Volume : 2 | Issue : 11 | Nov 2013 • ISSN No 2277 - 8160

Research Paper Engineering

Sequential Based Hypothesis Testing in Wireless Sensor Networks

G. UMA — Master Of Engineering, Dept.Of ECE, Parisutham Institute Of Technology And Science, Thanjavur, India.

A. DINESH — Assistant Professor,Dept.Of ECE, Parisutham Institute Of Technology And Science, Thanjavur, India.

ABSTRACT

In certain applications, the locations of events reported by a sensor network need to remain anonymous. That is, unauthorized observers must be unable to detect the origin of such events by analyzing the network traffic. We analyze two kinds of problems:

Communication overhead and computational load problem. In this paper, we present a new framework for modeling, analyzing, and evaluating anonymity in sensor networks. The novelty of the proposed framework is twofold: first, it introduces the notion of "interval indistinguishability" and provides a quantitative measure to model anonymity in wireless sensor networks; second, it maps source anonymity to the statistical problem We showed that the current approaches for designing statistically anonymous systems introduce correlation in real intervals while fake are uncorrelated. We show how mapping source anonymity to statistical hypothesis testing with nuisance Parameters leads to converting the problem of exposing private source information into searching for an appropriate data transformation that removes or minimize the effect of the nuisance information. By doing so, we transform the problem of analyzing real valued sample points to binary codes, which opens the door for coding theory to be incorporated into the study of anonymous networks. In existing work, unable to detect unauthorized observer in network traffic. But our work mainly based on enhances their source anonymity against correlation test. The main goal of source location privacy is to hide the existence of real events.

KEYWORDS : source location, privacy, anonymity, statistical hypothesis testing, nuisance parameters and coding theory.

1. INTRODUCTION

SENSOR networks are used in a wide range of applications such as military, health care, and animal tracking [3], [4], [5]. In monitoring networks, nodes are designed to transmit information only when a relevant event is detected**.** The location of real events reported by the node can be approximated within the sensing range. There are three parameters that can be associated with an event detected and reported by a sensor node: description of event, location and time taken for protecting the privacy because security is most important. While transmitting the description of sensed event, hiding the timing and spatial information cannot be achieved.

The Statistical Source Anonymity (SSA) problem in sensor networks is the study of techniques that prevent global adversaries from exposing source location by performing statistical analysis on nodes transmissions [11], [19], [20], [21], [22], [23], [24]. Practical SSA solutions need to be designed to achieve their objective under two main constraints: minimizing delay and maximizing the lifetime of sensors" batteries.

2 MODEL ASSUMPTIONS

In this section, we describe the network and adversarial assumption that will be used in this paper.

2.1 Network Model

Communication is assumed to take place in a network of energy constrained sensor nodes. Nodes are deployed to sense events of interest and report them with minimum delay. Consequently, given the location of a certain node, the location of the reported event of interest can be approximated within the node"s communication range at the time of transmission. When a node senses an event, it places information about the event in a message and broadcast an encrypted version of the message. Adversaries are unable to distinguish between the reports of events of interest and the fake transmissions by means of cryptographic tests.

2.2 Adversarial Model

The adversarial model used in this paper is similar to the one considered in [10], [11], in that it is external, passive, and global. An external adversary is an adversary who does not control any of the nodes in the network A global adversary is an adversary who can monitor the traffic of the entire network and can determine the node responsible for the initial transmission reporting an event of interest. The adversary, however, is not assumed able to break the security of the encryption algorithm and distinguish the report of event of interests via cryptographic tests.

3. PROPOSED FRAMEWORK FOR SSA

In this section, we introduce our source anonymity model for wireless sensor networks.

3.1 Interval Indistinguishability

Currently, statistical anonymity in sensor networks is modeled by the adversary"s ability to distinguish between real and fake transmissions by means of statistical analysis. Consider now an adversary observing a sensor network over multiple time intervals. Assume that, during a given time interval, the adversary is able to notice a change in the statistical behavior of transmission times of a certain node in the network. This distinguishable change in the transmission behavior of the node can be indicative of the existence of real activities detected and reported by that node during that interval, even if the adversary was unable to distinguish between individual transmissions. Consequently, in many applications, modeling source anonymity in sensor networks by the adversary"s ability to distinguish between individual transmissions is insufficient to guarantee location privacy.

