Presumptive Selection of Trust Evidence

Pierpaolo Dondio
Trinity College Dublin
School of Computer Science and Statistics
Westland Row 2.1 Dublin
0035318962730
dondiop@cs.tcd.ie

Stephen Barrett
Trinity College Dublin
School of Computer Science and Statistics
Westland Row 2.1 Dublin
0035318962730
Stephen.Barrett@cs.tcd.ie

ABSTRACT

1. This paper proposes a generic method for identifying elements in a domain that can be used as trust evidences. As an alternative to external infrastructured approaches based on certificates or user recommendations we propose a computation based on evidences gathered directly from application elements that have been recognized to have a trust meaning. However, when the selection of evidences is done using a dedicated infrastructure or user's collaboration it remains a well-bounded problem. Instead, when evidences must be selected directly from domain activity selection is generally unsystematic and subjective, typically resulting in an unbounded problem. To address these issues, our paper proposes a general methodology for selecting trust evidences among elements of the domain under analysis. The method uses presumptive reasoning combined with a human-based and intuitive notion of Trust. Using the method the problem of evidence selection becomes the critical analysis of identified evidences plausibility against the situation and their logical consistency. We present an evaluation, in the context of the Wikipedia project, in which trust predictions based on evidences identified by our method are compared to a computation based on domain-specific expertise.

Categories and Subject Descriptors

I.2.11 [**Computing Methodologies**]: Artificial Intelligence - *Distributed Artificial Intelligence, Intelligent Agents.*

General Terms

Experimentation, Security, Human Factors.

Keywords

Computational Trust, Presumptive Reasoning, Wikipedia

1. INTRODUCTION

Fig. 1 illustrates a typical computational trust solution, modeled after the high-level architecture of the Secure trust engine [8]. This paper is focused on the issue of evidence selection, (i.e. the selection of the trust computation inputs), proposing a general methodology to carry out this critical task. This analysis cannot be separated from the question of which trust model underlines the selection and which trust computation is supported, our discussion will partially cover these aspects.

In general, trust evidences encompass recommendation, reputation, past interactions, entity's properties having a trust meaning, credentials and so on. We further reduce our problem space by requiring that candidate trust evidences must be elements of the application (or domain) under analysis, discarding those trust solutions requiring infrastructures independent from application's core behavior, such as PKXI, or Recommendation Systems based on user cooperation.

Evidence selection is assumed to require an understanding of both the application domain and the structure of trust models. In Fig. 2 we divide the current solutions in three categories: solutions with a dedicate trust infrastructure, solutions based on user's collaboration and solutions where trust is computed directly on application elements.

In a dedicated trust infrastructure evidences are well-defined objects that an entity may have or not (eg. a certificate), and the selection of evidences is a well-bounded problem (retrieve the appropriate certificate) in the familiar territory of a trust infrastructure. The basic trust value of an entity is based on the validity of certificates. Derived trust values can be computed using transitivity (like in a PGP infrastructure), quorum counting (like Aberer in [6]) etc.

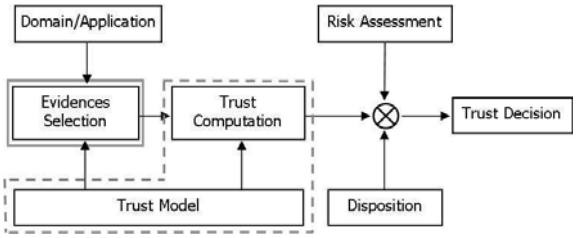


Fig. 1 A Computational Trust solution and the focus of the present work.

If a trust solution relies on users collaboration, we delegate the gathering of evidences (ratings) to users. The user rating, once collected, it is well defined and again the problem of the selection remains well-bounded. Of course, the mental processes that the user considers in order to give a rating could be complex and fuzzy, but they are out of the reputation system mechanisms. Derived trust value can be computed with any techniques. Golbeck used transitivity in her Social Network approach [16].

AAMAS'07, May 14–18, 2007, Honolulu, Hawai'i, USA.

978-81-904262-7-5 (RPS) ©2007 IFAAMAS

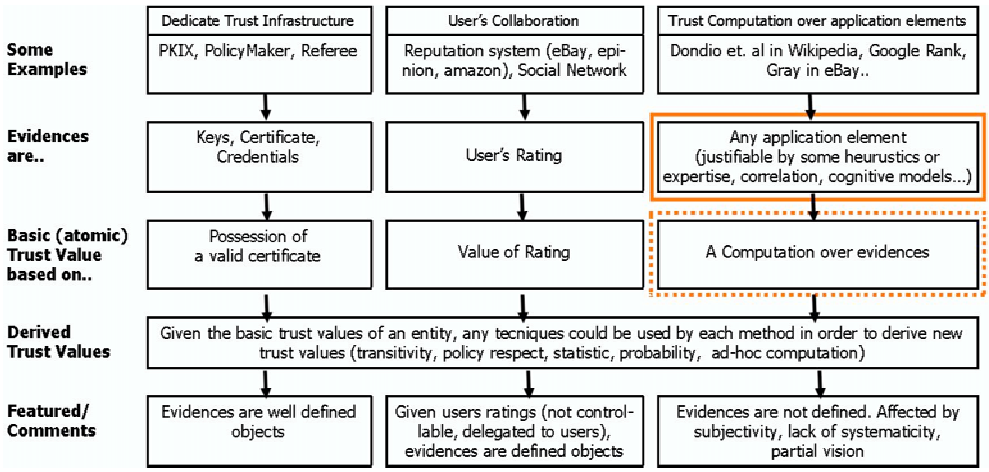


Figure 2 Different strategies for evidences selection.

