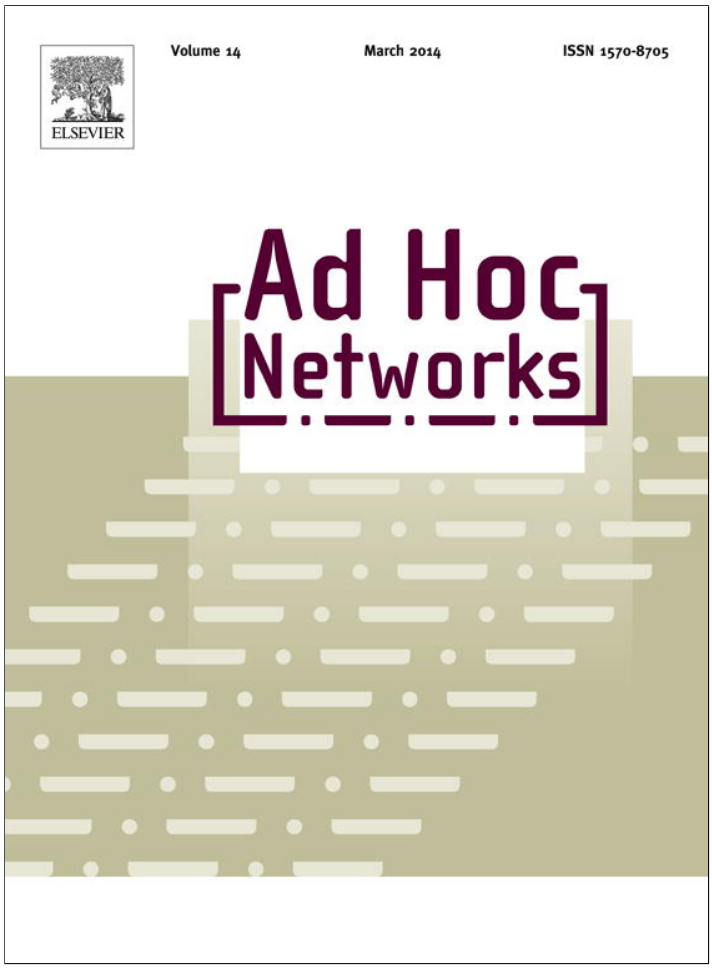


Ad Hoc Networks 14 (2014) 95–105

Contents lists available at ScienceDirect

Ad Hoc Networks

journal homepage: www.elsevier.com/locate/adhoc

Information theory and cryptography based secured communication scheme for cooperative MIMO communication in wireless sensor networks

Liang Hong [a,*], Wei Chen [b]

[a] Department of Electrical and Computer Engineering, Tennessee State University, Nashville, TN 37209, USA
[b] Department of Computer Science, Tennessee State University, Nashville, TN 37209, USA

ARTICLE INFO

Article history:
Received 16 November 2012
Received in revised form 5 September 2013
Accepted 12 November 2013
Available online 22 November 2013

Keywords:
Sensor networks
Physical layer security
Compromised nodes detection
Cryptography
Information theory

ABSTRACT

Emerging cooperative MIMO communication is a promising technology in improving communication performance for wireless sensor networks. However, the security problems inherent to cooperative communications also arise. In this paper, we propose a cross-layer secured communication scheme for cooperative MIMO communication in wireless sensor networks to overcome the external and active compromised nodes attacks. The scheme combines cryptographic technique implemented in higher layers with data assurance analysis at the physical layer to provide better communication security. An efficient key management system is proposed for the cryptographic processes. It provides secured communication and routing using a small number of keys shared between the clusters which cooperate on data transmission and reception. Although cryptography can ensure the confidentiality in the communications between authorized participants, it usually cannot prevent the attacks from compromised nodes. The situation where the cooperative nodes are compromised and try to corrupt the communications by sending garbled signals is also investigated in this paper. A novel information theory based detector that can identify the active compromised nodes and recover the symbols in transmission process at physical layer is proposed. When the compromised nodes are detected, the key management system calls the key revocation to isolate these nodes and reconfigure the cooperative MIMO network. Simulation results show that the proposed algorithm for compromised nodes detection is effective and efficient, and the accuracy of received information is significantly improved.

