Practical Provably Secure Communication for Half-Duplex Radios

Ahmed Elmorsy, Mohamed Yasser
Dept. of Comp. and Sys. Eng.
Alexandria Univ., Egypt
{a.elmorsy, m.yasser.nour}@student.alx.edu.eg

Mohamed Elsabagh
Wireless Intel. Net. Center (WINC)
Nile Univ., Egypt
mohamed.elsabagh@nileu.edu.eg

Moustafa Youssef
Dept. of Comp. Sc. and Eng.
Alex. University & EJUST, Egypt
moustafa.youssef@ejust.edu.eg

***Abstract*—In this paper, we present a practical and provably secure two-way wireless communication scheme in the presence of a passive eavesdropper. The scheme implements a randomized scheduling and power allocation mechanism, where each legitimate node transmits in random time slots and with random transmit power. Such randomization results in ambiguity at the eavesdropper with regard to the origin of each transmitted frame. The scheme is analyzed in a time-varying binary block erasure channel model and secrecy outage probabilities are derived and empirically evaluated. The scheme is implemented over an IEEE 802.15.4-enabled Sun SPOT sensor motes. The results show that the proposed scheme achieves significant secrecy gains with a vanishing outage probability, at the expense of slight decrease in throughput, even when the eavesdropper is equipped with a receive power based classifier and is located too close to the transmitter node.**

I. Introduction

The problem of secure wireless communication has undergone extensive research. Many of the security schemes used today, e.g. public key cryptography, are based on presumed [1] trapdoor one-way functions, where it is *hard* for an attacker to decipher the message without knowing the trapdoor (i.e. the secret key), in the sense that the attacker cannot decrypt the message in polynomial time given the present computational capabilities. Cryptosystems established upon that are typically called *computationally secure*. Such schemes, however, do not prevent a computationally unlimited attacker from decrypting the message without knowing the trapdoor since no one-way function has been proven to be so [1]. Therefore, these schemes are not *provably secure*.

From an information theoretic perspective, Shannon [2] proved that, using a shared secret key K, the achievability of *perfect secrecy* requires that the entropy of K be at least equal to the entropy of the message M (i.e., $H(K) \geq H(M)$). This introduced the possibility of having perfectly secure communication independently of the computational capabilities of the attacker. In [3], Wyner introduced the wiretap channel model and showed that it is possible to send perfectly secure messages at a non-zero rate, *without* relying on secret keys or any limiting assumptions on the computational power of the wiretapper, under the condition that the source-wiretapper channel is a degraded version of the source-destination channel. This was later extended to the non-degraded scenario in [4]. In [5], [6], the effect of fading on the secrecy capacity was studied and it was shown that distributing the message across different fading realizations actually increases the secrecy capacity. Following this, the authors in [7] developed a framework for sharing keys over multi-path fading channels using Automatic Repeat reQuest (ARQ) feedback. The authors showed that a positive key rate can be achieved under the assumption of a public and error-free ARQ feedback channel.

In [8], a randomized feedback approach was proposed which allows the receiver to transmit a random feedback signal over the same noisy channel used by the source. The authors proved that a non-zero secrecy capacity is achievable in both the full-duplex and half-duplex destinations. Our work in [9] extended this approach to the binary encoded half-duplex two-way Gaussian channel by employing randomized scheduling and power allocation in order to increase the ambiguity at the eavesdropper. It was assumed that error detection codes are not used, and thus the eavesdropper has to differentiate between the cases when Alice, Bob, or both are transmitting. We showed that a non-zero secrecy rate can be achieved at the expense of loss in the throughput due to the fact that legitimate nodes may be simultaneously idle (a silence period) or concurrently transmitting (which causes collision at the destinations).

Although information theoretic security schemes provide provable security, they are typically **not practical** due to the many simplifying assumptions they have to make to prove their security. Our work in [10], [11], [12], and based on the theoretical framework in [7], exploits the multi-path nature of the wireless medium to provide **practical** information-theoretic security in channels with feedback. The basic idea is to distribute the bits of a secret key among multiple ARQ frames. By employing this simple approach, we showed that the security levels in Wi-Fi and RFID networks can be significantly enhanced at the expense of slight loss in throughput.

