

# Node Localization in Wireless Networks Through Physical Layer Network Coding

Zhiwei Li

Department of SIS  
UNC Charlotte  
Charlotte, NC 28223  
Email: zli19@uncc.edu

Di Pu

Department of ECE  
Worcester Polytechnic Institute  
Worcester, MA 01609  
Email: dipu@wpi.edu

Weichao Wang

Department of SIS  
UNC Charlotte  
Charlotte, NC 28223  
Email: weichaowang@uncc.edu

Alex Wyglinski

Department of ECE  
Worcester Polytechnic Institute  
Worcester, MA 01609  
Email: alexw@ece.wpi.edu

**Abstract**—Previous research on physical layer network coding (PNC) focuses on the improvements in bandwidth usage efficiency. In this paper, we propose a PNC-based node localization mechanism. When two signal sequences collide at the receiver, the starting point of collision is determined by the distances between the receiver and senders. When the signal interference results from two receivers are combined together, we can determine a hyperbola with two senders as the respective focal points. In this way, by using multiple pairs of anchor nodes as senders, we can determine multiple hyperbolas and the node position will be at the intersection point of these hyperbolas. The proposed approach does not require the wireless nodes to be equipped with any special hardware such as synchronized clocks. We propose several schemes at the physical and network layers to transform the idea into a practical approach. We also investigate the overhead, localization accuracy, and safety of the approach.

## I. INTRODUCTION

With the proliferation of wireless networks and applications, the localization problem attracts a lot of research efforts. Various range-based and range-free localization algorithms have been designed, and the adopted techniques include Angle of Arrival, Received Signal Strength Indicator, Time of Arrival, Time Difference of Arrival, and Hop-based Reconstruction. Many of these approaches depend on some special hardware to estimate the positions of the nodes. The examples include the directional antennas [1], synchronized clocks [2], multiple signal sources [3], power level measurement devices [4], and frequency shift detectors [5]. Although the unit price of the hardware can be very low, the extra cost can still restrict the wide adoption of these methods.

In this paper, we propose a node localization algorithm for wireless networks based on the physical layer network coding (PNC) technique. PNC uses the additive nature of the electromagnetic waves to serve as the coding procedure. In our approach, we determine the position of a wireless node by letting the radio signals from two anchor nodes interfere with each other. The wireless node and another anchor node will capture the interfered signals. We can then determine a hyperbola on which the wireless node resides by comparing the starting points of collisions at the two nodes. When multiple hyperbolas are determined, the wireless node will be positioned at the intersection of these lines.

Although the basic idea is clear, we need to design schemes at both physical and network layers to make the approach

practical. At the physical layer, we need to carefully select data transmission parameters such as modulation. Schemes must be designed to accurately locate the starting points of the signal interference. At the network layer, we need to design a mechanism to reconstruct the data packets from the interfered signals and verify their authenticity. The bootstrap procedure in a network with a small number of anchor nodes will also be investigated.

The proposed approach has several highly desirable properties. First, since the mechanism uses only the starting points of interference to estimate the position of the node, we do not need the wireless nodes to synchronize their clocks. The approach can be conducted in a progressive manner: new hyperbolas can be determined for localization when the position accuracy does not satisfy the requirement of the application. Second, the proposed mechanism does not require any special hardware which will result in a lower node cost. Third, the proposed approach works in a distributed manner and does not depend on a centralized controller.

The remainder of the paper is organized as follows. In Section II we introduce the basic idea of the proposed approach. In Section III, we discuss the schemes in the network and physical layers to make the approach practical. The overhead and localization accuracy are studied in Section IV. Section V investigates the safety of the proposed approach. Finally, Section VI concludes the paper.

