A Collaborative Reputation System Based on Credibility Propagation in WSNs

Mohsen Rezvani and Aleksandar Ignjatovic
School of Computer Science and Eng.
University of New South Wales
{mrezvani,ignjat}@cse.unsw.edu.au

Elisa Bertino
Department of Computer Science
Purdue University
bertino@cs.purdue.edu

Sanjay Jha
School of Computer Science and Eng.
University of New South Wales
sanjay@cse.unsw.edu.au

***Abstract*—Trust and reputation systems are widely employed in WSNs to help decision making processes by assessing trustworthiness of sensor nodes in a data aggregation process. However, in unattended and hostile environments, more sophisticated malicious attacks, such as collusion attacks, can distort the computed trust scores and lead to low quality or deceptive service as well as undermine the aggregation results.**

In this paper we propose a novel, local, collaborative-based trust framework for WSNs that is based on the concept of credibility propagation which we introduce. In our approach, trustworthiness of a sensor node depends on the amount of credibility that such a node receives from other nodes. In the process we also obtain an estimates of sensors' variances which allows us to estimate the true value of the signal using the Maximum Likelihood Estimation. Furthermore, we augment the proposed trust framework with a novel collusion detection and revocation method. Extensive experiments using both real-world and synthetic datasets demonstrate the efficiency and effectiveness of our approach.

***Keywords*—reputation system, collusion attacks, data aggregation, credibility propagation, iterative filtering.**

I. Introduction

In many real-life distributed systems such as social networks, e-commerce platforms and wireless sensor networks (WSNs), the trustworthiness of participants has a significant role in the decision making processes. A reputation assessment at any given moment represents an aggregate of the behaviour of the participants up to that moment and has to be robust in the presence of various types of faults and malicious behaviour. There are a number of incentives for attackers to manipulate the trust and reputation scores of participants in a distributed system, and such manipulation can severely impair the outcome of such a system [1], [2].

In recent years, several approaches based on iterative filtering (IF) algorithms for trust and reputation systems have been proposed [3], [4], [5]. In these algorithms, the aggregate values are computed iteratively from the set of reports provided by the sources e.g. sensor nodes. Moreover, the sources whose reports often differ from other sources are assigned lower trustworthiness score [3]. Accordingly, these algorithms exhibit better robustness compared to the simple averaging techniques. However, they are still vulnerable against sophisticated collusion attacks. For example, we [6] recently showed how the IF algorithms can be compromised by a collusion attack when the adversary has enough knowledge about the trust computation method. Therefore, developing a robust reputation system by taking into account sophisticated collusion attacks against WSNs is an important challenge[1].

In order to address this problem, we propose a robust collaborative and local reputation model which is based on the novel concept of *credibility propagation* among the sensor nodes, called CrPr in this paper. In CrPr, each node obtains credibility from all other nodes according to the similarity between its reports and the reports of other nodes; the degree of such a transfer of credibility also depends on sensors' variances. We have developed an iterative procedure for simultaneously computing the trustworthiness of each node as well as the variance of its error. Subsequently, we use the variances of the sensor nodes to form a statistical estimator which, if the sensors' errors are independent and normally distributed, represents the Maximum Likelihood Estimator (MLE). Furthermore, we augment the reputation system with a novel collusion detection and revocation method which identifies the compromised nodes and eliminates their contributions according to the normality of the error behaviour in the nodes.

We provide a thorough empirical evaluation of the proposed reputation system using real world and synthetically generated datasets. The results show that our method provides both higher accuracy and better collusion resistance than the existing IF methods.

Our contributions can be summarized as follows:

- We introduce and formulate the concept of credibility propagation to measure the trustworthiness of sensor nodes in a sensor network;
- An iterative algorithm to estimate the true value of the signal based on an interdependency relationship among credibility, variance and aggregate values;
- A robust collaborative reputation system for distributed systems in general and WSNs in particular which is effective in a wide range of source faults as well as robust against collusion attacks;
- A novel collusion detection method based on an estimate of normality of sensor errors in the proposed robust reputation framework.

The rest of this paper is organized as follows. Section II formulates the problem and specifies the assumptions. Section III presents our novel reputation system. Section IV describes our experimental results. Section V discusses how the proposed

[1] In this paper, "trust" and "reputation" are used interchangeably.

method meets the requirements. Section VI presents the related work. Finally, the paper is concluded in Section VII.

II. Problem Description and Assumptions

A. Network Model

For the sensor network topology, we consider the abstract model proposed by Wagner [7]. The sensor nodes are divided into disjoint clusters, and each cluster has a cluster head which acts as an aggregator. Data are periodically collected and aggregated by the aggregator. We assume that each data aggregator has enough computational power to run an iterative filtering algorithm for data aggregation. The details of cluster formation, data collection and dissemination processes are out of the scope of this paper.

For simplicity, we also assume that all the sensor nodes in a cluster collect data for a single environment measurement, such as temperature. Note that, if there are multiple measurements monitored by sensors, the proposed reputation computation algorithm can be separately applied over each measurement.

B. Threat Model

In this paper, we use a Byzantine attack model, where the adversary can compromise a set of sensor nodes and inject false data through the compromised nodes [8]. We assume that when a sensor node is compromised, all the information which is inside the node becomes accessible by the adversary. Thus, we cannot rely on cryptographic methods for preventing the attacks, since the adversary may extract cryptographic keys from the compromised nodes.

Furthermore, we assume that the base station and the aggregators are not compromised; there are many approaches for dealing with the problem of compromised aggregators [9], [10], [11]. Thus, we limit our scope to the lower layer problem of false data being sent to the aggregator node by the compromised individual sensor nodes, as this problem has received much less attention. Moreover, we allow multiple adversaries to collusively report false data to deceive the reputation system, but we assume that the majority of the reports are good.

Finally, attacks via the communication channels and DoS attacks (e.g., eavesdropping, traffic jamming, etc.) are out of the scope of this paper.

C. Basic Concepts and Notation

Let us consider a WSN with n sensors S_i, $i = 1, \dots, n$. We assume that the aggregator works on one block of readings at a time, each block comprising of readings at m consecutive instants. Therefore, a block of readings is represented by a matrix $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ where $\mathbf{x}_i = [x_i^1 \ x_i^2 \ \dots \ x_i^m]^T$, $(1 \le i \le n)$ represents the i^{th} m-dimensional reading reported by sensor node S_i. Let $\mathbf{r} = [r_1 \ r_2 \ \dots \ r_m]^T$ denote the aggregate values for instants $t = 1, \dots, m$ which reflect a robust estimate of the true values of the signal for those time instants.

In the rest of this paper, we assume that the stochastic components of sensors' errors are independent random variables with a Gaussian distribution; however, if the error distributions of sensors are known, our method can be adapted to other random distributions to achieve an optimal performance. Accordingly, we assume that the reading of sensor s at time instant t is

$$x_s^t = r_t + e_s^t \tag{1}$$

where r_t is the latent true value of the quantity measured at time instant t, that is hidden to the external world, and e_s^t is a random variable with a zero-mean Gaussian distribution $e_s^t \sim \mathcal{N}(0, \sigma_s^2)$. TABLE I contains a summary of notation used in this paper.