In this paper, we present a **practical and provably secure** scheme for wireless communication. In particular, and based on our theoretical results in [9], we present a practical scheme that combines random transmission, erasure coding, random power allocation, and error detection to achieve practical provably secure communication in wireless networks. The

This work is supported in part by a grant from the Egyptian Science and Technology Development Fund (STDF).

basic idea is for the receiver to **clone** the identity of the transmitter and to randomly transmit in concurrent with the transmitter. This creates an ambiguity at the eavesdropper about the true identity of the transmitter. The legitimate receiver has the advantage of knowing its own transmission schedule, and hence has no ambiguity regarding the source of incoming frames. Power allocation is also used to further confuse an eavesdropper that uses an energy-based classifier to determine the source of the transmission. In order to analyze our system, we model it as a time-varying binary block fading channel where error detection codes are employed per frame. We derive analytical results and evaluate them experimentally. Furthermore, we implement our scheme over an IEEE 802.15.4-enabled sensor boards. Our results show that the proposed scheme achieves **significant** secrecy gains with a vanishing outage probability, without any limiting assumptions on the eavesdropper.

The rest of the paper is organized as follows. Section II describes our system model and scheme. Analytical results for the secrecy outage probability and the loss in throughput are provided in Section III. In Section IV, experimental results validating our theoretic analysis are reported. Finally, Section V offers some concluding remarks.

II. Randomization Scheme

A. System Model

Without loss of generality, we consider a three-terminal setup where two legitimate nodes (Alice and Bob) want to securely communicate a message M in the presence of a passive eavesdropper (Eve), as shown in Figure 1. We assume a time slotted based communication system. Therefore, the channel between Alice and Bob can be modeled as a slotted time-varying binary block erasure channel, where the time slots correspond to coherence intervals, and in which the channel is random from one time slot to another and fixed over each time slot. Alice and Bob are each equipped with a single half-duplex antenna. Each message, M, is subdivided into smaller frames, and each legitimate node is assumed to be able to transmit one frame per time slot. The frames originated from legitimate nodes are assumed to hold an error detection code, such as a CRC. There is no error detection capabilities *at the message level*. An erasure event over the channel between a transmitter and a receiver results in the frame being dropped at the receiver. Erasure events are assumed to be independent and the erasure probabilities are assumed to be independent and identically distributed random variables over time. The block erasure probability of the channel between transmitter X and receiver Y at time slot i is denoted by $\epsilon_{XY}(i) \in [0, 1]$. **All** nodes are assumed to a priori know the joint distribution of the erasure events, but not their instantaneous probabilities. We further assume that all system parameters are known to all parties, including the distributions (but not the instantaneous values) of the allocated transmit power of the legitimate nodes.

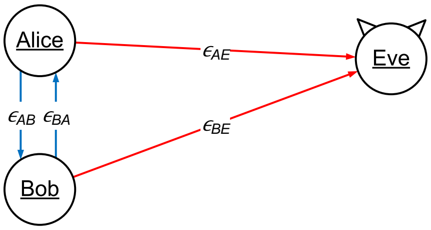

Fig. 1. System Model.

B. Basic Scheme

Without loss of generality, we assume that Alice and Bob are the legitimate transmitter and receiver, respectively. In order to transmit a message M, Alice subdivides it into m frames where each frame is selected for transmission in its corresponding time slot with probability q. In order to confuse Eve, Bob, independently from Alice, chooses to transmit a frame with random data in each time slot with probability q. When transmitting, Bob uses a frame header that corresponds exactly to the frames from Alice.

