The 5th International Workshop on Information Quality and Quality of Service for Pervasive Computing 2013, San Diego (22 March 2013)

TAF: A Trust Assessment Framework for Inferencing with Uncertain Streaming Information

Anthony Etuk and Timothy J. Norman
Department of Computing Science
University of Aberdeen
Scotland, UK
{aaetuk,t.j.norman}@abdn.ac.uk

Chatschik Bisdikian and Mudhakar Srivatsa
IBM Research
Thomas J. Watson Research Center
Yorktown Heights, USA
{bisdik,msrivats}@us.ibm.com

***Abstract*—Pervasive information consumers in open, loosely-coupled systems, such as in Internet of Things and crowd-sensing environment, will rely more and more often on streaming information from sensory sources with whom they have only ephemeral, transient relationships. In such settings, information uncertainties arise as the trustworthiness of the sources and their information become questionable. It is thus necessary to quantify the quality of inferences made with such information to aid more informed and effective decision making and action taking. One of the aspects of trust assessment systems is to provide for such quality metrics, however, these systems have been traditionally applied in static situations. In this paper, we introduce TAF, a trust assessment framework for streaming information that leverages the rich toolkit of subjective logic operators to estimate the quality of said inferences under information uncertainty. We present the system architecture, describe its components and provide some preliminary quality results for the framework.**

I. Introduction

Pervasive sensory systems are deployed to collect information whose analysis, processing, and fusion with other (background or not) information provides awareness to situations of interest. Expressed in the form of *hypotheses*, these situations of interest may range from simple (e.g., is it snowing?) to elaborate (e.g., are terrorist groups colluding to launch coordinated attacks at multiple locations?), and anything in between (e.g., does patient A show signs of deteriorating health, or is there a friend from the high-school years in the vicinity?). Making use of the information collected, *analysts* (humans or software agents) conclude (i.e., make *inferences*) in favor or against these hypotheses and drive subsequent decisions and actions in response to these inferences.

In order to make more informed decisions and take more prudent actions, the *quality of the inferences* made needs to be assessed and communicated appropriately. Specifically, based on the *quality of information* (QoI) attributes [1] of the information collected there is a need to assess the level of *trust* that can be placed on the inferences made using this information so that subsequent decisions are reflective of the uncertainties that may arise. The necessity of such inference assessments will further increase as the information collection and processing systems become more open and loosely-coupled, such as in Internet of Things (IoT), participatory-sensing, and other smart pervasive applications [2], [3], [4].

In this paper we further our initial work in [4], where we had defined trust as representing the degree of belief that an information consumer has that she can rely on the information that a provider has provided her; we also touched upon the concept of *quality of inference* (QoInf) [5] . Specifically, in this paper, using the former definition as a point of departure, we consider the QoInf of the inferences that could be made by an analyst analyzing the incoming *streams* of information reported by multiple sources of varying degrees of trust.

Investigating trust assessment over streaming data introduces new challenges and runtime design options (such as trade-offs between quality and latency) not encountered when dealing with trust assessment schemes typically considered in the literature (see next Section), which primarily deal with static and quasi-static data. In assessing trust with streaming information, the contributions in this paper are:

- The introduction of TAF, a trust assessment framework for streaming information with uncertainties;
- The use in TAF of subjective logic opinions and operators as measures of trust and QoInf based on the incoming information; and
- An evaluation of related design trade-offs for a set of hypotheses, source, and information report models.

This is an early work on this novel subject and the above contributions reflect initial developments in the related topics establishing the research area and presenting some preliminary results. We note here that in our current work, we leverage the power of subjective logic to represent trust and compute QoInf. We will discuss subjective logic more later in the paper.

The paper is organized as follows: Section II highlights related work in the area. Section III highlights subjective logic and pertinent terminology. Section IV introduces TAF, our trust assessment framework, while Section VI presents some preliminary performance results. Finally, we conclude in Section VII with concluding remarks.