TABLE I: Notation used in this paper.

Notion	Description
n	number of sensors
m	number of readings for each sensor
r_t	true value of the signal at time t
x_s^t	data from sensor s at time t
σ_s	standard deviation of sensor s t
$\mathrm{L}(j, i)$	the likelihood that sensor j could have made the readings of sensor i
$\mathrm{cr}(s)$	credibility of sensor s
$\mathrm{rep}(t)$	estimate aggregate value at time t
$\mathrm{var}(s)$	estimate variance of sensor s
c	number of malicious sensor nodes

D. Problem Statement

In an unattended WSN where some fraction of the nodes are subject to node compromise attack, false data injection by the compromised nodes can significantly skew the sink's estimate of the aggregate being computed using tiny aggregation functions, such as such as SUM, MIN or MAX, [11], [7]. Moreover, we [6] recently showed how the IF algorithms can be compromised by collusion attacks.

In this paper, our goal is to propose a robust trust-based data aggregation approach resilient to the false reports injected by compromised nodes. In particular, we aim to recover the aggregate values of r_t in Eq. (1), from the sensor readings by proposing a novel trust computation method with the following attributes:

(A1) the proposed algorithm is robust with respect to outliers. If only a very small fraction of readings x_s^t are far from some form of a "consensus" of other sensors, such readings have very little impact on the final aggregate values produced by the system;

(A2) the proposed algorithm is robust to collusion attacks, where a group of sensors tries collaboratively to skew the reputation values by an orchestrated effort;

(A3) the proposed algorithm is "statistically sound"; for example, if individual readings x_s^t provided by any sensor s are "correct values" plus some zero mean Gaussian noise independent for each instant, then the algorithm should produce an output close to the optimal output which is produced by the MLE;

(A4) sensors variances are not an input to the proposed algorithm, because such kind of information is unavailable in practice; the algorithm uses an adaptation procedure which automatically provides such output merely from the statistical features of the "raw" inputs.

III. Collaborative Reputation

We here assume that all sensors report readings for all time instants; we leave detailed extension of the proposed reputation system to sparse readings in [12].

A. Variance Estimation Principle

In the presence of stochastic errors for sensors readings in WSNs, a reputation system should produce estimates which are close to the optimal ones in the information theoretic sense. Thus, for example, if the noise present in each sensor is a Gaussian independently distributed noise with a zero mean, then the estimates produced by such reputation algorithm should have a variance close to the Cramer-Rao lower bound (CRLB) [13], i.e, in such a case the variance of our estimator should be close to the variance of the MLE. However, such estimation should be achieved *without* supplying the algorithm the variances of the sensors.

Thus, in the proposed reputation system, we first aim to achieve an optimal estimate of sensors variances. With such estimates, we employ an MLE-like technique to estimate the true value of the signal. We argue that this objective satisfies the above principle according to the following fact.

Proposition 1 (Minimum Variance Estimator)**.** *Assume that the readings of n sensors all have normal distributions but with different and* ***known*** *standard deviations $\sigma_1, \dots, \sigma_n$. Assume that using these sensors we have obtained measurements $X_1, \dots, X_n$ for a quantity r of interest. In this case, the MLE is the weighted average of readings of all sensors and it is expressed as follows:*

$$\hat{r} = \sum_{i=1}^{n} \frac{\frac{1}{\sigma_i^2}}{\sum_{k=1}^{n} \frac{1}{\sigma_k^2}} X_i. \tag{2}$$

and such an estimate has minimal possible variance.

We presented the proof of Proposition 1 in [12].

B. Definitions

The general idea of our proposed reputation system is based on an interdependency relationship among the credibility of sensors, the aggregate values at time instants, and the variance of sensors. In this work, we define:

- *Credibility* of a sensor node: it reflects how much other sensor nodes support the node based on the similarity among their readings.
- *Aggregate* of reports at a time instant: it is an estimate of the true value of signal in that particular time instant.
- *Variance* of a sensor: it is the square of the distance of the sensor's readings from the estimate of the true value of the signal.

In this paper, the credibility of a sensor node reflects the accumulated evidence from other sensors for how well-suited the readings of such a sensor are to serve as an estimation of the true value of the signal. In other words, the credibility of a node can be measured by the average of credibilities which such a node obtains from all other nodes.

Let us consider sensor node j with standard deviation σ_j; this sensor node "is going to assess" the credibility of the readings of another node i, by estimating the (normalised) likelihood that it could have made such readings itself, i.e., the credibility that sensor j confers to the readings of sensor i should be equal to

$$\mathrm{L}(j,i) = \left(\prod_{t=1}^{m} f(x_i^t; r_t, \sigma_j) \right)^{\frac{1}{m}} = \frac{1}{\sigma_j \sqrt{2\pi}} e^{\frac{-\frac{1}{m}\sum_{t=1}^{m}(x_i^t - r_t)^2}{2\sigma_j^2}} \tag{3}$$

We now define the credibility of a sensor i as the aggregate of credibilities conferred by other sensors, i.e., as normalised likelihood that all other sensors might have made readings of sensor i, i.e., as

$$\mathrm{cr}(i) = \left(\prod_{\substack{j=1 \\ j \neq i}}^{n} \mathrm{L}(j,i) \right)^{\frac{1}{n-1}} = \left(\prod_{\substack{j=1 \\ j \neq i}}^{n} \frac{1}{\sigma_j \sqrt{2\pi}} e^{\frac{-\frac{1}{m}\sum_{t=1}^{m}(x_i^t - r_t)^2}{2\sigma_j^2}} \right)^{\frac{1}{n-1}} \tag{4}$$

Clearly, Eq. (4) gives high credibility to nodes who report readings very close to other nodes ("the community sentiment"). In addition, the variance of each sensor has a significant role in amount of credibility which such a node can grant to other nodes: a sensor node with lower variance value grants higher credibility to sensors with readings close to the readings of such node.

Now we approximate the aggregate values $\mathbf{r} = \langle r_1, r_2, \dots, r_m \rangle$ using a weighted average, where the weight of a sensor node is the normalized value of its credibility. Thus, we compute the aggregate value at time instant t as follows:

$$r_t = \mathrm{rep}(t) = \sum_{i=1}^{n} \frac{\mathrm{cr}(i)}{\sum_{k=1}^{n} \mathrm{cr}(k)} x_i^t \tag{5}$$

where term $\sum_{k=1}^{n} \mathrm{cr}(k)$ is used to normalize the credibility values to make the sum of sensors weights equal to 1.

Given the aggregate values by Eq. (5), we can approximate the variance of a sensor as the average squared Euclidean distance of its readings to such aggregate values. Thus, we compute the variance of sensor i as follows:

$$\mathrm{var}(i) = \frac{1}{m} \sum_{t=1}^{m} (x_i^t - \mathrm{rep}(t))^2 \tag{6}$$

In the classical IF algorithm using the reciprocal function, the weight given to readings of a sensor i when an approximation of the aggregate vector is computed, is the reciprocal of its estimated variance. If in any iteration of the IF algorithm such approximate aggregate vector gets close to readings of a particular sensor, since the reciprocal has a pole at 0, the algorithm gives to that sensor weights converging to 1 and to all other sensors converging to 0, which results in suboptimal performance. Such situations cannot arise in our model because the credibility of a sensor is derived from a comparison with other sensors' readings, rather than the current value of the aggregate vector. Thus, regularisation, which degrades the performance of IF algorithms, is simply not needed in our method.

