A Probabilistic Model for Lifetime Measurement in Privacy-Aware Sensor Networks

Xice Sun, Zhi Wang
State Key Lab of Industrial Control Technology
Zhejiang University
Hangzhou, China
{xcsun, wangzhi}@iipc.zju.edu.cn

Honglong Chen, Wei Lou
Department of Computing
The Hong Kong Polytechnic University
Kowloon, Hong Kong
{cshlchen, csweilou}@comp.polyu.edu.hk

***Abstract*—Considering the resource-constraint metrics in sensor networks, it is necessary to develop the energy-efficient techniques to save the limited power while applying the security and privacy protection mechanisms during the data transmission. In this paper, we propose to define the probabilistic lifetime model of a single sensor node to measure the energy-efficient solutions towards the privacy-aware sensor networks. In addition, we tailor our model with three typical stochastic distributions in terms of typical application scenarios in reality. Finally, the probabilistic network lifetime model is formally defined and the experimental simulations validate our theoretical results, which shows its availability in sensor networks.**

I. Introduction

Sensor networks consist of a large amount of small and low-cost sensor nodes with wireless communication abilities, limited processing capacities and power supply, which operate in the collaborative way. It is feasible to complete the monitoring task of hotshots or data transmission by deploying the sensor nodes both in the deterministic and random manners. The software configuration, communication capability and physical structure of a sensor node should affect the quality of service (QoS) indexes such as real-time, fault-tolerance and network lifetime etc. In particular, resource-constraint is an essential characteristic for sensor networks [1]–[4], which imposes great challenges on prolonging the network lifetime both in terms of the energy efficiency and infrastructure limitation.

Commonly, the data flow is transmitted within the sensor networks featured by data-centric metrics via multiple hops and thus vulnerable to the content privacy threat and contextual privacy threat [5]. These attacks either aim to compromise the confidentiality, authentication and integrity by manipulating the exact content of the transmission packets across the network, or deduce the association relationship that is disclosed through data transmission pattern or accumulated sensitive information as quasi-identifiers. As shown in Fig. 1, users from different social communities are not willing to reveal their identities while entertaining the network-based services in an intelligent city. Hence, the anti-mechanisms are proposed to take against the potential threat, which will result in the energy consumption of sensor nodes. As the experiments reported in [6], applying security and privacy related techniques to the transmitted messages over the wireless links in sensor networks requires additional energy, which is based

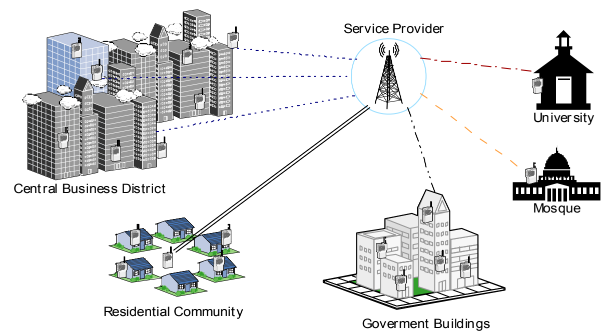


Fig. 1. An intelligent city based on privacy-aware sensor networks

on the down-to-earth experiments implemented on Crossbow and Ember sensor nodes. For instance, to protect the source privacy in sensor networks, researchers make great effort on designing the privacy-preserving techniques, such as flooding [5], random walk [7], dummy injection [8], fake data sources [9] etc, in which the cost of energy consumption is always considered as investigated in this paper. However, there are still open questions on formally defining the lifetime model under strict real analysis and probabilistic theory [10], [11], which can better illustrate the lifetime tendency as the time proceeds in privacy-aware sensor networks.

To tackle these challenges, we propose to numerically formulate the lifetime model of sensor networks based on the stochastic probability of any single sensor node. To the best of our knowledge, it is the first attempt to formulate the lifetime model of sensor networks under stochastic analysis, combined with the particular energy distribution models in the pervasive computing environment. The key contributions of this paper are summarized as follows:

- Define the probabilistic model for lifetime measurement of a single sensor node under stochastic analysis;
- Present the tailored models with three kinds of pervasive lifetime distributions, namely uniform distribution, exponential distribution and normal distribution;
- Propose the probabilistic lifetime model of a sensor network in a numerical way and the extensive simulations validate the theoretical results.