Research was sponsored by the U.S. Army Research Laboratory and the U.K. Ministry of Defence and was accomplished under Agreement Number W911NF-06-3-0001. The views and conclusions contained in this document are those of the author(s) and should not be interpreted as representing the official policies of the U.S. Army Research Laboratory, the U.S. Government, the U.K. Ministry of Defence or the U.K. Government. The U.S. and U.K. Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

978-1-4244-9529-0/13/$31.00 ©2013 IEEE

II. Related Work

There are several studies regarding trust computation models on static data such as the eBay recommendation system [6], Netflix movie ratings [7], EigenTrust [8], PeerTrust [9], etc. In [10], a non-streaming truth finder system attempts to assess the credibility of reported facts in social sensing applications, by applying a maximum likelihood estimator to information received from multiple sources. The real-time nature of Twitter and how sensory information (tweets) from multiple users (sensors) could be used to detect events is investigated in [11] but there is no consideration of trustworthiness. Trustworthiness of streaming data is considered in [12] but the emphasis is on trusting the data based on consistency expectations for the purpose of enforcing confidence policies in data-stream management systems with no consideration of subsequent inferences.

III. Subjective Logic Primer

We start with a brief primer on *subjective logic* (SL) to introduce the necessary terminology that is used later in the paper. For more details about SL, see [13].

Reports received from multiple sources can be unreliable, making it hard to extract precise information about the observed phenomena. This introduces uncertainty in inferencing which may impact subsequent decisions and, hence, needs to be addressed in a principled manner. Subjective logic is a form of probabilistic logic, extending it to explicitly account for uncertainty in *opinions* formed based on using evidences (the report content) received from sources. Specifically, an opinion ω_h^x regarding a proposition (i.e., a hypothesis) h is a statement about the degree of belief (b_h^x), disbelief (d_h^x), and uncertainty (u_h^x) about the validity of h held by the opinion *owner* x, where $b_h^x + d_h^x + u_h^x = 1$. In the special case of binary hypotheses spaces (see next section), where there is concern only with whether a hypotheses h is valid or not, we have corresponding binomial opinions which can be represented in subjective logic by the quadruple of nonnegative elements b, d, u, a:

$$\omega = \big(b, d, u, a\big), \text{ where: } b + d + u = 1 \text{ and } a \in [0, 1]. \quad (1)$$

where we drop the x and h from the notation when these are implied by the context. Note that disbelief represents the level of belief in the negation of h and uncertainty represents the level of ignorance regarding the validity of h. Finally, a, called the *base rate*, represents the *a priori* probability about the validity of h in the absence of any evidence.

SL provides an intuitive way to represent the belief an entity has in another, and a way to aggregate evidence to support such beliefs. Opinions about a hypothesis can be mapped to a *beta distribution* representing the distribution of the probability of the hypothesis to be true based on evidences observed in support of or against the hypothesis. SL provides a rich set of operators for combining evidences and deriving corresponding opinions including `discounting`, `consensus`, and `conjunction`. For example, the `discounting` operator receives as inputs: (a) the opinion ω_e^y of agent y about a piece of evidence e (a report in our case), and (b) the opinion ω_y^x expressing the level of trust that agent x has for agent y. It outputs $\omega_e^{x:y} = \omega_y^x \otimes \omega_e^y$ expressing the opinion that agent x develops regarding evidence e through his association with agent y, where

$$\begin{aligned} b_e^{x:y} &= b_y^x\, b_e^y, & d_e^{x:y} &= b_y^x\, t d_e^y, \\ u_e^{x:y} &= d_y^x + u_y^x + b_y^x\, u_e^y, & a_e^{x:y} &= a_e^y. \end{aligned} \quad (2)$$

