A Many Valued Representation and Propagation of Trust and Distrust

Martine De Cock[1] and Paulo Pinheiro da Silva[2]

[1] Ghent University
Dept. of Applied Mathematics and Computer Science
Krijgslaan 281 (S9), 9000 Gent, Belgium
Martine.DeCock@UGent.be
http://www.fuzzy.UGent.be
[2] Stanford University
Knowledge Systems
AI Laboratory
Stanford CA 94305, USA
pp@ksl.stanford.edu
http://iw.stanford.edu

Abstract. As the amount of information on the web grows, users may find increasing challenges in trusting and sometimes distrusting sources. One possible aid is to maintain a network of trust between sources. In this paper, we propose to model such a network as an intuitionistic fuzzy relation. This allows to elegantly handle together the problem of ignorance, i.e. not knowing whether to trust or not, and vagueness, i.e. trust as a matter of degree. We pay special attention to deriving trust information through a trusted third party, which becomes especially challenging when distrust is involved.

Keywords: network of trust, propagation, semantic web, intuitionistic fuzzy relation, interval valued fuzzy relation

1 Introduction

There is an increasing amount of information sources available to applications and users on the web. As information source breadth increases, users may find increasing challenges in trusting and sometimes distrusting sources. We expect a systematic support for trusting information sources to be one of the keys to a functional semantic web [2]. Trust in general has become an important interdisciplinary research area. We refer to [7] for a recently published collection of contributions which also shows an emerging interest in the notion of distrust. Existing computational models usually deal with trust in a binary way: they assume that a source is to be trusted or not, and they compute the probability or the belief that the source can be trusted (see e.g. [8], [10]). Besides full trust or no trust at all, in reality we also encounter partial trust. This is reflected in our everyday language when we say for example "this source is rather trustworthy"

or "I trust this source very much". In this paper we focus on (1) representing trust as a matter of degree, including the case that an agent may fully trust (or have blind faith) or distrust a source, and (2) on deriving trust information obtained through a trusted third party (TTP).

The first issue pertains to situations where sources can not be divided in the trustworthy ones and the malicious ones in a clear cut way, but they can be trusted to a certain extent. Think of trust as a matter of degree, i.e. instead of computing the probability that a source can be trusted, we are interested in the degree to which a source can be trusted. Whereas the existing probabilistic approach is suitable for problems where security is at stake and malicious sources need to be discerned from trustworthy ones, our approach leans itself better for the computation of trust when the outcome of an action can be positive to some extent, e.g., when provided information can be right to some degree, as opposed to being either right or wrong. In [6], it is argued that trust and distrust are distinct, opposite concepts. Trust and distrust can clearly coexist, e.g. among politicians who trust each other enough to cooperate, but at the same time maintain a "healthy level of distrust". In Section 2, we introduce a model that takes into account partial trust, distrust and ignorance simultaneously, as different but related concepts.

The second problem can informally be described as: if the trust value of source a in source b is p, and the trust value of b in source c is q, what information can be derived about the trust value of a in c? This problem of atomic trust propagation has been well researched in a probabilistic setting, where multiplication is used as the main operation to combine trust values. However, when distrust is involved as well, the need for a new, not necessarily commutative propagation operator arises. We discuss this in Section 3.

2 Trust Network Between Sources

Trust is a multi-faceted concept, it can be full or partial, it depends on the context, it depends on the purpose, etc. Developing a computational model forces us to make some initial simplifying assumptions. One aspect is the domain dependency of trust: e.g. we may trust the website of a store on information about location and opening hours, but that does not imply that we also take for granted everything they say in their advertisements. In this paper, we assume that we are dealing with trust in a single domain, expressed between a set of sources $\mathcal{A}$. Another aspect is the purpose of trusting a source: in this paper, we are not dealing with trust to support a decision. For instance, we do not provide or discuss the use of trust-related thresholds that along with trust values may be used for decision making.

Since trust may be a matter of degree, we use a number t between 0 and 1 to express the degree of trust of a in b. This value is not a probability nor a belief. In a probabilistic setting, a higher trust level corresponds to a higher probability that a source can be trusted, while in our interpretation it corresponds to a higher trust. Both approaches are complementary.