## II. THE BASIC IDEA

### A. System Assumptions

We adopt the unit disk graph model and assume that two wireless nodes are neighbors when the distance between them is shorter than  $r$ , where  $r$  is defined as the normal communication range. We assume that the wireless nodes share a secure, lightweight pseudo random bit generator (PRBG) [6]. The nodes need to exchange only the seeds for the PRBG to regenerate the same bit sequences. We assume that the wireless nodes use pairwise keys [7] to protect the data traffic amongst them. We assume there are anchor nodes in the network that already know their positions. How the density and distribution of the anchor nodes impact the proposed approach will be discussed in Section III.

### B. Introduction to PNC

In this part, we introduce the background of the physical layer network coding technique. Figure 1 illustrates the difference between the traditional approach and physical layer network coding. In the topology,  $A$  and  $C$  depend on  $B$  to forward the frames between them. In the PNC approach,  $B$  will receive the interference results of the two frames. It will rebroadcast the received signals to both  $A$  and  $C$ . They will use their knowledge about  $frame1$  and  $frame2$  respectively to separate the signals and recover the data. From this example, we can see that PNC has the potential to achieve higher bandwidth usage efficiency than digital network coding. Data transmission using PNC in more complicated network topologies can be found in [8], [9].

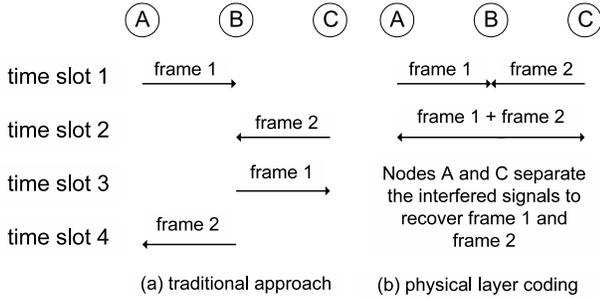


Fig. 1. Traditional approach and physical layer network coding.

### C. Basic Idea of the Approach

In this part, we introduce the basic idea of using PNC to calculate the position of a wireless node. We use  $d_{MN}$  to represent the distance between two nodes  $M$  and  $N$ . We use  $T$  to represent a specific moment and  $t$  to represent a time duration. If the radio wave propagates at the speed  $s$ , the transmission delay between  $M$  and  $N$  will be  $\frac{d_{MN}}{s}$ . In our analysis, we use the difference between the arriving time of two sequences. We must clarify that we are not using the system clocks in the wireless nodes to measure the actual time. On the contrary, we can locate the starting point in the sequence that the collision starts. Then we can translate this information into a time difference using the frequency of the radio signal.

Figure 2 illustrates an example of the signals colliding at the receivers. We assume that four nodes  $A$ ,  $C$ ,  $D$ , and  $E$  can hear each other. We also assume that nodes  $C, D, E$  are anchor nodes and they know their positions. Node  $A$  wants to determine its position based on the signal interference results. The two anchor nodes  $C$  and  $D$  send out long signal sequences that will collide at both  $A$  and  $E$ . Without losing generality, we assume that  $C$  starts sending at  $T_C = 0$  and  $D$  starts sending at  $T_D \geq 0$ .

Therefore,  $A$  will receive the signals from  $C$  at the time  $\frac{d_{AC}}{s}$ , and the signals from  $D$  at  $(T_D + \frac{d_{AD}}{s})$ . The difference between the arriving time of the two sequences is  $t_{diffA} = (T_D + \frac{d_{AD}-d_{AC}}{s})$ , as illustrated in Figure 2. In other words,  $A$  will first receive the sequence from  $C$  for  $t_{diffA}$  seconds, then the two sequences will collide at the node. If  $t_{diffA} < 0$ ,