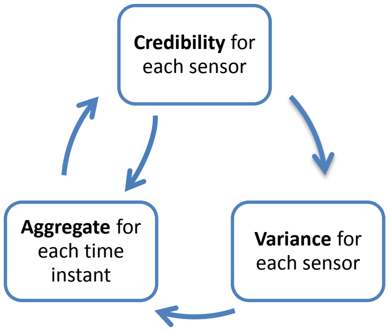

Credibility for each sensor
Aggregate for each time instant
Variance for each sensor

Fig. 1: Recursive relationship among credibility, aggregate and variance.

C. Computing Credibility and Variance

In this section, we describe an algorithm to find the credibility and variance values of sensors. Note that from equations (4), (5) and (6) we have recursive definitions for credibility, aggregate and variance concepts, as shown in Fig. 1. In this figure, arrows show the dependency between the concepts. Clearly, the credibility value of a sensor depends on both the aggregate values and variances of sensors; the aggregate values are computed using the credibility values of the sensors, and the variance of a sensor is measured by the distance of its readings from the aggregate values.

We denote the credibility, aggregate and variance values at iteration l by $\mathrm{cr}^{(l)}(i)$, $\mathrm{rep}^{(l)}(t)$, and $\mathrm{var}^{(l)}(i)$, respectively. Thus, the equations (4), (5) and (6) can be re-written as:

$$\mathrm{cr}^{(l+1)}(i) = \left(\prod_{\substack{j=1 \\ j \neq i}}^{n} \frac{1}{\sqrt{2\pi \mathrm{var}^{(l)}(j)}} e^{\frac{-\frac{1}{m}\sum_{t=1}^{m}(x_i^t - \mathrm{rep}^{(l)}(t))^2}{2\mathrm{var}^{(l)}(j)}} \right)^{\frac{1}{n-1}} \tag{7}$$

$$\mathrm{rep}^{(l+1)}(t) = \sum_{i=1}^{n} \frac{\mathrm{cr}^{(l+1)}(i)}{\sum_{k=1}^{n} \mathrm{cr}^{(l+1)}(k)} x_i^t \tag{8}$$

$$\mathrm{var}^{(l+1)}(i) = \frac{1}{m} \sum_{t=1}^{m} (x_i^t - \mathrm{rep}^{(l+1)}(t))^2 \tag{9}$$

We initialise the variances of all sensors with an identical value, computed as the mean of total variance of all readings.

Lemma 2 (Total Variance)**.** *Let $\bar{x}^t$ be the sample mean of readings in time t, then the statistic*

$$S(t) = \frac{n}{m(n-1)} \sum_{i=1}^{n} \sum_{t=1}^{m} \left(x_i^t - \bar{x}^t\right)^2$$

is an unbiased estimator of the sum of the initial variances of all sensors, $\sum_{i=1}^{n} v_i^{(0)}$.

We presented the proof of Lemma 2 in [12]. Using Lemma 2, the initial value of variance for all sensors is computed as follows:

$$\mathrm{var}^{(0)}(i) = \frac{1}{m \times (n-1)} \sum_{j=1}^{n} \sum_{t=1}^{m} \left(x_j^t - \frac{1}{n} \sum_{k=1}^{n} x_k^t \right)^2 \tag{10}$$

Given these identical initial values for sensors' variances, we obtain the initial aggregate values as the sample mean of the sensors' readings (see Proposition 1):

$$\mathrm{rep}^{(0)}(t) = \frac{1}{n} \sum_{i=1}^{n} x_i^t \tag{11}$$

In addition, the sensors' variances are obtained from Eq. (9) for all iterations $l > 0$. Thus, we transform Eq. (7) for computation of credibility of a sensor as follows:

$$\mathrm{cr}^{(l+1)}(i) = \left(\prod_{\substack{j=1 \\ j \neq i}}^{n} \frac{1}{\sqrt{2\pi \mathrm{var}^{(l)}(j)}} e^{\frac{-\mathrm{var}^{(l)}(i)}{2\mathrm{var}^{(l)}(j)}} \right)^{\frac{1}{n-1}} \quad \text{for all} \quad l > 1. \tag{12}$$

Note that for the first iteration, the credibility is computed using Eq. (7), because the initial variance is obtained from Eq. (10). However, we employed Eq. (12) for subsequent iterations to reduce the computational complexity of our iterative algorithm which is formally investigated in Section III-F.

Algorithm 1 shows our iterative algorithm for computing credibility and variance values.

Algorithm 1 CREDIBILITY AND VARIANCE COMPUTATION.

```
procedure CREDVARIANCECOMPUTATION($X = \{x_i^t\}$)
    Initialize $\mathrm{var}^0(i)$ using (10)
    Initialize $\mathrm{rep}^0(i)$ using (11)
    $l \leftarrow 0$
    repeat
        for each sensor node $i$ do
            if $l = 0$ then
                Compute $\mathrm{cr}^{(l+1)}(i)$ using (7)
            else
                Compute $\mathrm{cr}^{(l+1)}(i)$ using (12)
            end if
        end for
        for each time instant $t$ do
            Compute $\mathrm{rep}^{(l+1)}(t)$ using (8)
        end for
        for each sensor node $i$ do
            Compute $\mathrm{var}^{(l+1)}(i)$ using (9)
        end for
        $l \leftarrow l + 1$
    until reputations and variances have converged
    Return $\overrightarrow{\mathrm{var}}$ and $\overrightarrow{\mathrm{cr}}$
end procedure
```

D. Computing the Final Aggregate

Given the matrix $X = \{x_i^t\}$ where $x_i^t \sim r_t + \mathcal{N}(0, \sigma_i^2)$ and the estimated sensors' variances, we propose to recover $\mathbf{r} = \langle r_1, r_2, \ldots, r_m \rangle$ using an approximate form of the MLE. A similar approximation has been proposed in [6].

We now obtain an estimation which corresponds to the MLE formula for the case of zero mean normally distributed errors, but with estimated rather than true variances. Thus, in the expression for the likelihood function for normally distributed unbiased case, that is,

$$\mathcal{L}_n(r_t) = \prod_{i=1}^{n} \frac{1}{\sigma_i\sqrt{2\pi}} e^{\frac{-(x_i^t - r_t)^2}{2\sigma_i^2}}$$

we replace σ_i^2 by the obtained variance $\mathrm{var}(i)$ from the above iterative procedure. Moreover, by differentiating the above formula with respect to r_t and setting the derivative equal to zero we get

$$r_t = \sum_{i=1}^{n} \frac{\frac{1}{\mathrm{var}(i)}}{\sum_{k=1}^{n} \frac{1}{\mathrm{var}(k)}} x_i^t \quad \text{for all} \quad t = 1, \cdots, m. \tag{13}$$

Eq. (13) provides an estimate of the true value of the signal measured as a weighted average of sensors readings, with the readings given a weight inversely proportional to their estimated variances. See [12] for more details.

