Calculating the Trust of Event Descriptions using Provenance

Davide Ceolin, Paul Groth, Willem Robert van Hage
VU University Amsterdam
Amsterdam, The Netherlands
Email: dceolin,pgroth,wrvhage@few.vu.nl

***Abstract*—Understanding real world events often calls for the integration of data from multiple often conflicting sources. Trusting the description of an event requires not only determining trust in the data sources but also in the integration process itself. In this work, we propose a trust algorithm for event data based on Subjective Logic that takes into account not only opinions about data sources but also how those sources were integrated. This algorithm is based on a mapping between a general event ontology, the Simple Event Model, and a model for describing provenance, the Open Provenance Model. We discuss the results of applying the algorithm to a use case from the maritime domain.**

I. Introduction

The hijacking of a freighter in the Gulf of Aden, a goal not given in the semi-final of the World Cup and the sudden rise of the stock market, understanding these *events* requires the integration of data from multiple data sources using complex data integration routines. For example, to build a description of why a goal was not given there may be the report of the referee, the comments of managers and players, and video from different camera angles. The veracity of the resulting *description of the event* is dependent not only upon the trust one has in the original data sources (e.g. players, referees, cameras) but also in trust one has in the process used to create the event description.

Therefore, in this work, we investigate the generation of trust ratings for event descriptions. These trust ratings are calculated with respect to not only the original sources but also to the data integration process itself. Thus, the trust calculations consider the whole of an event description's *provenance*. The trust algorithms presented here rely on the novel combination of two existing representations, the Simple Event Model (SEM) for event representations and the Open Provenance Model (OPM) for representing the data integration process itself. Based on a mapping of these models, we develop a trust algorithm using subjective logic. We apply our trust algorithm to a use case from maritime shipping. The contributions of this paper are twofold:

1) A mapping of SEM to OPM.
2) An algorithm for computing trust ratings for event descriptions based on their provenance.

The rest of this paper is organized as follows. We begin with a description of a use case for data integration for event descriptions, which we use as a running example. This is followed by a discussion of both OPM and SEM and a presentation of the mapping between these models. Based on this mapping, we then present an algorithm for producing trust ratings for event descriptions. After this we present initial results applied to the use case. We end with a discussion of related work and a conclusion.

II. Use Case

Our use case comes from the maritime domain. It is of vital importance for the coast guard, harbors and ships to know where ships are and their vicinity to one another. Being able to track ships helps avoid collisions, manage traffic in crowded harbors, respond to emergency, and facilitate navigation. To enable this tracking, a common system has been developed called The Automatic Identification System (AIS) has been developed.[1] The International Maritime Organization requires that the system be installed on all ships over 300 tons. AIS works by exchanging messages between local ships and radar stations. This messages provide a range of information about the ship including its geoposition, navigation status, speed, radio call sign, the ship's unique registered id (MMSI - Maritime Mobile Service Identity), a permanent id (IMO - International Maritime Organization Number) and the ship's dimensions. Such messages are subject to manipulation, corruption, and errors impacting their reliability [1]. For example, the unique registered id may be falsely programmed into the system, the message may be corrupted during radio transmission, or users may fail to update their navigation status.

An AIS message or series of AIS messages describe the event of a ship's movement or change in status. Often, one would like to extract information about that event. Here, we use a simple example of extracting what nation the ship is registered to. This is known as the *flag* of the ship. This is actually a difficult problem as both the MMSI number as well as the IMO number report the country of origin and these may disagree because the MMSI can change when the ship is reregistered. Indeed, one report identified 26 vessels using the same MMSI number [1]. In addition, country information may be garbled or incorrectly entered. Thus, if part of the event description is a flag then it is important to be able to determine whether to trust that flag information based on the information sources and how those sources were combined. SEM is already

[1] http://www.uais.org

being used to represent ship movement events based on AIS messages [2]. However, we need to add additional information to represent the provenance of the description. For this, we turn to a model designed specifically for provenance, namely, OPM.