The reminder of this paper is organized as follows. Section

II presents the definitions of our probabilistic model with mathematical formulas. Section III tailors our model with three typical stochastic distributions while considering the initial condition. Section IV proposes the formal definition of the network lifetime based on our proposed probabilistic model. Section V performs simulations to illustrate the performance of our proposed lifetime models. Section VI concludes the paper with outlook on the future work.

II. Model Definition

Basically, a sensor network contains a mass of sensor nodes with capabilities of data collection, wireless communication and self-organization functions, and the network scale greatly depends on the scalability of a particular application scenario. As the structural unit of the whole network, the lifetime of a single sensor node could affect the network topology and performance. In this section, we propose to define the network lifetime of the whole network based on formulating the probabilistic lifetime model of a single sensor node.

The lifetime of a single sensor node is limited, which is subject to current energy supply mechanisms commercialized by Crossbow and Ember sensor nodes both in the academic research and prototype development [6]. We use X to denote the lifetime of a single sensor node, which is a bounded random variable to describe the outer influence exerted by ambient environments, such as the invalidation caused by the energy depletion, the uniform attack [12] or the inference attack [13]. We consider a sensor network designed for monitoring the traffic flow or context-aware location-based system, where the sensor nodes are powered by lithium batteries to provide the successive energy to complete the surveillance job. Suppose that the probability distribution function $\Theta(X)$ of X is continuous and the probability density function $\theta(X)$ also exists to guarantee the validity in the following mathematical analysis.

Definition 1. *Lifetime Distribution Function (LDF):* LDF is defined as $L(t) = P(X > t)$, $t \in [0, \infty)$, which indicates the probability that the lifetime of a single sensor node from the initial operation to the total burst-up is no less than t. The corresponding expectation is denoted as $E(X) = \int_0^\infty L(t)dt$.

Lemma 1. *Given non-negative random variable X and integer n, two LDFs are denoted as* $L(t) = P(X > t)$ *and* $\Theta(t) = P(X \leq t)$*, respectively. If* $E(X^n) < \infty$*, then we can get* $E(X^n) = \int_0^\infty nt^{n-1}L(t)dt$.

Proof: From the definition of Matrix Moment $E(X^n) = \int_0^\infty t^n d\Theta(t) < \infty$, if $M \to \infty$, then

$$\int_M^\infty t^n d\Theta(t) < \infty.$$

Hence, we can get

$$\int_M^\infty t^n d\Theta(t) \geq \int_M^\infty M^n d\Theta(t) = M^n L(M) \to \infty.$$

In addition, M is the continuous point of $\Theta(t) = P(X \leq t)$, thus we can further derive

$$\begin{aligned} E(X^n) &= \lim_{M\to\infty} \int_0^M t^n d\Theta(t) = -\lim_{M\to\infty} \int_0^M t^n dL(t) \\ &= -\lim_{M\to\infty} ([t^n L(t)]|_0^M - \int_0^M nt^{n-1} L(t)dt) \\ &= \lim_{M\to\infty} [-M^n L(M)] + \lim_{M\to\infty} \int_0^M nt^{n-1} L(t)dt \\ &= \int_0^\infty nt^{n-1} L(t)dt. \end{aligned}$$

The second order moment is $E(X^2) = \int_0^\infty 2tL(t)dt$, where $n = 2$ in this case. ■