E. Node Compromise Detection and Revocation

Upon computing the aggregate values by Eq. (13), we carry out a novel node compromise detection and revocation method based on an analysis of the features of error distribution of the sensor nodes. In the existing approaches, compromised nodes are usually detected as outliers from some form of average of all readings. Instead, we propose a finer analysis based on a sequence of sensor readings, by considering how differences between readings of the individual sensor nodes and the estimate obtained by Eq. (13) are distributed. The main idea behind our method is that, while faulty or compromised nodes might skew the estimate, their action can only make non-compromised sensor appear biased, but the variability of such sensors around such a value will still have a distribution close to a normal distribution; on the other hand, the difference between the values provided by the compromised nodes will have highly non-normal distribution, reflecting their essentially non-stochastic behaviour. Thus, we assume a node with a non-Gaussian error distribution is likely to be a compromised node.

Accordingly, we employ a hypothesis testing method to assess the normality of the obtained error values for each sensor node. Thus, let $\mathbf{e}_s = \{e_s^t \ : \ t = 1, \cdots, m\}$ be the vector of error terms for a sensor s, defined as

$$\mathbf{e}_s = \mathbf{x}_s - \mathbf{r}$$

The problem of deciding whether a sensor node s is compromised can be formulated as a hypothesis testing problem with null and alternative hypotheses as follows:

- Null hypothesis $\mathbf{H_0}$: The sequence of errors $\mathbf{e}_s$ is drawn from a Normal distribution.
- Alternative hypothesis $\mathbf{H_1}$: The sequence of errors $\mathbf{e}_s$ is not drawn from a Normal distribution.

In order to detect the compromised sensor nodes, we employ the Kolmogorov-Smirnov test (K-S test) [13] on sample errors of each node. Using the estimates for the sample mean and the sample variance we normalise the errors; the Kolmogorov-Smirnov statistic then quantifies a distance between the empirical distribution of such normalised samples of sensor errors $\mathbf{e}_s$ and the $\mathcal{N}(0,1)$ Normal distribution.

The proposed node compromise detection scheme classifies sensor nodes in two disjoint sets: the set of the compromised, and the set of the benign nodes. We can now re-apply Algorithm 1 on the benign sensors' readings to produce a more accurate estimation of the true value of the signal.

F. Algorithm Complexity

Algorithm 2 shows the CrPr reputation system including the credibility and variance computation (Algorithm 1), final aggregate computation, and node compromise detection.

Algorithm 2 CrPr.

```
 1: procedure CrPr($X = \{x_i^t\}$)
 2:     $\overrightarrow{\mathrm{var}} \leftarrow$ CredVarianceComputation($X$)   ▷ Section III-C
 3:     Compute $\mathbf{r}$ using (13)   ▷ Section III-D
 4:     $compromised \leftarrow \varnothing$
 5:     for each sensor node $s$ do   ▷ Section III-E
 6:         Compute $\mathbf{e}_s \leftarrow \mathbf{x}_s - \mathbf{r}$
 7:         Test the normality of errors $\mathbf{e}_s$ using K-S test
 8:         if $\mathbf{e}_s$ is non-normal then
 9:             $compromised \leftarrow compromised \cup \{s\}$
10:         end if
11:     end for
12:     if $compromised \neq \varnothing$ then
13:         $\hat{X} \leftarrow X - compromised$
14:         $\overrightarrow{\mathrm{var}} \leftarrow$ CredVarianceComputation($\hat{X}$)   ▷
    Section III-C
15:         Compute $\mathbf{r}$ using (13)   ▷ Section III-D
16:         for each sensor node $s \in colluders$ do
17:             $\mathrm{var}(s) \leftarrow \frac{1}{m} \|\mathbf{x}_s - \mathbf{r}\|_2^2$
18:         end for
19:     end if
20:     Return $\overrightarrow{\mathrm{var}}$ and $\mathbf{r}$
21: end procedure
```

We first evaluate the time complexity of Algorithm 1 in the worst case when all sensors report readings for all time instants. We must model the complexity of two main parts of the algorithm including the non-iterative computations (lines 2-3 and the first iteration) and the iterative part (lines 6-19). The complexity of the computation of initial variances, initial aggregate values, and all computations in first iteration can be modeled respectively by $O(n \times m)$, $O(n \times m)$, and $O(n^2 \times m)$. Thus, the complexity of computing initial values and the first iteration of the algorithm is dominated by the complexity of line 8 which is in $O(n^2 \times m)$. Although this looks expensive, for most real-world datasets such as the SensorScope [14], the number of edges is more or less linear in the number of time instants. Thus, the complexity can be reduced to $O(n \times m)$.

The complexity of the iterative part of Algorithm 1 depends on the complexity of credibility, aggregate, and variance computations which are in $O(n^2)$, $O(n \times m)$, and $O(n \times m)$, respectively. Accordingly, each iteration in our algorithm requires a total $O(n^2 + n \times m) = O(n \times m)$ time, and for k iterations, the total running time for the iteration part of the algorithm is $O(k \times n \times m)$. By accumulating the complexity of the two main parts of Algorithm 1, the complexity of the algorithm in the worst case is in $O(k \times n \times m)$.

In the worst case, the complexity of our MLE-like aggregation method is in $O(n \times m)$. Similarly, the complexity of the node compromise detection method is in $O(n \times m)$. Therefore, the CrPr algorithm totally runs in $O(k \times n \times m)$ in the worst case when the readings matrix is dense.

IV. Experiments

The objective of our experiments is to evaluate the robustness of our approach for estimating the true value of the signal based on the sensor readings in the presence of faults and attacks.

A. Experimental Environment

We conducted our experiments by using both the SensorScope [14] as a real-world dataset and generating synthetic datasets with parameters similar to the real-world dataset. The deployed WSN in the SensorScope project consists of 23 nodes that are capable of measuring temperature. We selected a dataset that was collected every 2 minutes over 43 days by 23 nodes.

For generating the synthetic datasets, we exploit the statistical parameters of one day readings from the SensorScope dataset. If not mentioned otherwise, we generate the synthetic datasets according to the parameters listed in TABLE II. The program code has been written in MATLAB R2012b.

TABLE II: Experimental parameters for synthetic datasets.

Parameter	Value
n	23
m	720
Convergence threshold	10^{-12}
Number of repeat	100
True value of the signal	$\mathrm{f}(t) = 20 + \mathrm{Sin}\left(2\pi\frac{t}{m} - \frac{\pi}{2}\right)$
Significance level in K-S test	$\alpha = 0.05$

In all experiments, we compare the CrPr algorithm against three other IF techniques proposed for reputation systems. For all parameters of other algorithms used in the experiments, we set the same values as used in the original papers. TABLE III shows a summary of aggregation and discriminant functions for all of these three different IF methods. We also consider the improvement for *dKVD-Reciprocal* recently proposed in [6], and we call it *Robust-Reciprocal*. Moreover, we use the Root Mean Square (RMS) error as the accuracy comparison metric in all experiments, which is as follows:

$$RMS\ Error = \sqrt{\frac{\sum_{t=1}^{m}(r_t - \hat{r_t})^2}{m}} \tag{14}$$

where r_j and $\hat{r_j}$ denote the true value and the estimated value for time instant t, respectively.

