CORA: Correlation-based Resilient Aggregation in Sensor Networks

Péter Schaffer   István Vajda
Laboratory of Cryptography and Systems Security (CrySyS)
Department of Telecommunications
Budapest University of Technology and Economics, Hungary
{schaffer, vajda}@crysys.hu

ABSTRACT

In this paper we consider the problem of resilient data aggregation, namely, when aggregation has to be performed on a compromised sample. We present a statistical framework that is designed to mitigate the effects of an attacker who is able to alter the values of the measured parameters of the environment around some of the sensor nodes. Our proposed framework takes advantage of the naturally existing correlation between the sample elements, which is very rarely considered in other sensor network related papers. The algorithms presented are to be applied without assumption on the sensor network's sampling distribution or on the behaviour of the attacker. The effectiveness of the algorithms is formally evaluated.

Categories and Subject Descriptors: D.4.6 [**Operating Systems**]: [Security and Protection]

General Terms: Algorithms, Design, Security

Keywords: Sensor networks, Resilient aggregation, Correlation, Attack detection

1. INTRODUCTION

Sensor networks are considered to become the most powerful monitoring applications ever. These networks consist of a huge number of tiny sensors which operate unattendedly. This, combined with the fact that the sensors are generally not tamper-resistant, results in a high vulnerability of these networks. Simply lighting a lighter (or flashing with a flashlight, etc.) near to a thermometer (or photometer) sensor node is enough to disturb the measurements of it. This is a serious threat that cannot be circumvented by cryptographical methods, since the nodes that measure the disturbed phenomena generate cryptographically sound messages. It is especially annoying if one wants to aggregate the measurements of the sensor nodes. As the deployer of the network is usually not interested in the measurements apiece, this can be considered as the general case.

This problem has been already noticed, e.g., by Wagner in [11], and afterwards considered by Buttyán et al. in [2]. However, both of these papers make the simplifying assumption of the sensor network to produce independent and identically distributed measurements. In reality, the measurements made by the sensors always have some kind of relationship among them. This relationship can be either temporal correlation, or spatial correlation. In this paper, we focus on spatial correlation (i.e., when the nodes' physical proximity is the basis of the relationship). Spatial correlation can be exploited to cross-check the sample, testing whether there is an (environment altering) attack or not. This naturally existing characteristic of the sample helps in improving the attack detection capabilities. Furthermore, adding correlation to the model of the sensor networks is a significant step towards having a realistic data processing model of these networks.

Hereinafter, we introduce our sensor network model that is able to track correlation, and we also introduce a novel resilient data aggregation scheme developed for sensor networks. In this scheme, we show how spatial correlation can be exploited. Moreover, we show how considering correlation improves the attack detection capabilities of our attack detection scheme. As far as we know, there exists no research paper that would exploit correlation in order to defend an attacker in sensor networks.

The rest of the paper is organized as follows: In Section 2, we present the related papers. In Section 3 we introduce our model of the sensor network. In Section 4, our novel correlation-based resilient aggregation approach is detailed and its efficiency is analyzed. Then, in Section 5, some emerging questions are answered. Finally, in Section 6, we conclude our work and propose some interesting future research topics.

2. RELATED WORK

Today's literature on sensor networks usually neglects an important feature: the correlation among the elements of the sample measured by the sensor network. However, in reality, these measurements are always correlated. This correlation can be exploited in many ways.

There are papers that deal with with data gathering considering correlation [8, 5, 4, 6, 13]. Another set of papers is related to both data aggregation and data gathering [10, 9, 1, 7]. Finally, there exist papers that aim at the intersection of data aggregation and data correlation [14, 12]. However, all these papers handle correlation in a simple fashion with the usual aim to lower the energy consumption of the sensor network. Nowadays, however, security in sensor networks is getting more and more important. Handling the security problem of messages that are cryptographically sound, but false in content is a serious problem in this area, as mentioned in Section 1.

Some researchers already considered the problem of such messages that cannot be filtered using solely cryptographical checks [11, 3, 2]. These papers consider the resilience of data aggregation in case of an attacker's activity, but do not consider data correlation. Therefore, as far as we know, there exists no paper that would consider both an attacker who can compromise messages in a non-cryptographical way, and the naturally existing correlation between the measurements of the sensors. In the following sections we introduce a proposal to fill this leak.

