Steganography in 802.15.4 Wireless Communication

Ankur M. Mehta
mehtank@eecs.berkeley.edu

Steven Lanzisera
slanzise@eecs.berkeley.edu

Kristofer S. J. Pister
pister@eecs.berkeley.edu

***Abstract*—A popular physical layer used in wireless sensor networks is the IEEE 802.15.4 standard, which provides for a single coding scheme with constant data rate regardless of channel properties and noise conditions. This paper proposes and simulates a simple steganography method to embed additional information in 802.15.4 data packets when link quality permits, with a modest increase in signal to noise ratio (SNR) required for the same error performance of the underlying 802.15.4 communication. By expanding the code set to include a cluster of 31 ancillary codes for each original 802.15.4 code word, 5 bits can be steganographically overlaid on each 4 bit legacy symbol, allowing this additional data to be transmitted without the knowledge of legacy 802.15.4 receivers over links of greater than 1.95 dB SNR. Increasing the information content by 3.5 dB in this manner can lower the overall energy per bit to noise ratio by 0.1 dB.**

***Index Terms*—802.15.4, Steganography, Wireless Communication**

I. Introduction

Wireless sensor networks (WSN) have received significant attention in recent years. The most common WSN physical layer (PHY) used is a narrowband 2.4 GHz wireless PHY standardized as IEEE 802.15.4. It is intended for low data rate communication while meeting stringent power and cost constraints [1]. The standard defines a coding scheme for error free communication in the presence of significant noise. In typical sensor networks, many links have a signal-to-noise ratio (SNR) sufficiently high such that not all of the error correcting performance is needed. This surplus error correction can be traded away to add extra information to the signal.

This paper will discuss how information is encoded in the existing standard and how additional information can be added without significant performance degradation. In particular, this additional information will be overlaid on existing communication using a steganography method; the data will be transmitted such that an unsuspecting or standards-compliant legacy observer will not even know that additional data is being sent. In contrast to other systems that send steganographic data using invalid legacy data via the link layer [2], this system hides information in the physical layer via the coding scheme itself; decoded data will thus yield no evidence of steganography.

The proposed system is an adaptation of steganography for wireless communication, and section II covers the required background information. Section III gives a functional overview, section IV discusses code selection and presents the final system, and its performance is evaluated in section V. Finally, section VI presents conclusions and discusses potential future work on the subject.

II. Background

A. Wireless Communication

In wireless systems, information is sent between nodes using a radio frequency (RF) link, with some probability of error as a function of interference, channel characteristics, modulation scheme, and signal coding. For the purposes of this work, interference will be neglected and the channel will be assumed to be an additive white Gaussian noise (AWGN) channel. The impact of the signal coding scheme on performance will be evaluated given a particular modulation. Signal coding is used to reduce the energy per bit to noise ratio E_b/N_0 required to achieve some defined error probability by encoding k data bits into code words of n symbol bits, or chips. By increasing the Hamming distance between allowable symbols, higher noise can be tolerated before introducing an error [3].

The IEEE 802.15.4 standard encodes $k = 4$ bits into $2^k = 16$ symbols that are each $n = 32$ chips long. Each symbol has a Hamming distance of at least 12 from any other codes resulting in the ability to tolerate at least 6 chip errors without a symbol error. The standard specifies that a 1% packet error rate is tolerable for 20 byte packets; including the required packet overhead, this translates to a maximum allowable symbol error rate (SER) of $1.9 \cdot 10^{-4}$ [4].

The symbol error performance of an actual 802.15.4 reciever can be simulated across signal-to-noise power ratios SNR = P_s/P_n, as in figure 1. This data simulates adding band-limited white noise to codewords selected uniformly at random, with a coherent QPSK detector demodulating the incoming baseband signal. This demodulated signal is correlated against the master code set with the best matching codes selected by the receiver. Symbol errors are registered when the receiver selected code words do not match the transmitted code words. Given the specification, this demonstrates a minimum acceptable SNR of -2.2 dB.