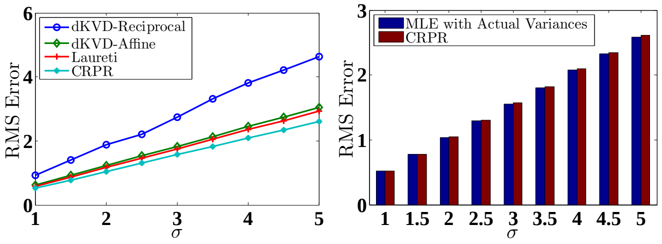
B. Accuracy without Attacks

In order to evaluate Property ((A3)) (see Section II-D) for the CrPr algorithm, in the first batch of experiments we assume that there are no malicious sensor nodes. Thus, the sensors' errors are fully stochastic; we generate the errors of sensors with zero-mean Gaussian distributions. In order to evaluate the performance of CrPr algorithm in comparison with the performence of existing IF algorithms, we consider unbiased errors with different variances for sensor nodes. We have chosen to present the case with the error of a sensor s at time t given by $e_s^t \sim \mathcal{N}(0, s \times \sigma^2)$, considering different values for the baseline sensor variance σ^2. Fig. 2(a) reports the performance of CrPr algorithm for estimating the true value of the signal as well as the performance of other IF algorithms. The results in this experiment show that, the performance of our approach is superior to other IF algorithms as it has a smaller RMS error.

TABLE III: Summary of different IF algorithms.

Name	Discriminant Function
dKVD-Reciprocal [3]	$w_i^{l+1} = \left(\frac{1}{m}\left\|\mathbf{x}_i - \mathbf{r}^{l+1}\right\|_2^2\right)^{-1}$
dKVD-Affine [3]	$w_i^{l+1} = 1 - k\frac{1}{m}\left\|\mathbf{x}_i - \mathbf{r}^{l+1}\right\|_2^2$
Laureti [5]	$w_i^{l+1} = \left(\frac{1}{m}\left\|\mathbf{x}_i - \mathbf{r}^{l+1}\right\|_2^2\right)^{-\frac{1}{2}}$
Robust-Reciprocal [6]	same as the *dKVD-Reciprocal*

Fig. 2(b) reports the accuracy results of the CrPr algorithm and the information theoretic limit for the minimal variance provided by the CRLB, achieved, for example, using the MLE with the *actual, exact variances* of sensors, which are NOT available to our algorithm. As one can see from these results, our proposed approach closely matches the minimal possible variance coming from the information theoretic lower bound; thus the CrPr algorithm meets Property ((A3)).



(a) Unbiased error (b) Unbiased error

Fig. 2: Accuracy for *No Attack* scenarios.

C. Robustness Against Simple Attacks

In order to evaluate the robustness of our algorithm against some simple attacks (Requirement (A1) in Section II-D), we use two types of malicious behaviour proposed in [3], [15] over a subset of the SensorScope dataset: random readings and a promoting attack. For both of the simple attacks, we selected a subset of readings from the SensorScope dataset as the baseline dataset which includes the temperature values measured by all 23 sensors for a day in October 23, 2007.

In both scenarios, we considered 20% of the sensor nodes as malicious nodes. For the random rating attack, we modify the readings of the malicious nodes by injecting uniformly random real values in the range of [-5,5]. In the promoting attack scenario [1], malicious nodes always report the lowest temperature value of -5 except for their preferred time instants, where they report the highest temperature value of +5. We also assume that the malicious nodes are promoting only for 10% of all time instants.

Let $\mathbf{r}$ and $\tilde{\mathbf{r}}$ be the aggregate vectors before and after injecting each scenario, respectively. In the proposed reputation system, the vectors are the results of Equation (13). TABLE IV reports the values of the 1-norm difference between these two vectors, $||\mathbf{r} - \tilde{\mathbf{r}}||_1 = \sum_{t=1}^{m} |r_t - \tilde{r}_t|$ as well as the percentage of the normalized absolute errors, $\frac{100}{m}||\mathbf{r} - \tilde{\mathbf{r}}||_1$ for CrPr algorithm along with the other IF algorithms. Clearly, all of the IF algorithms are significantly more robust than *Average*. In addition, the CrPr algorithm provides significantly higher accuracy than other methods for both attack scenarios. This can be due that the proposed algorithm effectively filters out the contribution of the malicious nodes. Note that the malicious nodes in the random attack can be detected by our node compromise detection module when they present a non-Normal error behaviour. On the other hand, since the malicious nodes in the promoting attack report outlier readings for 90% of the instants, the algorithm will assign low credibilities. As a consequence, the credibility values that they can achieve from the majority of nodes (benign nodes) is very low, according to our credibility computation. Clearly, the malicious nodes can only skew the aggregation results if they can establish a community larger than the number of benign nodes.

The results in TABLE IV also show that the accuracy of two algorithms, *dKVD-Reciprocal* and *Robust-Reciprocal* are identical. The reason is that the later algorithm which is a robust version of the former one [6], can consolidate the former one only if it stops in its pole point. However, in this experimental settings for random and promoting attacks, there is a unique stationary point for the *dKVD-Reciprocal* algorithm which leads to the convergence of the algorithm in such point regardless of the initial values. In other words, using a robust initial trustworthiness for the IF algorithms can only address the pole point problem for them [6].

TABLE IV: Absolute errors (percentage of the normalized absolute errors) for simple attacks

	$\|\mathbf{r}-\tilde{\mathbf{r}}\|_1 \left(\frac{100}{m}\,\|\|\mathbf{r}-\tilde{\mathbf{r}}\|\|_1\right)$	
Algorithm	Random Readings	Promoting Attack
Average	142.73 (20%)	797.83 (111%)
dKVD-Reciprocal	51.96 (7%)	76.90 (11%)
dKVD-Affine	102.29 (14%)	111.64 (16%)
Laureti	62.41 (9%)	197.09 (27%)
Robust-Reciprocal	51.96 (7%)	76.90 (11%)
CrPr	14.41 (2%)	25.73 (4%)

D. Robustness Against Collusion Attacks

Rezvani et al. [6] recently proposed a collusion attack against existing IF algorithms. In this attack, the adversary first compromises c ($c < n$) nodes. The attacker then uses the first $c - 1$ malicious nodes to report far readings in order to skew the simple average of all readings. The adversary also falsifies the last node's reading by injecting a value very close to this skewed average. This collusion attack makes the IF algorithm converge to a wrong stationary point.

In order to investigate the accuracy of the CrPr algorithm with respect to this collusion attack, we first implement the attack in the same settings presented in the previous experiments over the SensorScope dataset. TABLE V reports the results for this experiment. We also synthetically generate several datasets with various numbers of compromised nodes (c). Rezvani et al. [6] showed that the *dKVD-Affine* method is the least sensitive IF algorithm to the attack. Therefore in this paper we compare the accuracy of the CrPr algorithm along with the accuracy of the *dKVD-Affine* method in the presence of the collusion attack, as shown in Fig. 3.

