Privacy Preserving Node Selection: A Physical Layer Approach

Shalabh Jain, John S. Baras
Institute for Systems Research
University of Maryland, College Park, MD 20742
Email: {shalabh, baras}@umd.edu

***Abstract*—Distributed monitoring and sensing networks are ubiquitous in current technological systems. Security and efficiency of the data aggregation and fusion process are key determinants in the adoption of these systems. The principal thesis investigated in this paper is the role of privacy in providing security guarantees for such networks. We demonstrate that two-level hierarchy (or partition) in information fusion networks can be used to ensure security in several adversarial scenarios. We propose a privacy preserving physical layer framework to generate such a hierarchical structure. We illustrate the security of this framework and its application to low power sensor networks.**

***Keywords*—Privacy, Security, Information fusion, Physical layer watermarking**

I. Introduction

Recent advances in sensor technology have significantly increased the adoption of these devices in current systems. Distributed networks of low powered sensors with limited capability have been deployed in several critical systems such as cyberphysical systems for monitoring and regulation of power grids, large scale intrusion detection systems or civil infrastructure monitoring systems.

A typical requirement in several such applications is the ability to effectively and efficiently extract data from the network. Significant research efforts have been directed towards increasing the efficiency of the information retrieval process. This includes selection of a subset of nodes for observation, [1], [2], [3], utilizing compressive sensing techniques [4], [5], [6] or distributed signal processing techniques [7].

The critical nature of the deployment scenarios has made such systems a valuable target for adversarial action. The risk is compounded by the fact that typically, such networks are composed of unattended devices with limited protection. Additionally, such networks are expected to operate efficiently in dense environments and over long periods of time. Thus efficient security is a key requirement for such systems. Several techniques have been proposed in literature to ensure security in such systems, e.g. [8], [9], hop-by-hop encryption [10], end-to-end encryption [11], or secure data aggregation [12], [13]. However, such techniques introduce both, significant processing and communication overhead.

In this paper, we present an efficient framework to ensure security in the network. We investigate the role of privacy in ensuring security of the network. Intuitively, we select a subset of nodes to act as 'pseudo-adversaries', and inject malicious (noisy) data in the network. The creation of such a partition in a privacy preserving manner, i.e. such that the partitions are unknown to the adversary, can obfuscate the adversarial view. Systems where data acquisition requires a subset of measurements, e.g.: [1], [2], [4], we may create such a hierarchy by sacrificing only the communication efficiency.

Further, we utilize the physical layer watermarking in [14] to define a privacy preserving method to select nodes. This serves as an efficient method to define 'pseudo-adversarial' partitions and guarantee security.

The rest of this paper is organized as follows. In Section II, we discuss the systems under consideration and describe the overall scheme. In Section III, we describe the privacy preserving messaging scheme and illustrate its security properties. In Section IV we validate our results via MATLAB simulations.

II. System Description

Consider a network $\mathcal{M} = \{M_1, M_2, \ldots, M_N\}$, of N sensor nodes distributed uniformly over a region. Consider a central entity that, over a wireless interface, gathers data (or a function of the data) from the sensor nodes. We denote such an entity by FC (Fusion Center). The data acquisition (or state computation) is based on periodic measurements collected by the FC from the N nodes.

We restrict our study to systems with dynamics (partially known) that enable the FC to utilize techniques to reduce communication or processing overhead, e.g.: systems described in [1], [2], [4]. We assume that the optimization strategy is determined by the FC based on the system state. For optimization, the FC may partition the network based on significance of the information.

Consider the scenario where at sampling instance n, only observations from the set $\mathcal{M}_c(n) \subset \mathcal{M}$ are significant for the FC. To reduce system overhead, only the nodes $\mathcal{M}_c(n)$ may transmit the measurements and nodes $\mathcal{M}_d(n) = \mathcal{M} \setminus \mathcal{M}_c(n)$ do not transmit any data.

For simplicity, consider the scenario where the FC can communicate directly with the nodes $\mathcal{M}$, i.e.: one hop scenario. We may trivially extend the framework to multi-hop scenarios, by iterative application of the security strategy of the fusion center to cluster-heads (one hop neighborhoods).

A. Adversarial Model

We consider an external eavesdropping adversary A. Since the nodes $\mathcal{M}$ communicate over a wireless medium, we

978-1-4799-8428-2/15/$31.00 ©2015 IEEE

assume that the adversary may obtain complete data transmissions. e.g. for the single hop scenario, all observations of the FC.