III. Mapping SEM and OPM

In order to connect the description of an event to how that description was created, we need to be able to interpret the event description with respect to its provenance. To do so, we provide a mapping from the model used for event descriptions (SEM) to the model used for describing provenance (OPM). To facilitate the explaination of this mapping, we first briefly introduce both SEM and OPM.

A. SEM, the Simple Event Model

SEM [2], [3], [4] is a schema for the semantic representation of events. It does not deal with the way data about events is stored, but only with the events themselves. SEM focuses on modeling the most common facets of events: who, what, where, and when. These are represented respectively by the SEM core classes sem:Actor, sem:Place, sem:Object and sem:Time. SEM is a model that takes into account the inherent messiness of the Web by making as little semantic commitment (e.g. disjointness statements, functional properties) as possible. Every instance of one of the core classes can be assigned types from domain vocabularies. For example, the sem:Event instance ex:world_cup_2010 can be assigned a sem:eventType dbpedia:FIFA_Club_World_Cup. Any property of SEM, including the type properties, is optional and duplicable. SEM and Simple Knowledge Organization System (SKOS) [5] mappings to related models (DOLCE-Lite, CIDOC-CRM, SUMO, LODE, F, Dublin Core, FOAF, and the CultureSampo and Queen Mary's event models) can be accessed online.[2]. Additionally, through sem:View an event can have multiple, perhaps conflicting, descriptions.

B. OPM, the Open Provenance Model

OPM is a community developed model for the exchange of provenance information [6]. It stems from a series of interoperability challenges (Provenance Challenges) held by the provenance research community to understand and exchange provenance information between systems. While not as comprehensive as some other provenance models such as ProPreO [7] , OPM provides a common technology-agnostic layer of agreement between systems. OPM was used by 15 teams during the Third Provenance Challenge [6]. These teams used a variety of provenance management systems ranging from those focused on workflow systems to those concentrating on operating systems. Thus, by using OPM, we aim to be able to apply our trust algorithm to a variety of systems.

OPM represents the provenance of an object as a directed acyclic graph with the possibility for annotations on the graph. The graph is interpreted as being causal. An OPM graph captures the past execution of a process. The graph consists of three types of nodes:

- An *opm:Artifact*, which is an immutable piece of state, for example, a file.
- An *opm:Process*, which is perform actions upon artifacts and produce new artifacts. An example of a process would be the execution of the Unix command cat on two files to produce a new concatenated file.
- An *opm:Agent*, which controls or enables a process. An example of an agent would be the operating system that a process runs in or the person who started the process.

These nodes are linked by five kinds of edges representing dependency between nodes. An opm:Process used and generated opm:Artifacts, represented by opm:used and opm:wasGeneratedBy edges. These artifacts can be given an opm:Role with respect to an opm:Process distinguishing it from other artifacts. Note, an opm:Process can only produce one opm:Artifact. Dependency between opm:Artifacts is represented using opm:wasDerivedFrom while dependency between opm:Processes is represented using the opm:wasTriggeredBy edge. Finally, the control of an opm:Process by an opm:Agent is expressed using the opm:wasTriggeredBy edge.

Each part of an OPM graph can be labeled with an *account*, which allows the same execution to be explained from different perspectives. For example, one could describe the generation of an event description with more or less detail.

SEM	SKOS relation	OPM
sem:Event	skos:closeMatch	opm:Process
sem:Actor	skos:closeMatch	opm:Artifact
sem:Actor	skos:broadMatch	opm:Agent
sem:Place	skos:closeMatch	opm:Artifact
sem:Place	skos:broadMatch	opm:Agent
sem:Role	skos:closeMatch	opm:Role
sem:View	skos:closeMatch	opm:Account

TABLE I
Mapping between OPM and SEM

C. Mapping

Given an event description in SEM, we would like to determine how its facets should map to OPM so that we can describe the facet's provenance using OPM. For example, if an event occurred at a sem:Place, we could consider that place an opm:Artifact. This idea is in-line with the notion of sub-typing within OPM [6]. We could say that a particular opm:Artifact has a type of sem:Place. To represent the mapping, we use SKOS, a W3C standard for describing and mapping vocabularies (i.e. concept schemes). The use of SKOS follows the practice of the W3C Provenance Incubator Group in defining a set of Provenance Vocabulary Mappings [8]. We refer the readers to [5] for the exact definitions of skos:closeMatch, skos:relatedMatch and skos:broadMatch.