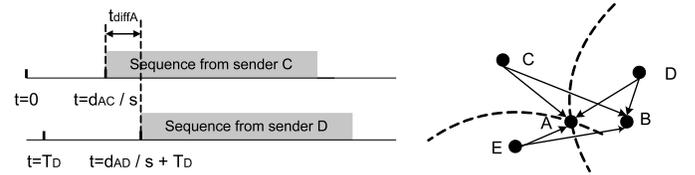
the sequence from  $D$  will arrive first at  $A$ . Similarly, we can derive the difference between the arriving time at node  $E$  as  $t_{diffE} = (T_D + \frac{d_{ED}-d_{EC}}{s})$ . Now let us look at the difference between  $t_{diffA}$  and  $t_{diffE}$ :

$$t_{diffE} - t_{diffA} = (T_D + \frac{d_{ED} - d_{EC}}{s}) - (T_D + \frac{d_{AD} - d_{AC}}{s})$$

We simplify this equation and will get:

$$d_{AD} - d_{AC} = (d_{ED} - d_{EC}) + s \times (t_{diffA} - t_{diffE}) \quad (1)$$

The values on the right side of Equation 1 can be determined as follows. Since nodes  $C, D$ , and  $E$  know their positions, they can figure out  $d_{ED} - d_{EC}$ . The propagation speed  $s$  of radio signals is a constant. Node  $A$  and  $E$  can count the number of symbols between the first sequence arrives and the collision starts. They can then translate the number of symbols to a time duration based on the frequency of the carrier signals. Therefore, we can use these values to calculate  $d_{AD} - d_{AC}$ . Since nodes  $C$  and  $D$  know their positions, node  $A$  will reside on one wing of the hyperbola that is jointly determined by the positions of  $C$  and  $D$  and the value of  $d_{AD} - d_{AC}$ .



Left:  $t_{diffA}$ : difference b/w arriving time of two sequences at  $A$ .

Right: Node  $A$  is at the intersection of the hyperbolas.

Fig. 2. Node localization through PNC.

Obviously, we need more hyperbolas to determine the position of node  $A$ . We can choose other pairs of anchor nodes to send out the signals and determine more hyperbolas. Node  $A$  will be at the intersection point of these hyperbolas, as shown in Figure 2. During this procedure, we need to pay attention to several issues. First, after every round of signal collision, we can calculate the number of intersection points of the hyperbolas to determine whether or not the next round is necessary. Second, when we choose the senders, we must make sure that the new hyperbola cannot be derived out from existing hyperbolas. For example, if we already have the values of  $d_{AD} - d_{AC}$  and  $d_{AD} - d_{AE}$ , choosing  $C$  and  $E$  as the senders will not bring any new knowledge to our localization procedure. One simple method to satisfy this requirement is to preserve one sender and keep changing the other one.

The proposed approach has several highly desirable properties. First, since the mechanism uses only the starting points of collisions to determine the hyperbolas and calculate the position of the node, we do not need the wireless nodes to synchronize their clocks. Second, the proposed mechanism does not require the wireless nodes to be equipped with any special hardware (e.g. directional antennas, very accurate clocks) which will result in a lower node cost. Third, the proposed approach works in a distributed manner and does not require a centralized controller. With these properties, the approach can be easily adopted by existing networks.

### III. PHYSICAL AND NETWORK LAYER ISSUES

In Section II, we introduce the basic idea of using physical layer network coding to calculate the position of a wireless node. However, several issues need to be solved before the idea can be turned into a viable solution. In this section we discuss the issues in the physical and network layers.

#### A. Physical Layer Issues

##### Modulation of Signals

To demonstrate the feasibility of physical layer network coding, existing approaches usually use simple modulation schemes such as MSK [8] and QPSK [9]. In our previous research [10], we use the MSK modulation to map the data bits to radio waves. MSK represents the data bits by varying the phase difference between consecutive complex signals. Specifically, a phase difference of  $\pi/2$  represents bit ‘1’, and a difference of  $-\pi/2$  represents bit ‘0’. We must emphasize that our approach does not depend on any specific modulation mechanisms and it can evolve with the advances in PNC.

##### Detection of Collision

The receiver needs to distinguish three states of the system: no signal, one incoming sequence, and two colliding sequences. To detect the arrival of the first data sequence, the receiver can monitor the incoming energy level since the received signal demonstrates a much higher energy level than that of the noises.

