An effective approach for accurate estimation of trust of distant information sources in the Semantic Web

Vangelis G. Bintzios, Thanasis G. Papaioannou, and George D. Stamoulis
Department of Informatics, Athens University of Economics and Business
76 Patision Str., Athens, GR 10434, Greece
{evbin, pathan, gstamoul}@aueb.gr

Abstract

To assess the trustworthiness of the information published in the World Wide Web referrals are often employed. This is due to the fact that most information sources are visited only occasionally by the same client, and thus, direct own experience rarely suffices. The accuracy of trust inference for unknown information sources may considerably deteriorate due to "noise" or to the intervention of malicious nodes producing and propagating untrustworthy referrals. In this paper, we describe an innovative approach for trust inference in the Semantic Web and in trust networks in general, referred to as FACiLE (Faith Assessment Combining Last Edges). Unlike all other approaches, FACilE infers a trust value for an information source from a proper combination of only the direct trust values of its neighbours. We evaluate the efficiency of our approach by means of a series of simulation experiments run for a wide variety of mixes of sources of untrustworthy information. FACiLE outperforms other trust-inference approach in the most interesting cases of population mixes.

1. Introduction

Nowadays an enormous volume of information of all kinds is published in the World Wide Web and accessed by users thereby. The freedom of publishing, the relative anonymity and the absence of any validation authorities raise important issues regarding trustworthiness of information. Indeed, the W3C Semantic Web activity has already identified the need for a trust layer in the stack of the Semantic Web [1].

In general, trust can be interpreted as the subjective probability, based on direct experience and prior or communicated belief, that a source publishes accurate information. Direct own (i.e. personal) experience is the most dependable means of inferring the trustworthiness of a source. However, since most information sources are visited very rarely by the same client, direct experience usually does not suffice. Thus, referrals are used for estimating the trustworthiness of occasionally visited sources. In the World Wide Web both means can be useful for trust assessment: a) direct trust based on direct own experience for sites that are often visited, and b) inferred trust based on referrals for occasionally visited sites. The effectiveness of the latter depends on the accuracy of referrals and the method of trust inference, i.e. estimation and propagation of trust.

Lately, there has been carried out considerable research on trust modelling and inference in the Semantic Web and in other contexts (ad hoc networks, peer-to-peer systems, etc.) The proposed approaches are generally classified into four categories:

1. Approaches based on a simple aggregation function; e.g. sum of positive/negative ratings [2].
2. Approaches based on linear algebra in the context of a Markovian model; trust inference is based on a probabilistic interpretation of the transition from host to host [3], [4], [5].
3. Approaches based on Path Algebra; the trust network is modelled as a directed-weighted graph; the weight of each edge connecting two nodes equals the value of the direct trust, while end-to-end trust of some path is inferred by calculating the weight of the path [6], [7], [8], [9].
4. Other specialized approaches; e.g. those involving multi-dimensional trust metrics [10], [11].

An important common characteristic of many of these approaches [6], [7], [4], [5], [3] is that trust is inferred on an *end-to-end* basis: That is, the calculation

Acknowledgements. This work was partly supported by the Network of Excellence EuroNGI (IST-2003-507613) funded by European Union.

of trustworthiness of the target-node by a requesting node incorporates the trust values of the intervening nodes. Thus, trust values may be seriously distorted along the way, especially if paths are not short and untrustworthiness and/or "noise" (e.g. due to subjectivity or malicious behaviour) are present in the system. As argued in [9], the longer the distance from the information source the more the uncertainty about trust and the worse the distortion that arises despite the small-world properties of the Web [12] (namely, high clustering but short average distance between two nodes). Even if a single node in the path is very "noisy", trust information can be completely distorted.