From *Definition 1*, $L(t) = 1 - P(X \leq t) = 1 - \Theta(t)$ if $\theta(X)$ is continuous. We can prove that

$$\begin{aligned} \beta(t) &= \lim_{\triangle t\to 0+} \frac{P(t < X \leq t + \triangle t \mid X > t)}{\triangle t} \\ &= \lim_{\triangle t\to 0+} \frac{P(t < X \leq t + \triangle t)}{P(X > t)\triangle t} \\ &= \lim_{\triangle t\to 0+} \frac{\int_t^{t+\triangle t} \theta(s)ds}{(1 - \Theta(t))\triangle t} \\ &= \frac{\theta(t)}{1 - \Theta(t)}. \end{aligned}$$

Definition 2. *Failure Probability Function (FPF):* We define FPF of a single sensor node as $\beta(t) = \frac{\theta(t)}{1-\Theta(t)}$, which shows the probability that the sensor node with expected lifetime t just suffers the functional failure (i.e., stop working) at t from the initialization time.

Combining *Definition 1 and 2*, we have shown that

$$\beta(t) = \frac{\theta(t)}{1 - \Theta(t)} = \frac{-dL(t)/dt}{L(t)}.$$

Then

$$\ln L(t) = -\int_0^t \beta(s)ds + C.$$

C is an integer constant and based on the initial condition $L(0) = 1$, $C = 0$ is acquired. Furthermore, we can get

$$L(t) = e^{-\int_0^t \beta(s)ds}.$$

From $\beta(t) = \frac{\theta(t)}{1-\Theta(t)}$, then

$$\theta(t) = \beta(t)(1 - \Theta(t)) = \beta(t)L(t) = \beta(t)e^{-\int_0^t \beta(s)ds},$$

which shows the relationship between LDF and FPF.

III. Typical distribution-based lifetime models

Towards the different application cases, the lifetime of a sensor network expresses diverse probabilistic distributions in particular, including the uniform distribution, the exponential distribution and the normal distribution. In this section, we tailor our proposed probabilistic model with above three stochastic distributions, which indicate the probabilistic idiosyncrasy of the lifetime distribution from the initial operation to complete failure.

A. Uniform distribution-based lifetime model

To complete the continuous monitoring task of the precious antiques in Louvre, the indoor privacy-aware sensor network is deployed at the predetermined locations, which is responsible to perform the surveillance protection over the time. As being known, the distribution of energy initialization is roughly considered as the uniform distribution and the lifetime of sensor nodes show the periodic metrics in different network operational phases. If the vital signs of the sensor nodes' lifetime comply with the uniform distribution, namely $X \sim U(a,b)$, we derive $\Theta(t) = \int_a^t \frac{1}{b-a}dt = \frac{t-a}{b-a}, (a < t \le b)$, where $\theta(t) = \frac{1}{b-a}$ and $(a < t \le b)$. Thus

$$L(t) = \frac{b-t}{b-a},\ \beta(t) = \frac{1}{b-t}.$$

B. Exponential distribution-based lifetime model

To defend against the global eavesdropping attacks [8] and save the limited energy of the data gathering sensor networks, the average rate of energy consumption seems less fluctuant and the tiny change of the available energy can be neglected to certain extent. Hence, the lifetime of such kind of sensor networks follow the exponential distribution, which accords with the non-memory characteristics of the electronics component. If the vital signs of the sensor nodes' lifetime comply with the exponential distribution, namely $X \sim E(\lambda)$, we derive $\Theta(t) = \int_0^t \lambda e^{-\lambda t} = 1 - e^{-\lambda t}$, where $\theta(t) = \lambda e^{-\lambda t}$ and $t \ge 0$. Thus

$$L(t) = e^{-\lambda t},\ \beta(t) = \lambda.$$

C. Normal distribution-based lifetime model

As for the common sensor nodes, it is natural to regard the lifetime distribution of the whole network as the normal distribution in the event-driven distributed networks. For instance, the sensor networks may suffer from the remote control attack or the wormhole attack during the localization phase [14], [15]. If the vital signs of the sensor nodes' lifetime comply with the normal distribution, namely $X \sim N(\mu, \sigma^2)$, the formulation is denoted as $\theta(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}}$ and $t_1 < t \le t_2$. We derive $\Theta(t) = P(X \le t) = P(\frac{X-\mu}{\sigma} \le \frac{t-\mu}{\sigma}) = \phi(\frac{t-\mu}{\sigma})$, where $\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$. Thus

$$L(t) = 1 - \phi(\frac{t-\mu}{\sigma}),\ \beta(t) = \frac{e^{-\frac{1}{2}(\frac{t-\mu}{\sigma})^2}}{\sigma(\sqrt{2\pi} - \int_{-\infty}^{\frac{t-\mu}{\sigma}} e^{-\frac{t^2}{2}} dt)}.$$