3. GENERAL ASSUMPTIONS

An attacker, who is able to alter the measured paremeters of the environment in the proximity of the sensor nodes, represents a serious threat against even the sensor network that is equipped with cryptographical tools. Handling the problem of such an attacker is a must in order to realize security in sensor networks.

3.1 The Attacker Model

The atacker we consider is able to produce some kind of "noise" that is added to the measurements of the sensors. This noise is completely under the control of the adversary, but it is considered to be independent and identically distributed. This attack can totally distort the aggregate considering the commonly used aggregation functions like the average and the min/max. We note that we do not restrict the adversary in the number of sample elements he is able to compromise, but we assume that the adversary's knowledge do not extend to the distribution of the sample produced by the sensor network, neither to the size of the sample gathered by the base station in a given query. Finally, we note that we do not consider any particular distribution for the attacker's noise.

3.2 The Data Model

In our envisioned application the base station collects a sample of measurements from the sensors and tries to aggregate them in a secure way. Each sensor contributes to this sample with its measurement by replying to the base station's query in an encrypted message. Upon reception of the messages the base station decrypts the messages and aggregates their information content. The aggregation is done in two steps: Firstly, the sample is analyzed and a decision is made whether it is compromised or not. Secondly, an aggregation step is performed depending to the previous decision. If there is no attack detected then usual aggregation is performed, otherwise the final output is calculated by extrapolation based on the previous outputs (see Figure 1). This separation of cases helps
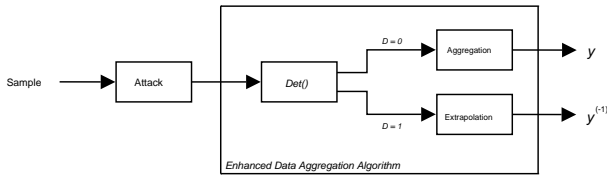
Figure 1: Resilient aggregation scenario including the attacker and the data processing part

us to obtain a significantly smaller distorted output of the aggregation function as having done the aggregation without attack detection.

In order to be able to measure the gain of our approach, we modelled the sensor network to produce measurements that can be represented by identically distributed random variables, but instead of assuming the independence of these random variables we exploit the correlation among them. Therefore, our sensor network data model consists of the following parts. Let n denote the number of sensor readings in the sample, and t denote number of readings compromised by the attacker. X_i is a random variable that stands for the ith uncompromised reading ($0 < i \le n$). The correlation coefficient between two uncompromised readings X_i and X_j is denoted by $r_{X_i, X_j} = r$ ($\forall i, j; \ i \ne j$). G_i is a random variable that denotes the additive noise produced by the attacker (G_i is independent of X_i, $\forall i$). Finally, $Z_i = X_i + G_i$ is a random variable that denotes the compromised sample elements ($0 < i \le t$).

We assume that the sensor network data is normally distributed. The choice of the normal distribution is a common assumption in practice when measurement data is considered. However, we note

that the algorithms we propose in the following sections are applicable to any kind of sampling distributions. The attacker's noise G_i does not have to fit to any parametric or non-parametric distribution.

Since this model handles the dependence of the sensor measurements, it can help us to quantify the power of correlation in attack detection. In the next section we will show how this quantification can be done in a simplified scenario consisting of only two nodes from which at most one is attacked.

4. EXPLOITING CORRELATION IN RESILIENT DATA AGGREGATION

As already mentioned in the previous sections, correlation among sample elements is a naturally existing phenomena which has never been considered so far in research papers related to resilient data aggregation in sensor networks. In this section, we show a method how this correlation can be exploited. We present our idea with the help of a scenario of two nodes which can be the basis of a general solution considering arbitrary number of nodes.

To demonstrate the attack detection possibilities provided by correlation we evaluate the case when there are only two sample elements (i.e., $n = 2$), and there is at most one element that is attacked (i.e., $t \le 1$). Our aim now is to pursue attack detection on this 2-element sample with a small error probability, and then, to pursue data aggregation with a remarkably lowered distorion. Our secondary aim is to show how correlation influences our results calculated for the distortion.