One can see from TABLE V that the CrPr is superior to other methods in terms of accuracy against the collusion attack. Moreover, both *dKVD-Reciprocal* and *Laureti* have been completely compromised by the attacker as they approximately provide the accuracy of the *Average*, with around 100% error rate. The reason is that both of the algorithms are using a kind of reciprocal discriminant function containing a pole. Furthermore, while the *dKVD-Affine* presents a better accuracy against the collusion attack, it still generated around 18% error. The table also shows that the CrPr is superior to the previously improved IF algorithms (*Robust-Reciprocal*). This can be explained by the fact that the previous improvement only consolidates the IF algorithm against the attack. In other words, the difference between the accuracy of the CrPr and *Robust-Reciprocal* can be described by the similar difference of their accuracy in *No Attack* experiments.

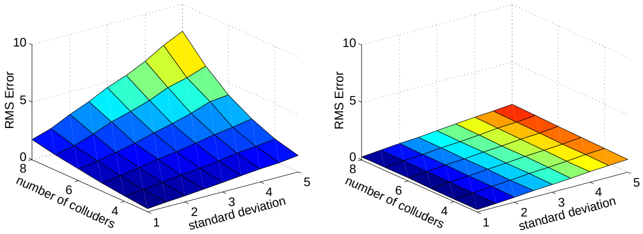
Fig. 3(a) shows that the accuracy of the *dKVD-Affine* significantly decreases as the number of compromised nodes and sensors' variances increase. In this case, the adversary has better chance to skew the simple average using a greater number of compromised nodes. Moreover, the first $c - 1$ malicious nodes have succeed report skewed values without being detected as as outliers using existing detection methods. On the other hand, the accuracy of the CrPr declines very slightly in the similar circumstances, as shown in Fig. 3(b). This may be due to the CrPr accurately filtering out the outlier reports by deriving the credibility of sensor nodes from both the current aggregate values and sensors' variances. Moreover, by comparing the accuracy of the CrPr algorithm in Fig. 3 with the results from *No Attack* experiment in Fig. 2, we can argue that our reputation system is robust against the collusion attack scenario. The reason is that our approach not only provides high accuracy against this attack, it also actually approximately reaches the accuracy of the scenarios without any false data by colluders.

TABLE V: Absolute errors (percentage of the normalized absolute errors) for collusion attacks

	$\|\mathbf{r}-\tilde{\mathbf{r}}\|_1 \left(\frac{100}{m}\,\|\|\mathbf{r}-\tilde{\mathbf{r}}\|\|_1\right)$
Algorithm	Collusion Attack
Average	730.25 (101%)
dKVD-Reciprocal	714.34 (99%)
dKVD-Affine	131.96 (18%)
Laureti	741.31 (102%)
Robust-Reciprocal	69.90 (9%)
CrPr	51.06 (7%)

E. Readings Resolution and Clustered Variances

Medo and Wakeling [16] showed that the data resolution has significant impact on the accuracy of the IF algorithms. Thus, we employ their methodology to investigate the accuracy of CrPr algorithm over the low resolution readings and different variance patterns. In this section, we employ a *clustered* pattern for sensors variances by uniformly randomly selecting sensors' variances from the distribution $U[\sigma_{min}; \sigma_{max}]$ by considering different values for σ_{min} and σ_{max}.

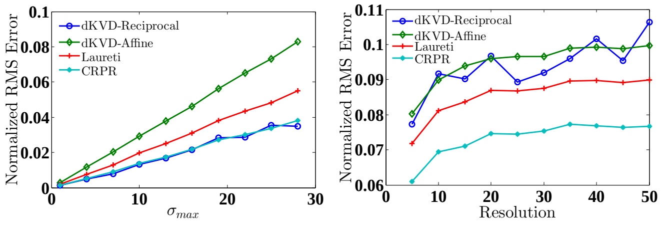


(a) *dKVD-Affine* (b) CrPr

Fig. 3: Accuracy with respect to the collusion attack.

For these experiments, we created synthetic datasets with parameters in TABLE II. The scale of the true value of the signal is in the range of $R = [5, 50]$. Thus, for each value of R, we generate the true value of the signal by using sine function $\mathrm{f}(t) = R + 5 \times \mathrm{Sin}\left(2\pi \frac{t}{m} - \frac{\pi}{2}\right)$. Moreover, we evaluate a normalized RMS error, $RMS/(R - 1)$ (see Eq. (14) for RMS) for each experiment.

For the first experiment, we set $R = 30$ and $\sigma_{min} = 0$, and vary the value of σ_{max} in the range of $[1, 29]$. By choosing such a range at the worst case, a highest noisy sensor with $\sigma_i = \sigma_{max} = 29$ could potentially report a very low temperature for the highest temperature environment circumstance, and vice versa. Fig. 4(a) shows the accuracy of CrPr algorithm along with the accuracy of the other IF algorithms for this experiment. We observe that CrPr is the least sensitive to the increasing error level, maintaining the lowest normalized RMS error. The results also validate the high sensitivity of both *dKVD-Reciprocal* and *dKVD-Affine* by increasing the sensors' error in general and particularly by changing the error pattern, as discovered by [16].



(a) (b)

Fig. 4: Accuracy with clustered variances and different resolutions.

In order to investigate the effect of changing the readings' resolution, we set $\sigma_{min} = \sigma_{max}/8$ and $\sigma_{max} = R - 1$, and vary the value of R in the range of $[5, 50]$, so that the maximum possible sensor errors cover the readings' scale. Fig. 4(b) shows the performance of CrPr algorithm along with the performance of other IF algorithms for this experiment. Medo and Wakeling [16] reported that *Laureti* algorithm is the least sensitive algorithm when the resolution scale changes. Comparing the trends of the RMS error in Fig. 4(b) shows that this experiment not only validates their results, but it demonstrates that CrPr algorithm exactly achieves the flexibility of *Laureti* while at the same time achieving the lowest RMS error.

F. Node Compromise Detection Performance

As we described, the aggregate values obtained in the first round of CrPr are used to classify the sensor nodes in two groups: compromised and benign nodes. Thus, the detection performance of the module is evaluated by its accuracy, precision, and recall measurements obtained from the confusion matrix, as shown in TABLE VI. The accuracy is the proportion of the total number of predictions that were correct; the recall or true positive rate is the proportion of compromised nodes that were correctly detected; precision is the proportion of the detected compromised nodes that were correct. These performance metrics are calculated based on a confusion matrix in TABLE VI as follows:

TABLE VI: Confusion matrix for collusion detection.

Actual Class	Predicted Class	
	Compromised	Benign
Compromised	True Positive (TP)	False Negative (FN)
Benign	False Positive (FP)	True Negative (TN)

TABLE VII: Performance of our collusion detection module.