Definition 1 (Interval indistinguishability). Let IF denotes a time interval without any real event transmission (called the "fake interval" for the rest of the paper), and IR denotes a time interval with real event transmissions (called the "real interval" for the rest of the paper). The two time intervals are said to be statistically indistinguishable if the distributions of intertransmission times during these two intervals cannot be distinguished with significant confidence.

3.2 Interval versus Event Indistinguishability

This section illustrates the relation between the traditional anonymity notions (i.e., individual event indistinguishability) and the proposed anonymity notion (i.e., interval indistinguishability).I n fact, the notion of interval indistinguishability is strictly stronger than the traditional notion individual event indistinguishability.

3.3 Mapping Statistical Source Anonymity to Binary Hypothesis Testing

In binary hypothesis testing, given two hypothesis, H0 and H1, and a data sample that belongs to one of the two hypotheses (e.g., a bit transmitted through a noisy communication channel), the goal is to decide to which hypothesis the data sample belongs. In the statistical strong anonymity problem under interval indistinguishability, given an interval of inter-transmission times, the goal is to decide whether the interval is fake or real (i.e., consists of fake transmissions only or contains real transmissions).

3.4 Quantifying Statistical Source Anonymity

Let σ denote any adversarial strategy for breaching the anonymity of the system. Let Pr [b0 = b]σ denote the adversary"s probability of winning using strategy σ. We

quantify the anonymity of a sensor network against the strategy σ by

$^{\wedge}\sigma=1-2(P_r[b"=b]\sigma-0.5]$; (1)

Now, let ∑ be the set of all possible adversarial strategies to

breach the anonymity of the sensor network. Then, we

define the anonymity of the system as:

$^{\wedge}=\min\ ^{\wedge}\sigma$

(2)

Definition 2 ($^{\wedge}$_-anonymity). A wireless sensor network is said to be $^{\wedge}$-anonymous if it satisfies two conditions

1. The anonymity of the system, as defined in (2), is at least

$^{\wedge}$,

2. There is no distinguishable transitional behavior between intervals.

3 STATISTICAL GOODNESS OF FIT TESTS AND THE SSA PROBLEM

In the literature, statistical source anonymity is shown to be achieved via the use of statistical goodness of fit tests [11], [19], [20], [21], [22], [23], [24]. In this section, we describe the current use of statistical goodness of fit tests in designing anonymous sensor networks. The challenge, however, is to come up with an appropriate model that captures all possible sources of information leakage and a proper way of quantifying anonymity in different systems.

TABLE 1
A List of Used Terms and Notations

SSA	Statistical Source Anonymity
E_i	The random variable representing the type of events reported in i^{th} transmission
X_i	Random variable representing time between i and $i+1^{th}$ transmission
M	Desired mean of X_i
l_F	Fake Interval
l_R	Real Interval

4.1 SSA Solutions Based on Statistical Goodness of Fit Tests

The statistical goodness of fit of an observed data describes how well the data fits a given statistical model. The following is a description of how statistical goodness of fit tests has been used to design anonymous sensor networks. Let sensor nodes be designed to transmit independent identically distributed (iid) fake messages according to a prespecified probabilistic distribution, D, with a desired mean, μ. Furthermore, let nodes store a sliding window of times between consecutive transmissions (intertransmission times), say Xi; Xi+1; . . .; Xk+i-1, where Xj is the random variable representing the time between the jth and the (j+1) st transmissions, and k is the length of the sliding window.