The algorithm that is designed to detect the attack is Algorithm 1. The $Det(x_1, x_2)$ Attack Detection Algorithm randomly chooses

---

Algorithm 1 $Det(x_1, x_2)$ Attack Detection Algorithm

1: Randomly select one element from the sample $\{x_1, x_2\}$ and let the selected element be denoted by x', the remaining one by x''
2: Calculate the $(1 - \alpha)\%$ confidence interval on x'' conditioned on x' according to the p.d.f. $p_{X_1|X_2}(\cdot|x')$
3: **if** x'' is inside this confidence interval **then**
4:     $D = 0$ (* no attack detected *)
5: **else**
6:     $D = 1$ (* attack detected *)
7: **end if**

---

one of the two elements from the sample and computes the $(1-\alpha)\%$ confidence interval for the remaining one conditioned on the chosen one. If the remaining one is inside this confidence interval, then the output of the algorithm is that there is probably no attack ($D = 0$), otherwise the algorithm signals that an attack is detected ($D = 1$).

This straightforward approach already exploits correlation by using the probability distribution function $p_{X_1|X_2}(\cdot|\cdot)$. The knowledge of this means the knowledge of the distribution of X_i and the correlation coefficient r. In most of the cases this can be a realistic assumption about the knowledge since the base station can perform data gathering and can establish an estimation of the sampling distribution (i.e., μ and σ) and the correlation coefficient r just after the deployment of the sensor network when the probability of being already attacked is negligibly small.

The output of Algorithm 1 can be applied in selecting the adequate way of data aggregation. If no attack is indicated then the sample can be handled in the usual way, e.g., its average can be calculated without the fear of obtaining a highly distorted aggregate. Otherwise, one can mitigate the effects of an attacker by handling the sample in a special way. Usually, dropping the compromised sample is the easiest method to apply, while extrapolating the current aggregate from the previous (unattacked) results can guarantee a small distortion without relying on other information. The type

of the extrapolation can suitably be chosen to the characteristics of the data one is going to measure.

This approach is formalized in the Enhanced Data Aggregation Algorithm (Algorithm 2), where output y is the aggregate of the input, while the output denoted by $y^{(-1)}$ is the minimum distortion output when we do not use outlier filtering. $y^{(-1)}$ is usually calculated as an extrapolation based on the output of the previous uncompromised outputs. For example, $y^{(-1)}$ can be the output of the last run of the data aggregation algorithm when the attack detection algorithm detected no attack.

Algorithm 2 Enhanced Data Aggregation Algorithm

1: Take both of the readings and apply the attack detection algorithm $Det(x_1, x_2)$
2: **if** $Det(x_1, x_2)$ indicates an attack **then**
3:     Output $= y^{(-1)}$
4: **else**
5:     Output $= y$
6: **end if**

The output of the Enhanced Data Aggregation Algorithm is interpreted as the aggregate value of the current round. Using the Attack Detection Algorithm and the Enhanced Data Aggregation Algorithm one can notably reduce the distortion of the aggregate compared to the case when aggregation is performed without prior analysis.

4.1 Analysis

To quantify the gain in the distortion of the output of Algorithm 2 we first have to evaluate the error probabilities of Algorithm 1. These probabilities are the false positive (α) and the false negative (β) probabilities. α is the probability of signalling an attack in the unattacked case, while β is the probability of not signalling the attack in the attacked case. In order to be able to define β we fix α to 0.1 (i.e., we tolerate 10% of false alarms). Moreover, for the evaluation we assume that the distribution of G_i is the Gaussian distribution with parameters $\tilde{\mu}$ and $\tilde{\sigma}$ (i.e., $G_i \sim \mathcal{N}(\tilde{\mu}, \tilde{\sigma})$). Here, the choice of the Gaussian distribution simplifies the analysis and its two parameters allows us to consider attacks of significantly different style. Without loss of generality, we further assume that the first sample element is compromised, i.e., $Z_1 = X_1 + G_1$. Since $t = 1$ we can set aside the lower indexes of the symbols corresponding to the attacker, thus $Z = X_1 + G$. Based on these, the β error probability can be determined by averaging the two particular false negative error probabilites corresponding to the two cases when (i) we select the compromised element as the condition (i.e., $x' = z$) or (ii) we select the uncompromised reading for the same role (i.e., $x' = x_2$). The averaging is justified by the fact that both of these events have a probability of 0.5 to occur because of the randomness of the selection. Formally,

