Security Breach Possibility with RSS-Based Localization of Smart Meters Incorporating Maximum Likelihood Estimator

Mahdi Jamei, Arif I. Sarwat, S. S. Iyengar, and Faisal Kaleem

1 Introduction

Recently, there has been a trend towards the Smart Grid (SG) to have secure and reliable electricity [1]. The SG is two-way data transfer in which the information plays a central role in energy dispatching [2]. Smart meter is one of the key components which enable the SG to involve the consumer engagement and demand response concepts.

Increasing rate of the electricity consumers necessitates a revolution in the conventional energy consumption metering and current methods of billing. In addition, human errors and operating costs and the need of improvement in the metering efficiency have encouraged utilities to employ Advanced Metering Infrastructure (AMI) systems [3]. In a report of the U.S. Energy Information Administration, about 37,290,374 AMI were installed by 493 U.S. electric utilities in 2011 [4].

AMI systems collect the energy usage data remotely to remove the inaccuracy and the reading cost. They can also provide both utilities and consumers with real-time consumption data to enable the grid for a better response to demand changes [5]. AMI transmitting and receiving methods have evolved gradually. Early model of AMI used telephone lines to send and receive the data. Power Line Communication (PLC), low power Radio Frequency (RF) and satellite-based communication are the next generations of the AMI systems.

However, smart meters have raised concerns over security and privacy issues [6]. Data collected and transmitted by the AMI systems can endanger the privacy of consumers. For example, robbers can receive and demodulate the sent data by smart meters to localize the unoccupied houses. People's daily routines can also be identified from the time and the amount of the energy consumption [7]. On the other hand, it is possible that home-owners hack the system and falsify the reported data through breaking into the communication channel or feeding the system by counterfeit consumption data [8].

To show the significance of the related security issues, this paper presents a method of localizing a smart meter using the Received Signal Strength (RSS) of the RF transmitted signals. This method shows the vulnerability of the AMI system's security since it allows hackers to identify the location of an intended owner via the energy usage data. It is important to recognize and study the attack scenarios and their effectiveness to find efficient anti-attack remedies in order to increase the impenetrability of the system.

RSS-based localization are mainly categorized as range-based and range-free methods. RSS is used as a distance reference in range-based techniques while range-free methods do not use distances in the localization procedure. In the range-based system, the location of the emitter can be identified by distances from a set of sensors with known positions [9]. For this purpose, at least three sensors are required to make the localization possible.

RSS-based localization problem requires estimation of an unknown parameter i.e. the coordination of the emitter, from a collection of observation data, x[n], by sensors in which additive noise, sensor inaccuracies, shadowing, multipath and path loss exponent have been included. The Maximum Likelihood (ML) estimator is an approach in estimating a parameter when the Probability Density Function (PDF) is known. With MLE, the unknown parameter is estimated by maximizing the PDF [10]. In this paper, RSS-based localization of a smart meter from the sent data by AMI system under the assumption of a log-normal path loss model and additive Gaussian noise has been proposed. ML estimator is employed to estimate the coordination and the reference power of the smart meter using the received signals by the sensors located at the

M. Jamei (✉) • A.I. Sarwat • F. Kaleem
Electrical and Computer Engineering Department, Florida International University, Miami, Florida, USA
e-mail: mjame044@fiu.edu; asarwat@fiu.edu; kaleemf@fiu.edu

S.S. Iyengar
School of Computing and Information Sciences, Florida International University, Miami, Florida, USA
e-mail: iyengar@cis.fiu.edu

H. Selvaraj et al. (eds.), *Progress in Systems Engineering: Proceedings of the Twenty-Third International Conference on Systems Engineering*, Advances in Intelligent Systems and Computing 1089,
DOI 10.1007/978-3-319-08422-0_20,