Our mapping focuses on the nodes within the OPM graph and not the edges, because our aim is to describe the provenance of both the event description and its facets. We now discuss the mapping shown in Table I in more detail.

[2]http://semanticweb.cs.vu.nl/2009/11/sem/

For sake of space, we report only a mapping at class level. A more comprehensive mapping detailed with justifications is available on the web.[3]

Each sem:Event is an action with some duration, this maps very closely with the notion of an opm:Process. SEM has the notion of an sem:Actor, the entities or people *who* take part or are involved in an event. If an sem:Actor is directly a cause or is vital for an event to take place, we would model this as an opm:Artifact used by an opm:Process. For people who were not directly involved but enabled the event to take place, the sem:Actor would be mapped to an opm:Agent. By way of example, the crew on board a ship would be modeled as opm:Artifacts while the CEO of the shipping company can be seen as an opm:Agent controlling the event of sending an AIS message. Similar reasoning applies to mapping sem:Place to OPM.

The sem:Role signifies the role a particular SEM facet plays in an event, just as an opm:Role signifies the role a particular opm:Artifact plays with respect to an opm:Process. Additionally, an sem:View allows for multiple descriptions of the same event, which maps naturally to an opm:Account describing different descriptions of the same execution. Finally, the time of an sem:Event can be easily mapped to the time annotations present on OPM edges.

IV. Trust Rating Algorithm

We now describe our trust rating algorithm. The algorithm works upon OPM graphs. We assume that the provenance of each facet of an event description is captured. Before applying the algorithm, the above mapping is applied in order to view the facets of the SEM event description in OPM.

A. Subjective Logic

Subjective logic [9] is a probabilistic logic that provides the basis for the evidential reasoning part of our trust model. Subjective logic's probabilities are based on the Beta probability distribution [10]. These probabilities represent the level of belief, disbelief and uncertainty about each proposition we encounter, according to the evidence we own and are represented by means of "opinions" about such propositions.

This logic provides also operators for combining such opinions in order to handle the combination of opinions that reflect the application of propositional logic operators to the proposition which are objects of such opinions.

B. Opinions

The key concept of Subjective Logic logic is the concept of "opinion", which is the probability of correctness of a proposition according to a certain source. An opinion according to source x about proposition y is represented as ω_y^x. More precisely, opinions are depicted as follows:

$$\omega_x^y(b, d, u, a)$$

[3]http://bit.ly/c8A3A7

which is a representation equivalent to the Beta probability distribution, where :

$$b = \frac{positive\_evidence}{total\_evidence + n} \; d = \frac{negative\_evidence}{total\_evidence + n}$$

$$u = \frac{n}{total\_evidence + n} \; a = \frac{1}{n}$$

b,d,u are, respectively, *belief*, *disbelief* and *uncertainty*. *a* is the *a priori probability*, that is the probability that the proposition is correct, in absence of evidence. *n* is the cardinality of the set of possible outcomes, so it may be equal to 2, in case of a boolean outcome, or higher.

The expected value of the probability distribution represented by an opinion is given by:

$$E = b + a \times u$$

The expected value E will be used as trust value about propositions. *E is the "trust value"*. Given the evidence that we have collected about a certain proposition, E represents the probability that the proposition is true. Therefore it numerically quantifies our trust in the proposition.

Consider the following example. There are 249 countries in the world. Thus, the number of possible outcomes for a flag is 249. For sake of simplicity, we consider the 35 most used flags, which cover 99% of ships.