	Accuracy	Precision	Recall
No Attack	95.27	0	0
Random Readings	99.19	96.24	100
Promoting Attack	99.22	96.44	100
Collusion Attack	98.71	96.35	98.32

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \tag{15}$$

$$Precision = \frac{TP}{TP + FP} \times 100 \tag{16}$$

$$Recall = \frac{TP}{TP + FN} \times 100 \tag{17}$$

TABLE VII shows the performance results of the node compromise detection module for the previous scenarios based on the average values of three metrics: accuracy, precision and recall for each experiment. We can see that the accuracy value for *No Attack* and the precision values for attack scenarios are lower than other metrics. It means that the node compromise detection provides some false positive error. We believe that though the current false positive rate (around 96%) is an acceptable rate for the module, choosing a better value for the significance level used in K-S test would improve the false positive rate of the module. Note that we can only investigate the accuracy metric for *No Attack* scenarios, because there is no compromised node for them and therefore $TP = 0$. Consequently, the precision and recall measurements are zero for all the cases in the scenarios.

G. Analysis of Error and Convergence

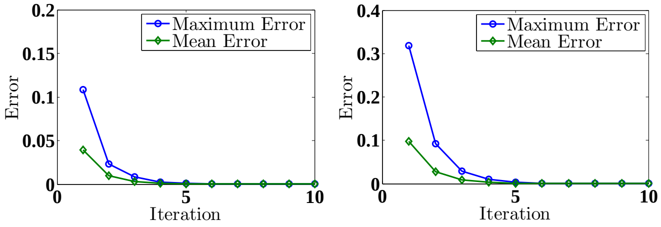
In this section, we perform a set of experiments to analyze the properties of our iterative algorithm in terms of error and convergence. Thus, we investigate two types of errors for both credibility and variance values computed in each iteration of CrPr algorithm over the SensorScope dataset. For each of credibility and variance values, we define the maximum error by choosing the worst-case error for all sensor nodes. Therefore, the maximum errors at iteration l is computed as follows:

$$error_{cr}^{(l)} = \max_i \left|\mathrm{cr}^{(\infty)}(i) - \mathrm{cr}^{(l)}(i)\right|$$
$$error_{var}^{(l)} = \max_i \left|\mathrm{var}^{(\infty)}(i) - \mathrm{var}^{(l)}(i)\right|$$

We also define the mean error of credibility and variance values over all sensors as follows:

$$error_{cr}^{(l)} = \frac{1}{n}\sum_{i=1}^{n} \left|\mathrm{cr}^{(\infty)}(i) - \mathrm{cr}^{(l)}(i)\right|$$
$$error_{var}^{(l)} = \frac{1}{n}\sum_{i=1}^{n} \left|\mathrm{var}^{(\infty)}(i) - \mathrm{var}^{(l)}(i)\right|$$

Fig. 5 illustrates how the aforementioned errors decline for both credibility and variance values. The figure is plotted only for the first 10 iterations because the algorithm mostly converged after that. For all experiments, we set convergence threshold with an error $\left\|\mathrm{var}^{(l+1)} - \mathrm{var}^{(l)}\right\|_2$ less than 10^{-12}. Fig. 5 shows that the error decreases exponentially in the CrPr algorithm.



(a) Credibility

(b) Variance

Fig. 5: Convergence and error of CrPr algorithm.

V. Discussion

Our reputation system uses the inherent redundancy of data collected by sensor nodes in WSNs in order to assure the trustworthiness of the collected data and identify the malicious nodes. In this section, we discuss how the CrPr algorithm meets four requirements introduced in Section II-D. We believe that, in general, every reputation system should be assessed with respect to those requirements.

For Requirement (A1), the experiment results of the random and promoting attacks show that the CrPr places the harshest sanction against outlier readings which diverge from the estimated aggregate values. Moreover, the results in the *No Attack* experiments validate that our method is statistically sound as it achieves the variance of CRLB which is theoretically minimal possible (Requirement (A3)).

Although we have assumed that the distribution of the noise in sensors is modeled by zero-mean Gaussian distribution, the CrPr has no prior knowledge about the sensors' variances. We note that the approach can be easily adopted for other distributions by substituting the new probability distribution function in Eq. (3) (Requirement (A4)).

Moreover, Our reputation system is motivated by the recent collusion attack introduced against the existing IF algorithms [6]. We have shown that the CrPr is robust against such an attack as well as a collusion promoting attack (Requirement (A2)). In the future, we plan to design a mathematical mechanism for finding the optimal strategy of an attacker, given the number of compromised nodes he has access to for distorting the aggregate values.

A large variety of WSN applications are characterized by real-time data streaming which introduces new challenges and runtime design options for trust and reputation computation [17]. We have successfully extended the CrPr to data streaming to achieve an online reputation system for WSNs [18].

It is to be noted that while our reputation system has been described with WSNs, we believe that it is straightforward to extend the solution to other distributed systems. Accordingly, we plan to extend the idea of credibility propagation for developing a robust reputation system to e-commerce applications [19] and participatory networks [20].

VI. Related Work

There are three areas of work related to our research: IF algorithms, reputation systems for WSNs, and secure data aggregation with node compromised detection in WSNs.

Several papers have proposed IF algorithms for trust and reputation systems [4], [3], [5]. Such IF algorithms consider simple cheating attacks. However, none of them take into account sophisticated malicious scenarios such as collusion attacks [6]. We compared the robustness of our approach with some of the existing IF methods as well as the method proposed in [6] against this collusion attack.

Our work is also closely related to the trust and reputation systems in WSNs. Lim et al. [21] proposed a cyclic framework based on an interdependency relationship between network nodes and data items for assessing their trust scores. Sun et al. [22] proposed a combination of trust mechanism, data aggregation, and fault tolerance to enhance data trustworthiness in wireless multimedia sensor networks which considers both discrete and continuous data streams. Tang et al. [23] proposed a trust framework for sensor networks in Cyber Physical System (CPS). An example of deployment of sensors in CPS is a battle-network system in which the sensor nodes are employed to detect approaching enemies and send alarms to a command center. CORE [24] and CONFIDENT [25] schemes have been proposed to Ad-hoc networks to enable nodes to detect cooperative and non-cooperative nodes. Although fault detection problems have been addressed by applying reputation systems in the above research, none of them take into account sophisticated malicious scenarios such as collusion attacks.

Reputation and trust concepts can be used to address the node compromised detection and secure data aggregation problems in WSNs. Ho et al. [26] proposed a framework to detect compromised sensor nodes in WSN and then apply a software attestation for the detected nodes. Ozdemir et al. [27] proposed a false aggregator detection scheme by the idea of using a number of monitoring nodes which are running aggregation operations and providing the integrity of the message using message authentication code (MAC) values. High computation and transmission cost required for MAC-based integrity checking in this scheme makes it unsuitable for deployment in WSN. Lim et al. [15] proposed a game-theoretical defense strategy to protect sensor nodes and to guarantee a high level of trustworthiness for sensed data. Moreover, a number of approaches have been proposed for detecting false aggregation operations [9], [10], [11], that is, when data aggregator nodes obtain data from source nodes and produce wrong aggregated values. Such approaches neither address the problem of false data being provided by the data sources nor the problem of collusion. Although the aforementioned research takes into account false data injection for a number of simple attacks, to the best of our knowledge, no existing work addresses this issue in the case of a sophisticated attack of colluding adversaries compromising a number of nodes in a manner which employs high level knowledge about the data aggregation algorithm used.