B. Steganography

Steganography is the science of hiding messages in media such that an unknowing observer will be unaware of even the existence of the hidden information. It has mostly been applied to digital images; changes to the least significant bit of selected pixels cannot be detected by a casual eye, but a confederate knowing what to look for can recover a useful message. Because steganography changes the data in the underlying message, it will be seen as errors or noise by an unsuspecting receiver, and so can only be applied to error-tolerant or redundant communication. The higher the

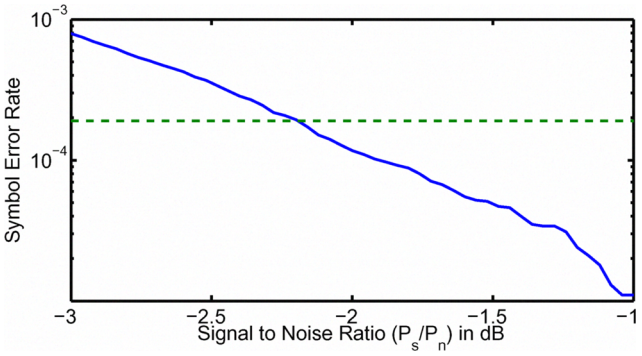

Fig. 1. Error performance of 802.15.4 communication over an AWGN channel. The dashed horizontal line represents the maximum symbol error rate tolerable by the 802.15.4 standard.

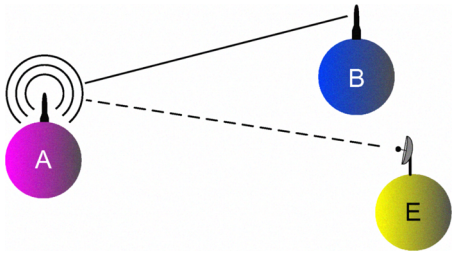

Fig. 2. A sample section of a wireless network. Node A is communicating to node B using 802.15.4 over a high SNR link. Node E can also overhear A's signal, and so A can send steganographic data to E via the communication with B.

error correction capabilities, the more data can be embedded steganographically [5].

III. Functional Overview

Consider a few nodes in an 802.15.4 wireless sensor network, as shown in figure 2. Node A is primarily communicating with node B, although node E is within range and can overhear messages sent by A. A notices that the RF link with B is very reliable; symbols received contain very few chip errors. A can now exploit this excess SNR to steganographically overlay additional information to E.

The primary consideration for such steganography is that B will continue to receive radio traffic as 802.15.4 packets, without noticing the additional communication from A to E. This means that A must send data in 32 chip long symbols that resolve to one of the original 802.15.4 codes by B, the legacy 802.15.4 receiver. This can be accomplished by expanding each 802.15.4 code c_i to a cluster of codes $\{c_{i.0}, c_{i.1}, ..., c_{i.2^k-1}\}$ with $c_{i.0} \equiv c_i$. As long as each code in that cluster is closer to c_i than any other $c_{i'}, i' \neq i$, B will resolve all of those codes to the corresponding 802.15.4 symbol, while E, a specialized receiver, can further resolve within that set to extract an additional k bits of data per symbol.

Legacy receivers resolve incoming symbols as any other 802.15.4 receiver; by computing the cross correlations with the each of the 2^4 original 802.15.4 code words, the incoming symbol can be resolved to that which it is closest to. In order to extract the additional steganographic bits, a specialized receiver has two options. A hierarchical receiver can compute the nearest 802.15.4 code word as a legacy receiver would first, and then further resolve the symbol to an element of the corresponding 2^k code steganographic cluster. Alternately, the symbol can be correlated against the complete 2^{k+4} code list by a full receiver, resolving all $k+4$ bits together.

Each new code $c_{i.j}$ is necessarily closer to some other code $c_{i'}$ than c_i was, and so fewer chip errors are required to result in a symbol error. This raises the required SNR at the receiver P_s/P_n for the communication to meet the 802.15.4 specification. Communications between A and B should be minimally impacted; to keep this SNR penalty low then, each code $c_{i,j}$ should be sufficiently close to the original c_i. However, in order to maintain robustness to noise for the steganographically transmitted data, all of $c_{i.j}$ should be sufficiently far apart, allowing E to resolve amongst the codes in the cluster despite the presence of noise-driven chip errors in its received symbol. These constraints will drive the selection of the expanded code set as described in the next section.