We assume that to effectively attack the network, the adversary, A, requires at least the information obtained by FC, i.e. the system view of A should be the same as FC. We assume the existence of a pre-shared secret (key k), unknown to the adversary. Further, we assume that the randomness (specific instantiations) in the optimization strategy of the FC is unavailable to the adversary.

B. System Operation

At a collection instance n, the FC selects a set $\mathcal{M}_c(n) \subset \mathcal{M}$ to query. The particular set may depend on the particular technique being used by FC and state of the system, i.e. [1], [2], [4], and is unknown to the adversary. The Fusion Center utilizes a privacy preserving framework (described in Section III) to covertly query (message) the selected nodes $\mathcal{M}_c(n)$. (Note: For scenarios where $|\mathcal{M}_d(n)| < |\mathcal{M}_c(n)|$, we query $\mathcal{M}_d(n)$.)

Upon receiving the query message, the nodes $M_c(n)$ reply with the true network measurements $\mathcal{O}_c(n) = \{M_i^o(n) \mid i \in \mathcal{M}_c(n)\}$. The remaining nodes transmit decoy measurements $\mathcal{O}_d(n) = \{g(M_i^o(n)) \mid i \in \mathcal{M}_d(n)\}$.

Consider the operation by FC to be a function of the observations, i.e. the view of FC is $\mathcal{V}^{FC} = v(\mathcal{O}_c(n))$. The adversary, in the absence of knowledge of $\mathcal{M}_c(n)$, possesses the view $\mathcal{V}^A = v(\mathcal{O}_c(n) \cup \mathcal{O}_d(n))$. By careful selection of $g(\cdot)$, we ensure $\mathcal{V}^{FC} \neq \mathcal{V}^A$, i.e. the system view of the adversary is different from the FC, thus fulfilling our security requirement.

For example, consider the scenario where $v(\cdot)$ is the averaging function. Even a simple selection of $g(\cdot)$, i.e. $g(x) = \Delta = 0$, is sufficient to distort the view of the adversary.

It should be observed that our scheme incurs a transmission overhead due to the decoy transmissions, thus decreasing the system efficiency due to optimizations by the FC. However, we achieve security guarantees without the use of cryptographic primitives. For low capability nodes, as is the case in typical sensor networks, that lack a cryptographic co-processor, this leads to a significant reduction in the energy overhead.

C. System Example

Consider the state estimation system described in [1] consisting of m sensor nodes ($\mathcal{M}$). The FC estimates the state of the system $x_n \in \mathbb{R}^s$ from $c \geq s$ linear measurements corrupted by Gaussian noise. The FC uses a greedy approach to select c nodes ($\mathcal{M}_c$).

Utilizing our scheme, the FC covertly queries the c nodes to obtain the sensor measurements. Nodes that have not been queried sample a Bernoulli random variable with success probability p_d. Upon a successful outcome, the node transmits a decoy measurement $g(x) = x + \epsilon$, where x is the true measurement and $\epsilon \sim \mathcal{N}(\Delta, \sigma_d^2)$. This causes the adversary to converge to a false system state (similar to the adversarial noise injection scenario demonstrated in [15]).

The system incurs an average communication and processing overhead, $d \approx p_d(m-k)$, due to the decoy measurements. In the absence of any knowledge of k, at each step, the search space for the adversary grows by $\mathcal{O}(2^{k+d})$. We select p_d based on the acceptable overhead vs. adversarial effort tradeoff for the given application.

III. Privacy Preserving Messaging Scheme

We describe the messaging scheme to covertly convey information to the selected nodes. We utilize the idea of low power tagging developed in [14]. Here we briefly describe important notation and aspects of the scheme relevant to our discussion. For details, constraints and performance metrics of the single tag scheme, the reader is referred to [14]. The goal here is to describe our framework based on [14] and the corresponding security properties.