D. Model extension with initial condition constraint

Till now, we have discussed the probabilistic lifetime model from the initialization to complete failure, which is due to the running out of the power supply. Another point needed to notice is that, the above mathematical deductions rely on the assumption that the initial condition is zero for the ease of presentation, which is the ideal lifetime case from the network initialization to final expiration.

In reality, the practical network operational phases impose great impacts on the dynamic topology changes, caused by the disequilibrium of energy consumption, the sudden departure of disabled sensor nodes or the reunion of recharged active sensor nodes. We further extend the probabilistic lifetime model with initial condition constraint to formulate the practical case in this section. Here, the target sensor nodes have operated for a time period ξ and we also employ $\varphi(\xi) = X - \xi$ to denote the predictable lifetime of any sensor node, which is a bounded random variable to describe the probabilistic lifetime model after time ξ. By assuming that the power supply of the sensor node is uniquely affected by the energy module, we can derive the probability distribution function and the probability density function as

$$\begin{aligned}\Theta_\xi(t) &= P(X - \xi \le t \mid X > \xi) = \frac{P(X \le \xi + t, X > \xi)}{P(X > \xi)} \\ &= \frac{P(X > \xi) - P(X > \xi + t)}{P(X > \xi)} = 1 - \frac{L(\xi + t)}{L(t)}, \\ \theta_\xi(t) &= \frac{d\Theta_\xi(t)}{dt} = \frac{-dL(t)/dt}{L(t)} = \frac{\theta(\xi + t)}{L(t)}.\end{aligned}$$

The corresponding LDF and FPF are shown as

$$L_\xi(t) = e^{-\int_0^t \beta(\xi+s)ds},\ \beta_\xi(t) = \frac{\theta_\xi(t)}{1 - \Theta_\xi(t)} = \beta(\xi + t).$$

Based on above statements, if the sensor node has operated for a time period ξ and continues to work for another time period t, the overall operating time sums up as $(\xi + t)$. It is common in the energy-efficient link scheduling [16] or operation scheduling [17], localization-based tracking algorithms [18] in sensor networks, where partial of nodes are activated upon demand from the application necessity. From inducing the relationship between these sensor nodes and those who operate consecutively from the very beginning of network initialization, we can interpret the probabilistic correlation as follows if $q \in [0, \infty)$,

$$P(\varphi(\xi) > q + t \mid \varphi(\xi) > t) = P(\varphi(\xi + t) > q).$$

Therefore, the LDF of the sensor nodes who firstly operate for a time period ξ and then live for another t after certain break holds the same result as that of the sensor nodes who operate for $(\xi + t)$ time period consecutively. For $\forall q \in [0, \infty)$, we can show the brief proof as follows

$$\begin{aligned}&P(\varphi(\xi) > q + t) \mid \varphi(\xi) > t) = P(X - \xi > q + t \mid X - \xi > t) \\ &= P(X - \xi - t > q \mid X > \xi + t) = P(\varphi(\xi + t) > q).\end{aligned}$$

Hence, we can safely derive at the conclusion that if we focus on the power supply which is affected uniquely by the energy module of sensor nodes, the probabilistic lifetime model can be deduced from the original probabilistic model introduced in Section II. In addition, it earns the attention that if more influence is taken into account, such as the active attacks towards the sensor nodes or denial of service attacks, the probabilistic model does not remain the same with new assumptions and lifetime models.