Since inferred trust is, in general, considerably less informative than the direct one, it should be employed in an efficient way so as to result in high overall accuracy and speed of trust assessment. In this paper, we propose an approach referred to as FACiLE (Faith Assessment Combining Last Edges) that significantly improves the accuracy of trust inference avoiding considerably the distortion described above. The underlying idea is as follows: in order to make a trust assessment over an unknown entity, simply ask its neighbours and adopt their assessment in a way that is based on their own relative inferred trust values. The approach is motivated from social networks [12], where every person knows her neighbours (i.e., persons with whom she socializes more frequently) better than anyone else. In the Semantic Web, a *neighbour* of a certain client can be defined as a source of information frequently visited by this client, while a *transaction* is the transfer of some information, which may be: a) a piece of content directly provided by a node, on the trustworthiness of which each neighbouring node has a certain belief from direct experience with this source node, or b) a trust value for another node. All nodes in the Semantic Web are information sources. Thus, both trust and belief measure the trustworthiness of information, and are not distinguished hereafter.

Our approach is based on Path Algebra yet in an innovative way: Trust is inferred for the neighbouring nodes of the target-node yet aiming at comparative (rather than absolute) evaluation of their trustworthiness. Thus, we introduce an innovative last step, referred to as *combination*. In this step, trust to the target node is deduced, either (i) by adopting the direct trust to the target node of its neighbour that is inferred by the requesting node as the most trustworthy, or (ii) by employing the weighted average of the direct trust values to the target node of all its neighbours, with weights proportional to the respective trust value of each of them as inferred by the requesting node. In [8], it was shown that it is preferable to rely on the referrals of friends (i.e., nodes for which there exist direct own trust values) rather than on those of the distant node's neighbours. Thus, just employing the referrals of the neighbouring nodes does not suffice for accurate trust inference without employing the neighbours' inferred trust, as opposed to our approach. Finally, our approach does not require the presence of pre-trusted nodes, as opposed to [5] and [9].

2. End-to-End and FACiLE

2.1. End-to-End approaches

Under the trust inference approaches based on Path Algebra, the trust network is represented by a directed weighted graph. We assume that each node of the graph corresponds to a web site, or to a subset of the content published therein for which a client can make a unified trust assessment (e.g. the sports news of a certain portal). Each directed edge has a weight ranging from 0 (total distrust) to 1 (full trust). This represents a trust value assigned by the node that requests information to the node that provides it, based on the direct experience of the former. When a node cannot assess the trustworthiness of another, then the corresponding edge is missing from the graph. Thus, distrust is distinguished from unawareness of trust. As already mentioned, trust for distant sources of information is inferred "end-to-end" in most of the proposed approaches. For instance, when node q needs to assess the trustworthiness of a certain statement or content published by another node s in the Semantic Web, the following steps are undertaken: First, q identifies nodes (e.g. by means of flooding with limited hop count) having formed a direct trust-value for s; i.e. nodes n_i with some edge $n_i \rightarrow s$ terminating to s. Then, for every n_i, node q calculates the inferred trust of s via n_i, that is the weight t_{q,n_i} of a path: $q \rightarrow \ldots \rightarrow n_i \rightarrow s$. This calculation involves [4], [7]:

1. The *concatenation* of the trust weights of the successive edges of each path by means of an operation; usually multiplication, harmonic mean, max or min.
2. The *aggregation* of alternative paths by means of an operation; usually addition**,** average, or max.

2.2 The FACiLE approach

As already explained, the inferred trust is a much weaker means than the direct trust to assess the

trustworthiness an information source. Ideally, only direct trust should be used in trust assessment. This is accomplished under our approach as follows: For assessing the trust of an unknown information source, one should ask its neighbours and adopt either the opinion of the most trustworthy of them or use the weighted average of all the direct trusts of the source's neighbours, with the weights being equal to their respective trustworthiness, which is inferred using Path Algebra. That is, we use the path-inferred trust values only as a relative measure of trustworthiness of the neighbours of the target entity. Thus, when some node q has to assess the trustworthiness of a statement of content s placed in the Semantic Web, the following steps are first undertaken:

1. Find and ask the nodes that have a direct trust on the source s, i.e. nodes n_i with some edge $n_i \rightarrow s$ terminating to s.
2. Calculate the inferred trust for every n_i, as the weight t_{q,n_i} of the path $q \rightarrow \ldots \rightarrow n_i$.