Each 802.15.4 symbol encodes 4 bits of data, and so the steganography scheme encoding k additional bits per symbol will increase the information sent during communication by a factor of $(k+4)/4$. That means increasing the total energy in the communication by a factor of $(k+4)/k$ to meet the higher SNR required by the receiver will hold E_b/N_0 constant for no effective penalty over legacy 802.15.4 communication.

IV. Code Selection

A. Selection Algorithm

The expanded steganographic code set comprises 16 clusters of 2^k codes, so the full set contains 2^{k+4} codes. In order to generate this code set, conditions can be sequentially imposed on all 2^{32} chip sequences. However, the entire code set does not need to be generated this way. 802.15.4 code words contain a symmetry: The first 8 code words are just a circular shift of each other, as are the last 8. The last 8 codes are generated by flipping every other chip of the first 8 codes. Thus, there is a single unique transformation that will generate all 16 code words starting from any of the 16. This implies that all 16 clusters can be too generated by applying that transformation on the codes in a single cluster $\{c_{0.0}, c_{0.1}, ..., c_{0.2^k-1}\}$, selected using the relevant constraints.

The first constraint is that the codes be balanced; each 32 chip word must contain 16 of each 0 and 1 chips. Of these balanced symbols, acceptable steganographic codes must correctly resolve to the original code c_0 by a legacy receiver, and so only those symbols are kept that do so with a negligible error rate through a simulated noisy channel. This symbol set can further be culled by picking from them only the codes which are the farthest from the original code c_0 to enhance steganographic robustness while also being far from the other codes $\{c_i\}$ to preserve legacy performance. The result of these constraints is a set of balanced symbols with a Hamming distance of 6 from c_0 but at least 14 from any other code.

TABLE I
STEGANOGRAPHIC CODE CLUSTER

$c_{0.0} \equiv c_0$	=	d9c3522e	$c_{0.16}$	=	5bd350aa
$c_{0.1}$	=	19e3da2a	$c_{0.17}$	=	5ce3d02e
$c_{0.2}$	=	31f3522e	$c_{0.18}$	=	91d2da2e
$c_{0.3}$	=	45c35a2f	$c_{0.19}$	=	c1b3d22e
$c_{0.4}$	=	4983da6e	$c_{0.20}$	=	c3c3d06e
$c_{0.5}$	=	49d74a2e	$c_{0.21}$	=	d0d3d42e
$c_{0.6}$	=	515b526e	$c_{0.22}$	=	d1c38b2e
$c_{0.7}$	=	51935a3e	$c_{0.23}$	=	d1d3d2a2
$c_{0.8}$	=	51d3512f	$c_{0.24}$	=	d1eb1a2a
$c_{0.9}$	=	51e7c22e	$c_{0.25}$	=	d1f35246
$c_{0.10}$	=	53e3126e	$c_{0.26}$	=	d5c3926a
$c_{0.11}$	=	55c3546e	$c_{0.27}$	=	d913d82e
$c_{0.12}$	=	58f35a26	$c_{0.28}$	=	d9c7c82a
$c_{0.13}$	=	59c3196e	$c_{0.29}$	=	d9d39a24
$c_{0.14}$	=	59d35c2c	$c_{0.30}$	=	d9f2580e
$c_{0.15}$	=	59e34a4e	$c_{0.31}$	=	ddd3d00c

$2^k - 1$ symbols must then be drawn from this set to generate the cluster of steganographic codes.

B. Code List

A cluster of 31 codes can be selected as described above such that each of them is distant from each other, yielding $k = 5$ steganographic bits per 4 bit 802.15.4 symbol. Representing each 32 chip symbol as a 32 bit hexadecimal number, these codes $\{c_{i.0}, c_{i.1}, ..., c_{i.31}\}$ are listed in table I for $i = 0$. The clusters for $i = 1...7$ can be generated by circularly shifting each code by 4 bits, and the clusters for $i = 8...15$ can be generated as in the original 802.15.4 codes by flipping every other chip, that is, XORing the codes with $c_0 \oplus c_8 =$ 0x55555555. Then, to steganographically send a five bit symbol j while sending a four bit symbol i over an 802.15.4 link, send code $c_{i.j}$.