The `consensus` operator fuses (independent) opinions about a hypothesis, $\omega_h^{x \diamond y} = \omega_h^x \oplus \omega_h^y$, where ($\kappa = u_h^x + u_h^y - u_h^x\, u_h^y$):

$$\begin{aligned} b_h^{x \diamond y} &= \frac{b_h^x\, u_h^y + b_h^y\, u_h^x}{\kappa}, & d_h^{x \diamond y} &= \frac{d_h^x\, u_h^y + d_h^y\, u_h^x}{\kappa}, \\ u_h^{x \diamond y} &= \frac{u_h^x\, u_h^y}{\kappa}, & a_h^{x \diamond y} &= \frac{a_h^x\, u_h^y + a_h^y\, u_h^x - (a_h^x + a_h^y) u_h^x\, u_h^y}{\kappa - u_h^x\, u_h^y}, \end{aligned} \quad (3)$$

We will use the `discounting` and `consensus` operators later on in TAF for normalizing incoming reports; for more on SL operators see [13].

IV. System Model

Fig. 1 summarizes the general set-up under consideration in this paper. It comprises an entity of interest (abstractly noted as a *region* in the figure) for which an observer (the *analyst*) wants to know (i.e., *infer*) if situations of concern (the *hypotheses*) have occurred, and assess the quality of this inference (QoInf). To aid her inference, the analyst collects information *reports* that are streamed from multiple sensory sources in the field capable of observing aspects of the situation of interest. The sources may be owned by a number of different organizations, and, as a result, the analyst treats reports received from them with different levels of trust.

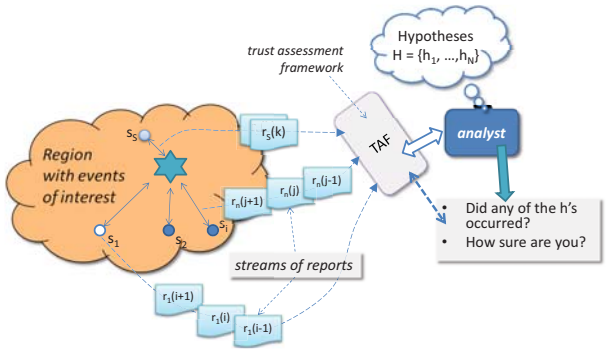

Fig. 1. The system model.

The purpose of TAF, our trust assessment framework, is to aid the analyst's inference process and the assessment of QoInf. In the rest of this section, we present the key aspects of the set-up in the figure, and then TAF will be presented in the next section.

A. The hypothesis model

The analyst collects sensory information to ascertain the validity of a hypothesis $h \in \mathcal{H}$, e.g., $h =$ "snowfall at location l, at time t." We assume that h can be described by a finite set of sub-events (also referred to as *event features*) $\{f_1^e, f_2^e, \ldots, f_n^e\}$, where the f_i^e's are assumed as perceivable and measurable. For

example, a `snowfall` event may be described by sub-events such as `temperature`, `precipitation`, etc. Information about such features describing an event can be derived from a knowledge base. Generally, the various hypotheses in $\mathcal{H}$ could be regarded as a set of (possibly) contending hypotheses, with overlapping sets of features. In this paper, we consider a simple binary hypothesis set $\mathcal{H}$ comprising a hypothesis h and its negation (the *null hypothesis*).

B. The report model

As previously stated, an analyst in her assessment of the validity of a hypothesis h is aided by a stream of reports $r_1, r_2, r_3, \ldots$ coming from a collection of sources $\mathcal{S}$. A report is annotated with attributes, including *time*, *location*, *event*, and *source*, that aid in associating it with a specific hypothesis. Furthermore, a report from a source s is assumed to contain the source's opinion (as defined in SL) $\omega^s_{f^e_i}$ about the event feature f^e_i of h.

C. The source model

The analyst makes inferences about hypotheses based on reports that she receives from a set $\mathcal{S}$ of sources having varying degrees of trustworthiness. Information sources are described by a set of observable features $\{f^s_1, f^s_2, \ldots, f^s_l\}$, which could be used to infer their pattern of behavior. We take as a working assumption, that there may be some correlation between source features and source behaviors. For instance, features such as *ownership*, *expertise*, *location*, etc., may influence the way a source behaves in a particular context. Here we take source behavior to mean the kind of report or opinion they provide.