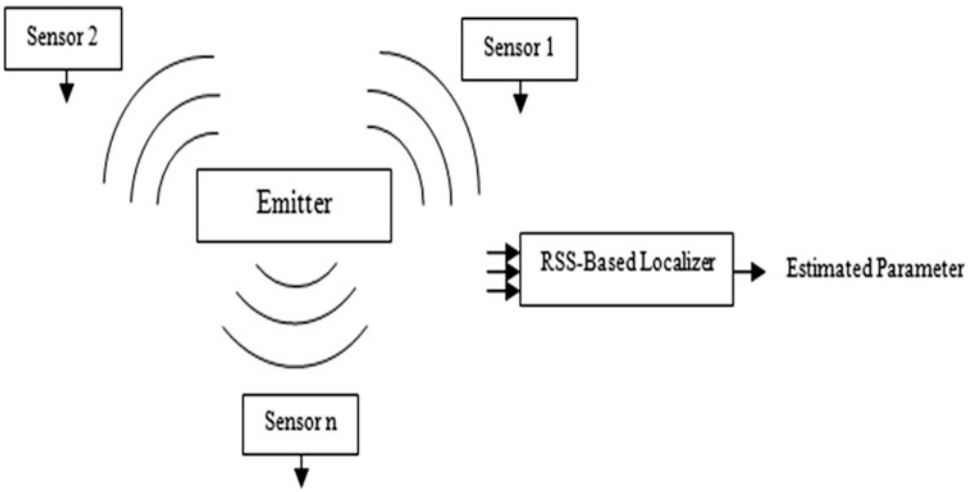

Fig. 1 General Schematic of the Localizer

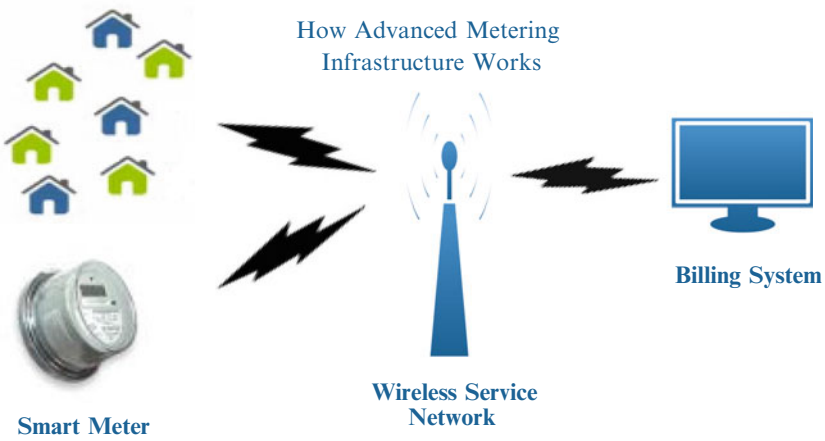

Fig. 2 Outline of the AMI System

known positions. The general schematic of the paper is shown in Fig. 1. The effectiveness of the proposed method is investigated through MATLAB simulation considering the FSK modulation and demodulation. PSO has been implemented to maximize the likelihood function in the ML estimator. Finally, the effect of the variance, the number of the sensors and the path loss exponent has been studied on the average Miss Distance Error (MDE).

The rest of this paper is organized as follows. Section 2 introduces the AMI main framework architecture. Section 3 presents the ML estimator concepts and the proposed method of the RSS-based localization problem. Simulation data and results are presented and discussed in Section 4. Finally, the conclusion is given in Section 5.

2 AMI Architecture

AMI system collects the energy usage data and transmits to the central collector for billing and analysis purposes. Most of the new AMI systems have been equipped with the RF or satellite communication. RF technology is broadly in use since it is more cost effective and easier to be implemented. The RF method is considered to be the communication infrastructure in this paper. Fig. 2 illustrates the outline of the AMI system.

AMIs are mainly classified as AMI meters that collect the energy usage data and AMI readers which receive and send it to the central processing offices [5].

2.1 AMI Meter

AMI meters gather the electricity, gas and water consumption. RF-based meters include Encoder, Receiver and Transmitter (ERT) which has a microprocessor as well as a low power transmitter [5]. The microprocessor processes the meter reading periodically and feed it into the transmitter to be sent along with the information such as the meter ID.

2.2 AMI Reader

AMI readers are used to receive the data from AMI meter and send it to the utilities main data collectors. There are three different types of readers in the current electricity grid [5], [11].

i. Handheld instruments for walk-by data collection
ii. Mobile data collection for drive-by data reading
iii. Permanent infrastructure for real-time data transfer

The first two classes require personnel to gather data so the information can be updated periodically but the permanent infrastructure can provide real-time energy data. On the other hand, the initial start-up cost of installing the fixed data transfer system is considerable so benefit-cost analysis must be conducted before choosing one of these two mentioned categories.

2.3 Communication Protocol

The communication module provides a robust connectivity in the Neighborhood Area Network (NAN). The data rates for the NAN communication is 100 kbps with the transmitter output of 27-30 dBm (500 mW - 1 W). It supports the frequency range from 902-928 MHz with 915 MHz ISM band. The module also include a 2.4 GHz radio for the Home Area Network (HAN) which supports the ZigBee protocol to communicate with the smart devices inside the house. The transmitter output of the HAN communication is 20 to 23 dBm (100 - 200 mW). On-Off Keying (OOK) and FSK are usual types of modulation used in the transmitter and the receiver.