A. Tagging Scheme

Consider the sender selecting the nodes by transmitting a query message, $\mathbf{s} = \{s_1, s_2, \ldots, s_L\}$, of length L symbols, to a set $\mathcal{M}_c \subset \mathcal{M}$ of nodes. The sender assigns a tag,

$$\mathbf{t}_i = f(k, \mathbf{s}, i) \quad \forall i \in \mathcal{M}_c,$$

to each node in $\mathcal{M}_c$. Here i denotes the identity of the node and k denotes the common shared key in the network. $f(\cdot)$ represents a 'secure' tagging scheme (e.g: keyed hash function). We require that for $f(\cdot)$, the distribution of the output is uncorrelated to the input. This requirement can typically be satisfied by cryptographic one way functions. Further, via proper encoding, we ensure $\mathbf{t}_i \in \{-1, 1\}^L$. The intuition is to tag the message $\mathbf{s}$ such that a node j with the key can identify whether $j \in \mathcal{M}_c$.

The sender superimposes the tag on the signal waveform to transmit as

$$\mathbf{x} = \rho_s \mathbf{s} + \rho_t \sum_{i \in \mathcal{M}_c} \mathbf{t}_i,$$

where $\rho_s, \rho_t \in (0,1)$ represent the power allocation to the signal and tag. Let us assume that for the system,

$$\max |\mathcal{M}_c| = K_{max}.$$

If we consider the number of tags superimposed at any time instant to be a random quantity $K = |\mathcal{M}_c|$ with expected value $\bar{K} = \mathbb{E}[K]$. We ensure that the total average power is maintained, i.e. $\rho_s^2 + \bar{K}\rho_t^2 = 1$.

Assume a Rayleigh block fading (slow fading) channel. The channel for the transmitted block is denoted by $h \sim CN(0, \sigma_h^2)$. CN denotes a circularly symmetric complex Gaussian variable. The receiver observes the block $\mathbf{y} = h \cdot \mathbf{x} + \mathbf{w}$, where $\mathbf{w} = \{w_1, \ldots w_L\}$ and $w_k \sim CN(0, \sigma_w^2), \forall k$. Using the pilot-based MMSE estimator highlighted in [14], a receiver recovers the transmitted signal $\hat{\mathbf{s}}$ and its expected tag $\hat{\mathbf{t}}_i = f(k, \hat{\mathbf{s}}, i)$. The receiver determines if it is one of the 'selected' receivers by verifying the presence of its tag in the residue

$$\mathbf{r} = \frac{1}{\rho_t}(\hat{\mathbf{x}} - \rho_s \hat{\mathbf{s}}) = \sum_{i \in \mathcal{M}_c} \mathbf{t}_i + \frac{1}{\rho_t} \frac{h^*}{|h|^2} \mathbf{w}. \qquad (1)$$

The receiver obtains the test statistic τ_j by applying a matched filter to the residue with the estimated tag, $\tau_j = \hat{\mathbf{t}}_{\mathbf{j}}^H \mathbf{r}$. The receiver performs a threshold test with hypotheses

$$\begin{aligned} H_0 &: \hat{\mathbf{t}}_{\mathbf{j}} \text{ is not present in } \mathbf{r} \\ H_1 &: \hat{\mathbf{t}}_{\mathbf{j}} \text{ is present in } \mathbf{r}. \end{aligned} \tag{2}$$

Assuming perfect channel estimation ($\hat{h} = h$) and tag estimation ($\hat{\mathbf{t}} = \mathbf{t}$), we obtain the statistic for the two scenarios when the tag $\mathbf{t}_j$ is present vs. not present as follows

$$\tau_j = \sum_{i \in \mathcal{M}_c} \mathbf{t}_j^H \mathbf{t}_i + \frac{1}{\rho_t} \frac{h^*}{|h|^2} \mathbf{t}_j^H \mathbf{w} \tag{3}$$

$$= \sum_{i \in \mathcal{M}_c} \mathbf{t}_j^H \mathbf{t}_i + w_t. \tag{4}$$

Since $t_j = \{-1, 1\}^L$, the noise term w^t can be viewed as a sum and difference of L components, w_i, of $\mathbf{w}$. As these are assumed to be iid Gaussian, we see that

$$w^t \sim \mathcal{CN}\left(0, L \frac{1}{\rho_t^2} \frac{\sigma_w^2}{\sigma_h^2}\right).$$

For the first term, firstly we consider the scenario where $t_j \in \mathcal{M}_c$. Clearly,

$$\begin{aligned} \sum_{i \in \mathcal{M}_c} \mathbf{t}_j^H \mathbf{t}_i &= \mathbf{t}_j^H \mathbf{t}_j + \sum_{\{i \in \mathcal{M}_c, i \neq j\}} \mathbf{t}_j^H \mathbf{t}_i \\ &= L + \sum_{\{i \in \mathcal{M}_c, i \neq j\}} \mathbf{t}_j^H \mathbf{t}_i. \end{aligned}$$