VII. Conclusions

In this paper, we introduced a novel collaborative reputation system which not only performs accurately in the presence of different types of faults and simple attacks such as random ratings and promoting attack, but also is robust against the sophisticated collusion attacks to which most of the existing IF algorithms are vulnerable. In the future we plan to implement our approach in a deployed sensor network. We also plan to extend the idea of credibility propagation to propose a decentralized and privacy-preserving reputation system.

References

[1] K. Hoffman, D. Zage, and C. Nita-Rotaru, "A survey of attack and defense techniques for reputation systems," *ACM Computing Surveys (CSUR)*, vol. 42, no. 1, pp. 1:1–1:31, Dec. 2009.

[2] A. Jøsang and J. Golbeck, "Challenges for robust trust and reputation systems," in *Proceedings of the 5 th International Workshop on Security and Trust Management*, Saint Malo, France, 2009.

[3] C. de Kerchove and P. Van Dooren, "Iterative filtering in reputation systems," *SIAM J. Matrix Anal. Appl.*, vol. 31, no. 4, pp. 1812–1834, 2010.

[4] A. Galletti, G. Giunta, and G. Schmid, "A mathematical model of collaborative reputation systems," *International Journal of Computer Mathematics*, vol. 89, no. 17, pp. 2315–2332, Nov. 2012. [Online]. Available: http://dx.doi.org/10.1080/00207160.2012.715641

[5] P. Laureti, L. Moret, Y.-C. Zhang, and Y.-K. Yu, "Information filtering via Iterative Refinement," *EPL (Europhysics Letters)*, vol. 75, pp. 1006–1012, 2006.

[6] M. Rezvani, A. Ignjatovic, E. Bertino, and S. Jha, "Secure data aggregation technique for wireless sensor networks in the presence of collusion attacks," *IEEE Transactions on Dependable and Secure Computing*, vol. 12, no. 1, pp. 98–110, January 2015.

[7] D. Wagner, "Resilient aggregation in sensor networks," in *Proceedings of the 2nd ACM workshop on Security of ad hoc and sensor networks*, ser. SASN '04. New York, NY, USA: ACM, 2004, pp. 78–87. [Online]. Available: http://doi.acm.org/10.1145/1029102.1029116

[8] B. Awerbuch, R. Curtmola, D. Holmer, C. Nita-rotaru, and H. Rubens, "Mitigating byzantine attacks in ad hoc wireless networks," Department of Computer Science, Johns Hopkins University, Tech. Rep., 2004.

[9] H. Chan, A. Perrig, and D. Song, "Secure hierarchical in-network aggregation in sensor networks," in *Proceedings of the 13th ACM Conference on Computer and Communications Security*, ser. CCS '06. New York, NY, USA: ACM, 2006, pp. 278–287.

[10] Y. Yang, X. Wang, S. Zhu, and G. Cao, "SDAP: a secure hop-by-hop data aggregation protocol for sensor networks," in *MobiHoc*, 2006, pp. 356–367.

[11] S. Roy, M. Conti, S. Setia, , and S. Jajodia, "Secure data aggregation in wireless sensor networks," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 3, pp. 1040–1052, 2012.

[12] M. Rezvani, "Trust-based data aggregation for wsns in the presence of faults and collusion attacks," Ph.D. dissertation, UNSW Australia, 2015.

[13] L. Wasserman, *All of statistics : a concise course in statistical inference*. New York: Springer, 2010.

[14] "The SensorScope lausanne urban canopy experiment (LUCE) project," *Data set available at: http://sensorscope.epfl.ch/index.php/LUCE*, 2006.

[15] H.-S. Lim, G. Ghinita, E. Bertino, and M. Kantarcioglu, "A game-theoretic approach for high-assurance of data trustworthiness in sensor networks," in *Data Engineering (ICDE), 2012 IEEE 28th International Conference on*, April 2012, pp. 1192 –1203.

[16] M. Medo and J. R. Wakeling, "The effect of discrete vs. continuous-valued ratings on reputation and ranking systems," *EPL (Europhysics Letters)*, vol. 91, no. 4, p. 48004, 2010.

[17] A. Etuk, T. J. Norman, C. Bisdikian, and M. Srivatsa, "TAF: A trust assessment framework for inferencing with uncertain streaming information," in *5th International Workshop on Information Quality and Quality of Service for Pervasive Computing*, 2013, pp. 475–480.

[18] M. Rezvani, A. Ignjatovic, E. Bertino, and S. Jha, "A trust assessment framework for streaming data in wsns using iterative filtering," in *Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2015 IEEE Tenth International Conference on*. IEEE, April 2015, pp. 1–6.

[19] F. Fouss, Y. Achbany, and M. Saerens, "A probabilistic reputation model based on transaction ratings," *Information Sciences*, vol. 180, no. 11, pp. 2095–2123, Jun. 2010. [Online]. Available: http://dx.doi.org/10.1016/j.ins.2010.01.020

[20] K. L. Huang, S. S. Kanhere, and W. Hu, "Are you contributing trustworthy data?: The case for a reputation system in participatory sensing," in *Proceedings of the 13th ACM International Conference on Modeling, Analysis, and Simulation of Wireless and Mobile Systems*, ser. MSWIM '10. New York, NY, USA: ACM, 2010, pp. 14–22.

[21] H.-S. Lim, Y.-S. Moon, and E. Bertino, "Provenance-based trustworthiness assessment in sensor networks," in *Proceedings of the Seventh International Workshop on Data Management for Sensor Networks*, ser. DMSN '10, 2010, pp. 2–7.

[22] Y. Sun, H. Luo, and S. K. Das, "A trust-based framework for fault-tolerant data aggregation in wireless multimedia sensor networks," *IEEE Transactions on Dependable and Secure Computing*, vol. 9, no. 6, pp. 785–797, Nov. 2012.

[23] L.-A. Tang, X. Yu, S. Kim, J. Han, C.-C. Hung, and W.-C. Peng, "Tru-Alarm: Trustworthiness analysis of sensor networks in cyber-physical systems," in *Proceedings of the 2010 IEEE International Conference on Data Mining*, ser. ICDM '10, 2010, pp. 1079–1084.

[24] P. Michiardi and R. Molva, "CORE: A collaborative reputation mechanism to enforce node cooperation in mobile ad hoc networks," in *Proceedings of the IFIP TC6/TC11 Sixth Joint Working Conference on Communications and Multimedia Security: Advanced Communications and Multimedia Security*, 2002, pp. 107–121.

[25] S. Buchegger and J.-Y. Le Boudec, "Performance analysis of the CONFIDANT protocol," in *Proceedings of the 3rd ACM International Symposium on Mobile Ad Hoc Networking and Computing*, ser. MobiHoc '02. New York, NY, USA: ACM, 2002, pp. 226–236.

[26] J.-W. Ho, M. Wright, and S. K. Das, "ZoneTrust: Fast zone-based node compromise detection and revocation in wireless sensor networks using sequential hypothesis testing," *IEEE Transactions on Dependable and Secure Computing*, vol. 9, no. 4, pp. 494 –511, July-August 2012.

[27] S. Ozdemir and H. Çam, "Integration of false data detection with data aggregation and confidential transmission in wireless sensor networks," *IEEE/ACM Transactions on Networking (TON)*, vol. 18, no. 3, pp. 736–749, Jun. 2010.