3 Problem Formulation

In this section, the concept of the ML estimator has been introduced first. The localization problem has also been formulated and related equations are given to be implemented in the Smart Meter security issue.

3.1 ML Estimator

The ML estimator is overwhelmingly the most famous approach to obtain practical estimators. In most of the cases, by large enough data observation, the ML estimator would be the optimal one. It can be said that the ML estimator is approximately Minimum Variance Unbiased Estimator (MVUE) [10].

Generally, the ML estimator defined as the value of θ which maximizes the likelihood function. The MLE is said to be asymptotically optimal which means that the properties of unbiasedness and achieving the Cramer-Rao Lower Bound (CRLB) can only be obtained by recording large enough data. The ML estimator of a vector parameter θ is the value maximizing the likelihood function which is already a function of the components of θ.

The ML estimator formulation is given as follows [10]:

$$\hat{\theta}_{MLE} = \underset{\theta}{argmax}\, P(X;\theta) \tag{1}$$

Where $\hat{\theta}$ is the estimated vector parameter of θ, X is the observation data vector, $P(X; \theta)$ is the PDF of X depending on θ which has been parameterized on it.

3.2 Proposed Method

Suppose that the smart meter is located at an unknown position (x, y) and there are n receiver sensors located at known coordination (x_l, y_l), $1 \leq l \leq n$. The RSS localization method use the distance between the smart meter and sensors as the main measurements to obtain the position of the emitter. Friis Transmission formula is the most common one used in this model [12] as shown in (2).

$$\Omega_l = P_t + G_t + G_r + 10\alpha \log_{10}\left(\frac{\lambda}{4\pi d_l}\right) \tag{2}$$

where Ω_l is the RSS in (dBm) measured by the l-th sensor, P_t denotes the transmission power, G_t is the transmit gain, G_r is the received gain, α is the signal carrier frequency, called path loss exponent and d_l is the propagation distance. (2) can be rewritten as (3) if the received power is known at a close distance to the smart meter, called the log-distance path loss model [9], [12].

$$\Omega_l = C - 10\alpha \log_{10}\left(\frac{d_l}{d_0}\right) + n_l \tag{3}$$

where C is the received power at distance d_0, and d_l is the Euclidean distance between the smart meter and the receiver at (x_l, y_l), i.e.,

$$d_l = \sqrt{(x - x_l)^2 + (y - y_l)^2} \tag{4}$$

The effects of the shadowing is included in (3) denoted by n_l and considered to be a zero-mean Gaussian random variable with known covariance matrix. Let define the column vector θ and Ω as follow:

$$\theta = [x \quad y \quad c]^T \Omega = [\Omega_1 \quad \Omega_2 \quad . \quad . \quad . \quad \Omega_n]^T \tag{5}$$

Assume that $f(\theta|\Omega)$ is representative for the likelihood function of θ based on the RSS measurements, Ω_l, $1 \leq l \leq n$. Using the path loss model in (3), and considering a diagonal covariance matrix with standard deviation σ_l, $1 \leq l \leq n$, $f(\theta|\Omega)$ can be obtained in the following form [12]:

$$f(\theta|\Omega) = c_p \exp\left\{ -\sum_{l=1}^{n} \frac{\left\{ \Omega_l - C + 10\alpha \log_{10}\left(\frac{d_l}{d_0}\right) \right\}^2}{2{\sigma_l}^2} \right\} \tag{6}$$

where c_p is a positive constant and independent from the vector parameter θ. $\hat{\theta}$ which represents the ML estimator of θ can be obtained by solving the optimization problem as given below [9]:

$$\begin{aligned} \hat{\theta}_{MLE} &= \arg\max_{\theta} f(\theta|\Omega) \\ &= \text{argmin} \left\{ \sum_{l=1}^{n} \frac{\left\{ \Omega_l - C + 10\alpha \log_{10}\left(\frac{d_l}{d_0}\right) \right\}^2}{{\sigma_l}^2} \right\} \end{aligned} \tag{7}$$

The PSO algorithm is implemented to solve this non-linear optimization problem. The ML estimator returns the estimated position and the reference transmission power of the smart meter, i.e,

$$\left[\hat{\theta}_1 \quad \hat{\theta}_2 \quad \hat{\theta}_3 \right] = \left[\hat{x}_{rss} \quad \hat{y}_{rss} \quad \hat{c}_{rss} \right] \tag{8}$$

Without any prior knowledge about the position and the reference power, optimization problem must be done through Euclidean R^3 but most of the practical problems contain some prior information about the vector parameter. In this case, Bayesian philosophy can be employed to find an estimation of the unknown vector parameter.

