11-2005

An Asymptotic Secrecy Model and its Applications in Sensor Networks

Bo Yuan
Rochester Institute of Technology

An Asymptotic Secrecy Model and Its Applications in Sensor Networks

Bo Yuan
Department of Information Technology
Thomas Golisano College of Computing and Information Sciences
Rochester Institute of Technology
Rochester, New York 14623
E-mail: bo.yuan@rit.edu

***Abstract*— This paper proposes a new model for secure communication channels between two parties. The new model assumes that adversaries are storage space bounded, but not computationally bounded. At the initial phase of the secret communication, both parties exchange a large amount of random bits so that adversaries are not able to save them due to the storage space limitation. Each party only saves received data. At the second phase, each party regenerates the random bits, combines them with received data, and generates an encryption key iteratively with a one-way hash function. The key is then used to encrypt the future transmissions from one party to the other. After each transmission, the key is updated iteratively based on data received. Applications of the new model are also discussed.**

I. Introduction

The basic problem in cryptography is secure communication over an insecure channel. Party *A* wants to send to Party *B* a secret message over a communication line which may be tapped by an adversary. The traditional solution can be illustrated in Figure 1. It requires two channels to realize secure communication between Party *A* and Party *B*. It is called the secret key approach. Through one channel with guaranteed secrecy, Party *A* and Party *B* exchange an agree upon an encryption method, *E*, and its associated secret key, *k*, and decryption method, *D*. On the other insecure channel, Party *A* sends a cipher text $c = E(m, k)$ to Party *B*. When Party *B* receives the cipher text c, it performs $D(c, k) = D(E(m, k), k) = m$ to decrypt the cipher text. The adversary *C* should not be able to decrypt the cipher text without knowing the encryption method *E* and the secret key *k*.

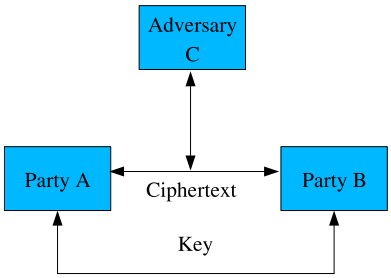

Fig. 1. Traditional Solution

In 1949, Shannon proved that to reach absolute security, the length of the key, *k*, needs to be at least as long as the message, *m*, itself [1]. It is assumed that the adversary has unlimited computational resources in Shannon's theory.

One problem with the traditional solution is the key distribution. It requires a guaranteed secure channel to exchange a common key to both Party *A* and Party *B*. This is very difficult to realize in many real world applications, especially for today's online applications of e-commerce.

In 1976, Diffie and Hellman solved this key distribution problem in their seminal paper [2] and started so-called modern cryptography. Diffie and Hellman introduced a public key distribution system to eliminate the need of a secure key distribution channel. It is based on assumptions that the adversary is computationally bounded. That is, it is computationally infeasible for the adversary to decrypt cipher text. More specifically, the public key encryption is based on assumptions that some one way functions are "easy" to compute but "hard" to invert. Examples include one way functions which factoring a very large integer, the discrete logarithm, RSA functions, etc. For detailed discussion of these functions refer to [3]. Many applications have been developed for public key cryptography. In fact, public key cryptography has enabled private data transactions on the Internet; online shopping, online banking, and e-commence have become reality.

However, the public key cryptography also has its shortcomings. First, public key cryptography is susceptible to the man in the middle attack [4]. Thus a public key infrastructure has to be established to authenticate public keys themselves. Secondly, the computationally infeasibility of some problems may be temporary. As a matter of fact, in [5], Shor has shown that factoring large integers and discrete logarithms can be performed in polynomial time on a quantum computer. This leaves a possibility that a passive eavesdropper can record all secret communications between two parties, and later, with more advanced computational algorithms and hardware, the adversary could decipher the messages.

To overcome the temporary nature of the computational infeasibility assumed in public key based cryptography, Aumann, Ding and Rabin introduced an bounded storage model in [6]. In the bounded storage model, it is assumed that an

adversary is computationally unbounded, but is bounded by the amount of storage available to store the output of computation. The authors also proved information-theoretic security in this model. However, the storage bounded model does require a shared secret key by both parties, which may be the reason why the storage bounded model is not as successful as the public key model in real world applications.