$$\beta = \frac{1}{2}(\beta^{(1)} + \beta^{(2)}) \tag{1}$$

where

$$\beta^{(1)} = \int_{-\infty}^{\infty} \int_{b_1(z)}^{b_2(z)} p_{X_2,Z}(u,v) \mathrm{d}u \mathrm{d}v \tag{2}$$

$$\beta^{(2)} = \int_{-\infty}^{\infty} \int_{b_1(x_2)}^{b_2(x_2)} p_{Z,X_2}(v,u) \mathrm{d}v \mathrm{d}u \tag{3}$$

The $b_1(z)$, $b_2(z)$, $b_1(x_2)$ and $b_2(x_2)$ integration bounds are defined

with the help of the previously fixed false positive probability as

$$\int_{-\infty}^{b_1(z)} p_{X_1|X_2}(u|z)\mathrm{d}u = \frac{\alpha}{2} \;\; ; \;\; \int_{b_2(z)}^{\infty} p_{X_1|X_2}(u|z)\mathrm{d}u = \frac{\alpha}{2}$$

$$\int_{-\infty}^{b_1(x_2)} p_{X_1|X_2}(u|x_2)\mathrm{d}u = \frac{\alpha}{2} \;\; ; \;\; \int_{b_2(x_2)}^{\infty} p_{X_1|X_2}(u|x_2)\mathrm{d}u = \frac{\alpha}{2}$$

respectively. Additionally, the correlation coefficient in $p_{X_2,Z}(\cdot,\cdot)$ is calculated as

$$r_{X_2,Z} = \frac{E\big[(X_2 - \mu)(X_1 + G - \mu - \tilde{\mu})\big]}{\sigma\sqrt{\sigma^2 + \tilde{\sigma}^2}} \tag{4}$$

$$= r_{X_1,X_2} \frac{\sigma}{\sqrt{\sigma^2 + \tilde{\sigma}^2}} \tag{5}$$

and the correlation coefficient in $p_{Z,X_2}(\cdot,\cdot)$ is $r_{Z,X_2} = r_{X_2,Z}$.

With the help of β we can analyze our Enhanced Data Aggregation Algorithm from its distortion point of view. Since the most interesting aggregation function is the average because of its vulnerability (only one compromised measurement can totally mislead it) and its widespread usage we considered it in our analysis too. To evaluate the distortion of the output of Algorithm 2 we have to distinguish two basic cases: the case when an attack happens, and another one when there is no attack. In order to do this, we introduce the notations A as an indicator random variable denoting whether there is an attack or not (0 - no attack, 1 - attack), Y as the average of the sample, $Y^{(-1)}$ as minimum distortion output in case an attack is detected, and $\hat{Y}$ as the average of the sample elements when there is no attack. Still considering (wlog) the first reading to be compromised, the distortion in the first case can be expressed as

$$d(Y|A = 1) = E\big[|Y - \hat{Y}|^2 \big| A = 1\big] = \tag{6}$$

$$= E|Y^{(-1)} - \hat{Y}|^2 \cdot (1 - \beta) + \frac{1}{4}(\tilde{\mu}^2 + \tilde{\sigma}^2) \cdot \beta \tag{7}$$