With TAF, we exploit such correlations to enhance our situational assessment. In particular, TAF employs the *diversity model* described in [14] to stratify the sources into a set $\mathcal{G}$ of different groups according to their perceived similarity in features *and* behavior. The diversity model defines a function that maps all known sources in the system to a set of groups by exploiting their features and past reports. The analyst x maintains an opinion about the reliability of each group $g \in \mathcal{G}$, and uses this to determine the confidence placed on reports obtained from members belonging to the different groups. For example, the analyst may have learned over time that sources such as sensors with feature `battery:Low` typically provide unreliable opinions about an event, and therefore groups all such sources, and subsequently exploits this model in future encounters with members of the group. For more details on group formation based on the diversifying of sources, see [14].

1) The source reporting behavior: It is possible that sources may report opinions about events they have observed differently from their truly perceived opinion with the intention to mislead, or report erroneously due to a partial or imperfect knowledge of the environment. The ability of the analyst to minimize the effect of unreliable reports on her formed opinion about situations of concern would greatly enhance the confidence in the system. For instance, the rate of false alarms or undetected events could be greatly minimized when unreliable reports do not influence the opinion formed about a hypothesis h. We will assume that sources are either *reliable* providing truthful opinions about h, or *unreliable* providing untruthful opinions instead. It is assumed that sources belonging to the same group generally tend to report in a similar way. Therefore, for a group that has been identified as reliable, its members would also be expected to provide reliable reports. However, it is also assumed that group behavior is not 100% applicable to all sources in the group, which introduces some further uncertainty in the opinions maintained by the analyst about groups.

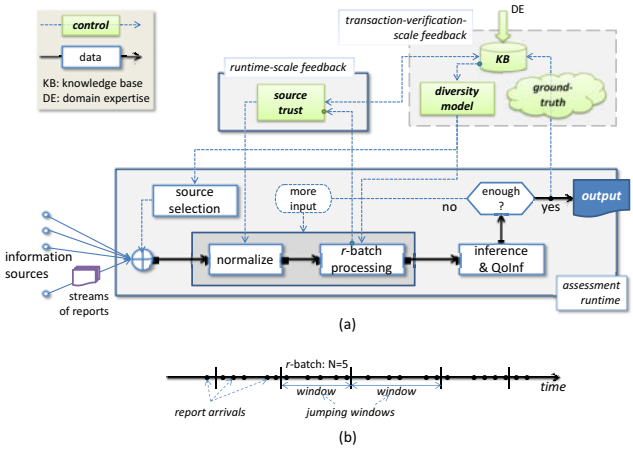

Fig. 2. (a) Trust Assessment Framework (TAF); and (b) Jumping window operation.

V. Trust Assessment Framework

The *trust assessment framework* (TAF) aids the analyst in managing and processing the incoming streams of reports to infer the validity of a hypothesis of concern as well as qualify this inference (QoInf). TAF makes use of SL opinions and operators to attain these.

Fig. 2(a) highlights the components of TAF. It comprises a report processing part that outputs the desired inferences; this is the lower part of the framework noted as the *assessment runtime*. It also comprises a collection of management components that adjust the operation of the assessment runtime based on the reports received–the solid-line block noted as *runtime-scale feedback*–and the correctness of the inferences made–the dashed-line block noted as *transaction-verification-scale feedback*. These two management components operate at different time scales, a faster one that reacts to reports collected at runtime to adjust to operational assumptions made at runtime and a slower one that reacts to operational assumptions that persist at large (including the assumed models themselves).

More specifically, the framework operates on the reports arriving from various sources as well as background knowledge (noted in the figure as *knowledge base* KB) about the assessed situation to form an opinion about the state of world. The formed opinion is constantly updated as more evidence is received about the assessed situation, until some stoppage condition is met. In TAF, incoming reports are first filtered out if they come from highly untrusted sources as may be

indicated by the operational state of the source diversity model (see Section IV-C) currently in effect. Subsequently, each admitted report is normalized (SL-discounting) based on current knowledge about the reporting source, as well as knowledge about the domain. Assessment can be made immediately with an input report or done in batches. These two alternatives are noted in the figure as r-*batch processing*, where an r-batch represents the collection of reports processed at each stage of the TAF inference making operations; an r-batch of size $N = 1$ implies inferences are updated after each report.