Here we consider three sources of information about the flag. Two sources say the flag is Italy. One source says the flag is the USA. Each of these opinions is secure according to each source, therefore they assume the pattern $\omega_y^x\left(1, 0, 0, \frac{1}{n}\right)$.

$$\omega_{italy}^{s_1}\left(1, 0, 0, \frac{1}{35}\right) \; \omega_{italy}^{s_2}\left(1, 0, 0, \frac{1}{35}\right) \; \omega_{usa}^{s_3}\left(1, 0, 0, \frac{1}{35}\right)$$

These are the opinions about the three sources, where $n = 2$ because, unlike previous opinions that represent the probability that a given value is correct (in a multivalued distribution), these opinions represent the probability that the source is reliable (therefore in this case the probability distribution is binomial):

$$\omega_{s_1}^x\left(\frac{8}{12}, \frac{2}{12}, \frac{2}{12}, \frac{1}{2}\right) \; \omega_{s_2}^x\left(\frac{9}{12}, \frac{1}{12}, \frac{2}{12}, \frac{1}{2}\right)$$

$$\omega_{s_3}^x\left(\frac{5}{12}, \frac{5}{12}, \frac{2}{12}, \frac{1}{2}\right)$$

Procedure *opinion_source(A_i)* of Algorithm Fig. 1 (Lines 26 - 30) builds opinions for given Artifact A_i.

C. Weighting (discounting) operators

Subjective Logic allows to build networks of opinions. The logic allows opinions to be transitive, but such opinions are weighted on the reputation of the source when evaluated by third parties. Given the opinion of z on y (ω_y^z), and the opinion of x on z (ω_z^x), the opinion that x derives from z about y is represented by $\omega_y^{x:z}$. The operator for weighting opinions is:

$$\omega_z^x \otimes \omega_y^z = \omega_y^{x:z}(b_z^x b_y^z, b_z^x d_y^z, d_z^x + u_z^x + b_z^x u_y^z, a_y^z)$$

Following the previous example, the weighted opinions become:

$$\omega_{italy}^{x:s_1}\left(\frac{8}{12},0,\frac{4}{12},\frac{1}{35}\right)\ \omega_{italy}^{x:s_2}\left(\frac{9}{12},0,\frac{3}{12},\frac{1}{35}\right)$$
$$\omega_{usa}^{x:s_3}\left(\frac{5}{12},0,\frac{7}{12},\frac{1}{35}\right)$$

All the disbeliefs have value zero as consequence of starting from secure opinions.

On line 31 of Algorithm of Figure 1, procedure opinion_sources(A_i) returns opinions about artifact A_i weighted on reputation of the sources.

D. Fusion operator

Finally, the logic provides a range of operators which allow us to combine opinions about the same proposition (fusion). The fusion of n opinions given by sources $x_1, ..., x_n$ about the same proposition y is represented as $\omega_y^{x_1 \diamond ... \diamond x_n}$. The operator works as follows:

$$\omega_y^{s_i} \oplus \omega_y^{s_j} = \omega^{s_i \diamond s_j}(\frac{b_y^{s_i} \times u_B + b_y^{s_j} \times u_y^{s_i}}{u_y^{s_i} + u_y^{s_j} - u_y^{s_i} \times u_y^{s_j}},$$
$$\frac{d_y^{s_i} \times u_y^{s_j} + d_y^{s_j} \times u_y^{s_i}}{u_y^{s_i} + u_y^{s_j} - u_y^{s_i} \times u_y^{s_j}}, \frac{u_y^{s_i} \times u_y^{s_j}}{u_y^{s_i} + u_y^{s_j} - u_y^{s_i} \times u_y^{s_j}}, a_y^{s_i})$$

Since s_i's and s_j's opinion have the same object, their a priori probability is the same ($a_y^{s_i} = a_y^{s_j}$).

$\oplus$ is an operator that returns cumulative fusion of opinions [11] (since we assume that they are independent opinions, evidence that these opinions resemble are cumulated).

Continuing our example, by merging the previous opinions regarding the two outcomes (Italy and USA), we obtain:

$$\omega_{italy}^{x:s_1 \diamond x:s_2}(0.77, 0.14, 0.09, 0.5)\ \omega_{usa}^{x:s_3}(0.42, 0.42, 0.16, 0.5)$$

Line 21 of algorithm of Figure 1 iteratively merges opinions about the Artifact of interest.

