2014 IEEE NINTH INTERNATIONAL CONFERENCE ON INTELLIGENT SENSORS, SENSOR NETWORKS AND INFORMATION PROCESSING (ISSNIP)
SYMPOSIUM ON SENSOR NETWORKS
SINGAPORE, 21–24 APRIL 2014

# Improving The Accuracy of Simulation Models for Localization Schemes

Walid M. Ibrahim
School of Computing,
Queen's University
Kingston, Ontario K7L 3N6
Canada
Email: walid@cs.queensu.ca

Najah Abu Ali
College of IT,
United Arab Emirates University
Ali-Ain
United Arab Emirates
Email:najah@uaue.ac.ae

Abd-Elhamid M. Taha
College of Engineering,
Alfaisal University
PO Box 50927, Riyadh 11533
Saudi Arabia
Email: ataha@alfaisal.edu

Hossam S. Hassanein
School of Computing,
Queen's University
Kingston, Ontario K7L 3N6
Canada
Email: hossam@cs.queensu.ca

***Abstract*—Localization plays a substantial role in the future Internet, especially within the context of the Internet of Things (IoT). Increased dependence on devices and sensed data presses for more efficient and accurate localization schemes. In the IoT environment the area covered is large making it impossible to localize all devices and Sensor Nodes (SNs) using single-hop localization techniques. A solution to this problem is to use a multi-hop localization technique to estimate devices' positions. Simulating localization techniques for wireless sensor networks is required in order to reduce cost and study the difference between localization techniques easily especially if the simulated environment is large. Thus a realistic model is required to simulate the localization process as accurately as possible. Many multi-hop localization techniques use Received Signal Strength Indicator (RSSI) to estimate the distance between SNs. Our interest in this work is to enhance the validation of these schemes prior to deployment. Specifically, we propose the use of a more realistic model for generating RSSI values. The model is based on practical measurements and is validated through extensive simulation.**

## I. Introduction

Wireless Sensor Networks (WSNs) are used in numerous applications such as military target tracking and surveillance, natural disaster relief, biomedical health monitoring, hazardous environment exploration and intrusion detection. Sending the sensed data to the central server is useless if the locations of the events are not associated with the position of the received data. It is important to locate the position of the event in order to take the correct action. A simple solution is to use a GPS for every Sensor Node (SN) to locate their positions, however this will increase the cost of the devices. Moreover, GPS cannot work indoors. Thus an accurate localization technique is required to estimate the location of SNs.

Unlocalized SNs estimate their positions with the aid of anchor nodes, i.e. SNs that have the knowledge of their absolute position. Anchor nodes know their positions by either using a global positioning system (GPS), or by being attached to predefined locations with known coordinates. In order to localize SNs, anchor nodes broadcast their location with the operating instructions to SNs, and SNs use the received instructions to estimate their own locations.

Depending on the application and size of the sensed environment, localization techniques can either be single-hop or multi-hop. In single-hop techniques, the SNs use one of the distance measurements: Received Signal Strength Indicator (RSSI), Time of Arrival (ToA), or Phase of Arrival (PoA) to estimate the distances between themselves and anchor nodes. The SN uses the estimated distance to localize its position. The un-localized SNs require a minimum of 3 anchor nodes in 2-D and 4 anchor nodes in 3-D within their transmission range in order to estimate their locations [1].

In large scale environments it is impossible to localize all SNs using single-hop localization techniques. A solution to this problem is to use a multi-hop localization technique to estimate SN positions. Multi-hop localization techniques are classified into two major categories: range-based and range-free. Range-based relies on the individual inter-sensor distance data. Range-free localization techniques do not depend on any of the distance measurement techniques mentioned earlier; this approach is mainly based on connectivity information to estimate the locations of the un-localized SNs [2], [3].

Evaluating multi-hop localization techniques in large areas is expensive and time consuming, especially if the experiments involve hundreds or thousands of SNs in area that covers hundreds of square meters. This requires the purchase of hundreds of SNs, spend weeks in network deployment and analysis of the results. A solution to this problem is to simulate the localization techniques and SNs. Therefore, a simulation environment that can capture what happens in the real environment is required. In this paper, we aim to propose a more realistic representation for distance measurement error.