Hence, there are four possibilities in each time slot: (a) Alice is transmitting alone. This happens with probability $p(1 - q)$. (b) Bob is transmitting alone. This happens with probability $(1-p)q$. (c) Alice and Bob are transmitting concurrently. This happens with probability pq. And, (d) no one is transmitting. This happens with probability $(1 - p)(1 - q)$. Eve can detect cases (c) and (d) clearly through the mismatched frames CRC and the lack of energy in the channel, respectively. For cases (a) and (b), when Eve receives a frame, she cannot determine whether Alice or Bob actually transmitted this frame, since the frames headers from both nodes are exactly the same. Bob, on the other hand, only receives frames from Alice when he is silent. This gives an advantage to Bob over Eve, which is the main property we exploit to achieve our provable security.

However, due to the random jamming of Bob to the frames transmitted by Alice, the set of frames received by Bob from Alice is not known a priori. To compensate for that, M is erasure-coded into m frames such that the reception of any $n < m$ frames at the receiver can be used to reconstitute M with high probability. Note that using erasure coding does not give any advantage to Eve as (a) she cannot determine the identity of the transmitter and (b) there is no message level error detection (only CRC at the frame level). We quantify the secrecy advantage and throughput of the proposed scheme in Section III.

C. Extended Scheme

The basic scheme assumes that Eve has no knowledge other than the frame header to determine the identity of the transmitter. However, Eve in practice can use an energy-based classifier to determine the identity of the transmitter. For example, if Eve is very close to Alice, she can assign frames received with high signal strength to Alice and those with weak signal strength to Bob. More formally, an energy-based classifier detects the source of the frame as:

$$\frac{p_{R|H}(r|A)}{p_{R|H}(r|B)} \underset{\hat{H}=B}{\overset{\hat{H}=A}{\gtrless}} \frac{P_B}{P_A}, \tag{1}$$

where R is a random variable represents the received frame signal strength, P_A and P_B are the prior probabilities of the received frame being from Alice and Bob, respectively, H is the true transmitter, and $\hat{H}$ is the detected transmitter.

To further enhance our scheme, we extend it to employ a randomized power allocation, where the transmit power levels $\mathcal{P}_A$ and $\mathcal{P}_B$ of Alice and Bob, respectively, are independently selected according to predefined distributions. This increases the ambiguity at Eve if she uses the received power to determine the identity of the transmitter.

In the following section, we analyze both the basic scheme and the extended scheme.

III. Security Analysis

We start by a common analysis of the required number of frames followed by the security analysis for both the basic and extended schemes. We end the section by an analysis of the throughput-secrecy tradeoff.

A. Common Analysis

Bob can decode a frame sent by Alice in time slot i only if he was idle while she was transmitting. This happens with probability $p(1-q)$. Therefore, given a number of time slots k, there are roughly $n_A = k \cdot p(1-q)$ time slots consumed by Alice only, $n_B = k \cdot (1-p)q$ time slots consumed by Bob only, $k \cdot p \cdot q$ jammed time slots, and $k \cdot (1-p)(1-q)$ silence periods. Let N be a random variable representing the number of received frames at Bob. Since a frame can be erased at Bob at time slot i with probability $\epsilon_{AB}(i)$, N has a binomial distribution with parameters given by $N \sim Bin(n_A, p(1-q)\mathbb{E}[1-\epsilon_{AB}])$. Therefore, the probability of Bob recovering the erasure-coded message M is given by

$$\Pr\{\text{message reconstruction at Bob}\} = \sum_{i=n}^{n_A} \Pr\{N=i\}. \tag{2}$$

The average number of time slots (k) required for the protocol to successfully deliver n frames from Alice to Bob can be calculated as

$$\mathbb{E}[k] = \frac{n}{p(1-q)\mathbb{E}[1-\epsilon_{AB}]}. \tag{3}$$

B. Basic Scheme (Random Classifier)

When the basic scheme is used, Eve is assumed to depend only on the frame header for determining the identity of the transmitter. From Eve's perspective, and among the frames that she needs to tell apart, a frame has a probability of $P_A = \frac{n_A}{n_A+n_B}$ or $P_B = \frac{n_B}{n_A+n_B}$ to originate from Alice or Bob, respectively. Hence, the probability of Eve to correctly guess that Alice is the originator of n frames is

$$P_{A^{(n)}} = \binom{n_A}{n} \Big/ \binom{n_A+n_B}{n} \tag{4}$$

$$= \left(\frac{n_A}{n_A+n_B}\right)^n \text{ when } n_A, n_B \text{ are sufficiently large.} \tag{5}$$

This results in a secrecy outage probability of

$$\Pr_{out} = P_{A^{(n)}} \cdot \prod_{i \in \mathcal{N}_A^{(n)}} (1-\epsilon_{AE}(i)), \tag{6}$$

where $\mathcal{N}_A^{(n)}$ is the set of the n overheard frames from Alice by Eve.