IV. Definition of network lifetime based on LDF

As one of the important network QoS measurements in sensor networks, the lifetime of network is used to indicate the fundamental ability of providing and maintaining the desirable service quality. In terms of certain confidence level as concerned, we specify this index as the time period from the beginning of network initialization to the complete failure time when the network QoS cannot be satisfied. In this section, we define the lifetime of network numerically based on our proposed probabilistic lifetime model of any single sensor node, which also takes the impact exerted by the disabled nodes into account.

Suppose there exist k sensor nodes $\{X_1, X_2, ..., X_k\}$ in a sensor network, the LDF follows the same probabilistic distribution $L(t) = P(X > t)$, where $t \in [0, \infty)$. The number of sensor nodes with expected lifetime X_0 is denoted as k_{X_0}. Therefore, the number of sensor nodes with practical lifetime X_0 in the whole network can be presented as

$$K(X_0) = \sum_{i=1}^{k} u_{\{X_i \geq X_0\}},$$

where

$$u_{\{X_i \geq X_0\}} = \begin{cases} 1, & X_i \geq X_0 \\ 0, & X_i < X_0 \end{cases}.$$

By deriving its expectation on both sides, we can get

$$k_{X_0} = E(K(X_0)) = \sum_{i=1}^{k} E(u_{\{X_i \geq X_0\}}) = kP(X_1 \geq X_0) = kL(X_0).$$

Furthermore, we can derive

$$\frac{dk_{X_0}}{dX_0} = \frac{d(kL(X_0))}{dX_0} = -\frac{k_{X_0}}{L(X_0)}\theta(X_0) = -k_{X_0}\beta(X_0).$$

If the lifetime of above k sensor nodes complies with $X_i \in [X_0, X_0 + t]$, where $i \in \mathbb{N}^+$, then we can define the lifetime of network based on our proposed probabilistic lifetime model of any single sensor node as

$$\int_0^t k_{X_0+s} ds = \int_0^t s k_{X_0+s} \beta(X_0 + s) ds + t k_{X_0+t},$$

where

$$\int_0^t k_{X_0+s} ds = -s k_{X_0+s} \mid_0^t + \int_0^t k_{X_0+s} ds + t k_{X_0+t},$$

$$\int_0^t s k_{X_0+s} \beta(X_0 + s) ds + t k_{X_0+t} = -\int_0^t s dk_{X_0+s} + t k_{X_0+t}.$$

The left side $\int_0^t k_{X_0+s} ds$ indicates the accumulated overall lifetime of sensor nodes whose lifetime falls in the range $[X_0, X_0+t]$, namely the lifetime of network. The first monomial of right side $\int_0^t s k_{X_0+s} \beta(X_0 + s) ds$ shows the expectation of the overall lifetime of sensor nodes who have turned down during the time range $[X_0, X_0 + t]$. The second monomial of right side $t k_{X_0+t}$ displays the overall lifetime of sensor nodes who operate for more than $(X_0 + t)$ time period.

V. Performance Evaluation

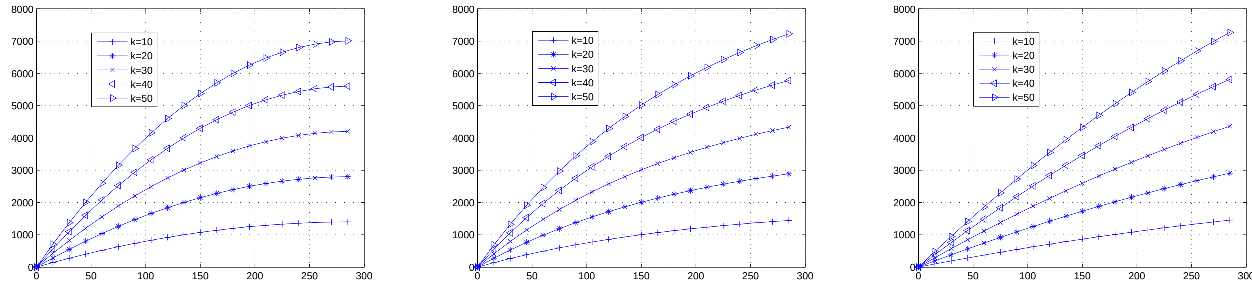
In this section, we evaluate the performance of the proposed probabilistic model based on three typical stochastic lifetime distributions. The program is implemented on a computer running MS Windows Vista OS, with 4.2 GHz CPU and 2 GB memory. To ease the presentation of the following simulations, we specify the model parameters as $a = 0$, $b = 100$, $\lambda = 0.5$, $\mu = 30$, $\sigma = 80$ towards three stochastic lifetime distributions, respectively. In particular, the parameter k and X_0 should place great impacts on the final probabilistic model of network lifetime and thus we conduct the experiments over these two indexes.