Previous works have assumed that the measurement error added to the estimated distance follows a normal distribution and they used such assumptions to simulate SN localization. However in this paper, we show that the simulation can be accurately represented by using Rayleigh distribution instead of using Gaussian distribution. We show through obtaining real measurements that using Rayleigh distribution gives a more realistic representation of the localization error. Moreover, we show, by using multi-hop simulation, the difference between using Gaussian and Rayleigh distribution.


978-1-4799-2843-9/14/$31.00 © 2014 IEEE

The remainder of this paper is organized as follows: the background is covered in section II. Section III presents the error components used to estimated the distance between SNs. The results and our findings are discussed in Section IV. The conclusion is given in Section V.

## II. Background

Estimating the distance between a pair of SNs is the main component for ranging techniques. The most common ranging techniques used in WSN localization are RSSI and ToA. Both techniques introduce noise to the estimated distance [4]. In ToA, the SN estimates the distance by measuring the time the signal takes to travel between the sender and itself, then multiplies the time taken by the speed of the signal by assuming that such speed is constant. While in RSSI, the SN estimates the distance between itself and the sender using the strength of the received radio frequency signal by either using RSSI profiling measurements or estimating the distance via the analytical model by mapping the RSSI to distance using the path-loss propagation model. In this case, the rate in which the signal attenuation over distance is assumed to be previously known and the distance is estimated using the following equation [3]

$$\hat{d}_{ij} = d_0 \left( \frac{P_{ij}}{P_0(d_0)} \right)^{-1/n_p} e^{\frac{\sigma^2}{2\eta^2 n_p^2}}, \qquad (1)$$

where $P_0(d_0)$ is a known reference power value at a reference distance $d_0$ from the transmitter, $P_{ij}$ is the RSSI measurement between a transmitter $i$ and a receiver $j$, $n_p$ is the path loss exponent that indicates the rate at which the RSSI decreases with distance, and $\eta = \frac{10}{ln(10)}$. RSSI is sensitive to channel noise, interference and reflections, all of which have significant impact on signal amplitude. While, ToA relies on the speed of the signal rather than the signal strength, it is relatively immune to most sources of noise including signal attenuation, refraction and reflection. However, the estimated distance is affected when there is no line-of-sight between SNs, the processing time and queuing time of the packets.

Previous works in localization that use RSS and ToA in their theoretical analysis or simulation usually adopt the noisy disk model. They use the noisy disk model to estimate the distance between SNs in order to: 1) evaluate and compare different localization techniques, 2) mathematically derive the maximum likelihood for localization and/or 3) study lower bounds on localization error. The noisy disk model has two components: node connectivity and error. The node connectivity component represents the actual distance between the two SNs, while the noise component represents the noise distribution of the estimated distance and the actual distance.

Whitehouse et al. [4], Savvides et al. [5], Chang et al. [6] and Sheng and Hu [7] study different localization problems and they all use the noisy disk model using the Gaussian noise that defines the estimated distance between the $i^{th}$ and $j^{th}$ SN is represented as follows:

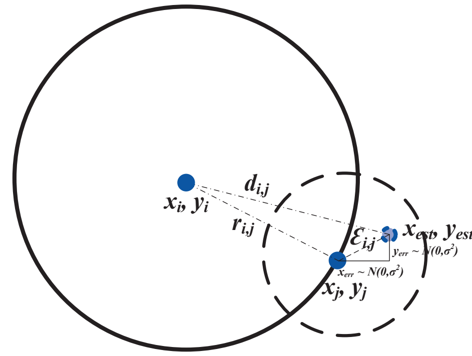



Fig. 1. The estimated distance between $SN_i$ and $SN_j$ is resulted from the displacement in both x and y. The SN can be estimated in any position inside the doted circle.

$$d_{i,j} = d_{j,i} = r_{i,j} + \varepsilon_{i,j} \qquad \forall i,j = 1,2,\ldots,M \qquad (2)$$

where $r_{i,j} = \|\mathbf{x}_i - \mathbf{x}_j\|$ is the noise free distance between node $i$ and $j$, and $\varepsilon_{i,j} \sim \mathcal{N}(0, \sigma_{i,j}^2)$ represents the uncorrelated noise, where $\sigma_{i,j}^2$ is assumed to be accurately estimated and is known a priori [5], [6].