C. Extended Scheme (Energy-based Classifier)

As discussed before, Eve can be equipped with an energy classifier to identify the origin of each received frame. The decision of the classifier is based on the received power level of each frame in each time slot. Since Eve knows a priori the probabilities P_A and P_B of the frame being from Alice and Bob, respectively, she is assumed to be equipped with a Maximum A Posteriori Probability (MAP) classifier with the following decision rule

$$\frac{p_{R|H}(r(i)|A)}{p_{R|H}(r(i)|B)} \underset{\hat{H}=B}{\overset{\hat{H}=A}{\gtrless}} \frac{P_B}{P_A}, \tag{7}$$

where $r(i)$ is the average received signal power in time slot i. The misclassification probability at Eve is thus given by

$$P_\epsilon = \int_{-\infty}^{\infty} \min\left(p_{R|H}(r|A)P_A,\ p_{R|H}(r|B)P_B\right) dr. \tag{8}$$

Consequently, the probability of secrecy outage is given by

$$\Pr_{out} = (1-P_\epsilon)^n \cdot \prod_{i \in \mathcal{N}_A^{(n)}} (1-\epsilon_{AE}(i)). \tag{9}$$

For the case of Gaussian power distributions, i.e. $\mathcal{P}_A \sim \mathcal{N}(\mu_A, \sigma_A^2)$ and $\mathcal{P}_B \sim \mathcal{N}(\mu_B, \sigma_B^2)$, and by employing a log-distance path loss model where the received power level at Eve from a node X in time slot i is given by

$$r(i)|X = \mathcal{P}_X(i) - (PL_o + 10\gamma \log_{10} d_{XE}), \tag{10}$$

where $\mathcal{P}_X(i)$ is the transmitted power (in dBm) of node X in time slot i, PL_o is the path loss in dB at the reference distance (1m from the transmit antenna of X), γ is the attenuation exponent which will be taken to 2 as in the free space propagation scenario, and d_{XE} is the distance in meters between the transmitter X and Eve; the received power at Eve will follow a Gaussian distribution with $R|A \sim \mathcal{N}\left(\mu_A - (PL_o + 20\log_{10} d_{AE}), \sigma_A^2\right)$ and $R|B \sim \mathcal{N}\left(\mu_B - (PL_o + 20\log_{10} d_{BE}), \sigma_B^2\right)$, for the received power from Alice and Bob, respectively.

Figure 2 shows the secrecy outage probability for different values of n using the basic and extended schemes. The plotted outage probability is obtained assuming that both transmit powers ($\mathcal{P}_A$ and $\mathcal{P}_B$) of Alice and Bob are drawn according

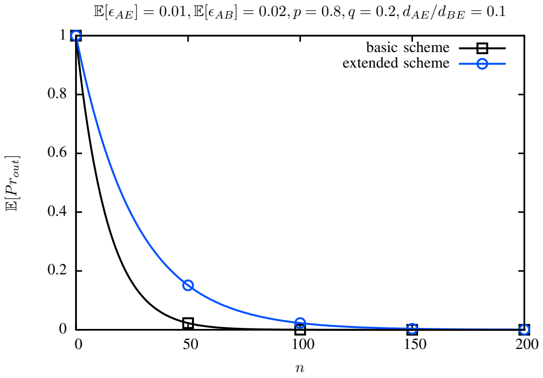

Fig. 2. $\Pr_{out}$ for different values of n when 1) Eve uses a random classifier, and 2) Eve is equipped with a MAP classifier, and she is closer to Alice than to Bob. $\mathcal{P}_A$ and $\mathcal{P}_B$ are drawn according to the same distribution.

to the same distribution, and that Eve is closer to Alice than to Bob, i.e. $d_{AE}/d_{BE} \ll 1$. We argue that this setup is most **favorable for Eve** since the received power levels from Alice and Bob will be more discriminant, which increases her classification accuracy.