We take that the inferred trust values to the neighbouring nodes of the information source is calculated according to Path Algebra approaches as those described in Subsection 2.1, although other approaches could be employed as well, e.g. max-flow. In particular, we employ the following alternatives for *concatenation* operations:

- *Multiplication*, henceforth denoted as MULTI, which is motivated as follows: if each edge-weight expresses the probability of a successful transaction, then the total probability of successful transaction over a path of two consequent edges equal their product [4], [7].
- The *harmonic mean*, to be denoted as HARM. This is motivated by viewing trust edges as electrical conductors [7]; recall that the conductivity of two serially connected conductors equals the harmonic mean of their conductivities.
- A *hybrid* of the above two operations, to be denoted as HYBRID. Its outcome equals the harmonic mean of the two weights unless the denominator exceeds 1; in this case, HYBRID gives the product of the two weights. Thus, for trustworthy nodes, the inferred trust does not diminish along the path, as the harmonic mean does because of the denominator's increase.

For *aggregation* we always employ the *maximum* operation due to its effectiveness for trust inference, as shown in [4]. Also, due to this choice, we can implement the trust inference to the neighbouring nodes of the information source using a generalized version of the Bellman-Ford algorithm. This is presented in Figure 1; $E^*(G)$ contains the permissible edges from each node (according to some selection policy) and $t[v]$ denotes the inferred trust for node v. Also, note that the results of the Bellman Ford algorithm could be cached for some time, thus further reducing computational and communication overhead.

for step=1 to $|V(G)|$ - 1
  for each edge $(u, v) \in E^*(G)$
    $t[v] \leftarrow \oplus [\, t[v], t[u] \otimes t_{u,v} \,]$

where $\oplus$: *aggregation* and $\otimes$: *concatenation*

Figure 1: Generalized Bellman-Ford algorithm

So far, we have explained how we compute the inferred trust of every issuer n_i of a direct trust-value for the information source s. Next, we employ the inferred trust of every n_i in order to *combine* their respective personal trusts and obtain a trustworthiness measure for s. This innovative step, referred to as *combination*, is implemented by means of either:

- the *maximum* operation, to be denoted as Max, thus adopting the most trustworthy personal direct trust, or
- the *weighted average*, to be denoted as WeiAvg, thus averaging the various direct trusts for s, each weighted by the trustworthiness of its issuer.

Note that the personal trust estimation of a neighbour n_i of s is only taken into account when the inferred trust to n_i is above a certain *threshold*, thus ignoring untrustworthy neighbours.

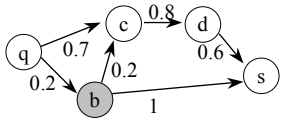

Figure 2: Example of a trust-graph

Next, we illustrate the effectiveness of our approach by means of a simple example. Consider the trust graph of Figure 2. Node q is interested in the content of site s and has to assess its trustworthiness. We use MULTI as concatenation operation and Max for aggregation. Since $t_{q,b} = 0.2$, b mostly provides untrustworthy referrals. On the other hand, d appears to be considerably more trustworthy. The inferred trust for s will be $t_{q,s} = t_{d,s} = 0.6$ (using Max in combination too). Had we applied end-to-end concatenation and aggregation, then the inferred trust for s would be $t_{q,s} = \max\{t_{q,c} \times t_{c,d} \times t_{d,s}, t_{q,b} \times t_{b,s}\} = 0.336$, which is actually almost half the trustworthiness assigned directly to s by d! Moreover, if both b and d were untrustworthy, then their trust values as inferred by FACiLE would be very low (i.e. below trust threshold). Thus, s would be

considered distrusted and could not be maliciously promoted by *b* and *d*.

