TIDY: A Trust-Based Approach to Information Fusion through Diversity

Anthony Etuk\textsuperscript{1}, Timothy J. Norman\textsuperscript{1}, Murat Şensoy\textsuperscript{1,3}, Chatschik Bisdikian\textsuperscript{2}, and Mudhakar Srivatsa\textsuperscript{2}

\textsuperscript{1}Computing Science Department, University of Aberdeen, UK
\textsuperscript{2}IBM T. J. Watson Research Center, NY, US
\textsuperscript{3}Computer Science, Ozyegin University, Istanbul, Turkey
\{aaetuk,t.j.norman,m.sensoy\}@abdn.ac.uk, \{bisdik,msrivats\}@us.ibm.com

Abstract—Trust and reputation are significant components in open dynamic systems for making informed and reliable decisions. State-of-the-art information fusion models that exploit these mechanisms generally rely on reports from as many sources as possible. Situations exist, however, where seeking evidence from all possible sources is unrealistic. Querying information sources is costly especially in resource-constrained environments, in terms of time and bandwidth. In addition, reports from multiple sources expose one to the risk of double-counting evidence, introducing an extra challenge of distinguishing fact from rumour. This paper describes TIDY (Trust-based Information fusion through Diversity), a trust-based approach to information fusion that exploits diversity among information sources in order to select a small number of candidates to query for evidence, and to minimise the effect of correlated evidence and bias. We demonstrate that reliable decisions can be reached using evidence from small groups of individuals. We show empirically that our approach is robust in contexts of variable trust in information sources, and to a degree of deception.

I. INTRODUCTION

Effective decision-making in large, open, and dynamic systems relying on information from sensors is plagued by uncertainty due to unreliable information from these sources. Sensory sources can be soft (e.g., humans) or hard (e.g., wireless sensors), and their reliability may be impaired either by extreme operating conditions and use of poor quality devices (wireless sensors), or through incompetence and deceptive behaviour (humans). While many existing information fusion protocols assume that all the information sources are reliable [1], other approaches exist, that adopt methods in trust and reputation systems to address uncertainty, and maximise confidence in fused information [2], [3]. However, using trust and reputation to improve the quality of fusion, requires paying close attention to challenges such as constraints in resources and correlated evidence, which are often overlooked.

A common approach for assessing trust in information is to rely on reports from as many sources as possible, as more evidence minimises the risk of biased opinions. In the physical world, capturing and distributing evidence can be costly. For instance, in distributed environments such as peer-to-peer networks, sensor networks, pervasive computing, each participant is responsible for collecting and combining evidence from others due to lack of central authority or repository. Also, in emergency response, a decision maker normally reasons with high volumes of streaming information, with strict real-time requirements. The major constraints in these systems are bandwidth, delay overheads, and energy, motivating the need to minimise the number of messages exchanged in order to arrive at a decision. Furthermore, there is often no guarantee that evidence obtained from different sources are based on direct, independent observations. Sources may likely provide (unverified) reports obtained from others (e.g., copying in social networks, ‘gossiping’ in sensor networks [4]), resulting in correlations and bias. Consequently, although combining reports from multiple sources to support decision-making is useful, effective sampling of information sources in order to improve the quality of information fusion requires answering some important questions: With limited capacity to query for evidence, how can reliable decisions be reached using evidence from small groups of individuals? How can one ensure that opinions from diverse sources based on their private experience are taken into consideration, without the risk of double-counting evidence?

Using evidence from small subsets of sources to support decision-making in highly dynamic and open systems is non-trivial. Known and trusted sources may leave the system at some point, and unknown and possibly unreliable sources may join the system, necessitating a certain trade-off between exploration of the population of (possibly unknown) sources and exploitation of known and trusted ones. In this paper, we present a trust-based approach to information fusion, that exploits diversity among information sources to sample them, in ways that maximise the quality of assessments and minimise the costs of arriving at decisions.

Specifically, our approach involves grouping homogeneous sources (likely to provide similar evidence) together in order to minimise the number of sources consulted for evidence. In this paper, we further our initial work in [5], where we assumed a known similarity metric for grouping the sources. We adopt methods from machine learning to identify complex behaviour patterns and learn suitable metrics for stratifying the population of sources, based on their features and evidence from their past reports. Candidates for fusion are subsequently sampled from diverse groups of trustworthy sources. Our work

---

This research was supported by the Petroleum Technology Development Fund (PTDF) Overseas Scholarship Scheme (OSS) Nigeria, and 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 here on.
is particularly oriented towards resource-constrained environments, where querying information sources is costly, and in environments where similar opinions may be shared among multiple sources, thereby making the application of existing fusion models unrealistic.