4 EXPERIMENTAL ANALYSIS OF SSA SOLUTIONS BASED ON STATISTICAL GOODNESS OF FIT

The use of statistical goodness of fit tests in designing anonymous sensor networks was pioneered by Shao et al. in [11] and followed by schemes that build on it or acknowledge its effectiveness in providing secure SSA for sensor networks, such as [19], [20], [21], [22], [23], [24]. In this section, we analyze schemes based on statistical goodness of fit tests using the ideas implied by the theoretical analysis of Section 4.2.

5.1 Converting Real-Valued Samples to Binary Codes

; if xi > =μ

if xi <=μ

i = 1,2,. . . n. (We use g to denote the indicator

instead of the commonly used notation, I, since I is used to denote an interval.) Then, the real-valued e, X, is transformed into a binary code as follows: $x_1,x_2; \ldots ; x_n\}=(g(x_1); \ldots ; g(x_n)\}$. That is, short-long

will be represented by the ordered sequence "01."

e describe the statistical measure that will be used in erimental analysis of SSA solutions based on l goodness of fit tests.

Let every intertransmission time that is shorter than the mean _ be represented by the binary digit "0," and every intertransmission time that is longer than the mean μ be represented by the binary digit "1." That is, given a sequence of real-valued intertransmission times X =f{x1; . . . ; xn}

6 EXPLANATIONS FOR DISCREPANCIES BETWEEN OUR RESULTS AND PRIOR STUDIES

In statistical decision theory, the term "nuisance parameters" refers to information that is not needed for hypothesis testing and, further, can preclude a more accurate decision making [30]. When performing hypothesis testing of data with nuisance parameters, it is desired (even necessary in some scenarios) to find an appropriate transformation of the data that removes or minimizes the effect of the nuisance information before performing the hypothesis testing [30]. That is, given a data sample X ={x1; . . . ; xn} that belongs to one of two possible hypotheses H0 or H1, the test is performed on a transformation of the data sample, f(X), rather than the original data itself, X.

7 IMPROVING SSA VIA INDUCED CORRELATION IN FAKE INTERVALS

Our analysis of SSA solutions based on statistical goodness of fit tests shows that the use of such statistical tools is insufficient to guarantee source anonymity. In particular, not only the real-valued intertransmission times must be indistinguishable from the desired distribution of fake transmissions, but also the binary codes representing the intertransmission times of fake and real intervals must have indistinguishable statistical properties. In what follows, we describe a modification to approaches based on statistical goodness of fit tests to improve their anonymity. The main idea behind the proposed approach is the attempt to induce the same correlation pattern of intertransmission times during real intervals into intertransmission times during fake intervals.

8 RELATED WORKS

The privacy problem in wireless sensor networks comes in different flavors. Proposals dealing with providing sink anonymity in wireless sensor networks have appeared in, e.g., [32], [33], [34], [35], [36]. Network coding-based approaches that protect against traffic analysis have appeared in, e.g., [37], [38], [39]. The privacy problem most relevant to this work is the source location privacy in wireless sensor networks. The source location privacy in sensor networks is part of a broader area, the design of anonymous communication systems. The source location privacy in sensor networks is part of a broader area, the design of anonymous communication systems. The foundation for this field was laid by Chaum in [40], Path confusion has also been proposed as an anonymity-preserving routing scheme by Hoh and Gruteser in [18]. In the first scheme, some sensor nodes act as fake sources by mimicking the behavior of real events. For example, if the network is deployed to track

an animal, the fake sources could send fake messages with a

distribution resembling that of the animal"s movements. This, however, assumes some knowledge of the time distribution of real events. In the second scheme, packets (real and fake) are sent either at constant intervals or according to a predetermined probabilistic schedule. Although this scheme provides perfect location privacy, it also introduces undesirable performance characteristics, in the form of either relatively high delay or relatively high communication and computational overhead. The scheme of [11] was proposed to address this delay/overhead tradeoff. In [11], Shao et al. introduced the notion of statistically strong source anonymity in which a global adversary with ability to monitor the traffic in the entire network is unable to infer source locations by performing statistical analysis on the observed traffic. In order to realize their notion of statistical anonymity, nodes are programmed to transmit fake events according to prespecified distribution. More specifically, after the transmission of every fake event, the node draws an exponentially distributed random variable $t \sim Exp(\lambda)$, where λ is the prespecified rate of the exponential distribution. The node then waits for t time units and then transmits another fake event. That is, in the absence of real event transmissions, an adversary monitoring the sensor node will observe intertransmission times that are iid exponentials with mean $\mu=\lambda$.