The effect of the location of Eve on the secrecy of the proposed scheme is evaluated in Section IV.

D. *Throughput-Secrecy Tradeoff*

Since there are cases when the channel is either idle or jammed, there is a loss in throughput. The throughput of our scheme is governed by the number of messages transferred from one node to the other in n time slots. Assuming that $k = 2n$, and that each message was divided into n frames, a non-randomized and fully synchronized scheme where the k slots are shared between Alice and Bob, would achieve a throughput of 1 message per n time slots. Following the randomization scheme, in n time slots, Alice sends a non-jammed signal with probability $p(1-q)$; similarly Bob sends with probability $(1-p)q$. Therefore, the overall throughput per n time slots degrades to $p(1-q) + q(1-p)$. Figure 3 captures the throughput-secrecy tradeoff for different instances of complementary p and q. It is clear that significant secrecy gains can be achieved with slight decrease in the throughput.

IV. Experimental Results

The proposed schema is implemented on the experimental Sun SPOT motes [13] which use the Squawk virtual machine [14] that runs directly on the microprocessor without an OS. The motes are built upon the IEEE 802.15.4 MAC standard on top of which Zigbee [15] is running, and have a built-in CC2420 chip module [16] along with an integrated antenna operating in the 2.4 GHz band. Our setup consists of three nodes: two free-range SPOTs for Alice and Bob, and one basestation SPOT for Eve. The basestation is connected to a PC where further processing and classification are accomplished by Eve. We had to introduce modifications to the implementation of the standard IEEE 802.15.4 protocol in order to change the source information of the frames sent by Alice or Bob. Also, the CSMA-CA mechanism from the CC2420 driver was disabled so as to allow both Alice and Bob to transmit concurrently. A single experiment run starts by transmitting a synchronization message with the protocol parameters (k, p, q, $\mathcal{P}_\mathcal{A}$, $\mathcal{P}_\mathcal{B}$, and time slot period) from the basestation to Alice and Bob. Next, a packet sniffer is activated at the basestation and the sniffed packets are tunneled to the host PC where a random-guessing classifier and a MAP classifier are running. Bob (or Alice) start jamming the channel upon receiving the Start of Frame Delimiter (SFD) with random content frames.

A. *Basic Scheme (Random Classifier)*

In the random guessing experiment, Alice and Bob were placed 1 foot apart (i.e., $d_{AB} = 1'$), and Eve was placed 10 feet apart from each (i.e., $d_{AE} = d_{BE} = 10'$). No random power allocation was employed, and Eve used random guessing to infer the source of the sniffed frames. p and q were taken to be 0.8 and 0.2, respectively. The channel parameters were first estimated, and the experiment was repeated for different values of $n \in \{10 \,..\, 50\}$, with a time slot period of 300 ms. The results of the experiment is plotted in Figure 4 against the analytical results. The figure shows that the experimental results match the analysis, especially for $n > 28$.

B. *Extended Scheme (Energy-based Classifier)*

Here, Eve was given the advantage of employing a trained energy-based MAP classifier in deciding the origin of a cap-

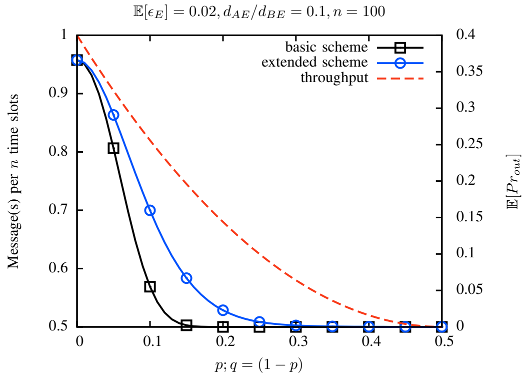

Fig. 3. Throughput-secrecy tradeoff for different instances of p and q. Eve is assumed to be much closer to Alice than to Bob, and n is set to 100.