The contributions of this paper are: First, we define the concept of diversity and present techniques for learning suitable similarity metrics for source diversification relevant to many application domains. Second, we propose TIDY, a trust-based fusion model that exploits diversity to sample a small group of sources for evidence in contexts of uncertain and deceptive information sources. Finally, we leverage the power of subjective logic as the underlying information fusion and trust assessments framework. We demonstrate the effectiveness of our approach against existing methods in both making more accurate assessments, and using fewer resources.

The rest of this paper is organised as follows. Section II provides some background on the concept of diversity and also highlights related work in the area. We formally introduce the problem statement in Section III, and present some necessary terminologies and formulae that are used later in the paper. In Section IV we introduce TIDY, our information fusion framework, and in Section V we present evaluation of our approach. Finally, we conclude in Section VI with a discussion and avenues for future research.

II. BACKGROUND

Humans and indeed any entity operating in a social space, are profoundly affected by the interaction of individuals and the social networks that link them together. Evidence of such pattern has emerged in literature, and is referred to elsewhere as herd thinking [6]. This describes a situation where no one is thinking, but everyone is just copying (or conforming with) everyone else. In most information dissemination environments such as social media (e.g., Facebook, Google+, and Twitter), individuals are linked together by different relations (features), that define how information may be accessed and propagated. For example, Facebook defines the concept of friends, whereby individuals can access and further propagate (possibly without acknowledgements) opinions obtained from their neighbours. Similar relation is defined in Google+ in its circles, and Twitter has the follower-followee relation. In other settings such as news media, opinions maintained by individuals may be influenced by their favourite news channels (e.g., CNN, BBC, Fox News). Furthermore, in coalition contexts, stakeholders may maintain different policies that determine the sort of information they share with others [7]. For example, sources from a certain country or organisation involved in a peacekeeping mission may add subjectivity to shared information in line with operating policies of their individual affiliations, especially if providing objective reports may be detrimental to their well-being. In these scenarios and similar ones that may be cited, revealing patterns in source features correlating with their reports may be likely discovered.

To generalise on this concept, one may infer that sources with similar features are likely to provide similar reports in certain contexts. For instance, individuals subscribing to the same news channel may likely provide similar opinions on a subject as a result of being informed by the same source. Similarly, it may be inferred that sensors located within the same physical space, or under the same administrative domain mostly provide similar reports. If such hidden relationships exist between features of sources and the reports they provide, one could potentially exploit diversity to limit the number of sources consulted for evidence, and protect oneself against correlations of evidence and bias. Our intuition here is that, if a group of sources is observed to consistently provide similar reports, then such group could be regarded as having similar view point, and such knowledge is taken into consideration in future interactions with the sources. We define diversity as a mapping $\Delta : 2^S \rightarrow G$ of different subsets of sources to a set of groups of homogeneous sources. This process involves learning appropriate similarity metrics that may be used to stratify the population of sources, and allows us to group sources likely to provide similar reports together based on their features and evidence from their past reports.

A. Related work

Trust and reputation have been active areas of research in multiagent systems, and have proved beneficial in addressing the problem of uncertainty in both human and electronic societies. The Beta Reputation System (BRS) [8], is based on theory of statistics, and uses beta probability density functions to derive reputation ratings by combining binary feedback from different sources. In [9], the BRS is extended to deal with the problem of misleading reports from malicious sources by using a majority-based filtering technique. This approach assumes that malicious sources are in the minority and proves ineffective otherwise. To address this shortcoming, reputation of the reporting source is taken into account in [10], in order to discount its opinion. This approach becomes problematic when evidence about past behaviour of a source is unavailable. In [13], correlation between features and behaviour of sources is exploited in the absence of concrete evidence about the trustworthiness of a source. However, the emphasis of the study is on task delegation to an agent. A bayesian approach is proposed in [14], to deal with unreliable reports when sources are known to consistently bias their reports. By exploiting features of a source, the evaluation function used by the source in reporting is learned over time and used to re-interpret its reports. This approach enables an agent to make use of all available reports, with limited need of discounting or discarding opinions considered to be misleading.

In order to deal with the problem of correlated evidence, the truth finder system in [11] employs an approach based on source diversification. Their idea is similar to ours, but the authors assume a static and prior knowledge of the metric that defines similarity among sources. In most cases, however, such prior knowledge may not be available especially in large, open and dynamic systems, where relationship between sources may be defined by complex feature combinations. In [12] an iterative approach is used to estimate the probability of dependence between sources. Their approach relies on knowledge of the ground truth, and therefore works on the assumption that sources providing the same false information are likely to be dependent.