For $i \neq j$, $\mathbf{t}_j^H \mathbf{t}_i = \sum_{r=1}^{L} b_r$, where b_r is a random variable, such that $\mathbb{P}(b_r = 1) = \mathbb{P}(b_r = -1) = 1/2$. Thus

$$\sum_{\{i \in \mathcal{M}_c, i \neq j\}} \mathbf{t}_j^H \mathbf{t}_i = \sum_{r=1}^{L(K-1)} b_r \sim \mathcal{N}(0, L(K-1)).$$

Note that the Normal approximation holds accurately for only large values of L, K, via the Central Limit Theorem, which will be true for most instances of our system. In the event that this is not the case, we may further add a small error term without much change to the analysis.

Proceeding as above, we may obtain the distribution for the case when $j \notin \mathcal{M}_c$ as

$$\sum_{i \in \mathcal{M}_c} \mathbf{t}_j^H \mathbf{t}_i \sim \mathcal{N}(0, LK).$$

Thus, conditioned on $\mathbf{t}_j$, the distribution of τ_j for the tagged and non tagged scenarios is

$$\begin{aligned} \tau_j | H_1 &\sim \mathcal{N}(L,\ L(K-1) + \gamma_t L) \\ \tau_j | H_0 &\sim \mathcal{N}(0,\ LK + \gamma_t L), \end{aligned} \tag{5}$$

where $\gamma_t = \frac{1}{2\rho_t^2} \frac{\sigma_w^2}{\sigma_h^2}$, since we use just the real component of w^t for decision making. The receiver performs a simple threshold test as

$$\tau_j \gtrless_{H_0}^{H_1} \tau_{th}. \tag{6}$$

Clearly, the scheme leads to a small degradation in the performance of transmission of the symbol s. However, most practical communications are conservatively designed to operate in a variety of environments. We claim, based on the application and the operating environment, we can tune ρ_s such that the perceivable degradation is negligible.

For the system, $K \in \{1, \ldots, K_{max}\}$. The value of K_{max} is determined by the acceptable probabilities of error in detecting the tags. For the above hypotheses, we may write the probability of false alarm (P_{fa}) and missed detection (P_{md}) as

$$P_{fa} = \Phi\left(-\frac{\tau_{th}}{\sqrt{LK + L\gamma_t}}\right) \tag{7}$$

$$P_{md} = \Phi\left(\frac{\tau_{th} - L}{\sqrt{L(K-1) + L\gamma_t}}\right), \tag{8}$$

where $\Phi(\cdot)$ is the Gaussian cdf function. Let ρ_s^{min} denote the minimum power allocation to the signal without perceivable QoS degradation. Thus we obtain

$$\rho_t^2 = (1 - \rho_s^{min^2})/K_{max}. \tag{9}$$

Thus, we select the value of K_{max} such that

$$\max_K P_{fa} \leq p_1, \text{ and } \max_K P_{md} \leq p_2. \tag{10}$$

The $\max$ constraint typically leads to a conservative selection of K_{max}. For system design, where we may calculate the distribution of K, (e.g. $K \sim \mathcal{U}(K_{max}^{-1})$), we may relax the constraints to

$$P_{fa} \leq p_1, \text{ and } P_{md} \leq p_2,$$

where the false alarm and missed detection probabilities are computed over distribution of K. We demonstrate via simulations in Section IV, that such criteria can be satisfied for several design parameters.

B. Security Properties

We emphasize that due to the low power of the tag, for an adversary, identifying the set $\mathcal{M}_c$ by decoding the tag components is difficult. Further, without knowledge of the key k, the adversary is unable to perform matched filtering on the residue to verify the presence of specific tags. The best strategy for the adversary is to perform statistical tests on (1). We prove that this yields insufficient information, even to accurately estimate the number of selected nodes. We highlight the security properties of the framework.