In our approach, 0 corresponds to total absence of trust. Roughly speaking, this can occur in either one of the following situations: (1) a has reason to distrust b fully, or (2) a has no information about b and hence no reason to trust b, but also no reason to distrust b. Taking into account the fundamental difference between the two situations, and the fact that distrust is no less important than trust in relying on a source, we propose to represent distrust d simultaneously with trust as a couple (t, d), in which both t and d are numbers between 0 and 1. Trust and distrust do not have to sum up to 1, but we assume that they satisfy the restriction $t + d \leq 1$. Omitting this restriction would result in allowing inconsistency — this is an interesting option for future development that is however not further considered in this paper. As a result, the network of trust between sources is represented by an intuitionistic fuzzy relation (IFR for short).

Intuitionistic fuzzy set theory [1] is an extension of fuzzy set theory that defies the claim that from the fact that an element x "belongs" to a given degree $\mu_A(x)$ to a fuzzy set A, naturally follows that x should "not belong" to A to the extent $1 - \mu_A(x)$, an assertion implicit in the concept of a fuzzy set. On the contrary, an intuitionistic fuzzy set (IFS for short) assigns to each element x of the universe both a degree of membership $\mu_A(x)$ and one of non–membership $\nu_A(x)$ such that

$$\mu_A(x) + \nu_A(x) \leq 1 \tag{1}$$

thus relaxing the enforced duality $\nu_A(x) = 1 - \mu_A(x)$ from fuzzy set theory. Obviously, when $\mu_A(x) + \nu_A(x) = 1$ for all elements of the universe, the traditional fuzzy set concept is recovered. Formally an IFS A in a universe X is a mapping from X to the lattice L^* defined by [3]:

$$L^* = \{(t, d) \in [0, 1]^2 \mid t + d \leq 1\}$$
$$(t_1, d_1) \leq_{L^*} (t_2, d_2) \Leftrightarrow t_1 \leq t_2 \text{ and } d_1 \geq d_2$$

An IFR in $\mathcal{A}$ is an IFS in $\mathcal{A} \times \mathcal{A}$.

Definition 1. A trust network is a couple $(\mathcal{A}, R)$ such that $\mathcal{A}$ is a set of sources and R is an IFR in $\mathcal{A}$. For all a and b in $\mathcal{A}$:

- $R(a, b)$ is called the trust value of a in b
- $\mu_R(a, b)$ is called the trust degree of a in b
- $\nu_R(a, b)$ is called the distrust degree of a and b
- $1 - \mu_R(a, b) - \nu_R(a, b)$ is the hesitation of a towards b

IFS theory has been shown to be formally equivalent to interval valued fuzzy set (IVFS) theory [4]. This is another extension of fuzzy set theory in which the membership degrees are subintervals instead of numbers from $[0, 1]$ (see [9]). A couple (t, d) of trust t and distrust d corresponds to the interval $[t, 1 - d]$, indicating that the trust degree ranges from t to $1 - d$. The hesitation degree from IFS theory corresponds to the length of the interval. The longer the interval, the more doubt about the actual trust value.

Table 1 illustrates this by means of some examples. $(0,1)$ and $(1,0)$ are respectively the smallest and the biggest element of L^*, corresponding to full distrust and full trust; obviously in these situations there is no hesitation. In the case of no knowledge, namely $(0,0)$, the hesitation is 1. The most wide spread approach (see column 2) only takes into account the degree of trust, and can not make the distinction between a case of full distrust and a case of no knowledge. In [5] the distrust degree d is subtracted from the trust degree t, giving rise to a trust value on a scale from -1 to 1. The examples $(0.2,0)$ and $(0.6,0.4)$ illustrate that valuable information is lost in this mapping process. Indeed $(0.6,0.4)$ expresses a strong opinion to trust a source to degree 0.6 but not more, while $(0.2,0)$ suggests to trust to degree 0.2 but possibly more because there is a lot of doubt in this case (the hesitation degree is 0.8). In [5], both cases are mapped to the same value, namely 0.2.

	trust	**trust and distrust**		
		IFS	IVFS	Guha[5]
	t	(t,d)	$[t,1-d]$	$t-d$
full trust	1.0	(1.0,0.0)	[1.0,1.0]	1.0
full distrust	0.0	(0.0,1.0)	[0.0,0.0]	-1.0
no knowledge	0.0	(0.0,0.0)	[0.0,1.0]	0.0
partial trust	0.2	(0.2,0.0)	[0.2,0.8]	0.2
partial trust and distrust	0.6	(0.6,0.4)	[0.6,0.6]	0.2
inconsistency		(1.0,1.0)		0.0