III. PRELIMINARIES

A. Sampling Budget

The aim of our model is to support the selection of reliable information sources within budgetary constraint in
order to improve the quality of information fusion. Formally, given some budget \( \Phi \), and a set of potential information sources \( S \), the objective is to select a subset of sources \( C \subseteq S \), such that the decision maker’s evaluation function \( E(C, \text{Cost}_C, q, \rho, R_C, \Phi) \) produces a sufficiently high quality of information. The optimal subset of sources to sample reports from, given a budget \( \Phi \), can be characterised as the subset that maximises the expected utility of a query \( q \):

\[
C_{\text{opt}} = \arg \max_{C \in S} E(C, \text{Cost}_C, q, \rho, R_C, \Phi),
\]

where \( R_C \) is the set of reports obtained from sources in \( C \) in response to a query \( q \) about some proposition \( \rho \). \( \text{Cost}_C \) is defined in terms of individual costs of the sources in \( C \), and \( \text{Cost}_C(q) = \sum_{s \in C} \text{Cost}_s(q) \). The sampling constraint is specified such that \( \text{Cost}_C(q) \leq \Phi \).

B. Subjective Logic

We present a brief primer on subjective logic (SL) used in this paper for evidence combination. More details about SL is presented in [15]. Subjective logic is a type of probabilistic logic that explicitly takes uncertainty and belief ownership into account. In general, SL is suitable for modelling and analysing situations involving uncertainty and incomplete knowledge. Arguments in SL are subjective opinions about propositions. A binomial opinion of an agent \( x \) about the truth of a proposition \( \rho \) is represented by the quadruplet \( \omega^x_{\rho} = (b, d, u, a) \), where: \( b \) is the belief that \( \rho \) is true; \( d \) is the belief that \( \rho \) is false; \( u \) is the uncertainty about \( \rho \); and \( a \) is the base rate, and represents the \textit{a priori} probability about the validity of \( \rho \) in the absence of any evidence. The default value of \( a \) is 0.5 [16], which means that before any positive or negative evidence has been received, both outcomes are considered equally likely, \( b + d + u = 1 \) and \( b, d, u, a \in [0, 1] \). A binomial opinion can be represented as a beta distribution, and opinions are formed on the basis of positive and negative evidence. The variables \( p \) and \( q \), represent the number positive and negative observations about \( \rho \) respectively, and can be used by \( x \) to obtain an opinion about \( \rho \) as follows:

\[
b = \frac{p}{p + q + 2}, \quad d = \frac{q}{p + q + 2}, \quad u = \frac{2}{p + q + 2}.
\]

The probability expectation value of an opinion is defined as:

\[
E(\omega^x_{\rho}) = b + u \times a.
\]

IV. SYSTEM DESCRIPTION

Fig. 1 illustrates the key components of our information fusion framework. It assumes there is one information consumer or decision maker, that uses reports from different sources to arrive at decisions. The sources may be owned by a number of different stakeholders, with varying degrees of trust. In addition, sources may not always report based on their direct, independent observation, and may only be relaying information obtained from others, or as specified by different information sharing policies.

The framework is equipped with a trust model based on subjective logic, that enables the discounting of reports to reflect the perceived reliability of the reporting source. The diversity model (DM) component uses suitable similarity metrics to stratify the source population, such that similar sources are grouped together. The source selection module, uses knowledge of source diversity to sample the population of sources for evidence according to the allowed budget. Issues of correlation are handled in the fusion process by exploiting knowledge provided in the DM. The knowledge base (KB) holds feedback after learning the ground truth with respect to the fusion output. KB also holds evidence about the behaviour of sources in different groups with respect to similarity of their reports. Based on the evidence obtained, both the trust and the diversity models are updated to reflect current knowledge. In the case of DM, this may trigger a learning process in order to keep the model consistent with current behaviour of sources. In the rest of this section, we present the formal description of aspects of the system, and also describe in detail how the different processes highlighted are implemented.

A. Information source

We denote as \( S \) the set of information sources accessible to the decision maker \( x \), and individual sources \( s_1, s_2, \ldots \in S \). An information source \( s \) is a tuple \( \langle Id, F, R \rangle \), where \( Id \) is a unique identifier, \( F \) is a set of features, and \( R \) is a set of past reports. A feature \( f \in F \) is an observable attribute of a source e.g., affiliation or location of the source. Some sources might be more similar based on their features, and over time, the decision maker learns the relative importance of features, represented by the vector \( \langle w^x_{f^1}, w^x_{f^2}, \ldots, w^x_{f^n} \rangle \), where \( w^x_f \) is \( x \)'s view of the importance of feature \( f \), in defining similarity among a group of sources, and \( w^x_f : \rightarrow [0, 1] \). Subsequently, \( x \) uses this metric to stratify the population of sources. A detailed description of this process is presented in Section IV-E.

B. Report

A report \( R \) is an opinion provided by a source \( s \), about a proposition \( \rho \) to the decision maker \( x \), in response to a query \( q \), \( s \) records its perceived opinion about \( \rho \) as \( \mathcal{R}_{s, \rho} \), and reports \( \mathcal{R}_{s \rightarrow x, \rho} \) when queried by \( x \). The variable \( t \) corresponds to a specific sampling round or time step associated with a report from \( s \), such that \( \mathcal{R}_{t, s, \rho} \) represents a report at time \( t \). \( \mathcal{R}_{s \rightarrow x, \rho} \) is the set of reports received by \( x \) from \( s \) in the interval \( t, t + k \). We assume that the same query \( q \) is made to all the sampled sources in a specific sampling round \( t \). A report is numerical value drawn from a Gaussian distribution and its interpretation depends on \( \rho \). For example, \( \mathcal{R}_{t, s, \rho} = 0 \) is interpreted differently for queries “is agent \( y \) trustworthy?” and “how many terrorists exist in the theatre?”.