As shown in Fig. 2, the network lifetime increases as the time proceeds. Subjected to a fixed X_0, if the number of sensor nodes k is increased, the network lifetime is prolonged according to our proposed model. If we enlarge the network scale, namely k, the network lifetime can be improved accordingly in all three stochastic lifetime distributions. It means that if more sensor nodes can be deployed in the monitoring field, the accumulated network lifetime of the whole sensor network will be improved based on more power-supply units. Thus, our proposed probabilistic network lifetime can best describe the practical cases and present the results in an efficient manner.

As reported in Fig. 3, the network lifetime also increases as the time proceeds. Subjected to a fixed k, if the predicted lifetime of sensor nodes X_0 is increased proportionally, the network lifetime can be improved within all three stochastic lifetime distributions. As the prevalent knowledge, the power supply is constraint under infrastructure limitation and the lifetime of any individual sensor node is approximately fixed as indicated in [6]. The network lifetime will decrease as time proceeds although the expected lifetime of the sensor node X_0 is increased, which also indicates that the energy is running out in the network operational phases based on typical LDFs.

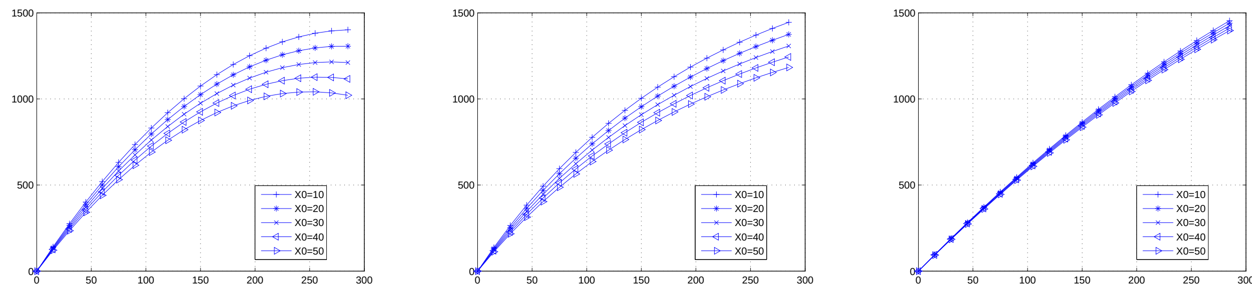
VI. Conclusion and Future Work

In this paper, we firstly propose the probabilistic model for lifetime measurement of a single sensor node under stochastic analysis and present the tailored models with three pervasive lifetime distributions. Then the probabilistic lifetime model of sensor network is formally defined in a numerical way, which complies with diverse kinds of application scenarios in sensor networks. In the future, our work intends to optimize the probabilistic model by considering the cooperative relationship among the sensor nodes and investigate the related issues in cyber-physical systems.



(a) Uniform Distribution (b) Exponential Distribution (c) Normal Distribution

Fig. 2. Impact of K towards the network lifetime model



(a) Uniform Distribution (b) Exponential Distribution (c) Normal Distribution

Fig. 3. Impact of X_0 towards the network lifetime model

ACKNOWLEDGMENT

This work is supported in part by grants National Natural Science Foundation of China (No.60873223 and 90818010), Shanghai Scientific and Technical Key Program (No.07DZ15012), PolyU 5236/06E, PolyU 5243/08E and ZJU-SKL ICT0903. Deeply grateful to the anonymous reviewers for their constructive comments for improving this paper.