While in the second case the distortion can be formalized as

$$d(Y|A = 0) = E\big[|Y - \hat{Y}|^2 \big| A = 0\big] \tag{8}$$

$$= E|Y^{(-1)} - \hat{Y}|^2 \cdot \alpha \tag{9}$$

To show how much gain our Enhanced Data Aggregation Algorithm induces compared to a scenario where no attack detection is employed, we define d_{imp} as the improvement in the distortion in case of an attack as follows:

$$d_{imp} = d(Y|A = 1, D = 0) - d(Y|A = 1, D = 1) \tag{10}$$

$$\cong \frac{1}{4}(\tilde{\mu}^2 + \tilde{\sigma}^2) \cdot (1 - \beta) \tag{11}$$

where we assume that $E|Y^{(-1)} - \hat{Y}|^2$ is close to zero. In Figure 2 one can see a plot of d_{imp} where the different curves belong to different correlation coefficients. The horizontal axis corresponds to the expected value of the attacker's distribution (i.e., $\tilde{\mu}$). The steeply ascending lines show that the improvement in the distortion grows with a growing difference between μ and $\tilde{\mu}$. The fact the line of $r = 0.5$ runs near to the line of $r = 0.95$ clearly indicates that our approach considerably exploits even correlations of moderate power. For the calculations we choose μ to be 0, σ to be 1, and $\tilde{\sigma}$ to be 1. We note, that the choice of $\tilde{\sigma}$ in the range $[0.5, 1.5]$ does not alter the results significantly.

Figure 2 clearly shows that correlation has a significant influence on the attack detection capabilities of Algorithm 1 and therefore on the distortion the attacker is able to cause in the output of Algorithm 2. Compared to the independent case (i.e., when $r = 0$), considering the naturally existing correlation between the sample elements results in smaller distortion, in other words, the attacker's abilities are more restricted when the base station maintains a correlation-based data model.
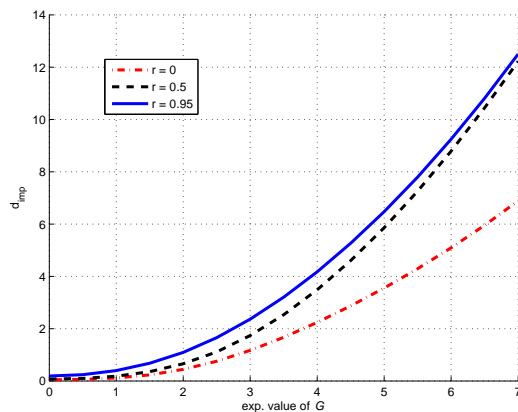
Figure 2: Calculated values for d_{imp} for different values of the correlation coefficient r (see (11)). The horizontal axis represents the expected value of the attacker's distribution (i.e., $\tilde{\mu}$).

Using the preliminary data model consisting of only two nodes from which one is possibly attacked we are able to quantify the "strength" of correlation. The results justify our suspicion: exploiting correlation can help in developing data aggregation algorithms for sensor networks that are more powerful from the resilience point of view than algorithms not considering correlation. The next step is to enable our algorithms to elaborate on arbitrary sized data sets. We already finished the elaboration of such an extended method, however, because of the space limitation, we do not detail it here.

5. DISCUSSION

Why not using standard statistical decisions instead of $Det(\cdot, \cdot)$? The two most prevalent statistical decisions are the Bayesian decision and the Maximum Likelihood decision. Informally, the Bayesian decision is about making a decision about the state of nature based on how probable that state is. Therefore, Bayesian decision theory plays a role when there is some *a priori* information about the states we are trying to classify. Since we do not rely on assumptions about the attacker's attacking frequency or distribution in time, the Bayesian decision that requires information about the attacking probability can not be applied in our case. On the contrary, the Maximum Likelihood approach decides about the state of nature based on conditional probabilities. The problem with this approach is that without assuming a concrete sampling distribution of the attacker's additive noise we cannot figure out one of the corresponding p.d.f.'s. We note that in our case the gaussian nature of the attacker is solely assumed in order to make the formal analysis feasible, the proposed algorithms (Algorithm 1 and 2) do not rely on this assumption. Therefore, regrettably, the Maximum Likelihood decision is not applicable either in our case.

How to relax the knowledge about the sampling distribution? Usually, in statistics, if the value of a parameter is not known then one tries to determine the confidence interval of that parameter. This can be figured out with higher or lower accuracy (i.e., confidence). Now, when the confidence interval is given, one can choose an arbitrary element from inside this confidence interval and consider it as the expected value. Without any other knowledge, choosing the element in the middle of the interval is the natural decision. The same method can be applied in order to determine an estimate for the value of the standard deviation. With this approximation one is still able to use the algorithms proposed without having the knowledge about the real parameters of the sampling distribution. We already performed the formal analysis of this model assuming

relaxed knowledge. However, due to space limitations, we do not detail the analysis here.