C. Group

Let \( G \) denote a stratification on \( S \), and \( g_1, g_2, \ldots \in G \). We define a group \( g_i \in G \) as a collection of homogeneous sources,
such that \( \{ s : s \in g_i, g_i \in G \} = S \). Groups do not overlap (i.e., \( g_i \cap g_j = \emptyset \), if \( i \neq j \)), and a valid stratification is one in which a source belongs in only one group. Groups are formed subjectively by an agent \( x \), who attempts to disambiguate what metrics lead to a better stratification of sources. Sources are grouped together based on how similar they are to one another, as specified by some similarity metric. It is important to emphasise here that group membership is not simply based on similar level of trustworthiness, rather it is a measure of the consistency of sources in giving similar reports in response to different queries. Therefore, it is possible for sources in different groups to have similar level of trustworthiness (e.g. sources from different, but equally reputable organisations).

\[
\begin{array}{c|ccc|ccc|ccc|ccc|ccc}
& s_1 & s_2 & s_3 & q_T & (s_1,s_2) & (s_1,s_3) & (s_2,s_3) \\
q_1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\
q_2 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\
q_3 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\
q_4 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\
q_5 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
\tau & 0.57 & 0.57 & 0.57 & & & & \\
\end{array}
\]

(a) Trust relationship. (b) Similarity relationship.

Fig. 2: Report matrix.

To illustrate this point, consider an interaction involving three sources \( s_1, s_2, s_3 \) in five sampling rounds \( q_1, \ldots, q_5 \) (Fig. 2), with reports in \([0,1] \). The trust relationship in Fig. 2a shows the trust score \( \tau \) for each of the sources after five sampling rounds. This is calculated using Equations 2 and 3. For instance, for source \( s_1 \), the number of positive evidence (complying with ground truth \( q_T \)) \( p \) is 3, and the number of negative evidence (conflicting with \( q_T \)) \( q \) is 2. Based on this evidence, the trust \( \tau \) for \( s_1 \) is computed. Fig. 2b represents the similarity relationship for the sources, and shows the report similarity score \( \varphi \) for each source pair after five sampling rounds, which is based on their reports in Fig. 2a. Similarly, Equations 2 and 3 is also used to compute the report similarity \( \varphi \), whereby positive evidence \( p \) represents instances a source pair gives similar reports, and negative evidence \( q \) are those instances the pair gives conflicting reports. Although the three sources are considered to be equally reliable with a trust score of 0.57 after the five sampling rounds, however, reporting pattern is not the same for all three sources in the sampling period. Sources \( s_1, s_3 \) appear more consistent in providing similar reports as indicated in their high similarity value (\( \varphi = 0.85 \)) than any other pair.

\( \frac{-0.0055 * f_1 + 0.0013 * f_2 + 0.0029 * f_3 + 0.8007 * f_4 + 0.4015 * f_6 + 11.0971}{\text{num: 1}} \)

\( \text{class} = \)

D. Trust computation

A trust score is maintained for each group \( g_i \), and is used by the decision maker as an expected reliability of members encountered in the group. First, the trust score \( \tau_{x,g_i}^p \) of each individual member of the group is calculated as demonstrated in the previous section, using Equations 2 and 3. The group trust \( \tau_{g_i:p}^x \) is then calculated using Equation 4 as average of the trust scores of group members.

\[
\tau_{g_i:p}^x = \frac{\sum \tau_{x,p}^s}{|g_i|}
\]

(4)

E. Learning Diversity

The primary goal in diversity is to find the best way of stratifying the population of sources, such that similar sources are grouped together. In diversity learning, our aim is therefore to identify a function \( \Delta : 2^S \rightarrow G \) that maps different subsets of sources to a set of groups. We take as a working assumption, that there may be some correlation between the features of sources and their reports. Where this exists, we could exploit information from observable features of sources, as well as evidence from their past reports to learn similarity metrics that may be subsequently used to diversify the source population.

To explain the learning process, we adapt our illustration in Fig. 2 and use as a running example in the rest of the paper. We assume a peace-keeping mission, with 100 sources each monitoring the level of conflict in a specific region. Reports on ground situation provided by the sources to \( x \) are expressed in terms of numeric values in the range \([0,1]\), where 0 implies no conflict state, and 1 high conflict state, which may warrant the deployment of troops to the region to manage the situation. Let’s also assume that \( x \) has access to the history of the sources reports, and therefore wishes to exploit this to learn the best way of diversifying the population, so as to carry out an effective sampling in future interactions. Each source is identified by 6 features, as perceived by \( x \), and these are labelled \( f_1, f_2, \ldots, f_6 \), where features may represent attributes such as country, location, expertise, and coded by numeric values.