REFERENCES

[1] O. Younis and S. Fahmy, "HEED: A hybrid, energy-efficient, distributed clustering approach for ad-doc sensor networks," *IEEE Transactions on Mobile Computing*, pp. 660–669, 3(4), 2004.

[2] Y. Yao and G. Giannakis, "Energy-efficient scheduling for wireless sensor networks," *IEEE Transactions on Communications*, pp. 1333–1342, 53(8), 2005.

[3] H. Kwon, S. C. T. Kim, and B. Lee, "A cross-layer strategy for energy-efficient reliable delivery in wireless sensor networks," *IEEE Transactions on Wireless Communications*, pp. 3689–3699, 5(12), 2006.

[4] R. Blum and B. Sadler, "Energy-efficient signal detection in sensor networks using ordered transmissions," *IEEE Transactions on Signal Processing*, pp. 3229–3235, 56(7), 2008.

[5] K. Kamat, Y. Zhang, W. Trappe, and C. Ozturk, "Enhancing source-location privacy in sensor network routing," in *Proceedings of 25th International Conference on Distributed Computing Systems (ICDCS)*, 2005.

[6] C. Chang, D. Nagel, and S. Muftic, "Balancing security and energy consumption in wireless sensor networks," *Mobile Ad-Hoc and Sensor Networks*, pp. 469–480, 2007.

[7] Y. Xi, L. Schwiebert, and W. Shi, "Preseving scource location privacy in monitoring-based wireless sensor networks," in *Proceedings of 20th IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, 2006.

[8] M. Shao, Y. Yang, S. Zhu, and G. Cao, "Towards statistically strong source anonymity for sensor networks," in *Proceedings of 27th IEEE International Conference on Computer Communications (INFOCOM)*, 2008.

[9] K. Mehta, D. Liu, and M. Wright, "Location privacy in sensor networks against a global eavesdropper," in *Proceedings of 15th IEEE International Conference on Network Protocols (ICNP)*, 2007.

[10] K. Athreya and S. Lahiri, "Measure Theory and Probability Theory," *Springer Press*, 2006.

[11] R. Dudley, "Real Analysis and Probability," *Wadsworth and Brooks/Cole Press*, 1989.

[12] C. Bettini, S. Mascetti, and X. Wang, "Anonymity in location-based services: Towards a general framework," in *Proceedings of 8th International Conference on Mobile Data Management (MDM)*, 2007.

[13] S. Mascetti and C. Bettini, "A comparision of spatial generalization algorithms for LBS privacy," in *Proceedings of International Workshop on Privacy-Aware Location-based Mobile Services (PALMS)*, 2007.

[14] L. Lazos and R. Poovendran, "SeRLoc: Robust localization for wireless sensor networks," *ACM Transactions on Sensor Networks*, pp. 73–100, 2005.

[15] H. Chen, W. Lou, and Z. Wang, "Conflicting-set-based wormhole attack resistant localization in wireless sensor networks," in *Proceedings of 6th International Conference on Ubiquitous Intelligence and Computing (UIC)*, 2009.

[16] J. Ma, W. Lou, Y. Wu, X.-Y. Li, and G. Chen, "Energy efficient TDMA sleep scheduling in wireless sensor networks," in *Proceedings of 28th IEEE International Conference on Computer Communications (INFOCOM)*, 2009.

[17] X. Sun, F. Cao, and Z. Wang, "An optimized sensor node scheduling algorithm for multiple target associated coverage in wireless sensor networks," *Information and Control*, pp. 29–36, 38(1), 2009.

[18] Z. Wang, H. Li, X. Shen, X. Sun, and Z. Wang, "Tracking and predicting moving targets in hierarchical sensor networks," in *Proceedings of IEEE International Conference on Networking, Sensing and Control (ICNSC)*, 2008.