In this paper, the asymptotic secrecy model is introduced to address both issues of the secret key sharing and computational infeasibility. Applications of the model are also demonstrated in sensor networks, where sensor nodes are usually storage bounded.

II. The Asymptotic Secrecy Model

In this model, the following assumptions are made.

1) The channel between Party *A* and *B* is noiseless and a passive adversary can capture all data exchanged between *A* and *B*.
2) The adversary is storage space bounded. That is, the storage space of the adversary is less than the total storage spaces available to Party *A* and *B*.

Many information based security theories assume a noise channel between two parties, which implies that an adversary is not able to obtain the exact the same information as the parties involved. With this assumption, researchers are able to prove that the information-theoretic security is able to be achieved with privacy amplification. See [7], [8] and [9]. In many real world applications, especially today's data networks, noiseless channels are required. It may be the reason why it is hard to find real world applications for information-theoretic based security theories. In the Asymptotic Secrecy model, however, the second assumption limits an adversary's capability of knowing all communications between *A* and *B*. Hence, the Asymptotic secrecy model can be regarded as a special case of information theoretic cryptography.

The protocol of the asymptotic secrecy model can be described in three phases.

A. Phase I: Initialization

Party *A* and Party *B* exchange unencrypted random texts. Both parties save only the received random texts. They exchange enough random texts so as to consume all of their storage spaces. Since the adversary's storage space is less than the total storage spaces of both parties, the adversary is not able to keep all random text exchanged between *A* and *B*.

Suppose $R_1, R_2, R_3, ..., R_p$ are p random bit streams exchanged between Party *A* and Party *B*. $l_i = |R_i|$ denotes the length of the bit stream i for $i = 1, ..., p$. Suppose again R_i is transmitted from *A* to *B*, when i is odd; from *B* to *A* when i is even. Thus the total number of bits transmitted between *A* and *B* is

$$Bits_{total} = \sum_{i=1}^{p} l_i. \tag{1}$$

Total bits Party *A* received is

$$Bits_A = \sum_{i=1, i|2}^{p} l_i, \tag{2}$$

and total bits received by Party *B* is

$$Bits_B = \sum_{i=1, i\nmid 2}^{p} l_i. \tag{3}$$

It is easy to see that $Bits_{total} = Bits_A + Bits_B$. Note that when l_i is 0, it means that no bits are exchanged between Party *A* and Party *B*.

The assumption that the adversary is storage space bounded, then, can be formulated as

$$Bits_{total} > M, \tag{4}$$

where M is the size in bits of the maximum storage space that the adversary possesses.

B. Phase II: Generating a shared secret

Both parties regenerate random bit streams, combine them with saved bit streams, and use the combination as the shared secret for both *A* and *B* to generate a key with a current timestamp. Since the adversary is not able to save all the random text exchanged in Phase I, it does not have the shared secret. The timestamp is public. The current timestamp can be synchronized with a message exchange. The timestamp is used to prevent the adversary from obtaining pre-calculated keys in order to save storage spaces. A reasonably good pseudo random number generator should be used with given seeds so that *A* and *B* can regenerate identical random bits. It should not be reversible, i.e., it should not be possible to reconstruct the pseudo random numbers, seeds, or algorithm from some amount of random bits.

With this the shared knowledge, an encryption key can be generated in the following manner.

Suppose $f : \{0,1\}^n \longrightarrow \{0,1\}^m$ with $m < n$ is a one-way hash function which is known to both parties and the adversary. Then a secret key k can be generated by the following equations.

$$k = f_p(R_p) \tag{5}$$

where

$$f_i(R_i) = f(R_i \parallel f_{i-1}(R_{i-1})) \tag{6}$$

for $i = 1, ..., p$; R_0 should be the current timestamp; and f_0 is the identity function. The symbol $\parallel$ stands for the concatenation of the two bit streams.

C. Phase III: Updating keys

Suppose $m_1, m_2, ..., m_t$ are t messages needed to be exchanged between Party *A* and *B*. Then, the cipher text should be generated by

$$c_i = E(m_i, f_i(m_{i-1})) \tag{7}$$

for $i = 1, ..., t$ and $m_0 = k$, where k is the key generated in Eq. 5. Then, decryption is performed by the following equation.

$$\begin{aligned} m_i &= D(c_i, f_i(m_{i-1})) \\ &= D(E(m_i, f_i(m_{i-1})), f_i(m_{i-1})) \end{aligned} \quad (8)$$

for $i = 1, ..., t$.