Liu et al. proposed an iterative least square to localize SN using small number of anchors [8]. They proposed an error control mechanism that uses error registry to prevent error from propagating and accumulating during the iteration process. They evaluated their algorithm using MATLAB and they simulated three different noise models. The first experiment, they did not add any noise to the distance. In the second experiment they added Gaussian noise to the distance similar to Eq 2 and fixed the $\sigma$ to 1.7 inches. Finally in the third experiment they used the following equation

$$z = \begin{cases} d + \varepsilon_1 & \text{if } d < d_0, \text{ where } \varepsilon_1 \sim \mathcal{N}(0, \sigma_1) \\ d_0 + \varepsilon_2 & \text{otherwise, where } \varepsilon_2 \sim \mathcal{N}(0, \sigma_2) \end{cases}$$

where $d_0 = 120$ inches and $\sigma_2 = K\sigma_1$ where $K$ is a large number ($10^6$). As they assume that the noise increases rapidly when the distance exceeds a certain threshold.

To take the range into consideration, Chan et al. [9] added a zero-mean white Gaussian process with the variance $\sigma^2 = d_m^2/\kappa$ to propose a new weighted multidimensional scaling for localization scheme, where $\kappa$ is a constant used to make longer distances have a larger measurement error. So and Chan [10], Wei et al. [11] and Qin et al. [12] take the quality of the channel into consideration and replaced constant $\kappa$ with the signal-to-noise ratio (SNR) in the equation of the variance of the zero-mean white Gaussian process with variance. The equation they used is as follows: $\sigma^2 = d_m^2/SNR$, where $SNR$ is the signal-to-noise ratio and $d_m^2$ is the actual distance.

## III. Error Modeling

The previous works, as discussed in the background section, tried to simulate the noise added to the distance as accurately

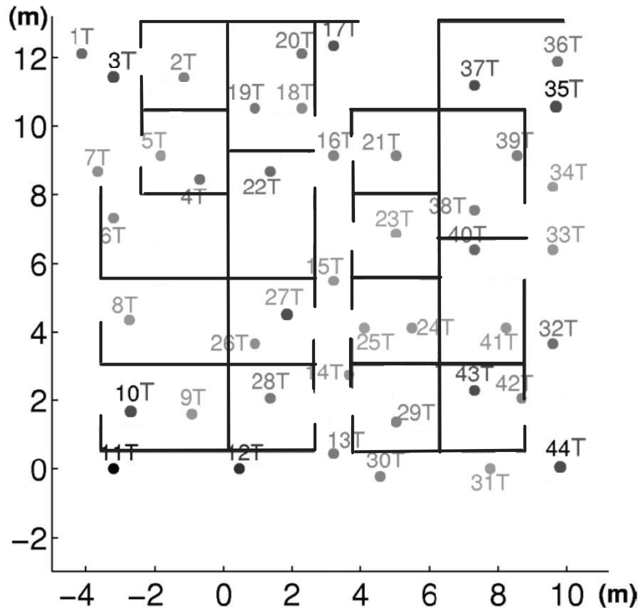
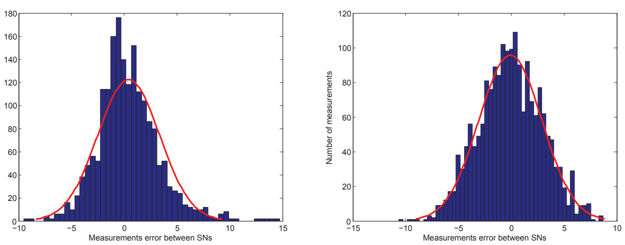



Fig. 2. Map locate the actual locations for SNs (● #T). The RSSI is used to estimated the distance between each SN pair. The distances are estimated by [13].