Fig. 3: Example model tree and classification rule for diversity.

The first stage of the process is an attempt at disambiguating what metrics lead to a better stratification of the population. A good metric in our estimation is one that produces the highest feature-behaviour correlation, such that the likelihood of sources in the same group giving similar reports is maximised. The second stage involves using the learned metric to partition the sources into semi-homogeneous subgroups. We make use of existing machine learning techniques in both instances and learn associations between sets of features, represented as numeric attributes, and a target value (similarity score) defining the degree of similarity between sources. Specifically, we employ the M5 model tree learning [17] algorithm\(^1\) to build a regression model to estimate similarity between sources, and

\(^1\)We use the M5 implementation of Weka [18], a popular open-source machine learning toolkit written in Java.
the hierarchical clustering [19] algorithm\(^2\) to form groups of similar sources based on the learned similarity model.

Fig. 3 shows an example model tree and classification rule for the diversification of the 100 sources. The model is constructed using the M5 algorithm for predicting similarity between sources, and generalises to new and unknown sources. The algorithm shares some similarities with decision tree classifiers (CART), which builds tree-based models for classification. However, unlike CART with class labels or values at the leaves, the leaves of a model tree are multivariate linear regression models, which are used to predict a target value (in our case, a similarity score \(\mathcal{M}\)). Input to the M5 algorithm is a collection of training instances. Each instance is specified by its value of a fixed set of attributes. For example, to add a training instance, we compute the similarity (1 - absolute difference of corresponding attributes i.e., \(1 - |f_{n1} - f_{n2}|\)) of each of the 6 attributes for each source pair \((s_1, s_2)\), and the class, which is the report similarity in a considered time interval \(T\) for the source pair. While computing the report similarity, we define a similarity threshold \(\eta\), and compute report similarity as a function of the number of similar reports \(p\), and conflicting reports \(q\) for the source pair for each time step \(t \in T\).

\[
D_{s_1,s_2}^{R_t} = |R_{s_1,p}^t - R_{s_2,q}^t|.
\]

\[
(p_{s_1,s_2}, q_{s_1,s_2}) = \begin{cases} (1, 0), & \text{if } D_{s_1,s_2}^{R_t} \leq \eta. \\ (0, 1), & \text{otherwise} \end{cases}
\]

\[
p_{s_1,s_2}^T = \sum_{t \in T} p_{s_1,s_2}^t, \quad q_{s_1,s_2}^T = \sum_{t \in T} q_{s_1,s_2}^t.
\]

The report similarity \(\varphi\) for \((s_1, s_2)\) pair is then computed using Equations 2 and 3. \(D_{s_1,s_2}^{R_t}\), in Equation 5 is the difference between reports of \(s_1, s_2\) for a specific time step \(t\).

The second stage of the process involves using the learned metric to stratify the population of sources. We employ the hierarchical clustering algorithm for this task, and define a stoppage criteria or diversity threshold, \(\delta\) for clustering. The \(\delta\) parameter which lies in the interval \([0, 1]\), specifies the maximum level of diversity required in the system. For instance, if the parameter is \(\delta = 1\), all the sources are assigned to singleton groups (extreme diversity). However, if \(\delta\) is too low (i.e., \(\delta = 0\)), all sources are assigned to one group (no diversity).

The clustering process uses the linear regression models constructed by the M5 algorithm to predict the similarity, \(\mathcal{M}\) for each source pair, and subsequently uses this measure to cluster the sources. Therefore, given the feature vector of any two sources as input, a similarity score \(\mathcal{M}\) is obtained, and this specifies the expected degree of ‘closeness’ of the pair. Specifically, \(\mathcal{M}\) is realised by using the most relevant linear regression model for the evaluated instance as illustrated in Fig. 3. With this information the clustering process proceeds as follows. Each source is initially regarded as belonging to a separate cluster, and the two clusters with the highest similarity score \(\mathcal{M}\) are continuously merged to form a new cluster until the stoppage condition \(\delta\) is met. The similarity score of a cluster is computed as the average similarity of all member pair in the cluster.

\[2\] Although there are various clustering techniques that can be used for this purpose, we select the hierarchical clustering as an algorithm of choice because it is well-known, and allows us to cluster into a set of groups the cardinality of which we do not know in advance.