3. Experimental Evaluation

3.1 The Simulation Model

In this section, we experimentally evaluate the accuracy of our approach for trust inference in comparison with end-to-end approaches. To simulate a part of the Semantic Web, we used a 100-node power-law directed weighted trust graph with some shortcuts to preserve small world properties [12] to a certain extent. Each node represents a source of information; the weight of each edge equals the respective direct trust value, which is continuously updated (see below). Nodes are classified in 3 types (namely "Good", "Bad", and "Ugly"). Associated with each type are: a) the probability of offering trustworthy content in a transaction, which is depicted with vertical bold arrows and the associated numbers in Figure 3, b) their behaviour regarding the accuracy of trust values they provide to others. Regarding the behaviour of nodes in offering content, we have considered two models already employed in the literature; namely a model with Ideal-world characters [6] and one with Real-world characters [4] in order to assess the effectiveness of our approach without and with "noise" respectively. In particular, in the Ideal-world model, "Bad" nodes always offer untrustworthy content, while in the Real-world model they do so with probability 0.9. "Good" nodes always offer trustworthy content in the Ideal-world model, while they do so with probability 0.9 in the Real-world model. "Ugly" nodes offer trustworthy content with probability 0.5 in both models. On the other hand, the behaviour of the nodes on reporting trust values is under both models as follows: "Bad" nodes exhibit a malicious behaviour, always reporting random trust values. "Ugly" nodes do so only with probability 0.5, while otherwise they report true trust values, thus exhibiting a dynamic behaviour. Finally, "Good" nodes always report true trust values. No pre-trusted nodes are assumed to exist. The allocation of reporting types of the nodes is determined randomly at the beginning of each experiment, yet for a predetermined population mix.

The scenario of each of the experiments is as follows: We simulate a network where various transactions are taking place among nodes. Without loss of generality we assume that a different user is attached at each of the nodes of the network. 90% of her transactions are done with randomly chosen neighbouring nodes and the remaining 10% with randomly chosen distant nodes. This justifies our consideration of a node's neighbour (resp. distant node) as one that is frequently (resp. infrequently) visited by the former node. Each transaction is classified as successful or not depending on the type of the visited node. Direct trust of a neighbouring node is calculated and constantly updated as the ratio of successful transactions over the total number of transactions with it. Using the various approaches for trust inference and running the algorithm of Figure 1, we estimate the trust for a distant visited node *s* after each transaction. Based on this, we classify node *s* as "Good", "Bad", or "Ugly" according to the intervals shown in Figure 3 for the Ideal- and the Real-world models. For example, under the Real-world model, if the estimated trust value for a distant node after a transaction is 0.85, then the distant node is categorized as "Good". If the type of the visited node is guessed correctly then we count a *hit*. The *hit ratio*, i.e. the number of hits over the number of estimations, is the metric used in order to assess the effectiveness of the various approaches. Note also that neighbours of the distant node *s* with inferred trust less than 0.1 are ignored in the combination step. If this applies to all neighbours of *s*, then we count a miss. We have considered in the experimental analysis all cases of population mixes (i.e. relative fractions of node types).

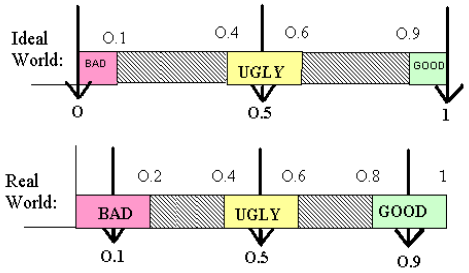

Figure 3. Node types and identification criteria

In the beginning, a node knows nothing (i.e. "cold start") about the type of each of its neighbours and nothing more about the rest of the graph. However, due to the high frequency of the local transactions, every node gets to know its neighbours' types, even if the latter were dynamically changed with some rate. Regarding distant nodes, no information concerning their quality is being stored except in the final experiments, where we study the effectiveness of combining it with inferred trust.

3.2. The Results

In this subsection, we investigate the most suitable pair of *concatenation,* and *combination* operators and the influence of the fractions of "Bad" and "Ugly" nodes to the accuracy of trust inference. Recall that *aggregation* is always performed by means of maximization. First, we compare the *concatenation* operations considered (i.e., "MULTI", "HARM" and "HYBRID") in terms of accuracy assuming the Ideal-world model in the graph. As depicted in Figure 4, MULTI and HYBRID are equivalent in terms of hit ratio, ending finally at hit ratio 1, while they both outperform HARM. In fact, HARM degrades as the percentage of "Good" nodes increases, because as already explained it is not well-suited for inferring trust in networks with highly trusted nodes. Note that all concatenation operators achieve high hit ratios for high fractions of "Bad" nodes. This can be explained as follows: Every node recognizes correctly the type of its neighbours and, since most of them are "Bad", it assigns to them and subsequently to the distant node (due to multiplication) a trust value of 0, which is very likely to be correct! Also, MULTI is the highest performing concatenation function for the end-to-end approaches in the case of the Real-world model. Thus, only MULTI is henceforth used in concatenation for both the end-to-end approaches and FACiLE.