6. CONCLUSION AND FUTURE WORK

In this paper we presented a serious threat against sensor networks which consists in altering the measured parameters of the environment around the sensor nodes. The aggregation functions that consider this attack are called 'resilient aggregators'. We pointed out that there exists no research paper that would consider correlation (which is nevertheless substantial in any measurements) in enhancing the resilience of aggregation functions.

We proposed a resilient data aggregation framework for sensor networks that considers correlation. Our approach neither depends on any particular distribution of the measured sample, nor on any distribution of the attacker's noise. The algorithms presented rely on correlation and exploit it in lowering the distortion of the aggregation function. We evaluated the effectiveness of the algorithms formally by characterizing its false positive and false negative probabilities along with the final distortion in the aggregate.

Until now, most of the papers that considered correlation in sensor networks usually considered correlated measurements equivalent. We tried to sophisticate this picture of correlation and tried to show its impact on resilient data aggregation.

Our intended future work consists in adopting the analysis presented in the paper to other kind of attackers, investigating the in-network aggregation capabilities of the proposed scheme, and applying it to sample filtering. We believe that with sample filtering we will be able to further reduce the distortion of the aggregate considering any kind of attacks.

7. ACKNOWLEDGEMENTS

8. REFERENCES

[1] I. F. Akyildiz, M. C. Vuran, and O. B. Akan. On exploiting spatial and temporal correlation in sensors networks. In *Proc. of WiOpt*, 2004.

[2] L. Buttyán, P. Schaffer, and I. Vajda. RANBAR: RANSAC-based resilient aggregation in sensor networks. In *Proc. of SASN*, 2006.

[3] L. Buttyán, P. Schaffer, and I. Vajda. Resilient aggregation with attack detection in sensor networks. In *Proc. of PerSeNS*, 2006.

[4] A. Coman, M. A. Nascimento, and J org Sander. Exploiting redundancy in sensor networks for energy efficient processing of spatiotemporal region queries. In *Proc. of CIKM*, 2005.

[5] R. Cristescu, B. Beferull-Lozano, and M. Vetterli. On network correlated data gathering. In *Proc. of INFOCOM*, 2004.

[6] H. Gupta, V. Navda, S. R. Das, and V. Chowdhary. Efficient gathering of correlated data in sensor networks. In *Proc. of MobiHoc*, 2005.

[7] S. Krishnamurthy, T. He, G. Zhou, J. A. Stankovic, and S. H. Son. RESTORE: A real-time event correlation and storage service for sensor networks. In *Proc. of INSS*, 2006.

[8] P. von Rickenbach and R. Wattenhofer. Gathering correlated data in sensor networks. In *Proc. of DIALM-POMC*, 2004.

[9] M. C. Vuran, O. B. Akan, and I. F. Akyildiz. Spatio-temporal correlation: theory and applications for wireless sensor networks. *Elsevier Computer Networks*, 45(3):245–259, 2004.

[10] M. C. Vuran and I. F. Akyildiz. Spatial correlation-based collaborative medium access control in wireless sensor networks. *TON*, 14(2):316–329, 2006.

[11] D. Wagner. Resilient aggregation in sensor networks. In *Proc. of SASN*, 2004.

[12] S. Yoon and C. Shahabi. Exploiting spatial correlation towards an energy efficient clustered aggregation technique (CAG). In *Proc. of ICC*, 2005.

[13] Y. Zhu, K. Sundaresan, and R. Sivakumar. Practical limits on achievable energy improvements and useable delay tolerance in correlation aware data gathering in wireless sensor networks. In *Proc. of SECON*, 2005.

[14] Y. Zhu, R. Vedantham, S.-J. Park, and R. Sivakumar. A scalable correlation aware aggregation strategy for wireless sensor networks. In *Proc. of WICON*, 2005.