Upon the occurrence of real events, the goal of a sensor node is to transmit them while maintaining the exponential distribution of the intertransmission times. Obviously, if nodes delay their transmission of real events to the next scheduled fake transmission, no statistical test can be used to distinguish between real and fake events (since intertransmission times are kept exponential iid"s with the same rate). The goal in [11], however, is to minimize the latency of reporting real events while maintaining statistical indistinguishability between real and fake transmissions. To reduce the amount of traffic in the network that is due to the transmission of fake events, techniques based on node proxies and data aggregation have been proposed [19], [44]. In such techniques, the overall communication overhead is reduced by making intermediate nodes act as proxies that filter out fake messages or by aggregating multiple messages in a single transmission.

9 CONCLUSIONS AND FUTURE WORK

In this paper, we provided a statistical framework based on binary hypothesis testing for modeling, analyzing, and evaluating statistical source anonymity in wireless sensor networks. We introduced the notion of interval indistinguishability to model source location privacy. We showed that the current approaches for designing statistically anonymous systems introduce correlation in real intervals while fake intervals are uncorrelated. By mapping the problem of detecting source information to the statistical problem of binary hypothesis testing with nuisance parameters, we showed why previous studies were unable to detect the source of information leakage that was demonstrated in this paper. Finally, we proposed a modification to existing solutions to improve their anonymity against correlation tests.

Future extensions to this work include mapping the problem of statistical source anonymity to coding theory by using sequential hypothesis testing and robust encryption algorithm in order to design an efficient system that satisfies the notion of interval indistinguishability.