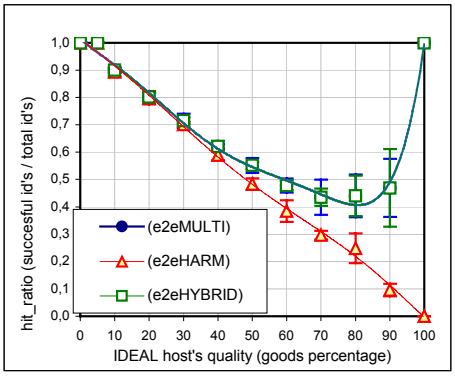

Figure 4. Identification for various concatenation operations (Ideal-world model)

Next, we compare the proposed FACiLE approach with the most accurate end-to-end one (denoted as e2eMULTI). Both Maximum and Weighted Average are used as *combination* operations. The Ideal-world model is considered first. FACiLE achieves much higher accuracy in trust estimation when the fraction of "Good" nodes is higher than 50%, for both combination operations as shown in Figure 5, unless all nodes are "Good". In the latter case, all three approaches are perfectly accurate. If the fraction of "Bad" nodes is greater than 50%, then it seems a better policy not to trust anyone, to which the end-to-end trust inference approach amounts.

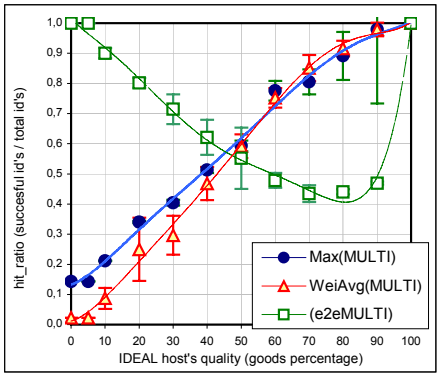

Figure 5. Identification accuracy for various combination operations and MULTI concatenation (Ideal-world model)

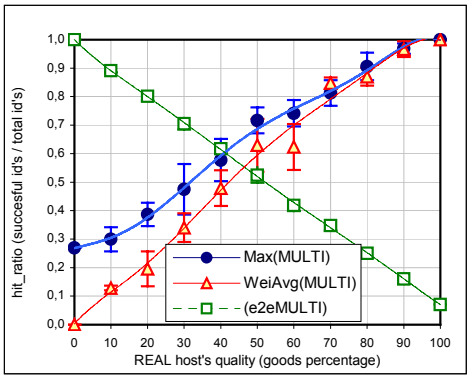

Figure 6. Identification accuracy for various combination operations and MULTI concatenation (Real-world model)

In the case of the Real-world model, significant distortion is induced to the inferred trust estimated by the end-to-end approach, which thus fails to discover "Good" nodes as their percentage increases; see Figure 6. On the other hand, FACiLE remains efficient with both combination operations achieving very high accuracy in inferred trust values Clearly, the cases where e2eMULTI outperforms our approach are not of practical interest because they have very small percentages of "Good" nodes. Also, we have conducted experiments having the population being changed at a high rate; the results have similar trends for the various concatenation and combination alternatives with somewhat lower accuracy due to the changing neighbours. (For brevity reasons we omit the relevant graphs.) Thus, our approach is applicable to

environments with changing population as well, e.g. peer-to-peer systems.

Next, we investigate how the effectiveness of FACiLE is affected by the presence of "Ugly" nodes. We consider MULTI and Max for concatenation and combination respectively in the Real-world model. As shown in Figure 7, our approach achieves high hit ratio even for large fractions of "Ugly" and "Bad" nodes. Hence, the effectiveness of our approach is not affected by such dynamic behaviours for reporting referrals. These results also apply for the Ideal-world model.