III. Analysis of the Protocol

The basic assumption of the protocol is that the adversary is storage space bounded. Its limit is less than the total storage spaces of Party *A* and Party *B*. When the adversary collects all data exchanged between *A* and *B* from the very beginning, the adversary does not have enough storage space to save all captured data, hence it does not have shared information between *A* and *B*. For *A* and *B*, they simply need to store all data received from the other party and they are able to regenerate their own data transmitted to the other party. If the adversary does not collect data from the very beginning, it does not know the shared data. This protocol does not exclude the old fashion security practice, i.e., exchanging a shared security through a guaranteed channel, which, in fact, provides more confidence in the privacy of the channel between Party *A* and Party *B*.

One concern is that both Party *A* and *B* do not know the storage space limit of the adversary, thus, Party *A* and *B* are not certain of the security of their channel. There are many ways to mitigate this risk. First, both parties should maximize the usage of the opposite party's storage space by transmitting random bit streams as much as possible. Secondly, Party *A* and *B* can exchange some apparently public meaningless information to discourage the adversary from saving the exchanged data. Third, party *A* and *B* should use the private channel only when it is necessary. Thus, data exchanged on the open channel can be used as common knowledge to generate secret keys, and the adversary may not save data on the open channel.

On the other hand, the same concern can be raised for computationally bounded assumptions. Adversaries are assumed to be resource limited individuals and not big organizations or governments.

The function f in the protocol is a one-way hash function. It can be any hash function with reasonable strength. From Eq. 5, we can see that the initial secret key is derived by applying the hash function f iteratively in same way that an iterative hash function is applied. Note that the function f may be an iterative hash function itself. According to [10], an iterative hash function is at least as secure as its underlying compression function, i.e., no iterative hash function. Thus, the initial key generated by Eq. 5 should be reasonably strong.

The worse case scenario is when the size of the storage space of the adversary is just one bit less than the total storage space of Party *A* and *B*. In this case, the adversary has all common knowledge between Party *A* and Party *B* except one bit in R_p. This requires that the hash function f should be sensitive to at least one bit differences in inputs. That is, two inputs with only one bit difference should yield very different hash values. On the other hand, the adversary can guess the missing bit, since it is just one bit. In general, the security of the protocol relies on how many bits the adversary is unable to save. The more bits the adversary misses, the smaller the probability it will The following proposition demonstrate this probability.

***Proposition** 1:* Suppose the maximum total storage space of the two parties *A* and *B* is N bits and the maximum storage space of the adversary is $M < N$ bits. Then the probability that the adversary obtains the initial key k is

$$P(k) = 2^{(M-N)}. \quad (9)$$

Proof: The adversary cannot store $N-M$ bits transmitted between *A* and *B*. Thus $2^{(M-N)}$ is the probability that the adversary guesses the missing bits correctly. ■

On the other hand, the protocol can be applied iteratively. That is, at the second round of the protocol, instead of transmitting plain random bit streams $R_1, R_2, R_3, ..., R_p$, both parties transmit encrypted random bit streams by Eq. 7.

***Proposition** 2 (**Asymptotic Security**):* Suppose the maximum total storage space of the two parties *A* and *B* is N bits and the maximum storage space of the adversary is $M < N$ bits. Then the probability that the adversary obtains the initial key k is

$$P(k) = 2^{K(M-N)}. \quad (10)$$

where K is the number of times the protocol is applied.

Proposition 2 indicates that even if *A* and *B* have just a few extra bits than the adversary, they can still achieve asymptotic secrecy for the communication channel by applying the protocol iteratively.

From Eq. 7, we can see that encryption keys are updated whenever new data is received. Encryption keys are never reused, thus minimizing the risk of key discovery attacks. This leaves the adversary the only choice of brute-force attacks. Since keys are hash values, dictionary attacks do not apply. With a reasonable length of hash values, such as 256 bits coupled with a strong encryption algorithm E, it will be very difficult for the adversary to crack keys. A symmetric encryption, E can also be selected as a strong algorithm without compromise the performance.