(a) Actual error measurements follow Normal distribution

(b) Estimated error measurements using Normal distribution

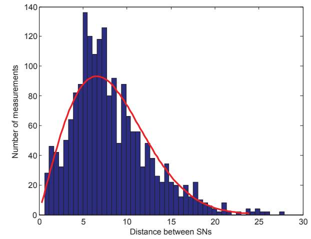
Fig. 3. The error measurement ($\varepsilon_{i,j}$) histogram and its distribution fit.

as possible in order to make their findings close to real experiments. All the previous work added the Gaussian noise to the actual distance similar to Eq. 2 with different variations to the variance. However, the error resulting from the estimated distance is resulting from the Gaussian noise is added to the displacement of SN position in the x and y co-ordinate, not to the absolute distance between the SNs. Fig. 1 shows that the error added to the estimated distance $d_{i,j}$ results from the displacement in both x and y of the SN position. If we assumed that the displacement in x and y follows the Gaussian distribution, therefore the estimated distance can be represented as follows:

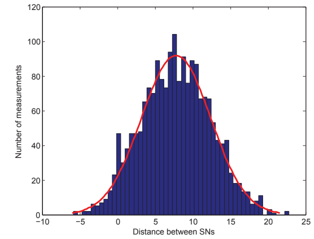
$$d_{i,j} = \sqrt{(x_{est} - x_i)^2 + (y_{est} - y_i)^2}. \tag{3}$$

If we assume that

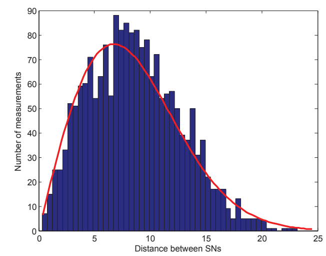
$$x_{est} = x_j + x_{err} \qquad where\ x_{err} \sim \mathcal{N}(0, \sigma^2) \tag{4}$$




(a) Actual distance measurements follow Rayleigh distribution




(b) Estimated distance measurements using Eq. 2 follow Normal distribution




(c) Estimated distances measurements using Eq. 6 follow Rayleigh distribution

Fig. 4. The distance measurement ($d_{i,j} = r_{i,j} + \varepsilon_{i,j}$) histogram and its distribution fit.

and

$$y_{est} = y_j + y_{err} \qquad where\ y_{err} \sim \mathcal{N}(0, \sigma^2). \tag{5}$$

by substituting eq. 4 and 5 in eq. 3, then we will have

$$d_{i,j} = \sqrt{(x_j - x_i + x_{err})^2 + (y_j - y_i + y_{err})^2} \\ where\ x_{err}\ and\ y_{err} \sim \mathcal{N}(0, \sigma^2). \tag{6}$$

From the definition of the Rayleigh, $\gamma \sim Rayleigh(\sigma)$ if $\gamma = \sqrt{X^2 + Y^2}$, where $X\ \ and\ \ Y \sim \mathcal{N}(0, \sigma^2)$ are independent normal random variables, which is the case in equation 6. Therefore $d_{i,j} \sim Rayleigh(\sigma_{i,j})$.

To validate this assumption, we used real data provided by Patwari et al. [13]. The reason we used this data set is they did many enhancements to estimate the distance between SNs and they reached 2-m location error using the RSSI. In their experiment, they used a wideband direct-sequence spread-spectrum (DS-SS) transceiver (Sigtek ST-515). They maintain the SNR $> 25$ db during the experiment to reduce the effect of the noise and ISM-band. They modeled the wideband radio channel impulse response as a sum of attenuated signal, phase-shifted and multi-path [14], [15].

Patwari et al. deployed 44 SNs within a 14 by 13 m area as shown in Fig. 2. The distance between each SN pair is estimated using RSSI measurements to have in total $44 \times 43 = 1892$ measurements. The histogram of the absolute noise (i.e. $\varepsilon_{i,j}$) resulting from estimating the distance between the SNs is plotted as shown in Fig. 3(a). The output of the histogram follows a Gaussian distribution with $\mu = 0.4$ and $\sigma^2 = 8.41$. The data can be replicated easily using the same values as