Google PageRank [19] represents the best well-know example of this *modus operandi*. Many authors, notably Massa [11] identified in Google PageRank all the elements of a trust metric In this case, the application is the whole Web, seen as an interconnection of mutually linked sites. Following the scheme of figure 2, PageRank selects as evidences the outgoing and ingoing links of a web page. The selection could be justified noting that linking a page is like an implicit act of recommending that page or recognizing its reputation [19].

Other examples encompass the work done by Gray with eBay [8] or Dondio [20] in the context of the Wikipedia project. In both the works evidence selection and computation was identified and justified by domain experts. These examples make clear the complex issues that such a method could imply: subjectivity, lack of sistematicity, loss of meaning of trust, lack of justification, loss of generality too much expertise dependency.

When no trust infrastructure is present, the selection of evidences and the trust computation are rarely obvious or intuitive. In general, domain-specific expertise is needed to justify the selection of particular evidences. We assume there may be circumstances where domain specific expertise cannot be used, perhaps because a domain has not been sufficiently investigated, or is too difficult to define. We wish to explore how we might reduce and limit the reliance on domain expertise.
Furthermore, trust computations need to be capable of interpretation if human decision-making is to be informed by them. The justification for the selection of evidences should be clear if resulting trust values are to have meaning.
We take as our starting point the presumption that exists a notion of trust is, independent of any particular application domain, based on intuitive techniques that humans adopt in their relationship and environment. We may wonder if, equipped only with this generic trust expertise, we might calculate reasonably good decisions that provide comparable results to those acquired by the application of domain expertise.
Thus, our research is motivated by the following two questions: *is it possible to have a general methodology that can help agents to identify which application elements have a meaning for trust, and why? Is domain-specific expertise necessary for this identification process, or can a generic, human-related trust expertise suffice?*

The paper is organized as follow: in section 2 we review some related work, in section 3 we present the theoretical framework of presumptive reasoning referring to Walton works. In section 4 we define our method requirements, in section 5 we presents our methodology, in section 6 we propose an evaluation of some key aspects of our method conducted on the Wikipedia project. Our research, along with the method proposed, poses interesting open questions for the future summarized in our conclusion section.

2. RELATED WORKS

This paper proposes a methodology to identify and justify the selection of certain application elements as trust evidences. The focus of our research is about "which" elements should be identified as trust evidence and "why". We provide a short review of the existing trust approaches underlying the central problems described above.

2.1 Evidence Selection in Expert-Based trust

In this approach trust computations are structured by an expert that identifies the evidences, knows how to gather them and can generally give explanation for his decision. However, it may happen that the expert's knowledge is not accessible (protected or exclusive) or it is strongly unconscious and hard to make explicit..
Gray [8] adapted the trust engine Secure to extend the eBay recommendation system, utilising the work done by Kauffman and Wood [9] regarding how to identify malicious behaviour during an auction. They identified clear rules to be applied requiring specific and defined inputs.
Expert-based trust solutions overlap expert system concepts adopting their methods and difficulties. Generally, experts can deliver ad-hoc high quality solutions in a specific domain, but there is a strong dependency on the expertise: its existence, its accessibility and complexity. The role of trust underlying the methodology is minimal. In Gray's example, the trust intelligence was encoded in the rules discovered by experts, not in the trust framework that she used to implement the solution. Finally, we should also consider that an expert opinion is not out of discussion. Walton [2] calls it argument from expert and identifies some critical questions one should consider to assign the correct validity to this argument.

2.2 Evidence in Probability-based trust

Probability-based trust performs trust calculation by applying the theory of probability over a relevant set of data. In general, the underlying idea is that trust is the probability that the trustee will fulfil the expectation of the trustor. This method requires that a sufficient set of triples (trustee, action, outcome) be collected in order to estimate the probability associated with a specific couple entity-action. Time is needed to fill the trustee memory while the interactions go on.

Examples include Despotovic [4], who used the maximum likelihood theorem to predict good outcomes in a p2p recommendation system scenario; Wang [5] applied the Bayesian probability to a file sharing scenario and Wuang [12] applied the same concept to eBay.

As Despotovic points out, the meaning of a percentage value is clear, but the justification for that probability is limited and not explicit. This approach stresses the nature of trust as a prediction-probability as in Gambetta classical definition [10].

The link between evidence selection problem and probability is all about the accessibility and definition of the outcomes. In the two example cited, outcomes were easy to define and measure. When this happens, the whole process does not require domain-specific expertise; it is strongly domain-independent and abstract.

When outcomes are more fuzzy and complicated to define (or measure), expertise or user collaboration is needed: the problem of evidence selection is still open. In this case, probability techniques have to be coupled with other approach able to identify evidence. For example, in the context of the Wikipedia project Zeng et al. [18] considered the history of revisions of a Wikipedia article as trust evidence and then they computed trust values using a Bayesian trust model.