IV. Applications in Sensor Networks

Security challenges for wireless sensor networks fall into the following areas. First, sensor nodes in typical wireless sensor networks are usually resource bounded. They have small computational power, low memory capacity and low transmission bandwidth. These limitations prevent certain popular security models, such as PKI and encryption algorithms, from being employed in sensor networks. Secondly, sensor networks are often deployed in hostile environments. Sensor nodes are subject to capture and tampering. Wireless transmissions among sensor nodes can be easily overheard by adversaries. Hence, resilience against node capture and eavesdropping are requirements for wireless sensor networks.

Several security schemes have been proposed for wireless sensor networks. Perrig et al proposed SPINS framework in

[11]. In SPNS, It is assumed that each node shares a secret key with a base station in the SPINS framework. Two sensors cannot establish a secret key directly. They need to use the base station as a trusted third party to establish a shared secret key. One advantage of this scheme is that each node only needs to store one shared key with its base station. The disadvantage, however, is that additional key exchange messages need to be transmitted between base stations and nodes. This scheme does not work in those sensor networks that do not have base stations.

Eschenauer and Gligor [12] originated the random key predistribution scheme based on random graph theory. Before they are deployed, sensor nodes are preloaded with some number of random keys from a large key pool. After deployment, a common random key that two neighboring nodes process, if it exists, is used as a shared secret key to encrypt communications between the two nodes. A shared random key is not guaranteed, but it is proven that as the number of total nodes in the network n increases, each node needs to preload $(n-1)(ln(n)+c)/n$ random keys from the key pool in order to ensure the network connectivity with a probability $P_c = e^{-e^{-c}}$, where c is a constant. The advantage of this scheme is that it does not require a key exchange protocol. There are no additional messages exchanged. A disadvantage of this scheme is that each node needs to store $(n-1)(ln(n)+c)/n$ number of random keys. Another disadvantage is that there is no mechanism to update the keys.

In LEAP, proposed by Zhu et al [13], each node is preloaded with a common master key. Once deployed, each node immediately identifies its neighbors and generates shared keys for all of its one-hop neighbors; then, the master key is erased from memory. One advantage of LEAP is its resilience to node capture. If one node is captured, only its neighboring nodes are potentially vulnerable to attack. One disadvantage of the LEAP, however, is that it does not have a key update mechanism. It is usually a not good practice to use a fixed encryption key for a long period of time. First, a single encryption will provide a large amount of cipher text for adversaries attempt to crack. Secondly, if the encryption key is compromised, all previous transmitted data with the same key is also compromised. In other words, the LEAP encryption scheme does not provide so-called "perfect forward secrecy."

Using the asymptotic secrecy model introduced previously, we can solve the following two problems in sensor networks.

1) updating secret keys to achieve "perfect forward secrecy"
2) establishing secure communication channels for non-neighboring nodes

V. Automatic Key Updating Scheme in Sensor Networks

To apply the asymptotic secrecy model to sensor networks, modifications are required. First, the scheme adopts the same initial master key distribution method as in LEAP. That is, before deployment all nodes are loaded with a shared secret master key. Once deployed, each node identifies its neighboring nodes and exchange some random text using the asymptotic secrecy model with the master key as the initial encryption key. As soon as this is done, each node stores a shared secret with each of its neighboring nodes, and the master key is erased from memory. In fact, a time can be set so that if a node cannot find a neighboring node with certain time, the master key is erased automatically. Thus, after a pre-set time, there is no node possessing the master key. This property will mitigate the node capture problem of the sensor networks.

The asymptotic secrecy model based key updating method can be illustrated in Figure 2. The basic idea of the automatic key updating is to XOR a hash value of previously transmitted secret messages with the current key to generate the next encryption key. More formally,

$$k_i = k_{i-1} \otimes h_i(m_{i-1}) \quad (11)$$

$$h_i(m_{i-1}) = h(m_{i-1} \| h_{i-1}(m_{i-2})) \quad (12)$$

where m_{i-1} is the secret message transmitted under key k_{i-1}, for $i \in \{1, 2, 3, ...\}$; and h is a one-way hash function; k_0 is a shared master key; m_0 is random data generated during the key agreement phase; $\|$ denotes string concatenation; and m_{-1} and h_0 are defined as null string and function, respectively. When Part *A* transmits a message to Part *B*, the first message from *A* to *B* is the cipher text of a random text, m_0, encrypted with the shared master key, k_0; the second message is the cipher text of m_1, encrypted with a new key that is the XOR of the previous key with the hash value of m_0; the third message is the cipher text of m_2 encrypted with another new key that is XOR of the previous key with the hash value of the concatenated string of m_1 and the hash value of m_0; and so on.