Since our approach does not require the wireless nodes to maintain synchronized clocks, there is a good chance that the two sequences will arrive at the receiver at different time points. Therefore, the receiver must be able to locate the starting point of the collision. Before this point, the receiver runs standard MSK decoding. After this point, the receiver needs to separate the interfered signals. To distinguish the two states, the receiver can measure the variance in the energy level of the incoming signals. Since MSK encodes the bits in the phase, the energy of a non-interfered signal is almost constant. When two signals collide at the receiver, the variance will become much larger [8]. Therefore, we can set up a threshold. When the variance becomes larger than the value, the sequence separation algorithm will be executed.

#### B. Network Layer Issues

##### Decoding the Interfered Sequences

As described in Section II, the wireless node depends on the signal interference results and the positions of the senders to jointly determine the hyperbolas. If the node cannot distinguish the real data sequences from some random signals, the accuracy of the positioning result cannot be guaranteed. Therefore, some mechanisms must be adopted to allow the node to decode the interfered signals and verify their authenticity. In this part, we introduce two mechanisms to accomplish this task.

In the first scenario, the anchor nodes send out pseudo random sequences for the positioning procedure. In Section II.A, we assume all nodes share the same pseudo random bit generator (PRBG). Therefore, each of the anchor nodes can generate a random seed for the PRBG and deliver the seed

to the wireless node using their pairwise key. The wireless node can then regenerate the sequences with the seeds. Using these sequences, the node can reconstruct the interference results and compare them to the received signals to verify their authenticity. Note that when the overlapping component of two sequences is reasonably long (e.g. 1000 bits), the probability that two random signal sequences have the same interference results as the sequences from the anchors is very low.

In the second scenario, we assume that the anchor nodes are sending out data packets and the receivers do not have any prior knowledge about the packet contents. Under this condition, a single receiver cannot separate the two interfered sequences. Fortunately, we have the observation that when the difference between two  $t_{diff}$  is large enough, the two receivers can combine their received signals to recover the two sequences. Below we use an example to show the idea.

Figure 3 shows the two sequences that are sent out by node C and D. Without losing generality, we assume that the collisions at node A and E happen at the fourth and seventh bits of sequence C respectively. If the interference results can be viewed as the sum of the two signals, Figure 3 also shows the received sequences at A and E. If nodes A and E combine their information, they can accomplish the data recovery task. As shown in Figure 3, since E already knows that the fourth bit in sequence C is a ‘1’, it can help A figure out that the first bit of sequence D is ‘0’. This will then help E determine that the seventh bit from C is ‘1’. This procedure will continue and A and E will recover both sequences. To allow the receivers to verify the authenticity of the packets, the senders can encrypt the hash results of the packets with the pairwise keys with the receivers and attach the encryption results to the packets.

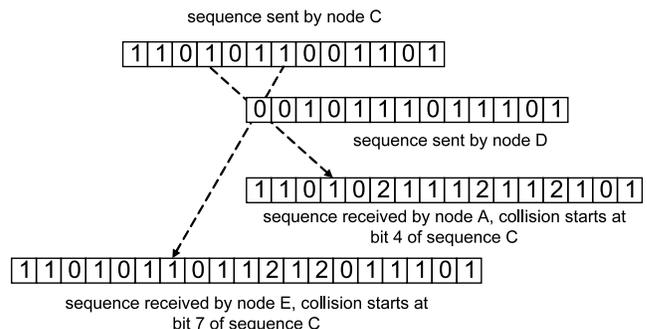


Fig. 3. Data recovery of the colliding sequences at the wireless nodes.

##### Impacts of the Anchor Node Distribution

One factor that may restrict the adoption of the proposed approach is the required number of anchor nodes. Since we need at least three anchor nodes to determine the position of a wireless node, it seems that the proposed approach requires a large percentage of nodes in the network to know their positions. However, below we will describe two scenarios in which the bootstrap condition of the proposed approach can be easily satisfied.