F. Sampling and Fusion of Reports

The diversity model offers rich context from which a fusion set may be derived. In general, a fusion set is made of candidates randomly selected from different groups, whose combined evidence is used to form an opinion about a proposition. Depending on the specific task requirements, richer contexts could be explored using the learned model of diversity. For instance, the cost and risk assessments of a potential transaction may influence the sampling process. Irrespective of the sampling strategy considered, a key objective is to maintain a reasonable view of the different sub-groups in the population, so as to have a competitive exploratory advantage of the population at a reduced cost. Given this requirement, we present two methods for selecting candidate sources to make up the fusion set. The choice between both methods depends on the available budget \(\Phi\), and the number of sub-groups identified.

We assume budget \(\Phi\) can be expressed in terms of the number of sources that may be sampled for a specific query \(q\). The fusion set \(\mathcal{C}\) comprises of candidates for fusion sampled from various groups in \(\mathcal{G}\). Let \(\mathcal{G}_i \subseteq \mathcal{G}\) denote the subset of groups in \(\mathcal{G}\) whose members are to be included in \(\mathcal{C}\), such that \(\{\mathcal{G}_i : \mathcal{G}_i \subseteq \mathcal{G}_i, \mathcal{G}_i \in \mathcal{G}\} = \mathcal{C}\). This implies that \(\mathcal{G}_i\) is the corresponding subset of sources sampled from group \(g_i\). We may have two different cases:

**CASE I (\(\Phi \geq |\mathcal{G}| \implies \mathcal{G} = \mathcal{G}\)):** This method has some similarities with the proportion allocation technique of stratified random sampling [20]. The number of candidates to be sampled from a group \(g_i\) based on the budget:

\[
\text{budget}(g_i) = |g_i| \times (\Phi / |\mathcal{G}|)\
\]

Representative candidates are then randomly selected from \(g_i\) according to \(\text{budget}(g_i)\), and these are contained in \(\mathcal{G}_i\). Applying this technique directly however, may lead to information exclusion in much smaller groups (e.g., not selecting from singleton groups), therefore, we constrain our specific implementation to the selection of at least one candidate from each group, by reducing the number of candidates to be sampled from much larger groups.

**CASE II (\(\Phi < |\mathcal{G}| \implies \mathcal{G} \subset \mathcal{G}\)):** Using this method, a single representative candidate is selected from each group, by ranking the groups in order of trustworthiness. Assuming groups are ordered in descending order of trust scores, then \(\text{budget}(g_i) = 1\), if group rank \(\mathcal{G}_i \geq \Phi\), and \(\text{budget}(g_i) = 0\), if group rank \(\mathcal{G}_i \leq \Phi\). The intuition here is that, although information is potentially lost from some of the groups, however, it is more beneficial to prioritize available resources to more trustworthy sources, and increase the scope of exploration as the budget increases.

We do not suggest these to be the only methods for sampling, but only that both techniques demonstrate possible realisation of our model, which we have used in our evaluation. Other sophisticated sampling techniques may be explored to meet specific application requirements.

Reports from sources in the fusion set \(\mathcal{C}\) are combined in order to form an opinion about \(\rho\). Specifically, the reports from the sources are first partitioned into different bins to correspond with their original groups. The consensus opinion
of each group is calculated by computing the mean report in the
group $\bar{R}_{g_i,p}$ as in Equation 9. The resulting opinions are then
discounted by the corresponding trust score $\tau_{g_i,p}$ of the groups.
Finally, the normalized opinions from all groups are combined
to obtain the overall opinion $E^p_{g}$ as also shown in Equation 9.
This fusion approach minimises the adverse effect of large
groups of unreliable sources working together to undermine
the trustworthiness of the fusion results.

$$\bar{R}_{g_i,p} = \frac{\sum_{s \in G_i} R_{s \rightarrow x,p}}{|G_i|}, \quad E^p_{g} = \frac{\sum_{g_i \in G} \bar{R}_{g_i,p} \times \tau_{g_i,p}}{\sum_{g_i \in G} \tau_{g_i,p}}$$

(9)

V. EVALUATION

In our evaluation, we focused on demonstrating (by simu-
lations) the effectiveness of TIDY in supporting the decision
maker in making reliable assessments of situations of interest.
In particular, we measure the robustness of the framework in
the presence of reports from varying degrees of unreliable
sources, who are particularly coordinated in their reporting.
We consider two scenarios. First, we study the effect of
different budget constraints in the presence of experts (honest
and malicious sources), who have knowledge of the ground
truth. In the second case, we assume there are no experts
in the system that can consistently provide reports about the
ground truth. In both scenarios, we explore the effect of
sources behaving similarly either due to their subjectivity being
conditioned by the same factor (e.g., organisation policy),
or by coordinating to supply similar evidence when queried
(e.g., collusion) in order to mislead the decision maker.
We are therefore interested in finding out how well our model
is able to discover the hidden groups of sources, and exploit
the knowledge to limit the number of sources queried for evidence,
and handle correlated evidence. We measure performance by
evaluating the predictive accuracy of the model to some ground
truth $gT$.