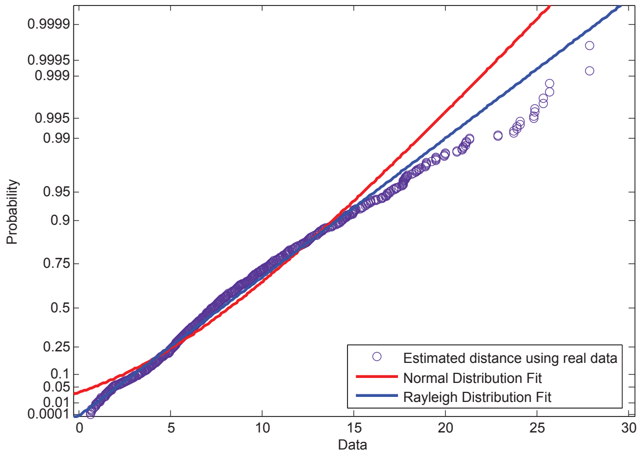
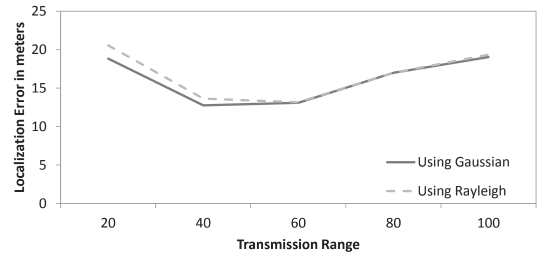
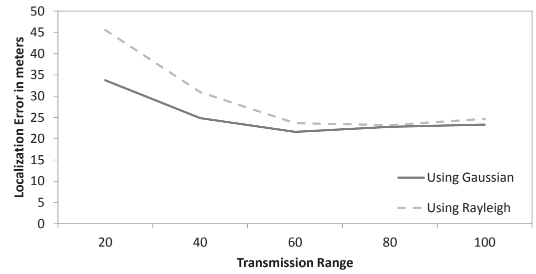



Fig. 5. Goodness of fitness for the actual distance using Gaussian distribution and Rayleigh distribution. It is clear that the actual distance follow the Rayleigh distribution not the Gaussian distribution.




(a) The effect of localization error when $\sigma^2 = 2$




(b) The effect of localization error when $\sigma^2 = 8$

Fig. 6. The relation between transmission range and localization Error. Number of anchors = 4 at the edge of the studied area.

shown in Fig. 3(b). Previous works have shown a similar finding. They used such finding and suggested that the added noise to actual distance follows the Gaussian distribution. Thus they added the generated noise to the absolute distance to represent the estimated distance. However, when the histogram of the estimated distance is plotted (i.e. the actual distance with the noise $d_{i,j} = r_{i,j} + \varepsilon_{i,j}$) using the real data, the result follows the Rayleigh distribution with $\sigma = 6.6$ as shown in Fig. 4(a).

When we replicate the estimated distance by adding Gaussian noise resulted from Fig. 3(b) to the actual distance using the following Eq. 2, the estimated distance follows the normal distribution with $\mu = 7.7$ and $\sigma^2 = 4.7$ as shown in Fig. 4(b); however by using Eq. 6, we get a Rayleigh distribution with $\sigma = 6.72$ as shown in Fig. 4(c). The histogram resulting using Eq. 6 gives a realistic representation of the error, as it gives an almost similar distribution resulting from using the estimated distances using real measurements. This means the added noise is not a pure Gaussian distribution and it is affected by the change in both $x$ and $y$ co-ordinates.

To test the validity of fitting the empirical histogram to the standard Raleigh distribution, we performed the chi square test on the estimated distance provided by Patwari et al. [13] as shown in Fig. 5, which shows that the Rayleigh distribution represents the data more accurately than the Gaussian distribution. Thus the estimated distance between SNs follows the Rayleigh distribution not the Gaussian distribution.

## IV. Simulation and Discussion

We preformed two different experiments using simulation to study the effect of adding Rayleigh distribution using Eq. 6 to the distance error between SNs instead of adding Gaussian using Eq. 2. In the first experiment we study the effect of the transmission range on localization accuracy, while in the second experiment we study the effect of changing the number of anchors with the localization accuracy.

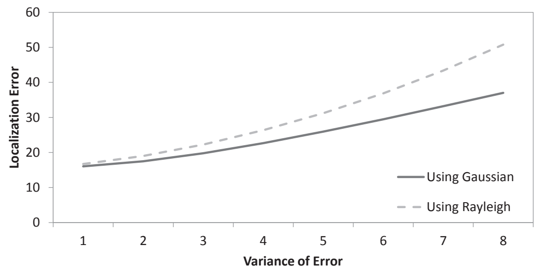
In the simulation, we use NS3 to study the effect of using Normal verses Rayleigh distribution on multi-hop localization scheme that uses DV-Deistance. A number of 500 nodes are randomly placed an area of $200 \times 200\ m^2$. In the first experiment we placed 4 anchor nodes at the edge of the simulated area, while in the second experiment we placed the anchor nodes randomly inside the simulated area. The same $\sigma^2$ is used for both Gaussian and Rayleigh distribution. All the result are the average of 10 runs.