At the receiver side, the decryption process is symmetric. Key updating is identical. Both sides share the same one-way hash function and the secret key.

Note that in Figure 2, a stream cipher is used. The key auto-updating scheme is not specific to any cipher type. However, a stream cipher is preferred in sensor networks as reported by Luo et al in [14].

An encryption algorithm paired with the proposed key auto-updating scheme has the following characteristics:

- Each message is encrypted with a different key. Keys are constantly updated with respect to messages.
- Encryption and decryption keys are self-synchronized, assuming the receiver received all cipher text.
- It has the "perfect forward secrecy" property. That is, if one key is compromised, messages encrypted under other keys are still secure.

This encryption scheme has also some weaknesses:

- Some encryption algorithms need an initialization time before encrypting message. The more key updates, the more initialization time is spent, thus prolonging the overall transmitting time.
- It is common in real world applications that some messages may never reach the intended receiver. If the

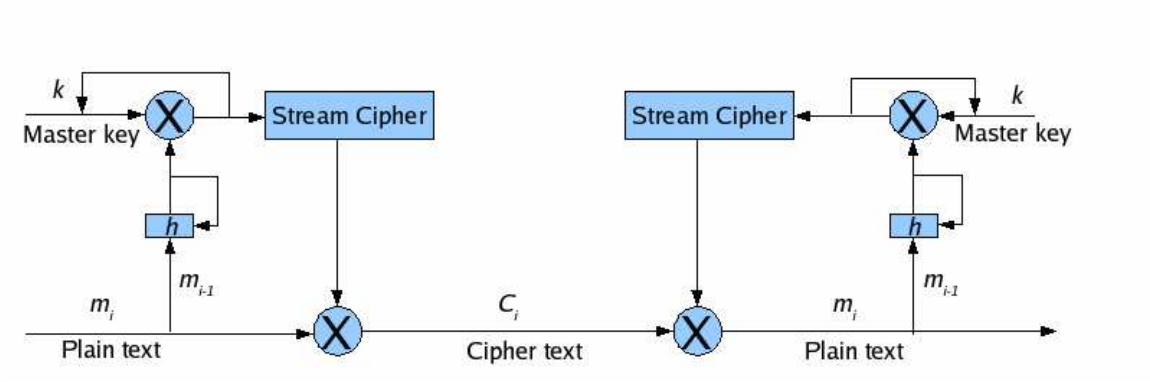

Fig. 2. An Encryption Scheme with Self-updating Keys

receiver misses some messages, the encryption and decryption keys are not synchronized.

To address the weakness of the key auto-updating scheme, first we can configure how often a key is updated according to the encryption algorithm employed. If the encryption algorithm needs a longer time to initialize a new key, less frequent updating should be specified. However, the hash value, which is used to generate the key, still needs frequent updates. More formally, we can use the following equations.

$$k_i = \begin{cases} k_{i-1} \otimes h_i(m_{i-1}), & \text{for } i \equiv 0 \mod T, \\ k_{i-1}, & \text{otherwise.} \end{cases} \quad (13)$$

where T is an integer that specifies how often to update the key. When $T = 1$, the encryption key is updated after every message; when $T = 2$, it is updated every other message, etc. This can be set according to the encryption algorithm so that best overall performance is reached.

To address the problem of message loss, the so-called efficient ACK method can be implemented. After transmitting T messages, the sender waits for an acknowledge message from the receiver. The acknowledge message can be very compact. It can be as little as T bits with bit 1 indicating a received message and bit 0 indicating a missed message. When the sender receives the ACK, it uses only those messages acknowledged to update the encryption key. Note that this ACK message is also encrypted with the current encryption key.

VI. Shared Secret Establishment for non Neighboring Nodes

A. multi-path approach

In [15], Chan, Perrig and Song introduced a multi-path key reinforcement scheme to mitigate the compromised node problem. First, two nodes,Party *A* and Party *B*, identity a common key, k, from the predistributed key pool. This method assumes that all disjoint paths between the two nodes are known. Part *A* generates n random text messages, $v_1, v_2, \cdots, v_n$ and sends them via n different paths to Party *B*. Then a new key can be generated by $k' = k \otimes v_1 \otimes v_2 \otimes \cdots \otimes v_n$. The secrecy of the new key is protected by all n random text messages, since the adversary has difficulty eavesdropping on all n paths. As pointed by the authors, the more paths used, the more security the new key provides for the channel between Parties *A* and *B*.