We compare our technique to two popular trust approaches
proposed in literature, referred to as endogenous and exogenous
methods [21]. The former attempts to identify unreliable
evidence by considering the statistical properties of the reported
opinions alone [e.g., [9],[3]], while the latter relies on the
reputation of the reporting source to discount its opinions [e.g.,
[22],[10]]. We refer to these two approaches in our evalua-
tion as observation-based sampling (OBS), and majority-based
sampling (MBS) respectively. The OBS computes a trust score
for the sources based on their past performance, as observed
by the decision maker through repeated interactions with the
sources, and subsequently uses this score to discount their
reports in future interactions. MBS assumes that reports from
multiple sources observing the mean of a physical phenomenon
independently will approximately follow normal distribution,
and thus filters out all reports that deviate more than one
standard deviation from the mean report. Also, we incorporate
random selection of sources as a baseline approach. We refer
to this as naive-based sampling (NBS).

Each information source in our simulations is assigned a
profile, which determines its reporting pattern in relation to
other sources in the system. Each profile defines a number
of features and the distribution from which feature values are
drawn for members of the profile. Specifically, each feature
value is drawn from a Gaussian distribution, with discriminat-
ing features, which are important in defining similarity for pro-
file members having very small standard deviation $N(\mu, 0.01)$,
and a uniform distribution $N(0, 1)$ for non-informative fea-
tures, which introduces noise to the diversity learning process.
In addition, each profile has a conformity parameter $P_c$
that specifies, in terms of a probability value, the degree of
conformity of a source to its profile. A source that does not
conform, deviates from mainstream opinion held by its group.
A low $P_c$ value means that more sources in a profile will report
independently, according to their individual reliability, and not
depend on their ‘neighbour’ (profile member) for an opinion.
A conforming source when reporting, first finds out about
opinions maintained by its neighbours about a proposition. If
any exists, it randomly selects one of such opinions to report,
discarding its own private opinion. The $P_c$ parameter adds
an extra challenge to the learning process, and allows us to
evaluate the ability of our model to deal with noise due to
uncorrelated feature to behaviour similarity. Each source has
a reliability parameter $P_r$ that determines its behaviour (i.e.,
honest, malicious). Report types are defined as follows.

- **Honest report**: This type of report is closer to the
ground truth, with small Gaussian noise $N(0, 0.01)$. Sources with high reliability ratio $P_r$ are more likely
to provide this type of report.

- **Malicious report**: Malicious sources with low $P_r$ are
more likely to provide this type of report, which,
if left unmanaged could potentially undermine
the fusion result. Reports in this category are significantly
deviated from the ground truth, with large Gaussian
noise $N(1, 0.01)$.

At the end of each round, the decision maker learns the ground
truth and updates the trustworthiness of the sources with new
evidence computed using Equation 10, which is based on the
intuition that information is still useful if it has a small amount
of noise or is slightly discounted [23].

$$\langle p^{t}, q^{t} \rangle = \begin{cases} 
(1, 0), & \text{if } |gT^t - \bar{R}_{g}^t| \leq 0.1 \\
(0, 1), & \text{otherwise}
\end{cases} \quad (10)$$

To keep things simple in our simulations, $\Phi$ is defined in terms
of the number of sources that may be sampled for evidence.
Consequently, we define a small budget with $\Phi = 5$ and
large budget with $\Phi = 75$, to indicate that fusion results may
be based only on evidence sampled from 5 and 75 sources
respectively. Since we are considering fusion of information
in large and open environments, sources can freely join and
leave the system. We simulate this property of the system
with the parameter $P_l$, which determines in each round the
probability of a particular source leaving the system. When
a source leaves, it is replaced with a new source of the
same profile, in order to keep the number of sources fixed
throughout the simulation. This property impacts on the ability
of the different fusion approaches to accurately learn the
reliability of the sources, and emphasis the need for a trade-
off between exploration and exploitation of the population.
However, dynamic activity is relaxed in all cases for the first
30 rounds of the simulations, in order to enable the different
approaches gather some information to build their individual
TABLE I: Experimental Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Test values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source population (S)</td>
<td>100</td>
</tr>
<tr>
<td>Profiles and Reliability probability ($P_{r}$)</td>
<td>$p_1 = 0.2$, $p_2 = 0.8$, $p_3 = 0.9$</td>
</tr>
<tr>
<td>Profiles and Conformity probability ($P_{c}$)</td>
<td>$p_1 = 0.8$, $p_2 = 0.8$, $p_3 = 0.8$</td>
</tr>
<tr>
<td>Population change probability ($P_{l}$)</td>
<td>0.1</td>
</tr>
<tr>
<td>Diversity threshold ($\delta$)</td>
<td>0.4</td>
</tr>
<tr>
<td>Report similarity threshold ($\eta$)</td>
<td>0.01</td>
</tr>
</tbody>
</table>

models. The parameter list with their default values is shown in Table I.