2.3 Evidence in Computational Trust models

In analytic trust models, an explicit model of trust derived from multidisciplinary studies is formalized and converted in computational procedures that are applied to situations where trust is required. In the field of Distributed Artificial Intelligence S. Marsh [1] proposed the first formal model of trust in 1994. Marsh had the clear intent to bring the human notion of trust in the digital world to exploit its benefits. The key idea is that trust can be defined as a formula, even complex, in an explicit analytic way.

In analytic trust models the problem of evidence gathering has not been studied as a central problem, while a lot of effort was put in defining the correct formal representation of trust. Seigneur [14], regarding this problem, writes that "there is the need for a clear process between trust models and trust evidences and there are a number of types of trust evidence that have not been considered in computational trust".

The notion of human trust has been formalized by many authors (see Carbone [15], Golbeck [16]), typically as an entity that is quantifiable, may be composed by transitivity, can be partially ordered, it is dynamic and usually decrease with time. Marsh recently has investigated and defined trust-related concepts like untrust, distrust, forgiveness and regret [17]. Evidence selection is not the key question. Many authors tend to consider evidence selection part of the domain analysis required to apply a trust engine to a domain, some others consider the selection of appropriate trust ingredients a subjective choice of the policy maker [15]. Although it is a crucial aspect to their correct working, analytic models leave the problem of evidence selection open. We believe that the lack of methodology in evidence selection is a serious limit to the range of application where trust models should be used.

A final mention regards the cognitive models of trust, referring to the well-known work of Castelfranchi and Falcone. In their model trust is a mental process with defined ingredients: the four basic beliefs of competence, disposition, dependence and fulfilment [13]. For evidence selection, the model, by making explicit the ingredients of trust, gives directions about which elements should be considered: elements related to one of the basic beliefs.

2.4 Evidence through Heuristics in Trust

Finally, the selection of trust evidence and its justification can be guided by the application of heuristics.

An example can be found in the work that McGuiness [3] did on the Wikipedia project. In order to assess the trustworthiness of a Wikipedia article, the author applied a heuristics based on a version of the PageRank algorithm. They considered the relative number of times an article's name appears as a link or not in the application. Thus, she selected the fact of being linked as an evidence for the trustworthiness of an article. PageRank algorithm can be considered a trust metric and thus the heuristics could be considered trust related. Applying heuristics is problematic because it could not be clear if the heuristic used has a relationship with a general notion of trust. It may be an efficient and intuitive way to predict good outcomes in that specific context that may not have clear connection to trust. By definition, heuristics are not systematic and, before answering the question why and which evidence are selected, we should answer why and which heuristics we should use. The choice could be affected by subjectivity and presumptions. Thus, a first concern with this kind of approach is the clear pertinence of a heuristics with trust.

Referring to the above example, that heuristics in that context may be severally argued: an expert Wikipedia user may argue that in Wikipedia there are automatic procedures that link articles, or that an author may link articles independently by the content of the linked article and so on. This shows the second concern regarding heuristics: they cannot be applied straightforward without a critical analysis of their applicability and relevance to the context.

3. PLACING OUR METHOD

The state-of-the-art review done in the previous section defines a multidimensional space in which we now place our approach. The task of identifying promising trust evidences is a component of a computational trust solution and it doesn't cover how to compute trust values, how to aggregate them, how to carry on the trust decision process. In fig 4, showing our method, the black line above the rectangle *trust computation* defines the limit of our interest.

Our methodology is a trust-based one. This means that the role and meaning of a human-related trust is central and justifies all the selection of evidences. In line with Castelfranchi and Falcone [13], we refuse the reduction of trust as a pure probability and we largely relies on multidisciplinary studies to define our trust-expertise. In this sense, we are close to the role that computational trust models assign to the (human) notion of trust. However, the method proposed here is not a cognitive model, but rather a more practical procedure composed by mechanisms that have justifications in some cognitive human-related activity that informs it.

We want our method to capture some strong points of the heuristics approach: the accessibility and low complexity of

knowledge, intuitiveness, and their generality. We want to set up more than a collection of simple heuristics: we want a systematic way of applying our trust expertise, a clear trust meaning and more objectivity in the selection process.
Our decisions should be sensitive to the context. Thus an important requirement is that our trust expertise is not applied without a critical analysis. This means that the focus of the method should be a dialectic analysis of candidate inputs, as Instead of producing an explicit and analytic trust model, we rather give a method to critically argue over the plausibility of evidences selection.. Domain specific expertise should play a marginal role in the method: trust is a decision that can be done without depth expertise knowledge, using only the specific expertise of trust. Expertise should be in general replace by some knowledge of the domain, or, if needed, used in a context where it doesn't build the trust solution but it rather helps to complete some stages of the method