1) Determination of elements of $\mathcal{M}_c$: Consider (1) to estimate the tag. As each component of the tag $\mathbf{t}_i = \{t_{i1}, \ldots, t_{iL}\}$ is independent, we may consider estimation of each component separately from (1). For the j'th component,

$$r_j = \sum_{\{i \in \mathcal{M}_c\}} t_{ij} + w'_j.$$

The adversary estimates a noisy version of the tags $\hat{T}_j \approx \sum_{\{i \in \mathcal{M}_c\}} t_{ij}$. Let us assume that the adversary estimates $\hat{T}_j$ perfectly and has knowledge of the number of components K. Even so, the probability that the adversary correctly reconstructs even a single tag, $t_j,\ j \in \mathcal{M}_c$, can be shown to be K^{-L}.

However, it is important to observe that the estimate $\hat{T}_j$ is obtained from a very noisy measurement. Thus the error in such an estimate would be large. Further, the number of components in the selected set, K, will be unknown to the adversary. This significantly reduces the probability of correctly decoding the tags.

2) Determination of K: We argue that given the adversarial observation, it is difficult to estimate even the number of tags embedded. We consider the parameter $K = |\mathcal{M}_c|$ to be the underlying parameter in our observation. The adversary observes

$$Y_j = \sum_{\{i \in \mathcal{M}_c\}} t_{ij} + w'_j = T_j + w'_j, \quad j = 1, \dots, L \tag{11}$$

It can easily be seen that the random variable T_j has the distribution

$$\mathbb{P}(T_j = K - 2t) = 2^{-K} \binom{K}{t} \tag{12}$$

As $w'_j \sim \mathcal{N}(0, \gamma_t)$, we may write the density of $Y_j, \ \forall j$ as

$$p_Y(x) = 2^{-K} \sum_{t=0}^{K} \binom{K}{t} \mathcal{N}(x; K - 2t, \gamma_t) \tag{13}$$

$$= \frac{2^{-K}}{\sqrt{2\pi\gamma_t}} \sum_{t=0}^{K} \binom{K}{t} \exp\left(-\frac{1}{2\gamma_t}\Big(x - (K - 2t)\Big)^2\right) \tag{14}$$

We can see that (14) essentially represents the Gaussian mixture model of K components with identical variances but different means. Such problems have been studied for several years in the context of biological systems or clustering problems in computer vision. However, for components that are close (as in our scenario), this estimate error is known to be large. A variety of methods [16], [17], [18] may be used to estimate K. We highlight some results in Section IV.

To see analytically the performance of the estimator, we may perform a coarse approximation and assume T_j to have Normal distribution, i.e. $T_j \sim \mathcal{N}(0, K)$. Thus the estimation problem reduces to estimating K from L observations of y_j with distribution $Y_j \sim \mathcal{N}(0, \gamma_t + K)$. From [19], the lower bound for the variance of the best estimator can be computed to $\mathrm{Var}_K[\hat{K}] \geq \frac{2(\gamma_t + K)^2}{L}$. Though this is was based on the Normal approximation of T_j, we can see that the order of the error is $(\gamma_t + K)$ when the number of components is K. Thus the adversary is unable to gain much information about the estimate of K.

Thus using the proposed tagging scheme, we can selectively identify a subset of nodes without leaking any information to the adversary.

Remark: Under the assumption of a shared key, this goal can be trivially achieved by symmetric key encryption, wherein the central node encrypts the identity of the nodes in N_s and transmits the signal. However, the current framework provides several advantages over the standard encryption methodology.

- Our method ensures that the total bandwidth and power per packet does not change. For encryption based methods, as the identities are transmitted with the same QoS as data, increased power and bandwidth are required. While for each packet, the gain may not be significant, for schemes where such packets are sent periodically, the savings over the lifetime of a node would be significant.
- Our method prevents leakage of even empirical data such as the number of paged nodes. To achieve a similar effect using encryption, each packet would have to be padded to a consistent length, thus incurring an increased overhead.
- Using our method, the signal can be overlapped over any existing protocol message, rather than requiring the design of new messages. For example, even periodic 'HELLO' or 'ALIVE' transmissions in a sensor network could be used to convey the desired information.
- Our method prevents the adversary from performing attacks like relaying or recording and replaying. This is due to the fact that re-transmission of the packets destroys the embedded identity information. This was further presented in [20].

We observe that the error guarantees provided by our scheme are not comparable to cryptographic methods. However, based on the specific application, the parameters of the scheme may be selected to ensure no degradation in the application metrics.

A similar framework was proposed by [21] to preserve the privacy in the paging channel in LTE systems. Though similar, the design criteria and constraints required for that system are significantly different. Our framework is more general. The system in [21] can be considered as a specific case of our framework.