### A. *Change in Transmission Range*

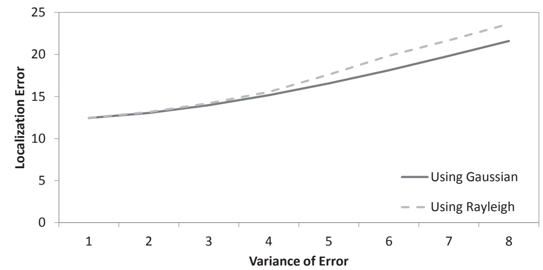
In the first experiment, we study the effect of localization error when we increase the transmission range for SNs. In order to minimize the effect of placing the anchor node on the transmission range, we placed 4 anchor nodes at the corner of the simulated area. The transmission range of the SNs are increased gradually from 20 m to 100 m with a step of 20 m.

Fig. 6 shows the relation between increasing the transmission range and localization error. When the error is small ($\sigma^2 = 2$), the localization error is the same for both Gaussian and Rayleigh distribution as shown in Fig. 6(a). The localization accuracy decreases as we increase the transmission range except when the transmission range is 20 meters. The reason that the localization error is high when transmission range = 20 is the density of the SNs is not that high, which leads the SN to take larger number of hops to reach the anchor node.

However, when the error is large ($\sigma^2 = 8$) and transmission range is small (20 meters), the difference between Gaussian and Rayleigh distribution is at maximum (12 meters). But as we increases the transmission range, the difference between Gaussian and Rayleigh distribution decreases, until both Rayleigh and Gaussian distribution have the same localization

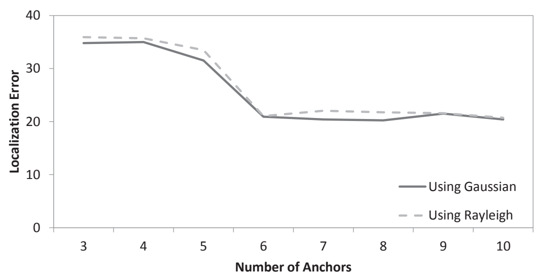



(a) The effect of localization error when $\sigma^2 = 2$

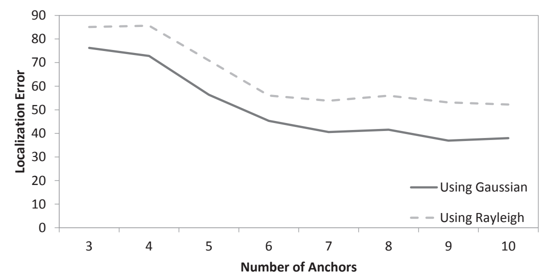



(b) The effect of localization error when $\sigma^2 = 8$

Fig. 7. The relation between $\sigma^2$ and localization Error using 4 anchor nodes located at the edge of the simulated area.

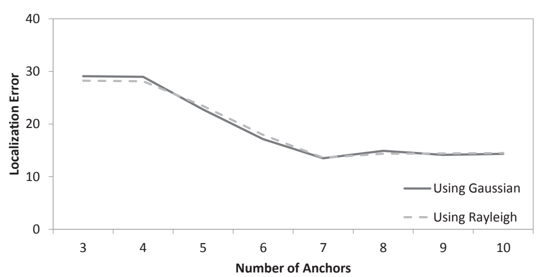



(a) The effect of localization error when $\sigma^2 = 2$

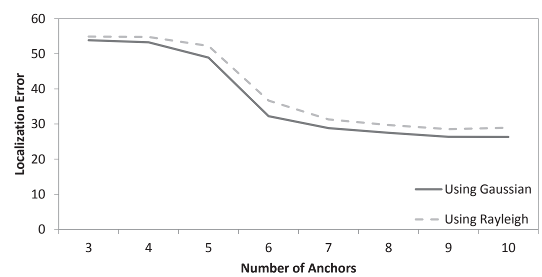



(b) The effect of localization error when $\sigma^2 = 8$

Fig. 8. The relation between number of anchors and localization Error when the transmission of the sensor node = 20 meters.