3.3 Particle Swarm Optimization (PSO)

The PSO algorithm, first proposed by Kennedy and Eberhart [13], considered to be a method for optimization based on the social behavior of flocks of birds or schools of fish. The standard PSO algorithm first will start by generating random positions for the particles, within an initialization region. Initializing the velocities can be done through using values within that region or they can be selected zero or small random values to prevent particles from leaving the search space during the first iterations. During the main loop of the algorithm, the velocities and positions of the particles are iteratively updated until a stopping criterion is met. The update rules are completely described in [13].

4 Simulation Results

This section investigates the effectiveness of the proposed method through MATLAB simulations. Three scenarios are considered to scrutinize the effect of the variance, the number of the sensors, and the path loss exponent on the localization accuracy and the average MDE. Noise is assumed to be Additive White Gaussian Noise (AWGN). MDE is defined in the following form:

$$MDE = \sqrt{\left(\hat{x}_{rss} - x \right)^2 + \left(\hat{y}_{rss} - y \right)^2} \tag{9}$$

PSO parameters are, iteration number =100, population size = 30 and $C_1, C_2 = 2.05$.

4.1 Scenario1

This scenario presents the effect of the variance on the localization problem. 6 sensors are located at vertices of a regular hexagon with 20 meters diameter and (0,0) center. The smart meter position is at (1,-1) and the AWGN has the $N(0, \sigma^2)$ distribution. From Fig. 3, it can be inferred that as the variance increases, the average MDE will go up. It is rational since higher variances for the noise means that the observation data deviates from the real value in a way that estimating the location of the emitter will contain more errors. Table 1 illustrates the estimated values of the vector parameter θ w.r.t the variance.

4.2 Scenario 2

The effect of the number of the sensors has been inspected in this part. The smart meter position is at (1,-1) and the AWGN has the $N(0, 4)$ PDF distribution. Sensors are located around the smart meter at known coordination on the vertices of a square, hexagon, octagon and ten-sided, respectively. Fig. 4 depicts the variation of the average MDE w.r.t the number of the sensors. As it can be seen, increasing the number of the sensors will decrease the MDE error since the number of the observations will increase. As a result, the estimated vector parameter will be more accurate and will return the results close enough to the real values. Table 2 shows the estimated values of the parameter and indicates the direct proportional relationship between the sensors and the average MDE.