IV. Simulations

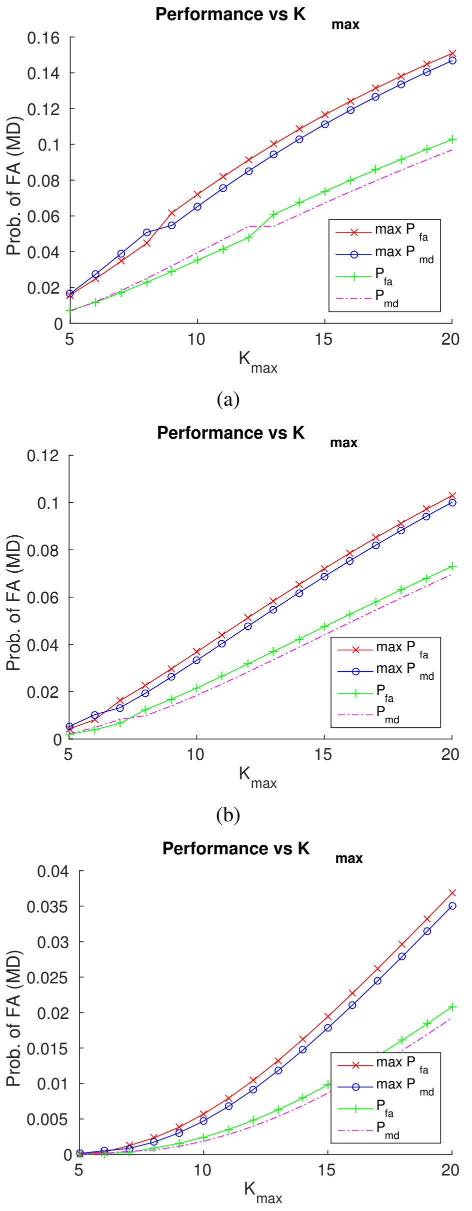
We demonstrate the security properties and the influence of design parameters of the scheme via MATLAB simulations. First we consider identification of critical parameters of system design. We then utilize the optimal range of those parameters and illustrate the security properties of the scheme.

A. Parameter selection

The performance of the scheme is highly dependent on the selection of the system parameters. We consider a system operating with a 10% power margin, i.e. selection of $\rho_s^2 \geq 0.9$ is sufficient to maintain the desired QoS. Thus from (9), we have $\rho_t^2 \cdot K_{max} = 0.1$. The selection of K_{max} is based on the system configuration and topology. An increase in K_{max}, while decreasing query latency, adversely impacts system performance.

In Fig. 1, we illustrate the variation in the probabilities of false alarm and missed detection with increasing K_{max}. We select the optimal decision threshold in (6) based on minmax rule (minimizing the $\max\{P_{fa}, P_{md}\}$), rather than fixed bounds. Assuming a 7% tolerable false alarm and missed detection, we observe that for tag length $L = 256$ symbols, we may select a maximum of 14 nodes. Assuming, that the application induces a uniform distribution over the number of selected nodes, i.e. $K \sim \mathcal{U}(K_{max}^{-1})$, we may relax the constraints to accommodate a maximum of 20 nodes.

Further, we observe that an increase in the tag length L allows for a higher number of selected nodes. However, this can adversely influence the security properties of the scheme as the adversary obtains a greater number of samples for prediction.



(a)

(b)

(c)

Fig. 1. Maximum and total probability of false alarm and missed detection with variation in maximum number of selected nodes (SNR, L, $K_{max} \cdot \rho_t^2$) = (a) (10, 128, 10%), (b) (5, 256, 10%), (c) (10, 256, 10%), (d) (5, 512, 10%),

B. Security properties

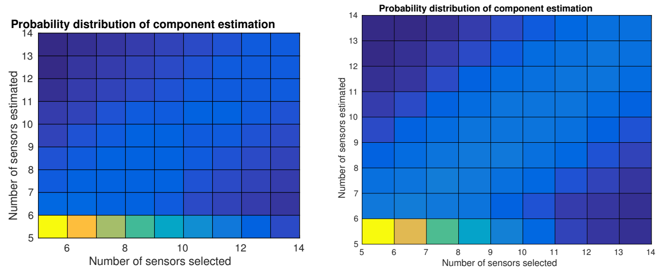
Clearly, determination of the individual nodes selected, without knowledge of the group key k is not feasible. Here, we demonstrate that the proposed method is robust to leakage of empirical information such as the number of nodes selected.