The first scenario is a wireless mesh network. Several special nodes in the network have high speed Internet access and know their positions. Since they are powered by the wall sockets, these special nodes can send out the signals at a much

higher power level than the normal nodes. We assume that the special nodes can hear each other and many normal nodes in the network are covered by multiple special nodes. The normal nodes will connect to the special nodes through multi-hop paths. To apply the proposed approach to this scenario, we can choose different pairs of special nodes to serve as the senders. Other special nodes will share the interference results that they receive with the normal nodes. The normal nodes can combine the information with their own interference results to calculate multiple hyperbolas to determine their positions. As an example, Figure 4.(a) shows four special nodes and the normal nodes covered by them.

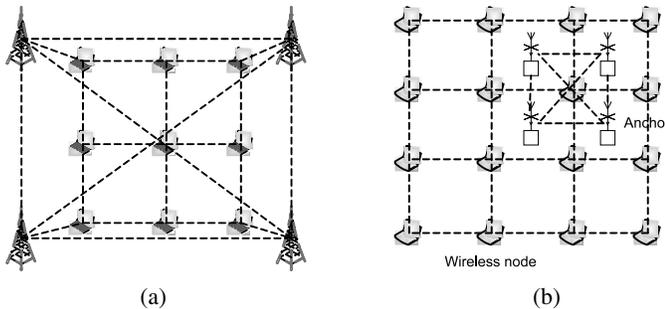


Fig. 4. Two scenarios of the anchor node distribution.

The second scenario is a self-organized ad hoc network. In this environment, every node has the same transmission power. To enable the localization procedure, we propose to deploy a small group of anchor nodes. These anchor nodes will be deployed very close to each other in the central area of the network. In this way, the common neighbors of the anchor nodes will be able to determine their positions based on the proposed approach. These nodes, after getting their positions, will become new anchors and help other nodes determine their positions. The localization procedure propagates as a growing circle until all nodes successfully calculate their positions, as shown in Figure 4.(b).

#### IV. OVERHEAD AND PERFORMANCE ANALYSIS

In this section we study the overhead of the proposed approach and the factors impacting the localization accuracy.

##### A. Overhead of the Proposed Approach

Our analysis will focus on the computation and communication overhead of the approach. The majority of the computation overhead is caused by solving the hyperbolas to determine their intersection points. Since a hyperbola can be represented as a second-degree equation in the Cartesian coordinates, determining the intersections of two hyperbolas can be viewed as a procedure to solve two second-degree equations. In [11] the authors propose a mechanism that uses only simple add and shift operations to calculate the intersection points. Therefore, it can be easily implemented in hardware or firmware. Research [11] has also shown that this method outperforms the traditional schemes in terms of the required number of operations for a specific accuracy level.

The communication overhead of our approach comes from two aspects: (1) exchanging signal interference results with the neighbors for sequence separation; and (2) serving as anchor nodes to help other nodes. The following analysis will focus on the self-organized ad hoc network scenario. We assume that each node needs  $p$  hyperbolas to locate its position. In most conditions,  $p$  has the value of 2 or 3. The signal sequences that are used for localization contain  $l$  bytes. We also assume that on average a node has  $n$  neighbors. Therefore, a node needs to exchange  $p$  signal interference results with the neighbors to locate its own position. The length of each interference result is at most  $2l$  bytes. At the same time, it will help at most  $n$  nodes for their localization procedures by sending out at most  $np$  sequences. Therefore, the total communication overhead at each node is  $2 \times p \times l + n \times p \times l = (n+2) \times p \times l$  bytes. If we assume that the average degree of connectivity is 10 and the packets have the length of 256 bytes, the proposed approach will incur the communication overhead of 9 KBytes, which can be easily handled by a laptop or a PDA.

##### B. Factors Impacting the Localization Accuracy

In this part, we investigate three factors that could have affected the localization accuracy of the proposed approach and introduce the methods to reduce their impacts.