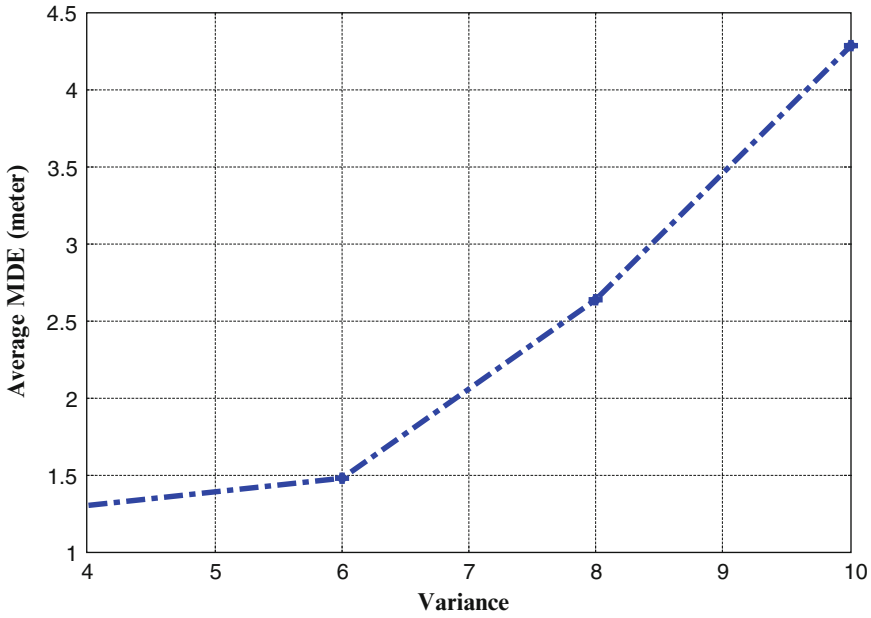

Fig. 3 Average MDE w.r.t the Variance

Table 1 Estimated parameters w.r.t the variance

Variance	$\hat{x}_{rss}$ (m)	$\hat{y}_{rss}$ (m)	$\hat{c}_{rss}$ (dBm)	MDE(m)
4	1.2777	-2.2742	27.0011	1.3041
6	0.0917	-2.1638	27.0599	1.4775
8	2.9358	-2.7934	26.9991	2.6381
10	1.8748	-3.1882	27.0037	4.2786

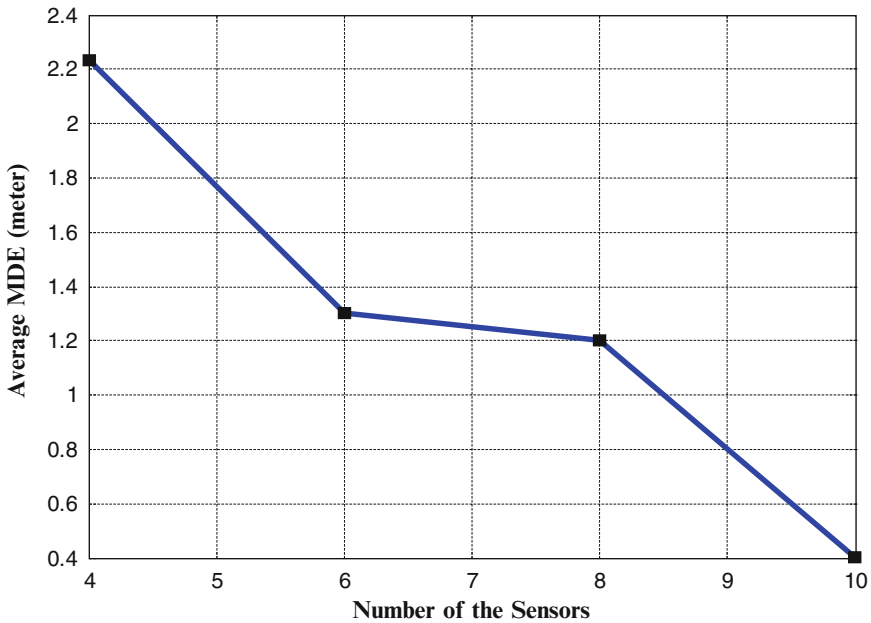

Fig. 4 Average MDE w.r.t the Number of the Sensors

4.3 Scenario 3

The results for the impact of the path loss exponent on the localization problem and the average MDE are given in this part. 6 sensors are located around the smart meter and the AWGN has the $N(0, 4)$ PDF distribution. Fig. 5. illustrates the variation of the average MDE w.r.t the path loss exponent. In the practical applications, the value of the path loss exponent ranges from 2 to 6.

It can be seen that the average MDE is inversely proportional to the path loss exponent. This is because as the path loss exponent increases, the PDF will get sharper so the estimation will be more accurate and probable. Table 3. represents the estimated values of the parameter θ in this scenario.

Table 2 Estimated parameters w.r.t the Number of Sensors

Sensors	$\hat{x}_{rss}$ (m)	$\hat{y}_{rss}$ (m)	$\hat{c}_{rss}$ (dBm)	MDE(m)
4	2.3278	0.7945	26.9971	2.2323
6	1.2777	-2.2742	27.0011	1.3041
8	0.5313	0.1052	27.0048	1.2005
10	1.3774	-0.8614	26.9988	0.4021

Table 3 Estimated parameters w.r.t the path loss exponent

Path Loss Exponent	$\hat{x}_{rss}$ (m)	θ(m)	MDE(m)
2	1.2777	-2.2742	1.3041
3	1.9293	-0.7430	0.9641
4	1.3646	-0.3421	0.7521
5	1.0373	-1.5144	0.5158