As discussed in Section III-B2, determination of K is equivalent to determination of the number of components (clusters) in a Gaussian Mixture Model, which is known to be difficult. However, we remark that in our scenario, the location of the cluster heads may be determined apriori based on the number of assumed clusters, thus reducing the parameter space.

Assuming the adversary has knowledge of the channel conditions and system parameters, we perform the maximum likelihood estimation of the number of selected nodes as

$$K^* = \arg\max_K \sum_{i=1}^{L} \log\left(2^{-K} \sum_{t=0}^{K} \binom{K}{t} \mathcal{N}(x_i; K-2t, \gamma_t)\right) \tag{15}$$

Using optimal parameter values determined in the previous section, we perform Monte Carlo simulations for varying number of selected nodes. In Fig. 2, we illustrate the probability distribution of estimated number of selected nodes in the different scenarios. The color gradient represents variation in the distribution. It can be seen that even in the high SNR scenario, for $L = 128$ and $L = 256$, the distribution of the estimate is close to uniform for any number of selected nodes.



(a)

(b)

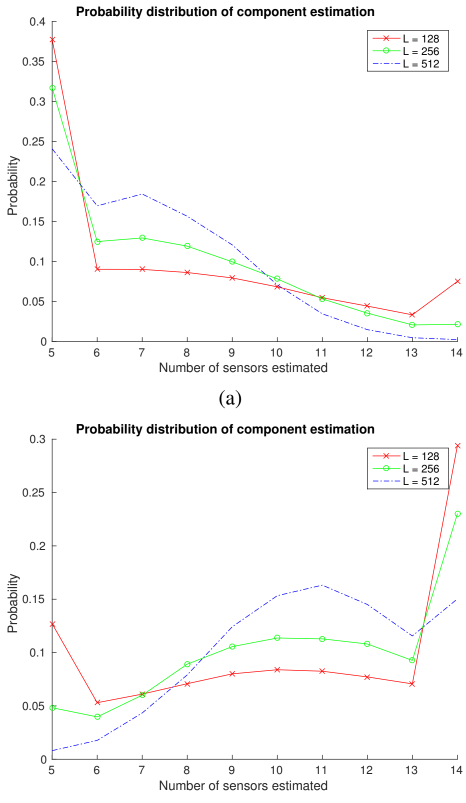
Fig. 2. Probability distribution for estimation of number of components for (a) L = 128, (b) L = 256

K	Mean	Var	MMSE	P_{err}
5	6.1635	3.1448	4.4987	0.4224
6	6.7254	4.4870	5.0131	0.8632
7	7.4605	5.9470	6.1592	0.8705
8	8.2211	7.0725	7.1214	0.8743
9	9.0445	7.7131	7.7151	0.8812
10	9.8384	7.9363	7.9624	0.8822
11	10.6622	7.4658	7.5800	0.8871
12	11.3598	6.6790	7.0889	0.8836
13	12.0006	5.5833	6.5823	0.8897
14	12.5451	4.2675	6.3842	0.4563

Further, we observe from Table IV-B, for $L = 256$, the estimation error is high for all possible choices of nodes. In Fig. 3, we illustrate the gain in adversarial advantage due to increase in the tag length. However, even in the scenario of $L = 512$, the advantage is insignificant to be of practical use.

V. Conclusion

We proposed an efficient method to covertly query a small subset of nodes from a large group of wireless nodes communicating over a wireless medium. This enables a central entity to partition the network into multiple levels without



(a)

(b)

Fig. 3. Probability distribution for estimation of number of components for different symbol lengths with variation in number of selected sensors (a) K = 7, (b) K = 11

the knowledge of an external adversary. We illustrated the application of this scheme to provide security guarantees in information fusion networks. In scenarios of information fusion, where information from a specific partition is sufficient, we utilize the remaining partitions to distort the view of the adversary. We utilize physical layer watermarking to design a privacy preserving method to query the nodes.

Acknowledgment

This material is based upon work partially supported by US Air Force Office of Scientific Research MURI grant FA9550-10-1-0573, by National Science Foundation grant CNS 1035655 and by NIST grant 70NANB11H148.