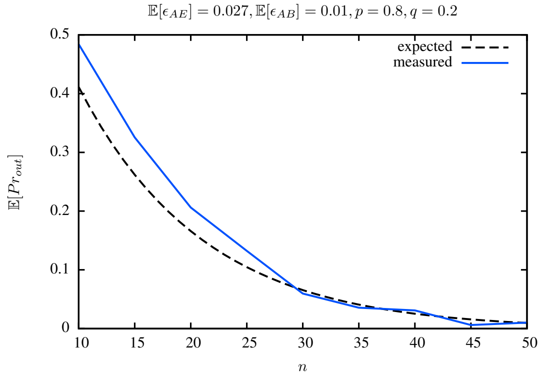

Fig. 4. Measured $\Pr_{out}$ for different values of n where Eve randomly guesses the origin of each frame.

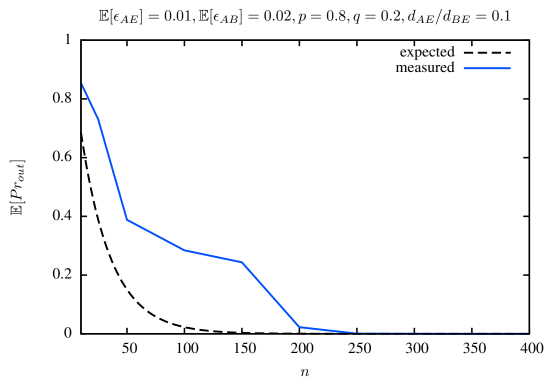

Fig. 5. Measured $\Pr_{out}$ for different values of n. Eve is equipped with a MAP classifier, and she is closer to Alice than to Bob. $\mathcal{P}_A$ and $\mathcal{P}_B$ are drawn according to the same Gaussian distribution of mean -8 dBm and variance 8.3.

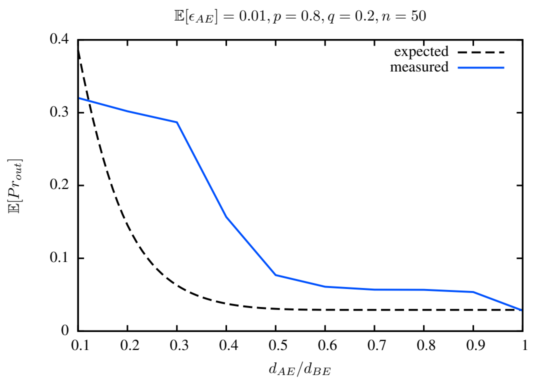

Fig. 6. Measured $\Pr_{out}$ against d_{AE}/d_{BE}. Eve is equipped with a MAP classifier, $\mathcal{P}_A$ and $\mathcal{P}_B$ are drawn according to the same Gaussian distribution of mean -8 dBm and variance 8.3, and n is set to 50.

tured frame. The classifier at Eve is trained on a set of readings taken by running the experiment in the same environment. In the training phase, the classifier was fed with p, q, the transmit power level selected at Alice and Bob, the measured signal strength readings at each time slot, and the original source address of the sender. The distribution of the receive power levels of the signal received from Alice and Bob, respectively, was then estimated. In the online phase, the classifier used the decision rule given by (1) to decide the origin of each received frame. Figure 5 shows the results for the case when Eve is located so close to Alice. In this configuration, $d_{BE} = 10'$ and $d_{AE}/d_{BE} = 0.1$. Alice and Bob send frames with randomized power that follows the same discretized Gaussian distribution of mean -8 dBm and variance 8.3. The experiment is repeated for different values of $n \in \{50 \,..\, 400\}$ with a time slot period of 300 ms. Under the same setup, Figure 6 shows the measured outage probability for different values of d_{AE}/d_{BE}. We fixed d_{BE} at $10'$, and varied d_{AE} to achieve a ratio $d_{AE}/d_{BE} \in [0.1, 1]$.