In the batch mode, a number N of input reports is first accumulated and assessment is made once on the accumulated reports. We are currently considering r-batches formed by (non-overlapping) *jumping windows* containing a fixed size N of reports, see Fig. 2(b) where $N = 5$. During r-batch processing, the reports within an r-batch are correlated with each other in an attempt to assess the situation of interest. We assume here that reports about the same physical property should be highly correlated with each other. When this is not the case, we may either infer a change in the underlying process observed, or the presence of unreliable sources in the set. However, we also recognize that sources may be correlated in some manner based on their features. In this case, the outcome of the information fusion may be wrongly skewed, and therefore not reflect the ground truth. As mentioned in Section IV-C, TAF tackles this issue by exploiting diversity modeling, which stratifies sources in groups and reduces the effect of overrated evidence supplied by similarly behaving groups of sources by using only a subset of the sources thusly identified. Finally, should the QoInf made using the evidence in an r-batch is not sufficient, processing of additional input may be required.

Next we describe how beliefs about hypotheses are updated based on received streams of reports.

A. The inferencing process

Dealing with streaming information entails inferring which of the hypotheses is active and updating the beliefs in them continuously. Hence, in her attempt to infer the prevailing hypothesis, an analyst will persistently be concerned with either basing her inferences on the reports received thus far or deferring the inferencing until additional evidences are accumulated in hope of improving QoInf.

The inference mechanism in TAF comprises two stages: (a) identifying when something of interest could have occurred, signifying the possible change of the prevailing hypothesis; and (b) determining the validity of such a change (the "inference") and subsequently computing its QoInf. Due to space limitations, we elaborate only the second stage. Regarding the first stage, we state here that TAF currently uses the information-theoretic *Kullback-Leibler* (KL) distance to determine whether the probability distribution of evidences in favor or against the hypothesis gathered from a collection of reports in one r-batch is significantly different from this distribution in the following r-batch.

By identifying a possible change of the underlying system dynamics, inference and trust assessment procedures could be adjusted to reflect the current situation thereby avoiding the use of stale information in these cases. Furthermore, such identification may cause *feedback* in the system, where for instance, the trust scores of reporting sources could be updated based on their reports in the last processed r-batch, see runtime-scale feedback in Fig. 2(a). However, due to the fact that there is plenty of uncertainty in the system, e.g., sources in an r-batch may report erroneously, we may employ additional steps in the process. For example, in the process of computing the KL distance in a window, we may also consider the trustworthiness of the sources reporting in the window.

We next describe the process of trust computation over r-batches determined by the aforementioned jumping windows, see Fig. 2(b), and the dynamic update of belief based on reported opinions.

B. Trust computation over jumping windows

To fuse opinions from different sources and make inferences based on the reports received, we use r-batches defined over non-overlapping jumping windows containing N successive reports from the incoming stream. For each such r-batch, TAF uses the source diversity model and SL opinion operators to conduct fusion of opinions based on the profiles of sources for which reports have been received in the window.

Specifically, the reports from the sources are first partitioned into different groups corresponding to the identified profiles of the sources. Then the SL `consensus` operator is applied to each of the groups in order to derive a consensus opinion for the reports in each group. Then, the SL `discount` operator is applied to the result obtained in each group to normalize the derived group opinion. The normalized opinions from all groups are then combined with each other using again the SL `consensus` operator to obtain the overall opinion from all the reports in the current window. This fusion approach minimizes the adverse effect of unreliable sources that may happen to be present in the majority in a given window.