Table 1. Examples of trust values

3 Trust and Distrust Propagation

As recalled in the introduction, in a probabilistic framework, trust is propagated by means of the multiplication operation. This can be straightforwardly adapted to a fuzzy setting by using a t–norm, i.e. an increasing, commutative and associative $[0,1]^2-[0,1]$ mapping that satisfies $\mathcal{T}(1,x)=x$ for all x in $[0,1]$. Hence

$$\mathcal{T}(\mu_R(a,b),\mu_R(b,c)) \tag{2}$$

is the trust degree of a in c, derived from the trust degree of a in b and the trust degree of b in c. Possible choices for $\mathcal{T}$ are $\mathcal{T}_{\mathrm{M}}(x,y)=\min(x,y)$, $\mathcal{T}_{\mathrm{P}}(x,y)=x\cdot y$ and $\mathcal{T}_{\mathrm{L}}(x,y)=\max(0,x+y-1)$.

However if, instead of only the trust degree, we consider the complete trust value, i.e. both the trust and the distrust degree, propagation is not straightforward at all anymore. In this case the propagation operator is an $(L^*)^2-L^*$ mapping Prop. An example shows that Prop is not necessarily commutative.

Suppose that a has full trust in b and b has full distrust in c, than intuitively we infer that a has full distrust in c, i.e.

$$\mathsf{Prop}((1,0),(0,1)) = (0,1) \tag{3}$$

However, if a has full distrust in b and b has full trust in c, more than one approach is possible. The full distrust of a in b might lead a to ignoring b, i.e. no knowledge is inferred

$$\mathsf{Prop}((0,1),(1,0)) = (0,0) \tag{4}$$

(3) and (4) together illustrate the non commutative behavior. However, the distrust of a in b might encourage a to take on the contrary of what b is saying, in other words to trust c fully, i.e.

$$\mathsf{Prop}((0,1),(1,0)) = (1,0) \tag{5}$$

For both approaches (4) and (5) a reasonable motivation can be given. This example only lifts part of the veil of the complex problem which propagation scheme to choose. Our aim in this paper is not to provide a clear cut answer to that question, but rather to provide some propagation operators that can be used in different schemes. Recall that a t–conorm $\mathcal{S}$ is an increasing, commutative and associative $[0,1]^2 - [0,1]$ mapping that satisfies $\mathcal{S}(0,x) = x$ for all x in $[0,1]$. Possible choices are $\mathcal{S}_{\mathrm{M}}(x,y) = \max(x,y)$, $\mathcal{S}_{\mathrm{P}}(x,y) = x + y - x \cdot y$, and $\mathcal{S}_{\mathrm{L}}(x,y) = \min(1, x+y)$. A negator $\mathcal{N}$ is a decreasing $[0,1] - [0,1]$ mapping satisfying $\mathcal{N}(0) = 1$ and $\mathcal{N}(1) = 0$. The most commonly used one is $\mathcal{N}_s(x) = 1 - x$.

Definition 2. The propagation operators Prop_1, Prop_2, and Prop_3 are defined by

$$\begin{aligned}
&\mathsf{Prop}_1((t_1,d_1),(t_2,d_2)) = (\mathcal{T}(t_1,t_2), \mathcal{T}(t_1,d_2)) \\
&\mathsf{Prop}_2((t_1,d_1),(t_2,d_2)) = (\mathcal{S}(\mathcal{T}(t_1,t_2), \mathcal{T}(d_1,d_2)), \mathcal{S}(\mathcal{T}(t_1,d_2), \mathcal{T}(d_1,t_2))) \\
&\mathsf{Prop}_3((t_1,d_1),(t_2,d_2)) = (\mathcal{T}(t_1,t_2), \mathcal{T}(\mathcal{N}(d_1),d_2))
\end{aligned}$$

for all (t_1,d_1) and (t_2,d_2) in L^*.

The following proposition shows that all three propagation operators copy the information given by a fully trusted third party. It also proves that Prop_1 and Prop_3 are in accordance with (4) because they derive no knowledge through a third party that they distrust, while Prop_2 takes on exactly the opposite information given by a distrusted source and hence is in accordance with (5). Prop_1 and Prop_2 derive no knowledge through an unknown third party, while Prop_3 displays a paranoid behavior in taking on some distrust information even from an unknown third party.