The first factor is the detected starting points of signal interference. As shown in Equation 1, the wireless node depends on the starting points of collisions to calculate  $t_{diff}$ . Considering the high propagation speed of the radio waves, if the detected collision is offset by several symbols, the introduced error can be large. To reduce the impacts of such errors, we can adopt the method described in [8]. Here each packet will start with a pilot bit sequence with known contents. Therefore, even when the detected collision has an offset of several symbols, we can still determine its correct starting point. Note that the pilot sequence has the length of 64 bits and it will not drastically increase the communication overhead.

The second factor we want to study is the frequency jitter. In our previous analysis, the carrier frequencies at the receivers and the senders are assumed to be the same. However, the carrier frequency is actually a time-varying variable. The jitter will impact the proposed approach from two aspects. First, it will cause an increase in the bit error rate (BER), which will harm the sequence separation procedure. Our previous research [10] shows that the BER caused by the frequency jitter can be compensated by introducing some redundancy into the data packets. Second, the frequency jitter will impact the accuracy of distance estimation. Fortunately, when the jitter is small and the carrier frequency is high (e.g.  $10^8 Hz$  or higher), the inaccuracy in the carrier signal wavelength will be very small.

Last but not least, we want to study the error accumulation during localization. In the self-organized ad hoc network scenario, if some errors are introduced into the positions of the nodes close to the anchors, these errors will propagate when they help other nodes determine their positions. To reduce the impacts of such errors, we can deploy multiple groups of

anchor nodes at different places in the network [12]. In this way, the wireless nodes can cross-reference the localization results from multiple sources to increase the accuracy.

## V. DISCUSSION

### A. Why Depend on PNC to Measure Time Difference

As shown in Section II, the proposed approach measures the starting point of interference of two sequences to estimate the distance between the nodes. Here we have to answer one question: why do not we directly use the system clocks to measure the difference between the arriving time of the two sequences? In that way, we can let the two senders send out their packets alternatively and still allow the receivers to estimate their distance.

One reason makes us use the physical layer network coding to measure the time difference. Previous research [13] has shown that wireless nodes have a maximum clock drift rate at microsecond level ( $10^{-6}$  second). At the same time, the deviations of clock drift rates are also at the microsecond level. Let us consider a wireless ad hoc network that has the radio range  $r$  of 250 meters. If the receiver is a neighbor of both senders, the difference between its distances to the senders is at most 250 m, which maps to  $250m \div 300,000 \text{ km/s} \approx 0.83 \times 10^{-6} \text{ sec}$  transmission delay. The measured duration and the clock drift are at the same level. Therefore, directly using the system clocks to measure the time difference will introduce a large error.

### B. Security of the Proposed Approach

In this part we investigate the security of the proposed approach. Specifically, we focus on two scenarios. First, we discuss the impacts on localization accuracy when some nodes share false position information to mislead other nodes. Second, we discuss the mechanisms to deal with jamming attacks.

In both the mesh and ad hoc network scenarios, we assume that the initial anchor nodes will always send out true information. However, in the ad hoc network scenario, after calculating their own positions, some malicious nodes may share false position information with other nodes to mislead them. To mitigate such attacks, we can adopt the mechanisms in [14] to reduce the impacts of localization anomaly. We can apply the clustering algorithms to the intersection points of different hyperbola pairs to detect any out-siders. In this way, we can not only detect the localization anomaly but also prevent the false information from propagating in an epidemic manner.

The malicious nodes can send out noises to disturb the localization procedures. Different from many anti-jamming scenarios, here we cannot directly adopt the frequency hopping technique since the senders and the receivers do not have synchronized clocks and they cannot guarantee that the interfered signals always have the same carrier frequency. However, the senders and the receivers can determine the carrier frequency of the signals through the secure communication channels among them before the localization is conducted. There are such transceivers on the market that allow the wireless nodes to adjust the carrier frequency within the range of 150MHz.