4. Presumptive Reasoning and Trust

In this section we introduce the theoretical framework that grounds our method Identifying trust evidences is a presumptive process: we presume that some domain elements are interesting for trust and we attribute them some trust meaning.
Our assumption is to use a presumptive approach where each element is not a right or wrong candidate, but plausible or fallacious depending on the results of a critical analysis on it. Walton [2] carried out an in depth study of presumptive reasoning. Presumptions, as the author writes, have their validity and correctness depending on the context of dialogue appropriate for that case. He defined a set of 28 presumptive arguments that can be used to carry on an investigation where uncertainty prevails, but a decision is practically useful and necessary. The 28 schemes identified are fixed, generic and a priori, they are mechanism to put an argument in a discussion, such as the *argument from analogy*, the *argument ad ignorantiam*, the *argument by example* or *the argument from expert opinion*.
For each scheme, Walton proposes a set of critical questions to test their sustainability. The role of these questions is crucial: they test the plausibility of an argumentation scheme in that context and help to deeper understand the specific problem.
Presumptive reasoning is carried out as a dialectical argumentation between two parties (the *proponent* and the *respondent*) following this pattern:

1. The topic is settled
2. each party select the appropriate argumentation scheme to sustain its thesis in specific situation and topic,
3. Each time an argument is proposed, the burden of Proof switched to the other party, that attacks the argument
4. the other party tests the argument using critical questions,
5. if the argument fails (i.e. the context makes impossible to satisfy critical questions or there is a counter argument), another scheme should be applied. If no other scheme can be applied to the specific situation, the presumption is not plausible.
6. If the context changes, arguments accepted can be rejected and rejected arguments can be re-used (i.e. presumptive reasoning is not monotonic)

Presumptive reasoning, continues Walton, *is a kind of lack-of-knowledge set of inferences, a guide to prudent action through uncertainty*.

We think that presumptive reasoning is a pertinent model to be considered in the context of computational trust. The concept of presumption fits the classic definition of trust as a subjective probability. There are many other similarity: trust itself is a presumption that is enforced or weakened by a critical analysis of evidences; it is clear a non-monotonic process (new evidences or situations could drastically change the trust decision), trust is usually a dialectic negotiation between two parties (the trustee and the trustor): one is trying to persuade the other to trust him, while the other is trying to find reasons to sustain its intuition of trusting/distrusting him. The argumentation should be carried on by some generic trust scheme; as a topic becomes plausible because a generic argumentation scheme is applied on it, an element becomes plausible trust evidence since a trust scheme can be applied on it. Generic trust schemes, becomes generic reason to trust or distrust, reason to be tested with critical questions exactly like Walton argumentation schemes.
In this context, we consider again the eBay example we introduced in our introduction. We want to understand if collusion is happening in an online auction. Gray faced the problem exploiting experts work in the field. We want an alternative not depending entirely on expertise. Instead, suppose to have a set of intuitive trust schemes, a set of generic reasons to trust an entity. One of this scheme, we could name *pluralism*, could state that the results of many entities activity should be less biased and manipulated than the results of one single or few entities. Applying *pluralism* to some elements of an auction, we may argue that if an auction has several bids from different bidders it could be difficult for a potential colluder to manipulate the price, grounding our reasoning on an intuitive generic trust scheme. This simple reasoning describes our key proposal. We might justify the selection of evidence by recognizing that it is sustained by a generic, human understandable trust scheme. In other words, the element represents an instance of this general trust scheme: it implicitly acts like that trust scheme or it fits its features.

Presumptive Reasoning and Trust Evidence Selection		
Proponent/Respondent		Trustee/Trustor
Argumentation Schema		Trust Schema
Critical Questions		Critical Questions
Sustain an argument in a generic topic		Sustain the selection of an element as trust evidence
Plausible prudent conclusions		Plausible trust decision

Figure 3. Similarities between Walton presumptive reasoning and Trust Evidence selection

5. PRESUMPTIVE SELECTION OF TRUST EVIDENCE

A diagram of our method is depicted in fig. 6. Our trust expertise is represented by a set of generic trust schemes TS_n. These trust schemes have been deducted from researches in the multidisciplinary study of trust. In table 1 there a list of some of them, along with references to researches in computational trust area that investigated their validity. Some trust schemes were seldom or never used in computational trust and the authors are carrying on separate evaluations of them.
Trust schemes looks like intuitive human-based reasons to trust. They are part of a trust expertise that is understandable and not domain-specific. Following our presumptive approach, they are reason to sustain the selection of an element as trust evidence. In

other words, an element is candidate for trust evidence if in its application it satisfies the properties of a trust scheme or it acts like it. For example, the action of linking Web pages is an implicit instance of *indirect trust* scheme).

We will use the trust scheme pluralism as an example for all this section. The grounding pattern of the scheme is this: an element X (being an entity, an action, a property) is affected by an action *A* done by a reasonable high number of entities Y_n. If so, the resulting element X could be considered trustworthy if it is the result of a collaborative activity of multiple entities and points of view. *Pluralism* should guarantee X to be less biased and more objective. Note that in order to identify the *pluralism* pattern we don't need to know if the contribution of an entity was good or not, we don't need a judgment over its actions. This is a core feature of *pluralism*; if we had to collect a set of judgments, we were talking about a recommendation system. The scheme presumes that when something is the result of many entities cooperation, this evidence, despite the different quality of the single contribution, should guarantee that the resulting entity, as a whole, is more objective and less biased. Of corse, some conditions make this assumption stronger or weaker.