References

[1] M. Shamaiah, S. Banerjee, and H. Vikalo, "Greedy sensor selection: Leveraging submodularity," in *Decision and Control (CDC), 2010 49th IEEE Conference on*, Dec 2010, pp. 2572–2577.

[2] R. Olfati-Saber and N. Sandell, "Distributed tracking in sensor networks with limited sensing range," in *American Control Conference, 2008*, June 2008, pp. 3157–3162.

[3] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *Signal Processing, IEEE Transactions on*, vol. 57, no. 2, pp. 451–462, Feb 2009.

[4] Y. Tang, B. Zhang, T. Jing, D. Wu, and X. Cheng, "Robust compressive data gathering in wireless sensor networks," *Wireless Communications, IEEE Transactions on*, vol. 12, no. 6, pp. 2754–2761, June 2013.

[5] J. Luo, L. Xiang, and C. Rosenberg, "Does compressed sensing improve the throughput of wireless sensor networks?" in *Communications (ICC), 2010 IEEE International Conference on*, May 2010, pp. 1–6.

[6] L. Xiang, J. Luo, and A. Vasilakos, "Compressed data aggregation for energy efficient wireless sensor networks," in *Sensor, Mesh and Ad Hoc Communications and Networks (SECON), 2011 8th Annual IEEE Communications Society Conference on*, June 2011, pp. 46–54.

[7] J. Chou, D. Petrovic, and K. Ramachandran, "A distributed and adaptive signal processing approach to reducing energy consumption in sensor networks," in *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, vol. 2, March 2003, pp. 1054–1062 vol.2.

[8] S. Ozdemir and Y. Xiao, "Secure data aggregation in wireless sensor networks: A comprehensive overview," *Computer Networks*, vol. 53, no. 12, pp. 2022 – 2037, 2009.

[9] H. Chan and A. Perrig, "Security and privacy in sensor networks," *Computer*, vol. 36, no. 10, pp. 103–105, Oct 2003.

[10] Y. Yang, X. Wang, S. Zhu, and G. Cao, "Sdap: A secure hop-by-hop data aggregation protocol for sensor networks," *ACM Trans. Inf. Syst. Secur.*, vol. 11, no. 4, pp. 18:1–18:43, Jul. 2008.

[11] Y. Sang, H. Shen, Y. Inoguchi, Y. Tan, and N. Xiong, "Secure data aggregation in wireless sensor networks: A survey," in *Parallel and Distributed Computing, Applications and Technologies, 2006. PDCAT '06. Seventh International Conference on*, Dec 2006, pp. 315–320.

[12] V. Kumar and S. Madria, "Secure hierarchical data aggregation in wireless sensor networks: Performance evaluation and analysis," in *Mobile Data Management (MDM), 2012 IEEE 13th International Conference on*, July 2012, pp. 196–201.

[13] 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, 2006, pp. 278–287.

[14] P. Yu, J. Baras, and B. Sadler, "Physical-layer authentication," *IEEE Trans. on Information Forensics and Security*, vol. 3, no. 1, pp. 38–51, Mar. 2008.

[15] S. Zheng and J. Baras, "Trust-assisted anomaly detection and localization in wireless sensor networks," in *Proc. IEEE Conf. on Sensor, Mesh and Ad Hoc Comm. and Netw (SECON)*, 2011, pp. 386–394.

[16] C. Fraley and A. E. Raftery, "How many clusters? which clustering method? answers via model-based cluster analysis," *The Computer Journal*, vol. 41, pp. 578–588, 1998.

[17] C. Biernacki, G. Celeux, and G. Govaert, "Assessing a mixture model for clustering with the integrated completed likelihood," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 7, pp. 719–725, Jul 2000.

[18] G. Celeux and G. Soromenho, "An entropy criterion for assessing the number of clusters in a mixture model," *Journal of Classification*, vol. 13, no. 2, pp. 195–212, 1996.

[19] V. Poor, *An Introduction to Signal Detection and Estimation*, ser. Springer Texts in Electrical Engineering. Springer, 1998.

[20] S. Jain and J. S. Baras, "Preventing wormhole attacks using physical layer authentication," in *Proc. IEEE Wireless Communications and Networking Conf. (WCNC)*, 2012, pp. 2712–2717.

[21] T. Ta and J. S. Baras, "Enhancing privacy in LTE paging system using physical layer identification," in *7th International Workshop on Data Privacy Mgmt. and Autonomous Spontaneous Security*, 2012, pp. 15–28.