A. Results

Each of our simulations is repeated 10 times, with each run made up of 100 sampling rounds. We present the average of our results, which are significant based on analysis of variance (ANOVA) with repeated measures, having a confidence interval of 0.95. Error bars show variation between means.

1) Robustness to deception with experts: Results presented in Fig. 4 show the performance of the different approaches in the presence of deceptive sources, and with different budget limits. The majority based approach MBS starts off better than other approaches in all cases when the ratio of malicious sources is low in the system. This is because MBS benefits from the large number of honest reports to filter out malicious ones. Also, it is not affected by the dynamic nature of sources in the system, as its filtering is based only on statistics on the reports and not on knowledge of the sources themselves. However, with increasing number of malicious sources, this approach degrades significantly. This result is expected since MBS algorithms are not robust in the presence of large number of malicious reports. The slight performance lag in the case of TIDY and OBS when fewer malicious sources are present in the system as compared to MBS is due to the discounting of opinion by learning the reliability of sources over time, which might take some time to converge, coupled with uncertainty introduced by source dynamism $P_{l}$. Therefore, lower weights might be assigned to reports from highly trustworthy sources and vice versa, and this impacts the fusion results. However, the benefits of observation based filtering becomes obvious as the ratio of malicious sources in the system increases. Effect of bogus reports is greatly mitigated due to acquired knowledge of behaviour of the sources. There is an interesting observation in the performance of both TIDY and OBS when the budget is increased in Fig. 4(B). Although performance might be expected to improve with increasing sampling allowance, however, our results show a different case. Increasing budget in the case of OBS gives a slight performance boost when malicious sources are in the minimum ($\leq 20$), however, performance degrades significantly when malicious sources increase. This can be explained by the effect of the $P_{l}$ parameter. As OBS is interaction based, and exploits only highly trustworthy sources, in situations that known and trustworthy sources leave the system, OBS assigns wrong weight to more unknown sources. Without any interaction, the trustworthiness of sources, even very reliable sources is computed as 0.5 (total uncertainty). Performance of TIDY remains relatively stable with changing budget, and outperforms all the other approaches. This is due to the fact that although known and trustworthy sources may leave the system, TIDY exploits its diversity component to select alternative sources of equal trustworthiness.

2) Robustness to deception with non-experts: Often times the degree of corroborations of evidence is used as an indication of trustworthiness, especially in systems where there are no clear experts. In such scenarios for example, one would likely believe an event reported by numerous sources more than conflicting evidence supplied by one or few sources. This is the case in applications such as crowdsourcing and citizen sensing, where information is often sought from numerous and mostly unreliable sources. If these sources are simply relaying what they heard from others, then this may lead to misinformation. In this set of experiments, we demonstrate the robustness of our approach to varying degree of source dependence (copying). There are no clear experts, and the decision maker only relies on the degree of corroborations of reports to determine the ground truth. In this case, we vary the proportion of sources depending on others for opinion from 0 to 90, and present the results in Fig. 5, with different sampling budgets as before. In this setting, sampling based on majority filtering MBS performs significantly worse since it generally assumes independence among sources, and copying sources do not necessarily obtain their report from reliable sources. Therefore, mainstream opinion becomes inadequate for filtering out outliers. NBS gains some performance improvement over MBS, but still does badly in general. OBS does significantly better than the other two approaches, but worse than TIDY, since without the presence of consistency of source behaviour, it is unable to model the reliability of sources effectively to discount their opinions. TIDY on the other hand performs better, because although its trust component does little in learning source trustworthiness, however, its diversity component is able to identify dependencies among the sources and uses this to inform source selection.

VI. CONCLUSION AND FUTURE WORK

We have presented a framework for sampling information sources and fusing information in environments where resource limitations and restrictions need to be taken into consideration. Our approach is particularly useful in scenarios where a decision maker is exposed to the risks of deception and double-counting of evidence. Deception and report correlations have been shown to greatly degrade the quality of information fusion, and therefore needs to be mitigated to enhance confidence in fusion results. Where hidden networks or patterns defining correlated behaviour exist in the population, our approach is able to uncover such, and subsequently exploit it in order to limit the number of information sources sampled. Where a naive approach of information fusion would perform poorly under these conditions as revealed in the simulation results, our model shows positive outcomes that outperform classical trust-based information fusion approaches within budgetary constraints. We have identified the need to incorporate more robust decision-theoretic mechanism to handle complex source selection strategies, so as to meet different information needs. Our current work operates on static information, where the assessed situation is assumed to remain fairly stable over time. We intend to extend our study to address dynamic settings where information is streamed by a number of sensory sources of uncertain trustworthiness. This is a non-trivial scenario involving timely assessments of trust in fused information under a rapidly changing world, and requires determining appropriate stopping criteria to stop processing incoming reports.
and inferring on a proposition. Finally, we intend to apply our work to real-life applications like crowdsourcing and sensor networks.

REFERENCES