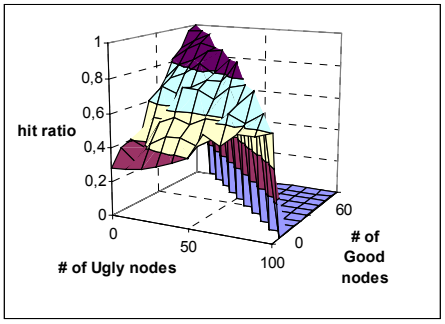

Figure 7. Identification accuracy in the presence of "Ugly" nodes (Real-world model)

Finally, we have performed experiments in which the direct experience after transacting with a distant information source is incorporated to the trust metric for that source. (For brevity reasons, we omit the relevant graphs.) In this variant, the trustworthiness of the remote node results from the weighted average of the inferred trust calculated by the various approaches and the direct trust resulted by experience of the querying node with this particular remote node. The weight of direct trust increases as more experience is gained. These experiments showed that a satisfactory improvement to the estimation accuracy of FACiLE can be thus attained. However, storing direct experience constitutes a considerable overhead, due to the huge number of sites visited. Therefore, it can only be beneficial for remote sites that are periodically visited over long time periods.

4. Concluding Remarks

Our proposed approach (i.e. FACiLE) uses the personal trust estimations of the neighbours of the distant node and combines them appropriately on the basis of the inferred trust values for the neighbours. It proved to be very effective in eliminating the distortion in trust inference introduced in paths of the Semantic Web or of trust graphs in general without relying on the existence of pre-trusted nodes. FACiLE can also be applied to other contexts such as grid, mobile ad-hoc networks etc. Such studies are left for further research. An important property of our approach is that its implementation does not require a centralized infrastructure for the discovery of paths to distant nodes; path discovery starts at each querying node that only needs to know its neighbours. This can be implemented effectively by means of a flooding mechanism that traverses either all edges or selectively only edges to nodes of the highest degree [12] and for a fixed number of hops away from the querying node. Our approach is compatible with the current specification of the Semantic Web, and, for example, it can be based on the extended "Friend-of-a-Friend" [6].

References

[1] M. R. Koivunen, E. Miller. W3C Semantic Web Activity. In *Proc. of the Semantic Web Kick-off Seminar*, Finland, November 2001. Available at: http://www.w3.org/2001/12/semweb-fin/w3csw

[2] eBay – The World's Online Marketplace. http://www.e-bay.com

[3] S. D. Kamvar, M. T. Schlosser, H. Garcia-Molina. The EigenTrust Algorithm for Reputation Management in P2P Networks. In *Proc. of 12th International WWW Conference*, Budapest, Hungary, May 2003.

[4] M. Richardson, R. Agrawal, P. Domingos. Trust Management for the Semantic Web, In *Proc. of* 2nd International Semantic Web Conference, Sundial Resort, FL, USA, October 2003.

[5] Z. Gyongyi, H. Garcia-Molina, and J. Pedersen. Combating web spam with TrustRank. In *Proc. of the 30th International Conference on Very Large Databases (VLDB),* Toronto, Canada, September 2004.

[6] J. Golbeck, B. Parsia, J. Hendler. Trust Networks on the Semantic Web. In *Proc. of Cooperative Intelligent Agents*, Helsinki, Finland, August 2003.

[7] G. Theodorakopoulos, J. S. Baras. Trust Evaluation in Ad-Hoc Networks. In *Proc. of Wireless Security*, Philadelphia, PA, USA, October 2004.

[8] S. Marti and H. Garcia-Molina. Limited Reputation Sharing in P2P Systems. In *Proc. of the ACM conference on Electronic commerce*, New York, NY, USA, May 2004.

[9] C.-N. Ziegler and G. Lausen. Spreading Activation Models for Trust Propagation. In *Proc. of e-Technology, e-Commerce, and e-Service*, Taipei, Taiwan, March 2004.

[10] J. Sabater, C. Siera. REGRET: reputation in gregarious societies. In *Proc. of Autonomous Agents*, Montreal, Canada, May 2001.

[11] A. Jøsang, S. Pope. Semantic Constraints for Trust Transitivity. In *Proc. of APCCM 2005*, Newcastle, Australia, January 2005.

[12] L. Adamic and E. Adar. How to Search a Social Network. *Social Networks*, vol. 27, no. 3, July 2005.