(a) The effect of localization error when $\sigma^2 = 2$




(b) The effect of localization error when $\sigma^2 = 8$

Fig. 9. The relation between number of anchors and localization Error when the transmission of the sensor node = 40 meters.

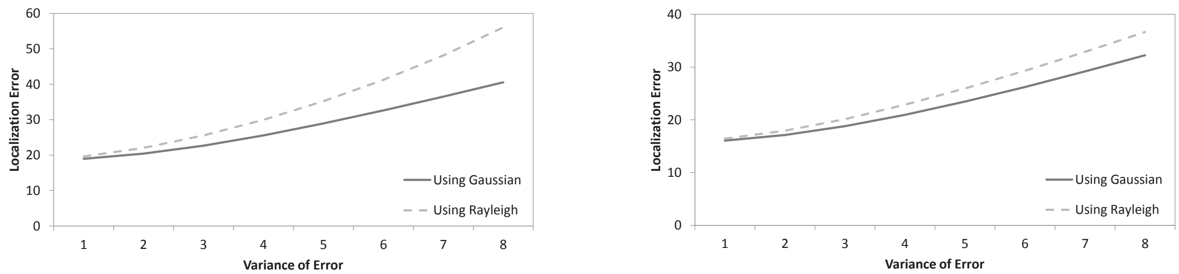
error when the transmission range = 60 meters as shown in Fig. 6(b).

This means when the transmission range is small the difference between using Gaussian and Rayleigh increases as the variance of the error increases. Also, when the transmission range is large both Rayleigh and Gaussian give the same localization error as the variance of the error increases. To validate such finding and check the effect of the variance on the localization error, we fixed the transmission range and increased the value of the variance gradually. When the transmission range = 20 meters the difference between using Gaussian and Rayleigh increases rapidly until the difference reaches 12 meters as shown in Fig. 7(a). However when the transmission range = 60 meters the difference between using Gaussian and Rayleigh increases slowly until the difference is 2 meters as shown in Fig. 7(b). Results in Fig. 7 validate the findings in Fig. 6.

## *B. Change in number of anchors*

In the second experiment, we study the effect of localization error when we increase the number of anchors. The anchor nodes are placed randomly in the simulated area and the number of anchor nodes are increased gradually. In the first experiment, we find that as we increase the transmission range the Gaussian and Rayleigh converge to give the same localization accuracy. So we repeated the experiment using two different transmission ranges 20 and 40 meters respectively.

Fig. 8 and 9 shows the relation between the number of anchor nodes and localization error when the transmission range of the SNs is fixed to 20 and 40 meters respectively. As expected as we increase the number of the anchor nodes the localization error decreases. This behavior is the same when we add Gaussian or Rayleigh distribution to the actual distance between nodes. However, when the transmission range is small (20 meters) and the variance is small ($\sigma^2 = 2$), both Gaussian




(a) The effect of localization error when $\sigma^2 = 2$

(b) The effect of localization error when $\sigma^2 = 8$

Fig. 10. The relation between Variance and localization Error using 7 anchor nodes located randomly in the simulated area.

and Rayleigh distribution give the same localization error as shown in Fig. 8(a). But when the Variance is large ($\sigma^2 = 8$), the difference between the Rayleigh and Gaussian increases to be 12 meters on average as shown in Fig. 8(b). When the transmission range increases, the difference between using Gaussian and Rayleigh decreases as shown in Fig. 9(a) when $\sigma^2 = 2$ and Fig. 9(b) when $\sigma^2 = 8$.

To see the effect of the variance on the localization error, we fixed the number of anchor nodes to 7 and increased the value of $\sigma^2$ gradually. Results in Fig. 10 show similar to Fig. 7. When the transmission range $=$ 20 meters the difference between using Gaussian and Rayleigh increases rapidly until the difference reaches 16 meters as shown in Fig. 10(a). However when the transmission range $= 40$ meters the difference between using Gaussian and Rayleigh increases slowly until the difference is 5 meters as shown in Fig. 10(b).