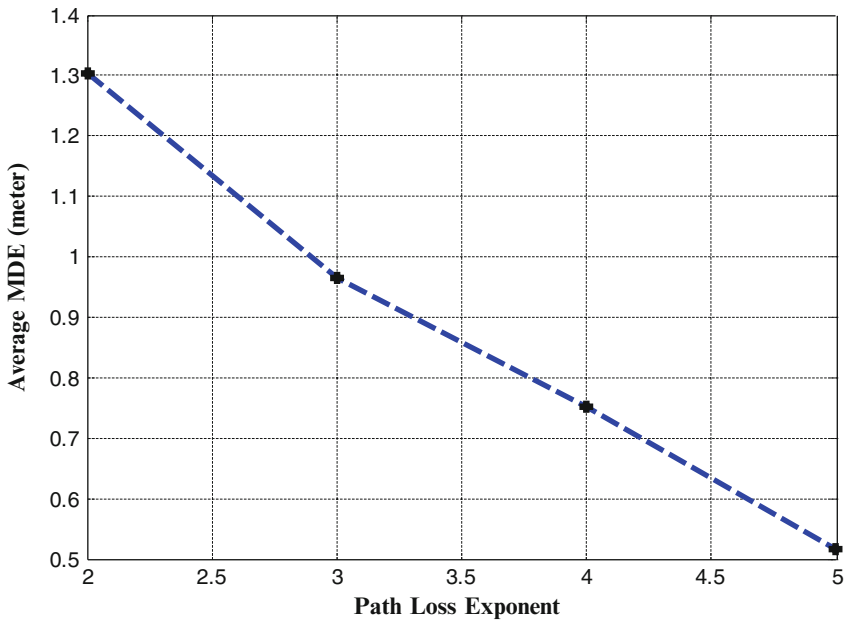

Fig. 5 Average MDE w.r.t the Path Loss Exponent

5 Conclusion

In this paper, the RSS-based localization of a smart meter is investigated as one of the AMI system's security aspects. The architecture of the studied AMI system is introduced and the included components are explained. Under the assumption of a log-normal path loss model and the AWGN shadowing, the ML estimator is implemented to estimate the position and the reference transmission power of the smart meter. Proposed method is verified through MATLAB simulations and the effect of the variance, number of the sensors and the path loss exponent are also inspected. Results show that the average MDE increases as the variance goes up and reduces as the number of the sensors increase. Additionally, it can be deduced that the average MDE is inversely proportional to the path loss exponent.

References

1. I. H. Cavdar. A solution to remote detection of illegal electricity usage via power line communications. IEEE Transactions on Power Delivery, 19(4):1663–1667, Oct. 2004.
2. F. Cleveland. Cyber security issues for advanced metering infrastructure (AMI). In Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, 2008.
3. Aravinthan, Visvakumar, et al. "Wireless AMI application and security for controlled home area networks." *Power and Energy Society General Meeting, 2011 IEEE*.
4. E. I. A, U.S. Energy Information Administration, Independent Statistics & Analysis, Available Online: http://www.eia.gov/.
5. Rouf, Ishtiaq, et al. "Neighborhood watch: Security and privacy analysis of automatic meter reading systems." Proceedings of the 2012 ACM conference on Computer and communications security. ACM, 2012.
6. Ishtiaq Roufa, et al. "A Practical Study of Security and Privacy Issues in Automatic Meter Reading System." *IEEE Spectrum*, October 2010.
7. M. Lisovich and S. Wicker, "Privacy concerns in upcoming residential and commercial demand-response systems," in *2008 Clemson* University Power Systems Conference. Clemson University, 2008.
8. P. McDaniel and S. McLaughlin, "Security and privacy challenges in the smart grid," *IEEE Security and Privacy*, no. 3, pp. 75–77, 2009.
9. Taylor, R. C. (2013). "Received Signal Strength-Based Localization of Non-Collaborative Emitters in the Presence of Correlated Shadowing" (Doctoral dissertation, Virginia Polytechnic Institute and State University).
10. Kay, Steven M. "Fundamentals of Statistical signal processing," Vol 2: Detection theory. Prentice Hall PTR, 1998.
11. Sargolzaei, Arman, Kang K. Yen, and M.N. Abdelghani. "Time-Delay Switch Attack on Load Frequency Control in Smart Grid." Advances in Communication Technology, Vol.5 (2013), 55–64.
12. Sichun Wang, Robert Inkol, and Brad R. Jackson. "Relationship between the maximum likelihood emitter location estimators based on received signal strength (rss) and received signal strength diference (rssd)". In Communications (QBSC), 2012 26th Biennial Symposium on, pages 64–69.
13. J. Kennedy and R.C. Eberhart (1995), Particle swarm optimization. In: Proceedings of the IEEE International Conference on Neural Networks, Perth, Australia, IEEE Service Center, Piscataway, NJ, 4, pp. 1942–1948.