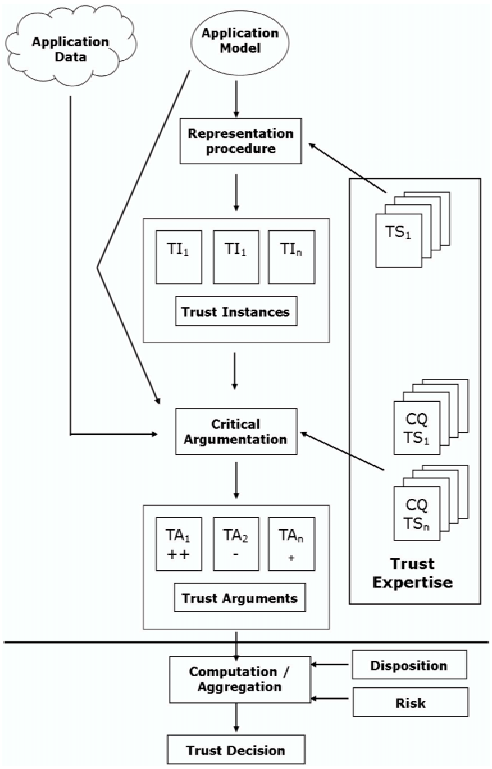


Figure 4. Presumptive selection of evidences

As Walton argumentation schemes, each trust schemes has a set of critical questions CQ attached to it, that are both inherent to the a priori assumption present in the scheme and both dependent on the specific situation (that explains the arrow from application data to argumentation in the diagram). For example, some critical questions attached to the TS "pluralism" are: *Is the number of entities Y_n significant in the context? Are entities independent from each other? How are the contribution distributed.*

Trust schemes and the corresponding critical questions represent our Trust expertise. In our methodology we select any domain elements that may be interpreted as an instance of some of our generic trust schemes. Trust schemes justifies why an element should be selected as a trust instance: the element could be an input needed for a trust scheme, or be a complete instantiated version of it. Given that we do a presumption over trust meaning of elements, we call these instances presumptive trust instances. These elements can be seen as acting like trust instances, general trust schemes instantiated with domain elements. Trust instances keep a rich notion of trust encoded in the schemes.

Table 1. Some Trust Schema

Trust Scheme	Intuitive meaning	Reference
Pluralism	I trust what is the result of many entities cooperation.	Aberer in p2p identification [9].
Indirect trust	I trust considering what others trust	Transitive Trust [Golbeck] and Recommendation Sys.
(auto)Similarity / Analogy / Categorization	I propagate trust on the basis of similarity with me (auto) or among categories or situations	See Ziegler [10] Categorization in Falcone [7]; Collaborative Filtering.
Persistence	I trust what has been stable and active for a reasonable period of time	Reputation in trust [11]. Ongoing evaluation by authors. Persuasive labs. [9]
Stability/Activity	I trust what has been stable and active for a significant long period of time	Dondio [4] Standford Persuasive Labs [9]
Standard compliance	X trusts Y in relation to its similarity to a standard/normal value in the domain	Ziegler in [10]
Mutual dependency	X trust Y because of their mutual dependency	Castelfranchi Falcone [7]
Common Goal /Common Risk	I trust an entity that have a common goal/share a common risk with me	Marsh [1]

We identify trust instances by starting form a representation of the application/domain, represented in fig. 4 as the *Application Model*. We augment this representation with a *representation procedure* informed by our test schemes. The *representation procedure* represent a less developed stage in our work and a detailed description is beyond the scope of this paper, that focuses of evidence selections. The *representation procedure* requires adding some relevant information on a basic domain representation. Having defined the a priori trust schemes, this description is actually a pre-matching of these schemes over domain elements, matching that will be quantified and tested in the next stage using the critical questions. We may wonder why is the procedure needed and if it is not enough to try directly to match our schemes on a domain representation. We see four rationales: (i) setting up a general procedure requires less or none knowledge of each of the schemes; (ii) direct matching can be more prone to error and partial visions (iii) a general procedure can be processed by chain-reasoning whose conclusions could be hard to identify with a direct matching; (iv) a *trust-aware* domain representation is by itself a useful new contribution to trust studies as discussed in the future works section. The goal of the modelling phase is to gather all the information needed to support the presumptive identification and the critical analysis of trust

instances. The information, already present in the starting representation or to be added includes analysis of entities, their properties and actions, ease of communication, environmental constraint, memory restriction and time and available historical data. Using the augmented application model, the identification of trust evidences is a two-stage process: the selection of an appropriate trust scheme and its critical argumentation. We represent a trust scheme as having a set of logical conditions that the application elements should satisfy to be considered potential trust instance of a specific scheme. The matching of these conditions guarantees the first level of matching, the *applicability* of the scheme. *Applicability* guarantees only that the element under analysis (like *the linking page* action) has the "shape" of a certain trust scheme. In table 2 the *pluralism* trust scheme is represented in a prolog notation. In order to be applied, *pluralism* needs to satisfy three conditions: a) a group of entities Y_n do a certain action *A* within a recognizable entity X, b) the outcome of the action *A* is observable and c) the outcomes are measurable.

Table 2. The pluralism trust scheme.