E. Trust Rating Algorithm

Here we present an algorithm for calculating the trust value of an event facet, represented by artifacts. However, because of its recursive nature, the algorithm is directly applicable to event descriptions.

Given an artifact to calculate the trust value of, our first step is determine the opinion of any source that directly generates the artifact's value. The following steps are:

- take the amount of evidence given by each source about each possible value for the artifact. Usually each source gives one output, but if more are available, then the resulting opinion is stronger (see subsect. IV.B).
- weight the opinions given by the sources according to the opinion on the source itself (in turn, based on previous evidence about its trustworthiness, see subsection IV.C)
- merge all the opinions (see subsection IV.D)

Generalizing, we can say that:

- given an artifact A;
- given a set of sources: s_1, ... s_n

```
(1)  proc tv (A_i) ≡
(2)    res := null
(3)    for P_k : A_i opm : wasGeneratedBy P_k do
(4)      for A_j : P_k opm : used A_j do
(5)        if A_i opm : wasDerivedFrom A_j
(6)          then
(7)                if res = null
(8)                  then res := tv(A_j)
(9)                  else res := F(P_k)(res, tv(A_j))
(10)               fi
(11)       fi
(12)     od
(13)   od
(15)   comment: res = ω_{v(A_i)}^{∀_{A_j} x:tv(A_j)}
(16)   for s_i : ∃v_{s_i}(A_i) ≠ ∅ do
(17)     if res = null
(18)       then res := opinion_sources(A_i)
(19)       else res := res ⊕ opinion_sources(A_i)
(20)     fi
(21)   od
(22)   return res
(23) end
(24) proc opinion_source(A_i)
(25)   for s_i : v_{s_i}(A_i) ≠ null do
(26)     record_evidence(v_{s_i}(A_i))
(27)   od
(28)   return ω_{v(A_i)}^{x:s_i}
(29) end
(30) proc π(t, s_i, A_i)
(31)   e : e ∈ domain ∧ dist(e, v_{s_i}(A_i)) =
(32)   = min_{∀e'∈domain}(dist(e', v_{s_i}(A_i)))
(33)   d := dist(e, v_{s_i}(A_i))
(34)   record ω_{v_{s_i}(A_i)=e}^{s_i}(b'_{s_i} · 1/d, 0, (d'_{s_i} + u'_{s_i}) · (1 − 1/d), a'_{s_i})
(35)   comment: b'_{s_i}, d'_{s_i}, u'_{s_i}, a'_{s_i} are the
(36)   comment: projections of b_{s_i}, d_{s_i}, u_{s_i}, a_{s_i}
(37) end
(38) proc dist
(39)   comment: distance between two points
(40)   comment: (e.g. Euclidean).
(41) proc record_evidence
(42)   comment: stores evidence in memory .
(43) proc record
(44)   comment: stores opinion in memory.
(45) proc ω
(46)   comment: returns an opinion
(47)   comment: based on stored evidence.
(48) comment: Possible values for F:
(49) F(concat) = ∧
(50) F(lookup(t)) = ∧ · π(t)
```

Fig. 1. Trust Rating Algorithm

- given a function $v(s_i, A) = v_{s_i}(A)$
- given opinions on the sources $\omega_{s_i}^x(b_{s_i}, d_{s_i}, u_{s_i}, a_{s_i})$

We compute the opinion on a event facet from each source:

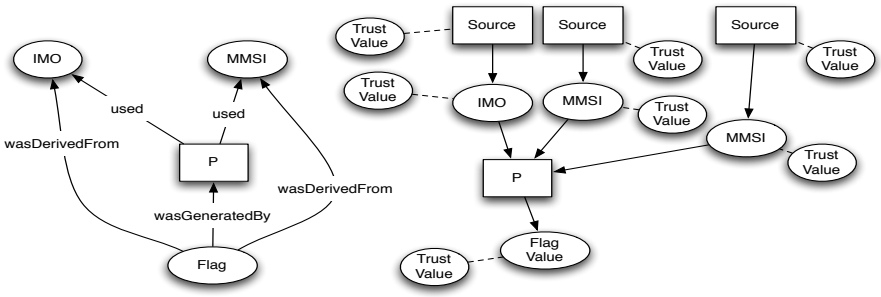

Fig. 2. Provenance and Trust graphs about the flag value of a ship. The left graph reconstructs the provenance of the flag field. The graph on the right, starting from the first ancestors of the flag field, collects all the evidence about all the artifacts involved in the provenance trail (of the left graph) and gradually merges them.

$$\omega^{x:s_i}_{v_{s_i}(A)}(b_{s_i}, 0, d_{s_i} + u_{s_i}, a_{s_i})$$

Once we have the opinions about the values from each source, we merge them in order to obtain an opinion for each value from all sources:

$$\bigoplus_{v_{s_i}} \omega^{x:s_i}_{v_{s_i}(A)}(b_{s_i}, 0, d_{s_i} + u_{s_i}, a_{s_i})$$

F. Integration process

We want to consider not only sources that directly provide the artifact value but also which process is used during integration to generate the artifact. Therefore, in case the artifact is not a leaf node, then we need to merge the (eventual) opinions computed taking into account the provenance of the artifact. For example, considering the example of Fig. 2, we see that the trust level of the root node depends on the trust levels of the leaf nodes, combined according to how the process manipulates them. Therefore, we should use a functor that, allows us to apply proper functions to the trust values of the input artifacts, according to the kind of process that manipulates them.

Two examples are provided in Algorithm 1: in case of a concatenation process (that takes as inputs two strings and outputs their concatenation), then all the trust value equally contribute to determining the outcome and therefore they are merged by conjunction. In case of a lookup process (that takes as inputs a key and a value table, and outputs the value in the table corresponding to the key), then before calculating the conjunction of the trust values, we project them into the space of the possible values, possibly smaller than the space of plausible ones. Moreover, in case the value we face does not fall into the range of possible values, then we consider the value or values closer to it and belonging to the sset of possible values. Clearly, we weight these contributions according to the distance to the given value.

V. Applying the algorithm

We now discuss how, by taking advantage of both provenance and background knowledge, the trust algorithm can produce more precise trust ratings.

One important feature of the algorithm is that, by means of provenance, we encorporate in our algorithm also semantic information.

This way, we restrict the domain of possible value for each field to the range of real, meaningful values. For instance, if the nationality field of a MMSI is a 3 digit code, then there are 10^3 possible values, since any cypher would be equally probable in each of the 3 positions. By taking into account the meaning (semantics) of the MMSI, the cardinality of the set of the plausible values would restrict to 35 (considering the countries which own 99% of the ships). This means that if we own 10 positive evidence and we restrict the plausibilty set from 1000 to 35, then the trust value rises from $E = \frac{10}{1010} + \frac{1}{1000} \times \frac{1000}{1010} = 0,0189...$ to $E = \frac{10}{45} + \frac{1}{35} \times \frac{35}{45} = 0,3143.....$ Note that the MMSI field is retrieved via traversing the provenance graph.

Another important feature of the algorithm is the usage of provenance information. Because of this, we enlarge the availability of evidence at disposal for calculating trust values. In fact, we don't limit to the use of direct evidence about the facets we have to evaluate, but we consider also evidence about elements used in the process that lead us to our facets. Therefore, we check whether these initial elements were correct and whether they were combined properly in order to produce the facet we are analyzing. Once we have this result, we can compare it with evidence directly referred to the facet we are evaluating, obtaining an improvement of the precision of the trust value.

Continuing the previous example, if we have also sources that provide a value for the nation, knowing that the national code is determined by looking it up into a trusted table, then by applying the Trust Ranking Algorithm, we obtain the following trust value: $E = \frac{20}{45} + \frac{1}{35} \times \frac{35}{45} = 0,4667.....$

If we adopt a conservative approach and accept only facets which trust value is above a certain threshold, then this change reduces the amount of errors due to false negatives.