One way of reaching an early estimate of the possibly prevailing hypothesis would be to exploit the features of the diversity model. Assuming momentarily that all the sources in a specific group were to maintain the same opinion over different windows with very little uncertainty, we could easily converge upon (i.e., reach) an inference earlier by exploiting domain knowledge about the distribution of the sources in the system. However, we recognize the fact that not all the sources in a specific group might maintain the same opinion as their group members. A possible cause of this might be in the accuracy of the underlying diversity model. To this end, we apply the following steps. After deriving the consensus in a specific group, we compare the belief b and disbelief d components of the derived opinion to find out what the group's view is about the situation. For instance, if $b > d$, then the group would be considered to maintain a belief in a hypothesis. Conversely, if $d > b$ the group would be taken to maintain a disbelief in this hypothesis. Based on this observation, the analyst uses a different pa-

rameter $\omega_{g,w}$ to associate the level of uncertainty in the *belief* of a group. This in a sense is a measure of the degree of conformity to the group opinion by its members. If the uncertainty is very high, then this might be an indication to recompute the underlying diversity model. However, if the uncertainty is observed to be very low in successive windows, the analyst could use this to inform its decision on the accuracy of its estimate. For example, assuming $b > d$ in the derived group's opinion, $\omega_{g,w} = (b_{g,w}, d_{g,w}, u_{g,w})$ is obtained as follows:

- $b_{g,w}$: percentage of reports having $b > d$;
- $u_{g,w}$: otherwise, i.e., $u_{g,w} = 1 - b_{g,w}$.

As we are only concerned with how members of a group conform to the group's opinion, $d_{g,w}$ is set to zero.

It should be emphasized here that the $\omega_{g,w}$ does not reflect the belief an analyst has in a hypothesis. It only serves to estimate the uncertainty involved in developing this belief based on the sources seen reporting thus far. If the uncertainty is sufficiently low between different windows, then this could serve as a good estimate for predictions on the expectation of reports yet to arrive in the stream.

VI. Numerical results

Our evaluation focused on the effectiveness of TAF in guiding the analyst in the decision making process. In particular, we measure the robustness of the framework in the presence of reports from unreliable sources who coordinated in reporting. We also measure the trade-offs in the use of different window sizes in the computation of trust and making inferences about different hypotheses.

A. The simulation setup

Our simulation setup consists of information sources that provide reports to the analyst in response to an event occurring at a random time. We consider the binary hypothesis space of whether or not the event occurred. The diversity model learns the best way to stratify sources, such that sources reporting in a similar manner are grouped together. The diversity model is instantiated in an offline mode, but used in real time in TAF. Sources provide reports based on the group or profile they are identified with. The reports from each profile are modeled after a truncated gaussian distribution with a *reliability* parameter that specifies its mean, and a *std* parameter that specifies its standard deviation. Based on this parameter, the SL opinion which serves as a source's opinion about an event is formed. Each profile is also assigned a *confProb* parameter that specifies the conformity probability of a source to its group. The system uses this parameter to select sources that do not conform to their profile behavior; these sources report randomly, without any pattern whatsoever. The analyst maintains *profOpinion*, a personal opinion about the different profiles, which is used as an a priori source trust to normalize the reports from sources in each of the profiles, and updated over the system's lifetime. The system at each point maintains the percentage of reliable and unreliable sources, and it does this by assigning different proportions to the different profiles, based on their reliability level. The trust assessment framework works in a window-based manner, and the parameter *windowSize* defines the number of reports the system operates on at any given time (i.e., the size of an r-batch). The *eventThreshold* is defined and used by the KL distance algorithm to prompt the system about underlying system evolution. The parameter list with their default values is shown in Table I.

TABLE I
Experimental Parameters

Parameter	Test values
Number of information sources	1000
Fraction of unreliable sources (%)	10, 20, …, 100
Number of profiles	3 profiles: p_1, p_2, & p_3
Profiles: reliability (*reliability*)	$p_1 = 0.9$, $p_2 = 0.2$, $p_3 = 0.9$
Profiles: standard deviation (*std*)	0.05
Profiles: conformity (*confProb*)	0.9
Window size (*windowSize*)	1, 10, 20, …, 100