Fact to be matched:	
E(x)	x is an entity (the trustee)
$E(y_0)$, $E(y_1)$,.., $E(y_n)$	y_0,..y_n are entities
E(z)	z is an entity (the trustor)
Re(x), $Re(y_n)$	Entities x and y_n are recognizable
$A(a,y_n,x)$	Entity y_n does the action *a* on X
Fa	Frequency of a
L_x, L_{yn}	Lifetime of entities X and y_n
Obs{(O_a,y_1),..(O_a,y_n)}	The outcomes of action *a* done by each y is observable (by Z).
$M(O_a,y_n)$	A metric M for the action *a* exists. That means that what an entity y_n did on X doing O_a is measurable
$S(x,M(O_a,y_n))$	Function estimating how the action *a* done by Y_n on X affects X status
Action a over X to be selected as trust instance if:	
Plur(z,x) :- $E(y_n)$, Re(x), $A(a,y_n,x)$, $Obs(O_a,y_n)$, $M(O_a,y_n)$	
Critical tests to be considered:	
1. $Re(y_n)$; 2. $Ind(E(y_n))$; (entities y_n are independent) 3. the value of the number *n*; 4. Fa >> $\max(L_{yn}, L_x)$; 5. $\Delta S(x,M(Oa,y_n))$ is significant; (action *a* is relevant to X) 6. the cardinality *m* of the set: $\{ y_n \mid \Delta S(x,M(O_a,y_n)) >$ significant threshold$\}$ 7. Sequence of Oa (analysis of the agreement, disagreement, rejections, dialectical process involved in bulding X)	

The output of this identification is the set of trust instances depicted in fig. 2. Obviously an element or action can be instantiated by more than one trust scheme and vice-versa: the perception of the same element depends on the particular point of view and the collection of all point of views can produce a better understanding of the unbounded problem. In the *Critical Argumentation*, the identified trust instances are tested against critical questions. These questions test the validity of the scheme applied on a particular entity in the context of interest. This stage may require information that comes from the application model but also it can consider actual data coming form the application. The critical questions test the efficiency of a trust instance, i.e. how it will work in that situation over the matched elements. But efficiency does not guarantee that the specific instance will be useful in trust computation. That's why critical questions consider also the relevance of the matched elements over the trustee, called the *strength* of the trust instance. A quite efficient scheme matched on critical domain element will be obviously stronger evidence than a perfectly efficient scheme represented by insignificant domain elements. The idea could be made clear with this example: if I have to trust Mary to follow the prescribed diet to reduce her weight, I don't care if she dresses in red while other patients in blue.

The relevance of the element is not its impact on trust, but rather a separate issue: does a modification of the element affect trustee status? What's the correlation between the elements and the trustee? Relevancy is a property inherent to domain structure, not to trust. Answering the questions does not mean to assess trustworthiness or define trust rules.

Critical questions are not to e intended as set of exact formulas to be computed over elements, but rather a precise about what to consider for testing the plausibility of a trust instance.

Referring to the trust scheme pluralism in table 2, which are its critical tests? First, entities should be recognizable (1), in order to be distinguished. The scheme is enforced when the cardinality *n* of the set is reasonable high in relation to our space problem (3) and the number *m* of entities whose contribution to X was not negligible is significant (6) (see evaluation), the action *a* happens with a frequency *Fa* that guarantees that many outcomes can be collected before entity X changes (Lx) (4), it is enforced if we can think that the entity are independent (think about the same information confirmed by different independent source) or there isn't an explicit relationship among them (2), and if action *a* is critical for X (5), that means that it can modify sensibly X status: if it can't, it becomes useless for our purposes.

Finally, further tests require analyzing how the process that produced entity X was carried out. We consider the sequence of action *a*'s outcomes in order to discover the significance of the dialectical process (agreement, disagreement, rejections…)

We point out that our focus is on how to justify the selection of an element as trust evidences. Thus, the computation of a trust value based on this evidence is beyond the core of our method. Anyway, our investigation, particularly the critical test to be satisfied, gives a clue and a quantification of how strong the plausibility of the selection.

After the critical questions our trust instances became trust arguments. Then, a trust arguments results composed by

1. a set of domain elements forming a trust instances
2. the reason why they were selected (the trust scheme)
3. the strength of their plausibility (*are the critical questions satisfied?*)
4. it represents an argument in favour or against the scheme

Deciding which elements have a trust meaning cannot be separated by a kind of computation. Again, we notice that we don't carry on a complete trust computation: we just evaluate if the element respect the condition imposed by the trust scheme or not. If it satisfies them, a trust instance is a positive, negative or undetermined argument for a subsequent trust computation, argument justified by its trust scheme.

6. EVALUATION

In this section we compare, in the same scenario of the Wikipedia project, the predictions of an expertise-based trust, carried on by

the authors in [20], and the ones based on the evidences found using our generic trust expertise. The evaluation focuses on how critically-tested trust scheme are effective in a trust computation compared to an expertise based method. However, it doesn't consider a formal procedure definition for a trust-aware augmented domain representation and automatic matching of schemes.

In order to compare the two experiments we completed the trust computation using a sum aggregation function among the identified evidences to produce a trust value, the same strategy used in the expert-based approach. Wikipedia is a global online encyclopedia, entirely written collaboratively by an open community of users and the problem of its articles trustworthiness has been strongly discussed. In order to evaluate our predictions we should know the actual quality of an article. Wikipedia gives its best articles (0.1% of them) an award called featured article status that guarantees these articles to represent the highest standard of the encyclopedia. The evaluation will success if our trust computation will indicate these articles among the most trustworthy. The set used for the both experimentations is composed of 7 718 articles downloaded on the 17th of March 2006, including all 846 featured articles plus the most visited pages with at least 25 edits. The set is significant, since it represents the majority of the editing activity of Wikipedia.