This scheme will not impact the positioning accuracy to a large extent. For example, if we change the wavelength of the signal from 1 meter to 0.9 meter, the carrier frequency will change for about 33.3 MHz. The change at this scale will have a good chance to avoid the jammer signals.

## VI. CONCLUSION

In this paper we propose a node localization mechanism for wireless networks based on physical layer network coding. The analysis shows that the difference between the starting points of interference at the receivers can be used to determine the hyperbolas based on the positions of the anchor nodes. We can then estimate the positions of the nodes by calculating the intersection points of the hyperbolas. To turn this mechanism into a practical approach, we study various problems in the network and physical layers. We also analyze the overhead, localization accuracy, and safety of the proposed approach.

Immediate extensions to our approach consist of the following aspects. First, the authors at WPI are implementing the proposed approach on a software defined radio platform so that we can test it in real network environments. Second, we will explore mechanisms to improve the efficiency of the proposed approach so that it can be applied to mobile networks. Finally, we will investigate using physical layer network coding to accomplish other tasks such as sender authentication in wireless networks.

## REFERENCES

- [1] D. Kumar and S. Varma, "An efficient localization based on directional antenna for wireless sensor networks," *Int. J. Computer & Electrical Engineering*, vol. 1, no. 5, pp. 542–549, 2009.
- [2] W. Qiu and E. Skafidas, "Distributed source localization based on toa measurements in wireless sensor networks," *Research Letters in Electronics*, pp. 1–4, 2009.
- [3] A. Bishop, B. Fidan, K. Doğançay, B. Anderson, and P. Pathirana, "Exploiting geometry for improved hybrid aoa/tdoa-based localization," *Signal Process.*, vol. 88, no. 7, pp. 1775–1791, 2008.
- [4] R. Chen, J. Park, and J. Reed, "Defense against primary user emulation attacks in cognitive radio networks," *IEEE JSAC*, vol. 26, no. 1, pp. 25–37, 2008.
- [5] B. Kusy, A. Ledeczi, and X. Koutsoukos, "Tracking mobile nodes using rf doppler shifts," in *Proc. of ACM SenSys*, 2007, pp. 29–42.
- [6] R. Jenkins, "Isaac," in *International Workshop on Fast Software Encryption*, 1996, pp. 41–49.
- [7] Y. Xiao, V. Rayi, B. Sun, X. Du, F. Hu, and M. Galloway, "A survey of key management schemes in wireless sensor networks," *Comput. Commun.*, vol. 30, pp. 2314–2341, 2007.
- [8] S. Katti, S. Gollakota, and D. Katabi, "Embracing wireless interference: analog network coding," in *ACM SigComm*, 2007, pp. 397–408.
- [9] S. Zhang, S. Liew, and P. Lam, "Physical-layer network coding," in *ACM MobiCom*, 2006, pp. 358–365.
- [10] W. Wang, D. Pu, and A. Wyglinski, "Detecting sybil nodes in wireless networks with physical layer network coding," in *IEEE/IFIP Int. Conf. on Dependable Systems and Networks (DSN)*, 2010, pp. 21–30.
- [11] E. Doukhnitch and M. Salamah, "General approach to simple algorithms for 2-d positioning techniques in cellular networks," *Computer Communications*, vol. 31, no. 10, pp. 2185–2194, 2008.
- [12] S. Yamada, J.-y. Takayama, and S. Ohyama, "Wireless sensor nodes localization based on multiple range data fusion," in *International Conference on Networked Sensing Systems*, 2008.
- [13] H. Song, S. Zhu, and G. Cao, "Attack-resilient time synchronization for wireless sensor networks," in *Proc. of IEEE MASS*, 2005, pp. 765–772.
- [14] K. Paladugu, "Dclad: Distributed cluster based localization anomaly detection in wireless sensor networks using single mobile beacon," Master's thesis, University of Cincinnati, 2007.