B. Robustness

We evaluated the robustness of the fusion scheme in the presence of malicious sources. As indicated in Table I, we ran test cases with the proportion of malicious sources increasing from 10 to 100, and in each case we kept the window size for which assessment is made fixed. We tested the robustness of the fusion scheme by first normalizing reports received in a given window by the trust of the reporting sources, before carrying out fusion of the collective opinions, and updating the belief of the analyst. We call this approach the *trustOnly* approach. We compared the *trustOnly* approach to diversity modeling, which takes the diversity among the sources reporting in the stream into consideration before carrying out fusion. The *diversity model* approach first carries out a local fusion based on identified profiles of the reporting sources before combining the (normalized) outcomes from the different groups. The result shown in Fig. 3 illustrates the performance of the different approaches with varying degrees of unreliable sources. The belief and uncertainty of opinion formed by the analyst is reported in each case; note $b+u < 1$.

At first glance, Fig. 3 seems to imply that the *trustOnly* approach outperforms the *diversity model* approach when the number of malicious sources are small. This is the case, because *trustOnly* assumes independence among the reporting sources, and therefore quickly converges on its uncertain parameter when more reports are received. However, this approach quickly falls over as more organized unreliable reports are provided by the sources, which misleads the analyst into settling towards the opposing hypothesis, when that is actually false. However, the *diversity model* manages this gracefully, because it uses the correlation among the sources to guide its inference. When the proportion of unreliable sources increases, instead of misleading the analyst, the approach reflects this anomaly in the high level of uncertainty attached to its opinion. This is an important and useful indicator, since instead of taking a contrary action, the system could be driven to seek further evidence to reduce the uncertainty. One way of achieving this could be by adjusting the bound of the window in real time to allow more reports to be assessed.

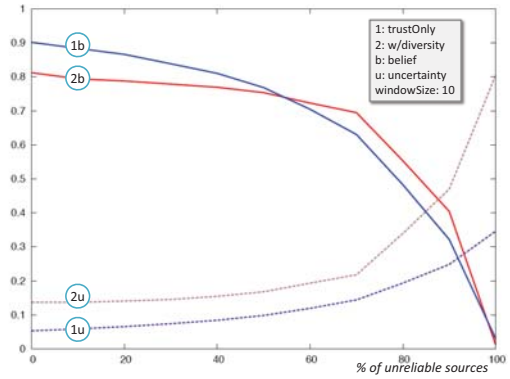

Fig. 3. Quality of inference ($b\&\bar{c}u$) with increasingly unreliable sources.

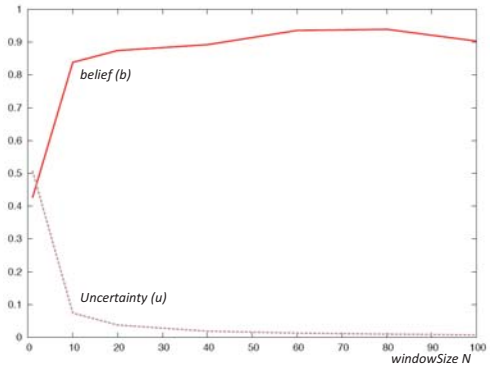

Fig. 4. Quality of inference ($b\&\bar{c}u$) with varying report window size.

C. Design trade-off

Next we studied the behavior of the system with different *windowSizes* N used in the assessment and the initial results are shown in Fig. 4. In each case the belief and uncertainty maintained in an inference is shown. In the figure, there is a clear distinction in the accuracy of estimate when the *windowSize* parameter is less than 15. This could be explained by the fact of the smaller number of pieces of evidence (reports) available, making it challenging for the inference engine to effectively make a good assessment of the underlying situation. This changes as the window size increases, and more reports are available to disambiguate the uncertainty in the environment. However, opting for large window sizes may still cause poor performance in cases where the underlying environment changes rapidly, thus, possibly having the system acting with stale information.