Graph 1: expertise-based trust computation

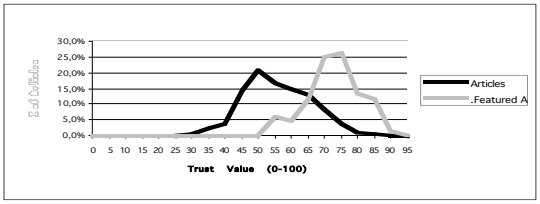


Table 3: expertise-based computation

Correlation	18.8 %		
	% of FA	% of SA	GAP
Bad: TV < 50	0	42.3%	42.3%
Average: TV in [50, 70]	22.2 %	54.7 %	32.5%
Good: TV > 70	77.8 %	13 %	64.8 %
Very Good: TV > 85	13.2 %	23 articles	13.2 %

Legend: TV = trust value [0-100], FA = featured articles, SA = standard articles

In [4] we derived explicit trust rules from expert researches in the field of collaborative editing and content quality. Our previous results are summarized in Graph 1, representing the distribution of the articles on the base of their trust value. We have isolated the featured articles (grey line) from standard articles (black line). Results obtained were positive and encouraging, summarized in table 3. In our new evaluation we start by considering a UML model of Wikipedia. We skipped the augmented representation of the domain and we directly matched a subset of trust schemes over the element *article*, enough to show how the method could work and test its validity. We therefore identified two trust instances: the pluralism of an *article* applied to the action *editing* and the *standard compliance* over a subset of properties of an *article*.

Pluralism. In Wikipedia we can identify the *pluralism* scheme on the action *editing* made by the entities *authors* on the recognizable entity *article*. Referring to the prolog-like formalism, the target entity *article* is done by many *authors* applying the *editing* action, and all the outcomes (=the editing contributions) are observable. So the trust scheme is applicable: article made by more than one authors should be less biased, with a correct point of view, complete. But how we test it? The scheme is more efficient when the number of different authors contributing the article is reasonable high, but we should consider authors that actually did a valuable contribution as the scheme prescribes (point 6 in table 2), discarding contributions of single words, grammatical correction and so on. Finally the impact (function S in table 2) of the *editing* action is obviously crucial, being the core functionality of a wiki-based application.

Standard Compliance. The presumption is that pages that show features compliant to a standard are trustworthy. Here we exploit a characteristic of Wikipedia that divide all the pages in category according to fixed subjects.

We present the output of a computation

1. based only on the *pluralism* scheme, to evaluate the effectiveness of a single strongly-matched trust scheme.
2. using both pluralism and *standard compliance* schemes to test how the value of the prediction is affected.
3. using pluralism and similarity after having tested the *standard compliance* scheme with critical questions.

Graph. 2: Pluralism based computation

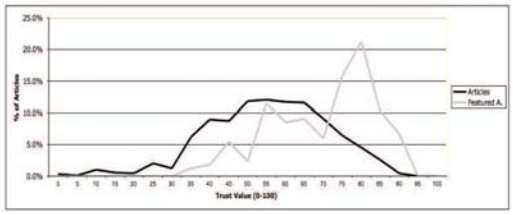


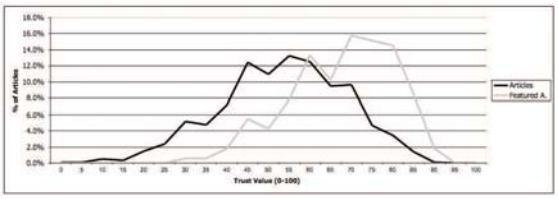
Table 4 Pluralism based computation.

Correlation	41.8 %		
	% of FA	% of SA	GAP
Bad: TV < 50	8.5 %	29.7 %	21.2%
Average: TV in [50, 70]	31.5 %	47.3 %	15.8 %
Good: TV > 70	60 %	23.2 %	36.8 %
Very Good: TV > 85	17 %	3.1 %	13.9 %

Graph. 2 and table 4 show the results obtained by applying only the *pluralism* trust scheme. The results can be considered less good than the expert-based but still valid: features articles have an interesting peak of distribution around 75 while standard articles have a wider distribution between 40 and 70.

In graph 3 we added to our estimation the *standard compliance* scheme in a straightforward manner, calculating it over the entire population of articles. As expected, given the broad variety of articles, the scheme has a controversial effect. The correlation increases, the results are deteriorated especially in the region of good articles. The application of the scheme was not plausible, as its critical questions could have spotted. Critical tests of the scheme suggest taking in consideration only sets of entities with small variance and to assure plausible standards.

Graph 3: Pluralism and standard compliancy combined



Thus, we reapplied the scheme calculating separated standard for each category of articles, subsets still numerous and more compact. Now the argument is more plausible, assertion that is confirmed by the final table 6. The shape of the graph is the same as Graph 2 and we gained 3% of less correlation and 2% of good predictions. The distinction between featured and standard articles returns to be clear and slightly better than the computation based only on pluralism, further reducing the gap with expert predictions.

