

[17] Mui, L., "Computational Models of Trust and Reputation: Agents, Evolutionary Games, and Social Networks", PhD Thesis, Massachusetts Institute of Technology, 2002.

[18] Platzer, C., "Trust-based Security in Web Services", Master thesis, Information Systems Institute, Technical University of Vienna, 2004

[19] Swartz, A., "The Semantic Web In Breadth", 2002 Available at: http://logicerror.com/semanticWeb-long.