1. Introduction

Due to recent advances in electronics, wireless communications and computing technologies, wireless sensor networks (WSNs) have been widely deployed in many applications, including military sensing and tracking, environment monitoring, smart home appliances management, and health-care [1]. WSNs are expected to be the basic building block of pervasive computing environments.

In WSNs, the sensor nodes are remotely deployed in harsh environments, where reliable communications links are usually not available and each sensor node must depend on its energy-limited battery for its operation. By exploiting spatial diversity with multiple antennas at the transmitter and receiver, the Multiple-Input Multiple-Output (MIMO) technique, which can provide significant increases in data rate and link range without additional bandwidth or transmission power, has attracted much attention in literature [2]. However, the physical

* Corresponding author. Tel.: +1 6159635364.
E-mail addresses: lhong@tnstate.edu (L. Hong), wchen@tnstate.edu (W. Chen).

1570-8705/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.adhoc.2013.11.008

implementation of multiple antennas at a sensor node may not be feasible. With limited physical size, a sensor node typically can only support a single antenna [3–5].

Recently, cooperative MIMO has been an emerging technique to achieve the benefits of the MIMO technique without the need of multiple antennas at each sensor node [6]. In cooperative MIMO WSNs, multiple single-antenna sensor nodes are physically grouped together to cooperatively transmit and/or receive. Due to the smaller distance, the sensor nodes within the same group can communicate with relatively lower power as compared to inter-group communication. Taking into account that the total energy consumption includes transmission energy, circuit energy, and signal processing energy consumption, it has been proved that by using diversity gain earned from MIMO technology, the cooperative MIMO based sensor networks may lead to better total energy optimization and smaller end-to-end delay than single input single output system [7,8].

Many WSNs have mission-critical tasks; however, the involvement of multiple nodes for transmission and/or receiving poses a challenge to the reliability of the information. Unfortunately, most schemes for traditional cooperative MIMO WSNs do not include considerations for potential security problems in communications at the design stage and are known publicly [9]. Therefore, attackers can easily launch attacks by exploiting security holes in those schemes. In general, the attacks in WSNs can be classified as external attacks and internal attacks. Cryptography can prevent some of the external attacks where the attacking nodes are not authorized participants of the sensor networks. However, in an adversarial environment, the nodes for cooperative MIMO communications could be compromised, leading to internal attacks. Node compromise is one of the most detrimental attacks to WSNs. In general, cryptography based approaches cannot prevent the attacks from compromised nodes because they can encrypt and decrypt the information. Therefore, compromised nodes can eliminate all the efforts to prevent attacks [10]. According to the operation mode, the attacks of compromised nodes can be uncooperative or active [9]. The uncooperative compromised nodes do not relay the information at all. The active compromised nodes will maliciously modify the relay information and inject falsified information. If there are active compromised nodes and the receiver treats them as trusted nodes, it will easily lead to symbol detection errors. Therefore, the impact of active attacks is more threatening than uncooperative attacks from the compromised nodes. Active attacks from just a few compromised nodes around the event would make the entire network fail due to the garbled information collected from these compromised nodes. Ref. [11] elaborates the impact of the attacks from active compromised nodes and shows how easily the garbling can lead to a failed data transmission through an example. The simulations in Section 6 will also illustrate this impact and show that the conventional system without compromised nodes detection will fail due to the high bit error rate.