Tab. 6 Pluralism and standard compliance based computation

Correlation	38.8%		
	% of FA	% of SA	GAP
Bad: TV < 50	8.5 %	48 %	39.5%
Average: TV in [50, 70]	28.3 %	32.4 %	4.1 %
Good: TV > 70	62.2 %	19.6 %	40.6 %
Very Good: TV > 85	15.2 %	**1.6 %**	12.6 %

7. OPEN ISSUES AND FUTURE WORKS

In this paper we proposed a generic methodology to identify elements of an application that can be trust evidence. We grounded the method on a set of intuitive trust scheme and a presumptive approach where each element is not a right or wrong candidate, but a plausible or not depending on the results of a critical analysis on its validity. The methodology presented in this paper has three key features: (i) it is general, reducing domain-specific expertise reliance; (ii) it is capable of offering justification for why it selects certain elements and discards others, being based on a vision of trust strongly human-based; (iii) it supports trust decisions. Our evaluation sections showed that, only equipped with a set of intuitive trust schemes, good results.

Switching the burden of proof between Trustee and Trustor. A way of implementing our method in a multi-agents scenario could be by defining a protocol of communication in which the trustee and the trustor are engaged. The idea is that the trustee itself have to prove to the trustor its trustworthy, by providing, in a trial-like fashion, the appropriate information to satisfy critical questions.

The problem of aggregation. Our outputs are a set of trust arguments to be aggregated that may contradict each other. A strategy to solve the conflict and support a subsequent decision could be the focus of an extension of the present work. Our anticipated solution is to consider the logical consistency of the trust evidence selected, instead of aggregating them with a weighted average. The latter approach seems to lose information and mild contradictions, as noted by Massa [11]. We seek to use argumentation theory to understand the defensibility of each argument in respect to the others, considering elements selected, trust scheme and critical questions answered.

Expertise dependency. An open problem partially solved is the dependency from a domain expert in answering the critical questions. A first answer is that a clear difference exists. In expert-based trust the expert delivers a solution, in our approach expertise, if needed, is invoked in the framework of a methodology to gather specific information on the domain elements not regarding trust.

8. REFERENCES

[1] Marsh, *S. Formalizing Trust as a Computational Concept*. PhD thesis, University of Stirling, D. of Mathematics 1994

[2] Walton, D. *Argumentation Schemes for Presumptive Reasoning*. Lawrence Erlbaum Associates, 1996, USA, 1998

[3] McGuiness, D. et al. *Investigations into Trust for Collaborative Information Repositories: A Wikipedia Case Study*. MTW 06, Edinburgh, Scotland, 2006

[4] Z. Despotovic, *Maximum Likelihood Estimation of Peers performance in a P2P network*. CIA, Erfurt, Germany, 2006

[5] J. Wang, Vassileva J. *Bayesian Network Trust model in P2P Networks*. IEEE WIT, Halifax, Canada, 2003.

[6] K. Aberer. *Efficient, handling of Identity in a P2P network*. IEEE Transaction on Knowledge Engineering, Vol.16, Issue 6, 2004

[7] C. Ziegler, Golbeck J. *Investigating correlation between Trust and Similarity*. Decision support system 2005.

[8] L. Gray. *Trust-Based Recommendation Systems*. PHD *Thesis, Trinity College Dublin, Ireland*

[9] R. Kauffman, C. Wood. *Detecting, predicting and preventing reserve price shilling in online auctions*. International Conference on E-Commerce, Pittsburgh,USA, 2003.

[10] D. Gambetta, *Trust: Making and Breaking Co-operative Relations* Basil Blackwell, Oxford 1988.

[11] Massa P. *Controversial Users demand local trust Metrics*. American Association for Artificial Intelligence, 2005.

[12] Wuang Y. et al. *Bayesian Network Based Trust Management*, IEEE ATC, Wuhan, China, 2006

[13] Castelfranchi, C., Falcone, R.. *Trust is much more than subjective probability: Mental components and sources of trust*. 32nd Hawaii Int. Conference on System Sciences, 2000

[14] Seigneur J.M. *Ambitrust? Immutable and Context Aware Trust Fusion*. Technical Report, Univ. of Geneva, 2006

[15] Carbone, Nielsen M., M. Sassone V. *A Formal model of Trust in Dynamics Nerwork*. iTrust 2005, 3rd conference on Trust Management, Roquefort, France, 2005

[16] Golbeck, J., Hendler, J., and Parsia, B., *Trust Networks on* the Semantic Web. ,University of Maryland, USA, 2002

[17] S. Marsh, Trust, Untrust, Distrust and Mistrust - An Exploration of the Dark(er) Side, iTrust 2005, 3rd conference on Trust Management, Roquefort, France, 2005

[18] H. Zeng et al. *Computing Trust from Revision History*, PST 2006, Privacy, Security and Trust, Canada, 2006

[19] L. Page, S. Brin et al., *The PageRank citation Ranking: bring Order to the Web*, Standford University, US, 1999

[20] P. Dondio et al. *Extracting trust from domain analysis: a study on Wikipedia*, IEEE ATC, Wuhan, China, 2006