VI. Related Work

Trust is a widely explored topic within a variety of areas within computer science including security, intelligent agents, software engineering and distributed systems. Here, we focus on those works directly touching upon the junction of trust, provenance and the Semantic Web. For a readable overview

of trust research in artificial intelligence, we refer readers to Sabater and Sierra [12]. For a more specialized review of trust research as it pertains to the Semantic Web see Artz and Gil [13]. Finally, Golbeck provides a longer review of trust research as it relates to the Web [14].

Our work is closest to the work on using provenance for information quality assessment on the Semantic Web. In the WIQA framework [15], policies can be expressed to determine whether to trust a given information item based on both provenance and background information expressed as Named Graphs [16]. Hartig and Zhao follow a similar approach using annotated provenance graphs for a given information item to perform the quality assessment and thus generate a trust value [17]. However, their work uses a more complex provenance representation similar to OPM that captures not only the data origins but also the processing steps involved. Similarly, IWTrust generates trust values for answers produced by a question answering system based on a combination of source data, provenance information, and user ratings [18]. Our work differs from these approaches in three respects: First, we concentrate on event descriptions and not generic data items. Second, our work takes advantage of a priori knowledge about the likelihood of data items in order to correct for possible data errors. Finally, we use Subjective Logic to allow for multiple (possibly conflicting) opinions about data sources to be taken into account, but unlike [19], we use it in combination with provenance.

Recent work has focused on querying trust using SPARQL [20]. We see our work as complementary in that it could facilitate the population of the trust values to query over. Finally, other work has considered using provenance and ontologies to determine the trust in electronic contracts [21]. Our work differs, in that they use provenance as a source of experience for calculating opinion values whereas we focus on the combination of current opinion values to produce a final trust value.

VII. Conclusion

Here, we presented a trust algorithm for determining trust of event descriptions based on provenance. We provide a novel mapping between an event ontology and a widely used provenance ontology. Secondly, we show how Subjective Logic can be used in combination with provenance to generate improved trust values in a maritime data integration domain. In the future, we will perform a comprehensive evaluation of the model and extend Subjective Logic to handle a contextualization of opinions and address some of its limitations (see [22]). Additionally, we aim to expand our work applying it to the problem of determining trust of event descriptions produced from data integrated from the Web.

References

[1] A. Harati-mokhtari, A. Wall, P. Brooks, and J. Wang, "Automatic identification system (ais): A human factors approach," 2008. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=?doi=10.1.1.127.1049

[2] W. R. van Hage, V. Malaisé, G. de Vries, G. Schreiber, and M. van Someren, "Combining ship trajectories and semantics with the simple event model (sem)," in *EiMM '09: Proceedings of the 1st ACM international workshop on Events in multimedia*. New York, NY, USA: ACM, 2009, pp. 73–80.

[3] N. Willems, W. R. van Hage, G. de Vries, J. Janssens, and V. Malaisé, "An integrated approach for visual analysis of a multi-source moving objects knowledge base," *IJGIS (to appear)*, 2010.

[4] W. R. van Hage, V. Malaisé, G. de Vries, and A. T. Schreiber, "Abstracting and reasoning over ship trajectories and web data with the simple event model (sem)," *MTAP (to appear)*, 2010.

[5] S. Bechhofer and A. Miles, "SKOS Simple Knowledge Organization System Reference," W3C, W3C Recommendation, Aug. 2009.