V. PERFORMANCE

Figure 3 shows a plot of symbol error rates seen by various receivers with a transmitter sending a randomly selected steganographic code $c_{i.j}$. We can compare this plot, simulated as described in section II-A, with the corresponding plot in figure 1 to evaluate the relative performance of the steganographic scheme presented above. The red line with square markers is the SER seen by a legacy 802.15.4 receiver resolving the underlying 4 bits against the original code list, showing that the minimum acceptable SNR has increased to 1.95 dB to meet the specified 802.15.4 packet error rate.

A hierarchical steganographic receiver, shown by the green line with diamond markers, will see the same errors as a legacy receiver, plus some additional symbol errors when resolving the 5 steganography bits. The full receiver performs far better, requiring only 1.5 dB minimum SNR to resolve all 9 bits per symbol at the same SER. If using this scheme to send more bits per symbol to a full receiver, however, fewer symbols are needed to send a complete packet, increasing the maximum

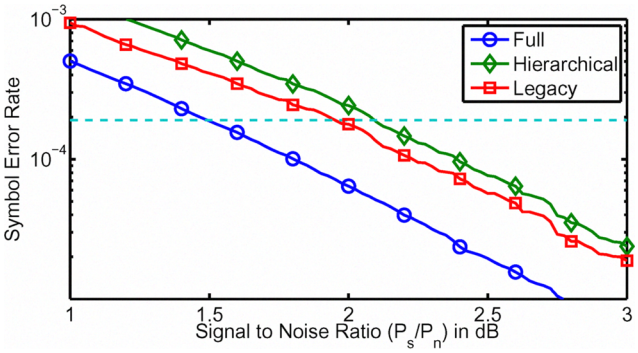

Fig. 3. Noise performance of various receivers for steganographic communication over an AWGN channel. The dashed horizontal line represents the maximum 4 bit symbol error rate tolerable by the 802.15.4 standard.

allowable SER to $3.3 \cdot 10^{-4}$ for the specified packet error rate. This lowers the required SNR to 1.2 dB; the sensitivity requirement is 3.4 dB higher than legacy 802.15.4 while the data rate is 3.5 dB higher. The system presented above thus decreases the required energy per bit to noise ratio E_b/N_0 by 0.1 dB.

VI. CONCLUSIONS AND FUTURE WORK

This work defines and simulates a steganography system that can be used to encode additional data over existing 802.15.4 communication links with greater than 1.95 dB SNR. Links in wireless sensor networks are often at least this quality, and so this system can be used to communicate at greater data rates with 0.1 dB reduction in E_b/N_0. More interesting is the ability to communicate this additional data to a specialized receiver while still speaking 802.15.4 to a legacy receiver.

Steganography in wireless systems is a significantly unexplored research area, and there many avenues for improvement and further research. A more methodical search for the optimal cluster of codes could yield better performance. In addition, it would be interesting to develop a variable size steganographic code set based on available SNR, sending less data over lower quality links but increasing the additional bits per symbol at higher SNR. On the flip side, the system proposed in this paper can be used by a malicious or compromised node in a sensor network. It would be relevant to be able to detect the presence of steganography, perhaps by comparing the received signal strength (RSS) against the chip error rate or link quality.

REFERENCES

[1] "IEEE std 802.15.4-2006: Wireless medium access control (MAC) and physical layer (PHY) specifications for low-rate wireless personal area networks (WPANs)," http://ieeexplore.ieee.org/servlet/opac?punumber=11161.

[2] K. Szczypiorski, "HICCUPS: Hidden communication system for corrupted networks," in *The Tenth International Multi-Conference on Advanced Computer Systems*, Oct. 22-24, 2003, pp. 31–40.

[3] J. Proakis, *Digital Communication*, 3rd ed. McGraw Hill, 2001.

[4] S. Lanzisera and K. S. J. Pister, "Theoretical and practical limits to sensitivity in IEEE 802.15.4 receivers," in *IEEE International Conference on Electronics, Circuits and Systems*, Dec. 11-14, 2007.

[5] T. Jamil, "Steganography: the art of hiding information in plain sight," *IEEE Potentials*, vol. 18, no. 1, pp. 10–12, Feb./Mar. 1999.