REFERENCES

[1] B. Alomair, A. Clark, J. Cuellar, and R. Poovendran, "On Source Anonymity in Wireless Sensor Networks," Proc. IEEE/IFIP 40th Int"l Conf. Dependable Systems and Networks (DSN "10), 2010. | [2] B. Alomair, A. Clark, J. Cuellar, and R. Poovendran, "Statistical Framework for Source Anonymity in Sensor Networks," Proc. IEEE GlobeCom, 2010. | [3] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless Sensor Networks: A Survey," Computer Networks, vol. 38, no. 4, pp. 393-422, 2002. | [4] T. Arampatzis, J. Lygeros, and S. Manesis, "A Survey of | Applications of Wireless Sensors and Wireless Sensor Networks," Proc. IEEE 13th Mediterranean Conf. Control and Automation (MED "05), pp. 719-724, 2006. | [5] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless Sensor | Network Survey," Computer Networks, vol. 52, no. 12, pp. | 2292-2330, 2008. | [6] L. Schnauzer and V. Gligor, "A Key-Management Scheme for Distributed Sensor Networks," Proc. Ninth ACM Conf. Computer and Comm. Security (CCS "02), pp. 41-47, | 2002. | [7] A. Perrig, R. Szewczyk, J. Tygar, V. Wen, and D. Culler, "SPINS: Security Protocols for Sensor Networks," Wireless Networks, vol. 8, no. 5, pp. 521-534, 2002. | [8] D. Hong, J. Sung, S. Hong, J. Lim, S. Lee, B. Koo, C. Lee, D. Chang, J. Lee, K. Jeong, H. Kim, J. Kim, and S. Chee, "HIGHT: A New Block Cipher Suitable for Low-Resource Device," Proc. Eighth Int"l Workshop | | cryptographic Hardware and Embedded Systems (CHES | "06), pp. 46-59, 2006. | [9] A. Bogdanov, L. Knudsen, G. Leander, C. Paar, A. | Poschmann, M. Robshaw, Y. Seurin, and C. Vikkelsoe, "PRESENT: An Ultra-Lightweight Block Cipher," Proc. Ninth Int"l Workshop Cryptographic Hardware and Embedded Systems (CHES "07), pp. 450-466, 2007. | [10] K. Mehta, D. Liu, and M. Wright, "Location Privacy in | Sensor Networks against a Global Eavesdropper," Proc. | IEEE 15th Int"l Conf. Network Protocols (ICNP "07), pp. | 314-323, 2007. | [11] M. Shao, Y. Yang, S. Zhu, and G. Cao, "Towards Statistically Strong Source Anonymity for Sensor Networks," Proc. IEEE INFOCOM, pp. 466-474, 2008. | [12] P. Kamat, Y. Zhang, W. Trappe, and C. Ozturk, "Enhancing Source-Location Privacy in Sensor Network Routing," Proc. IEEE 25th Int"l Conf. Distributed Computing Systems (ICDCS "05), pp. 599-608, and 2005. | [13] C. Ozturk, Y. Zhang, and W. Trappe, "Source-Location Privacy in Energy-Constrained Sensor Network Routing," Proc. Second ACM Workshop Security of Ad Hoc and | Sensor Networks (SASN "04), pp. 88- 93, 2004. | [14] Y. Ouyang, Z. Le, G. Chen, J. Ford, F. Makedon, and U. Lowell, "Entrapping Adversaries for Source Protection in Sensor Networks," Proc. IEEE Seventh Int"l Symp. World of Wireless, Mobile and Multimedia Networks (WOWMOM | "06), pp. 32-41, 2006. | [15] X. Wang, X. Li, Z. Wan, and M. Gu, "CLEAR: A | Confidential and Lifetime-Aware Routing Protocol for Wireless Sensor Network," Proc. IEEE 20th Ann. Int"l Symp. Personal, Indoor and Mobile Radio Comm. (PIMRC "09), pp. | 2265-2269, 2009. | [16] Y. Li and J. Ren, "Source-Location Privacy through Dynamic Routing in Wireless Sensor Networks," Proc. IEEE INFOCOM,pp. 1-9, 2010. | [17] Y. Xi, L. Schwiebert, and W. Shi, "Preserving Source Location Privacy in Monitoring-Based Wireless Sensor Networks," Proc. IEEE 20th Int"l Parallel & Distributed | Processing Symp. (IPDPS "06), pp. 1-8, 2006. | [18] B. Hoh and M. Gruteser, "Protecting Location Privacy | Through Path Confusion," Proc. IEEE/CreatNet First Int"l | Conf. Security and Privacy for Emerging Areas in Comm. | Networks (SecureComm "05),pp. 194-205, 2005. | [19] Y. Yang, M. Shao, S. Zhu, B. Urgaonkar, and G. Cao, "Towards Event Source Unobservability with Minimum Network Traffic in Sensor Networks," Proc. First ACM Conf. Wireless Network Security (WiSec "08), pp. 77-88, | 2008. | [20] N. Li, N. Zhang, S. Das, and