Noticeably, Eve is mostly bewildered when $d_{AE}/d_{BE} = 1$ since the received power levels of the signals propagating in the free space will approximately match, which maximizes the detection error. Also, it is observed that the measured outage probabilities are slightly higher than the numerical results. We believe that the reason is Eve's enhanced ability to distinguish the sources of the transmitted frames due to the discrete nature of the transmit power levels in the actual motes. In spite of that, the experimental results attest on the practical capability of our scheme to achieve provable secrecy, even when the distance between Eve and the legitimate nodes is so small, and without any restricting assumptions on the computational capabilities of her.

V. Conclusion

This paper developed a novel, and practical, approach to achieve **provably secure** communication using a time-slotted MAC protocol in the presence of an eavesdropper (i.e. Eve). The key idea is to introduce ambiguity at Eve about the source of the overheard frames by randomizing the transmission schedule and transmit power. Using a block erasure channel model, we analyzed the secrecy and performance of our scheme under practical conditions. Our theoretical results were validated under different configurations by doing minimal modifications to Sun SPOT motes. Experimental results show the achievability of a vanishing outage probability, even when Eve is so close the transmitter node and is equipped with a receive power based classifier, and without any limiting assumptions on the computational capabilities of Eve.

References

[1] R. Keon, "RSA Laboratories' Frequently Asked Questions About Today's Cryptography, Version 4.1," *Published online: http://www. rsa. com/rsalabs/faq/February*, 2001.

[2] C. E. Shannon, "Communication theory of secrecy systems," *Bell System Technical Journal*, vol. 28, pp. 656–715, 1949.

[3] A. D. Wyner, "The wire-tap channel," *Bell Systems Technical Journal*, vol. 54, no. 8, pp. 1355–1387, January 1975.

[4] U. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. on Information Theory*, vol. 39, no. 3, pp. 733–742, may 1993.

[5] P. K. Gopala, L. Lai, and H. E. Gamal, "On the secrecy capacity of fading channels," *IEEE Trans. on Information Theory*, 2006.

[6] X. Tang, R. Liu, and P. Spasojevic, "On the achievable secrecy throughput of block fading channels with no channel state information at transmitter," in *CISS*. IEEE, 2007, pp. 917–922.

[7] Y. Omar, M. Abdelatif, M. Youssef, A. Sultan, and H. Elgamal, "Keys through ARQ: Theory and practice," in *IEEE Transactions on Information Forensics & Security*, To appear.

[8] L. Lai, H. E. Gamal, and H. V. Poor, "The wiretap channel with feedback: Encryption over the channel," *IEEE Trans. on Information Theory*, vol. 54, no. 11, pp. 5059–5067, 2008.

[9] A. El Gamal, M. Youssef, and H. El Gamal, "Randomization for security in half-duplex two-way gaussian channels," in *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, nov. 2009, pp. 1–6.

[10] Y. Omar, M. Youssef, and H. El Gamal, "ARQ secrecy: From theory to practice," in *IEEE Information Theory Workshop ITW 2009*, October 2009, pp. 6–10.

[11] M. Elsabagh, Y. Abdallah, M. Youssef, and H. El Gamal, "ARQ security in Wi-Fi and RFID networks," in *Forty-Eighth Annual Allerton Conference*, September 2010.

[12] M. Elsabagh, M. Youssef, and H. El Gamal, "ARQ security in rfid networks," in *IEEE ICC '11 Workshop on Physical Layer Security*, June 2011.

[13] Sun Microsystems, "Sun SPOT theory of operation, red release 5.0," June 2009.

[14] "Squawk virtual machine," http://research.sun.com/projects/squawk.

[15] S. C. Ergen, "ZigBee/IEEE 802.15.4 summary," September 2004.

[16] "Chipcon CC2420 2.4 GHz IEEE 802.15.4/ZigBee-ready rf transceiver," March 2007.