## V. Conclusion

In this paper, the error model introduced to estimate the distance between SNs using RSSI is investigated in order to enhance the simulation modeling for localization. There has been a belief in the literature that the Gaussian noise is added directly to the distance, which makes the estimated distance to follow the Gaussian distribution for the distance. We assess such belief by showing that the introduced error follows the Gaussian distribution, but the estimated distance follows Rayleigh distribution. This Rayleigh distribution is introduced by adding the introduced error to the x and y coordinated to the SN position while calculating the distance. After that we compared the difference between representing the estimated distance using Gaussian and Rayleigh distributions. Our results show that as we decrease the transmission range of the SNs, the difference between using Gaussian and Rayleigh increases. The same effect also appears when we increase the $\sigma^2$. Thus, it is recommended to add noise to the x and y co-ordinate (Rayleigh distribution) not to the whole distance (Normal distribution) to have an accurate estimation for the distances between SNs especially in high dense environment.

## Acknowledgment

This work was made possible by NPRP grant # NPRP4-553-2-210 from the Qatar National Research Fund (a member of The Qatar Foundation). The statements made herein are solely the responsibility of the authors.

## References

[1] J. Wang, R. Ghosh, and S. Das, “A survey on sensor localization,” *Journal of Control Theory and Applications*, vol. 8, pp. 2 – 11, 2010.

[2] I. Amundson and X. D. Koutsoukos, “A survey on localization for mobile wireless sensor networks,” in *Proc. of the 2nd international conference on Mobile entity localization and tracking in GPS-less environments*, 2009, pp. 235 – 254.

[3] G. Mao, B. Fidan, and B. D. Anderson, “Wireless sensor network localization techniques,” *Computer Networks*, vol. 51, no. 10, 2007.

[4] K. Whitehouse, C. Karlof, A. Woo, F. Jiang, and D. Culler, “The effects of ranging noise on multihop localization: an empirical study,” in *Fourth International Symposium on Information Processing in Sensor Networks.*, 2005, pp. 73–80.

[5] A. Savvides, W. Garber, S. Adlakha, R. Moses, and M. Srivastava, “On the Error Characteristics of Multihop Node Localization in Ad-Hoc Sensor Networks,” in *Information Processing in Sensor Networks*, ser. Lecture Notes in Computer Science, 2003, vol. 2634, pp. 317–332.

[6] C. Chang and A. Sahai, “Estimation bounds for localization,” in *Sensor and Ad Hoc Communications and Networks, 2004. IEEE SECON 2004. 2004 First Annual IEEE Communications Society Conference on*, 2004, pp. 415–424.

[7] X. Sheng and Y.-H. Hu, “Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks,” *Signal Processing, IEEE Transactions on*, vol. 53, no. 1, pp. 44–53, 2005.

[8] J. Liu, Y. Zhang, and F. Zhao, “Robust distributed node localization with error management,” in *Proceedings of the 7th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, ser. MobiHoc ’06, 2006, pp. 250–261.

[9] F. Chan and H. So, “Efficient weighted multidimensional scaling for wireless sensor network localization,” *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4548–4553, 2009.

[10] H. So and F. Chan, “A generalized subspace approach for mobile positioning with time-of-arrival measurements,” *IEEE Transactions on Signal Processing*, vol. 55, no. 10, pp. 5103–5107, 2007.

[11] H. Wei, Q. Wan, Z. Chen, and S. Ye, “A novel weighted multidimensional scaling analysis for time-of-arrival-based mobile location,” *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3018–3022, 2008.

[12] S. Qin, Q. Wan, and Z. X. Chen, “Fast subspace approach for mobile positioning with time-of-arrival measurements,” *IET Communications*, vol. 5, no. 14, pp. 2035–2039, 2011.

[13] N. Patwari, A. Hero, M. Perkins, N. Correal, and R. O’Dea, “Relative location estimation in wireless sensor networks,” *Signal Processing, IEEE Transactions on*, vol. 51, no. 8, pp. 2137–2148, 2003.

[14] H. Hashemi, “The indoor radio propagation channel,” *Proceedings of the IEEE*, vol. 81, no. 7, pp. 943–968, 1993.

[15] T. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Upper Saddle River, NJ, USA: Prentice Hall PTR, 2001.