[6] L. Moreau, B. Clifford, J. Freire, J. Futrelle, Y. Gil, P. Groth, N. Kwasnikowska, S. Miles, P. Missier, and J. Myers, "The Open Provenance Model core specification (v1.1)," *Future Generation Computer Systems*, Jul. 2010. [Online]. Available: http://dx.doi.org/10.1016/j.future.2010.07.005

[7] S. S. Sahoo, A. Sheth, and C. Henson, "Semantic provenance for eScience: Managing the deluge of scientific data," *IEEE Internet Computing*, vol. 12, pp. 46–54, 2008. [Online]. Available: http://doi.ieeecomputersociety.org/10.1109/MIC.2008.86

[8] S. Sahoo, P. Groth, O. Hartig, S. Miles, S. Coppens, J. Myers, Y. Gil, L. Moreau, J. Zhao, M. Panzer, and D. Garijo, "Provenance Vocabulary Mappings," 20010. [Online]. Available: http://www.w3.org/2005/Incubator/prov/wiki/Provenance_Vocabulary_Mappings

[9] A. Jøsang, "Probabilistic logic under uncertainty," in *Theory of Computing 2007. Proceedings of the Thirteenth Computing: The Australasian Theory Symposium (CATS2007). January 30 - Febuary 2, 2007, Ballarat, Victoria, Australia, Proceedings*, ser. CRPIT, J. Gudmundsson and C. B. Jay, Eds., vol. 65. Australian Computer Society, 2007, pp. 101–110.

[10] Wikipedia. (2010, Jun.) Beta distribution. [Online]. Available: http://en.wikipedia.org/wiki/Beta_distribution

[11] A. Jøsang and D. McAnally, "Multiplication and comultiplication of beliefs," *Int. J. Approx. Reasoning*, vol. 38, no. 1, pp. 19–51, 2005.

[12] J. Sabater and C. Sierra, "Review on Computational Trust and Reputation Models," *Artificial Intelligence Review*, vol. 24, no. 1, p. 33, 2005. [Online]. Available: http://portal.acm.org/citation.cfm?id=1057866

[13] D. Artz and Y. Gil, "A survey of trust in computer science and the Semantic Web," *J. Web Sem.*, vol. 5, no. 2, pp. 58–71, 2007. [Online]. Available: http://www.isi.edu/~gil/papers/artz-gil-jws07.pdf

[14] J. Golbeck, "Trust on the world wide web:a survey," *Foundations and Trends in Web Science*, vol. 1, no. 2, pp. 131–197, 2006. [Online]. Available: http://portal.acm.org/citation.cfm?id=1373449

[15] C. Bizer and R. Cyganiak, "Quality-driven information filtering using the WIQA policy framework," *Web Semantics: Science, Services and Agents on the World Wide Web*, vol. 7, no. 1, pp. 1–10, Jan. 2009.

[16] J. J. Carroll, C. Bizer, P. Hayes, and P. Stickler, "Named graphs, provenance and trust," *International World Wide Web Conference*, 2005. [Online]. Available: http://portal.acm.org/citation.cfm?id=1060745.1060835

[17] H. Olaf and J. Zhao, "Using Web Data Provenance for Quality Assessment," in *Proceedings of the 1st Int. Workshop on the Role of Semantic Web in Provenance Management (SWPM) at ISWC*, Washington, USA, 2009.

[18] I. Zaihrayeu, P. Pinheiro˜da Silva, and D. L. McGuinness, "IWTrust: Improving User Trust in Answers from the Web," in *Proceedings of 3rd International Conference on Trust Management (iTrust2005)*. Paris, France: Springer, 2005, pp. 384–392.

[19] D. Ceolin, W. R. van Hage, and W. Fokkink, "A trust model to estimate the quality of annotations using the web," in *WebSci10: Extending the Frontiers of Society On-Line*, 2010. [Online]. Available: http://journal.webscience.org/315/

[20] O. Hartig, "Querying Trust in RDF Data with tSPARQL," in *Proceedings of the 6th European Semantic Web Conference (ESWC)*, Heraklion, Greece.

[21] P. Groth, S. Miles, S. Modgil, N. Oren, M. Luck, and Y. Gil, "Determining the Trustworthiness of New Electronic Contracts," in *Proceedings of the Tenth Annual Workshop on Engineering Societies in the Agents' World, (ESAW-09)*, Utrecht, The Netherlands, 2009.

[22] R. Guha, R. Kumar, P. Raghavan, and A. Tomkins, "Propagation of trust and distrust." ACM Press, 2004, pp. 403–412.