VII. Concluding Remarks

Trust assessment is expected to become an important component in open, loosely-coupled pervasive environments such as the Internet of Things, machine-to-machine systems, and crowd-sensing. The information needed by these dynamically evolving systems will be streamed by a number of sensory sources of questionable and uncertain trustworthiness. As a result, inferences that are made using such information will be of questionable and uncertain quality as well. For these environments, we have proposed TAF, a trust assessment framework, which, unlike traditional trust assessment systems, operates on dynamic, multi-class and multi-attribute source systems that stream uncertain information to potential consumers.

The framework exploits the subjective logic toolkit of operators to process opinionated reports about observed events that give credence in favor or against situations (hypotheses) of interest and quantify the quality of inference that can be made. Instances of TAF also make use of the diversity model of sources to discriminate sources in groups based on their features and behavior. It then discounts and filters-out seemingly unreliable or colluding sources to improve the quality of derived inferences.

This is an early work in the area. Our study so far has identified the key components of the trust assessment framework and introduced the subjective-logic-based toolkit to handle the computational aspects of trust assessment with uncertain streaming information. In the course of our study, we have also identified a number of issues that will need further investigation. These include more extended performance analysis, the consideration of dynamically adapting window sizes at runtime, the study of the performance trade-offs with respect to the windowing process, the system sensitivity to the accuracy of the underlying models used, investigating various stopping criteria to stop processing incoming reports and inferring instead with regard to a hypothesis, and so on.

References

[1] C. Bisdikian, L. M. Kaplan, M. B. Srivastava, D. J. Thornley, D. Verma, and R. I. Young, "Building principles for a quality of information specification for sensor information," in *12th Intl Conf. on Information Fusion (FUSION'09)*, Seattle, WA, USA, July 2009.

[2] N. Gershenfeld, R. Krikorian, and D. Cohen, "The Internet of Things," *Scientific American*, vol. 291, no. 4, pp. 76–81, Oct. 2004.

[3] J. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, and M. B. Srivastava, "Participatory sensing," in *World Sensor Web Wksp (in ACM Sensys'06)*, Boulder, CO, USA, October 31, 2006.

[4] C. Bisdikian, M. Şensoy, T. J. Norman, and M. B. Srivastava, "Trust and obfuscation principles for quality of information in emerging pervasive environments," in *IEEE PerCom IQ2S Wksp*, Lugano, Switzerland, Mar. 2012.

[5] N. Roy, A. Misra, S. K. Das, and C. Julien, "Quality-of-inference (QoINF)-aware context determination in assisted living environments," in *1st ACM Int'l Wksp on Medical-Grade Wireless Networks (WiMD'09)*, New Orleans, LA, USA, May 18, 2009.

[6] J. B. Schafer, J. Konstan, and J. Riedl, "Recommender systems in e-commerce," in *ACM Conference on Electronic Commerce*, 1999.

[7] Netflix, "Netflix prize," http://www.netflixprize.com/.

[8] S. Kamvar, M. Schlosser, and H. Garcia-Molina, "Eigentrust: Reputation management in p2p networks," in *WWW Conference*, 2003.

[9] L. Xiong and L. Liu, "Supporting reputation based trust in peer-to-peer communities," *IEEE Trans. on Knowledge and Data Eng.*, July 2004.

[10] D. Wang, L. Kaplan, H. Le, and T. Abdelzaher, "On truth discovery in social sensing: A maximum likelihood estimation approach," in *11th ACM Int'l Conf. on Information Processing in Sensor Networks*, 2012.

[11] T. Sakaki, M. Okazaki, and Y. Matsuo, "Earthquake shakes twitter users: real-time event detection by social sensors," in *19th ACM Int'l Conf. on World Wide Web*, 2010.

[12] H. S. Lim, Y. S. Moon, and E. Bertino, "Assessing the trustworthiness of streaming data," Tech. Rep. TR 2010-09, CERIAS, Tech. Rep., 2010.

[13] A. Jøsang, "Subjective logic," (draft book), http://folk.uio.no/josang/papers/subjective_logic.pdf.

[14] A. Etuk, T. J. Norman, and M. Şensoy, "Reputation-based trust evaluations through diversity," in *15th AAMAS TRUST Wksp*, June 2012.