The multi-path approach by Chan, Perrig and Song is used as a reinforcement to the secrecy of a link between Parties *A* and *B*. It is assumed that Party *A* and Party *B* already share a key from the predistribution key pool. In the Asymptotic Secrecy model, however, there is no shared key between Parties *A* and *B*. The initial shared master key should have been erased from memory a short time after the deployment. Also note that it may not be easy to find multiple disjoint paths between the two nodes.

B. single path approach

Suppose a path between non-neighboring nodes *A* and *B* is known. The protocol of the asymptotic secrecy can be used to establish a shared secret between *A* and *B*. At the initial stage of the protocol, both *A* and *B* transmit random text to each other and exhaust storage spaces of both nodes. Since all nodes should have more or less the same storage space in the same sensor network, the total amount of random text is about twice the storage space of any compromised nodes on the path. Thus, compromised nodes are not able to store all random text. Moreover, hop to hop transmissions on the path are all encrypted with shared secrets of neighboring nodes. Nodes that are not on the path should not be able to understand the traffic on the path. When A and B send encrypted data, even nodes on the path between A and B do not understand the traffic passing between them.

VII. Conclusions

The proposed asymptotic secrecy model is based on the assumption that an adversary is storage bounded. It is founded on the idea of privacy amplification in information-theoretic security. The protocol of the model is introduced and applied to solve two problems in sensor networks. One is the "perfect forward secrecy" problem; the other is the problem of the shared secret establishment between non neighboring nodes.

REFERENCES

[1] C. E. Shannon, "Communication theory of secrecy systems," *Bell System Technical Journal*, vol. 28, pp. 657–715, 1949.

[2] W. Diffie and M. E. Hellman, "New directions in cryptography," *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 644–654, Nov. 1976.

[3] S. Goldwasser and M. Bellare, *Lecture Notes on Cryptography*. http://www.cs.ucsd.edu/users/mihir/papers/gb.html, 1996.

[4] S. M. Bellovin and M. Merritt, "An attack on the interlock protocol when used for authentication," *IEEE Transactions on Information Theory*, vol. 41, pp. 273–276, January 1994.

[5] P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer," *SIAM J. Computing*, vol. 26, pp. 1484–1509, 1997.

[6] Y. Aumann, Y. Z. Ding, and M. O. Rabin, "Everlasting security in the bounded storage model," *IEEE Transactions on Information Theory*, vol. 48, pp. 1668–1680, June 2002.

[7] U. Maurer and S. Wolf, "Secret-key agreement over unauthenticated public channels part i: Definitions and a completeness results," *IEEE Transactions on Information Theory*, vol. 49, pp. 822–831, April 2003.

[8] ——, "Secret-key agreement over unauthenticated public channels part ii: The simulatability condition," *IEEE Transactions on Information Theory*, vol. 49, pp. 832–838, April 2003.

[9] ——, "Secret-key agreement over unauthenticated public channels part iii: Privacy amplification," *IEEE Transactions on Information Theory*, vol. 49, pp. 839–851, April 2003.

[10] B. Schneider, *Applied Cryptography*. New York: Wiley, 1996.

[11] A. Perrig, R. Szewczyk, J. D. Tygar, V. Wen, and D. E. Culler, "SPINS: Security protocols for sensor networks," *Wireless Networks*, vol. 8, pp. 521–534, 2002.

[12] L. Eschenauer and V. D. Gligor, "A key-management scheme for distributed sensor networks," in *Proceedings of the 9th ACM Conference on Computer and Communications Security*, Washington, DC, November 2002, pp. 41–47.

[13] S. Zhu, S. Setia, and S. Jajodia, "LEAP: Efficient security mechanisms for large-scale distributed sensor networks," in *CCS'03*, October 2003, pp. 62–72.

[14] X. Luo, K. Zheng, Y. Pan, and Z. Wu, "Encryption algorithms comparisons for wireless networked sensors," in *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, 2004, pp. 1142–1146.

[15] H. Chan, A. Perrig, and D. Song, "Random key predistribution schemes for sensor networks," in *IEEE Symposium on Security and Privacy*, May 2003, pp. 197–213.