B. Thuraisingham, "Privacy Preservation in Wireless Sensor Networks: A State- of-the-Art Survey," Elsevier J. Ad Hoc Networks, vol. 7, no. | 8, pp. 1501-1514,2009. | | [21] H. Wang, B. Sheng, and Q. Li, "Privacy-Aware Routing in Sensor Networks," Elsevier J. Computer Networks, vol. | 53, no. 9, pp. 1512-1529, 2009. | [22] M. Shao, W. Hu, S. Zhu, G. Cao, S. Krishnamurthy, and T. La Porta, "Cross-Layer Enhanced Source Location Privacy in Sensor Networks," Proc. IEEE Comm. Soc. Sixth Ann. Conf. Sensor, Mesh and Ad Hoc Comm. and Networks (SECON "09), pp. 324-332, 2009. | [23] B. Carbunar, Y. Yu, W. Shi, M. Pearce, and V. Vasudevan, "Query Privacy in Wireless Sensor Networks," ACM Trans. Sensor Networks, vol. 6, no. 2, pp. 1-34, 2010. [24] Y. Zeng, J. Cao, S. Zhang, S. Guo, and L. Xie, "Random-Walk Based Approach to Detect Clone Attacks in Wireless Sensor Networks," IEEE J. Selected Areas in Comm., vol. 28, no. 5, pp. 677-691, June 2010. | [25] S. Goldwasser and S. Micali, "Probabilistic Encryption," | J. Computer and System Sciences, vol. 28, no. 2, pp. 270- | 299, 1984. | [26] T. Anderson and D. Darling, "Asymptotic Theory of Certain „Goodness of Fit" Criteria Based on Stochastic Processes," The Annals of Math. Statistics, vol. 23, no. 2, pp. | 193-212, 1952. | [27] F. Massey Jr., "The Kolmogorov-Smirnov Test for Goodness of Fit," J. Am. Statistical Assoc., vol. 46, no. 253, pp. 68-78, 1951. | [28] C. Jarque and A. Bera, "A Test for Normality of | Observations and Regression Residuals," Int"l Statistical | Rev./Revue International deStatistique, vol. 55, no. 2, pp. | 163-172, 1987. | [29] S. Golomb and G. Gong, Signal Design for Good | Correlation.Cambridge Univ., 2005. | [30] L. Scharf, Statistical Signal Processing: Detection, Estimation, and Time Series Analysis. Addison-Wesley, | 1991. | [31] H. Karl and A. Willig, Protocols and Architectures for | Wireless Sensor Networks. Wiley, 2005. | [32] Q. Gu, X. Chen, Z. Jiang, and J. Wu, "Sink-Anonymity Mobility Control in Wireless Sensor Networks," Proc. IEEE Fifth Int"l Conf. Wireless and Mobile Computing, Networking and Comm. (WiMob "09),pp. 36-41, 2009. | [33] E. Shakshuki, T. Sheltami, N. Kang, and X. Xing, "Tracking Anonymous Sinks in Wireless Sensor Networks," Proc. IEEE 23rd Int"l Conf. Advanced Information networking and Applications (AINA "09), pp. 510-516, 2009. [34] E. Ngai and I. Rodhe, "On Providing Location Privacy for Mobile Sinks in Wireless Sensor Networks," Proc. 12th ACM Int"l Conf.Modeling, Analysis and Simulation of Wireless and Mobile Systems(MSWiM "09), pp. 116-123, | 2009. | [35] E. Ngai, "On Providing Sink Anonymity for Sensor Networks,"Proc. Fifth Int"l Conf. Wireless Comm. and Mobile Computing:Connecting the World Wirelessly (IWCMC "09), pp. 269-273, 2009. | | [36] Z. Liu and W. Xu, "Zeroing-In on Network Metric Minima for Sink Location Determination," Proc. Third ACM Conf. Wireless Network Security (WiSec "10), pp. 99-104, | 2010. | [37] E. Ayday, F. Delgosha, and F. Fekri, "Location-aware Security Services for Wireless Sensor Networks Using Network Coding," | Proc. IEEE INFOCOM, pp. 1226-1234, 2007. | [38] Y. Fan, Y. Jiang, H. Zhu, and X. Shen, "An Efficient | Privacy- Preserving Scheme against Traffic Analysis Attacks | in Network | Coding," Proc. IEEE INFOCOM, pp. 19-25, 2009. | [39] Y. Jiang, Y. Fan, X. Shen, and C. Lin, "A Self-Adaptive | Probabilistic Packet Filtering Scheme against Entropy Attacks in Network Coding," Computer Networks, vol. 53, no. 18, pp. 3089-3101, 2009. | [40] D. Chaum, "Untraceable Electronic Mail, Return | Addresses, and Digital Pseudonyms," Comm. ACM, vol. 24, | no. 2, pp. 84-90, 1981. | [41] M. Reed, P. Syverson, and D. Goldschlag, "Anonymous | Connections and Onion Routing," IEEE J. Selected Areas in | Comm., vol. 16,no. 4, pp. 482-494, May 1998. |