Proposition 1. For all (t,d) in L^* it holds that

$$\begin{aligned}
&\mathsf{Prop}_1((1,0),(t,d)) = (t,d) \qquad &&\mathsf{Prop}_1((0,1),(t,d)) = (0,0) \\
&\mathsf{Prop}_2((1,0),(t,d)) = (t,d) \qquad &&\mathsf{Prop}_2((0,1),(t,d)) = (d,t) \\
&\mathsf{Prop}_3((1,0),(t,d)) = (t,d) \qquad &&\mathsf{Prop}_3((0,1),(t,d)) = (0,0)
\end{aligned}$$

$$\mathsf{Prop}_1((0,0),(t,d)) = (0,0)$$
$$\mathsf{Prop}_2((0,0),(t,d)) = (0,0)$$
$$\mathsf{Prop}_3((0,0),(t,d)) = (0,d)$$

Using $\mathcal{T}_{\mathrm{P}}$ and $\mathcal{S}_{\mathrm{P}}$, Prop_1 and Prop_2 take on the following form

$$\begin{aligned}\mathsf{Prop}_1((t_1,d_1),(t_2,d_2)) &= (t_1 \cdot t_2, t_1 \cdot d_2)\\ \mathsf{Prop}_2((t_1,d_1),(t_2,d_2)) &= (t_1 \cdot t_2 + d_1 \cdot d_2 - t_1 \cdot t_2 \cdot d_1 \cdot d_2,\\ &\qquad t_1 \cdot d_2 + d_1 \cdot t_2 - t_1 \cdot d_2 \cdot d_1 \cdot t_2)\end{aligned}$$

This particular form of Prop_1 has previously been proposed in [8] to combine pairs of beliefs and disbeliefs. Subtracting the distrust degree from the trust degree, the operations above reduce respectively to $t_1 \cdot (t_2 - d_2)$ and $(t_1 - d_1) \cdot (t_2 - d_2)$, which are the two distrust propagation schemes put forward in [5].

4 Conclusion

In this paper we have introduced a many valued approach for a network of trust between sources. We represent trust values as couples (t, d) in which t corresponds to a trust degree, d to a distrust degree, and $1-t-d$ to an ignorance degree. As such, to our knowledge, we are the first to introduce a model that takes into account partial trust, distrust and ignorance simultaneously. We have also presented a collection of three operators used for atomic propagation of trust, distrust and ignorance. These operators are generic enough to be used in several "trust" schemes, including those where trust, distrust and ignorance are either full or partial, and those where propagation is commutative or not. The ability to take into account ignorance and to propagate trust become extremely meaningful in a large web where the trustworthiness of many sources is initially unknown to a user, which does not imply that the user distrusts all of them, but that the user may eventually gather evidences to trust or distrust some sources and still ignore others.

The representation and propagation solutions presented in this paper are preliminary since there is a lot more to computing trust on the web, such as further propagation (longer chains) and aggregation (combining the trust information received from several TTP's). Yet one step further is to update the trust network. Another aspect not yet mentioned in the paper is that it is important to be able to calculate trust in a distributed manner taking into consideration both efficiency and privacy.

Acknowledgments

Martine De Cock would like to thank the Fund for Scientific Research–Flanders for funding her research, and the members of the Knowledge Systems Lab at Stanford University for their hospitality and the inspiring cooperation, leading to the current paper.

References

1. Atanassov, K. T.: Intuitionistic Fuzzy Sets. Fuzzy Sets and Systems **20** (1986) 87–96
2. Berners-Lee, T., Hendler, J., Lassila O.: The Semantic Web, Scientific American, May 2001.
3. Deschrijver, G., Cornelis, C., Kerre E. E.: Intuitionistic Fuzzy Connectives Revisited. Proceedings of IPMU2002 (9th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems) (2002) 1839–1844
4. Deschrijver, G., Kerre, E. E.: On the relationship between some extensions of fuzzy set theory. Fuzzy Sets and Systems **133** (2003) 227–235
5. Guha, R., Kumar, R., Raghavan, P., Tomkins, A.: Propagation of Trust and Distrust. Proceedings of WWW2004 (2004) 403–412
6. Harrison McKnight, D., Chervany, N. L.: Trust and Distrust Definitions: One Bite at a Time. Lecture Notes in Artificial Intelligence **2246** (2001) 27–54
7. Herrmann, P., Issarny, V., Shiu, S.: Trust Management, Third International Conference, iTrust 2005. Lecture Notes in Computer Science 3477 (2005)
8. Jøsang, A., Knapskog, S. J.: A metric for trusted systems. Proceedings of the 21st National Security Conference, NSA (1998)
9. Türksen, I. B.: Interval Valued Sets Based on Normal Forms. Fuzzy Sets and Systems **20** (1986) 191–210
10. Zaihrayeu, I., Pinheiro da Silva, P., and McGuinness, D. L.: IWTrust: Improving User Trust in Answers from the Web. Lecture Notes in Computer Science 3477 (2005) 384–392