A variety of techniques has been proposed to secure WSNs' communication. In [12,13], the threats and vulnerabilities to WSNs, security requirements and secured communication solutions are summarized. Cryptography based approaches, either using public key cryptography or using symmetric key cryptography, are widely used for communication security in WSNs [9,14,15]. For the WSNs of small size sensor nodes, symmetric key cryptography is more time and energy efficient. On the other hand, physical-layer secured communication techniques are more promising, since they can be more effective in resolving the boundary, efficiency, and link reliability issues [16]. Li and Hwu [16] and Kim and Villasenor [17] exploited signal randomization which, when combined with channel diversity, effectively randomizes the eavesdropper's signals but not the authorized receiver's signals. In [18] the security of communications is enhanced by adding artificial noise to the transmission process in the physical layer with extra MIMO antennas. Their scheme assumes a key management system in a higher layer and the artificial noise is generated by the keys shared with neighboring nodes. However, none of these schemes detects and defends against node compromise. Moreover, all these schemes need extra MIMO antennas to achieve data assurance, which largely reduces the advantage of MIMO technique. On the other hand, Mao and Wu proposed a cross-layer scheme that uses pseudo-random tracing symbols at the physical layer and direct sequence spread spectrum symbols at the application layer for tracing and identifying the compromised nodes [11]. However, the insertion of tracing symbols will increase the overhead of the transmission and reduce the data rate. The complexity of the system and the power consumption will also be increased for tracing symbol transmission and extraction.

In this paper, we proposed a cross-layer secured communication scheme for cooperative MIMO communication in wireless sensor networks to overcome the external and active compromised nodes attacks. The scheme combines a cryptographic technique implemented in higher layers and data assurance analysis at the physical layer to provide better communication security. An efficient key management system is proposed for the cryptographic processes ensuring data confidentiality, message authentication, etc. It provides secured communication and routing using a small number of keys shared between the clusters which cooperate on data transmission and reception. The situation where some of the cooperative nodes are compromised and try to corrupt the communications by sending garbled signals is also investigated. A novel information theory based detection approach is proposed to identify the compromised nodes and recover the symbols in transmission process at physic layer. It can identify all the compromised nodes in the latest configured cooperative transmission groups, if the compromised nodes are less than half in the cooperative cluster and the number of the nodes in this cluster is not larger than that of its all neighboring clusters. If the number of the nodes in the cluster under detection is larger than that of its all neighboring clusters, the proposed detection approach could detect all the compromised nodes if the number of the compromised nodes is less than one third of the number of the nodes in the largest neighboring cluster. After the compromised nodes are detected, the key management

system can call the key revocation to isolate these nodes and reconfigure the cooperative MIMO network.

Comparing with existing schemes [7,9–18], our proposed scheme detects and defends against compromised nodes without the need for extra MIMO antennas or the tracing symbols. Moreover, the proposed scheme requires much smaller number of pre-loaded keys for key establishment and prevents the compromised nodes to pretend to be trustworthy nodes. Furthermore, by adjusting the security level, the proposed scheme can achieve different tradeoffs between energy and communication efficiency and the credibility of the received data.

The rest of this paper is organized as follows. Section 2 provides the system model and the proposed framework of the cross-layer secured communication scheme. Section 3 elaborates the cluster formation and the cooperative relay scheme. Section 4 presents the security key management scheme. Section 5 first presents in detail the algorithms for compromised nodes detection and then shows the symbol recovery method in the physical layer that eliminates the garbled symbol, and finally gives the key revocation and network recovery scheme. Section 6 shows the performance of the proposed secured communication scheme for cooperative MIMO networks through computer simulations. Finally, a conclusion is drawn in Section 7.

2. System model and the proposed cross-layer secured communication scheme

2.1. System model

In this paper, a multi-hop cooperative wireless sensor network that relays multiple source data back to the sink is considered. As shown in Fig. 1, the WSN consists of a set of sensor nodes that are equipped with a single-antenna radio. These single-antenna nodes are called primary nodes. Information collected by multiple local sensors needs to be aggregated and relayed to a remote sink. The sensor nodes between the source nodes and the sink will form into clusters and serve as relay nodes to improve the communication quality using the benefit of the MIMO technique. These clusters are also called virtual MIMO nodes in the rest of the paper, such as nodes 1, 2 in Fig. 1. The transmission link between two virtual MIMO nodes is called virtual MIMO link.

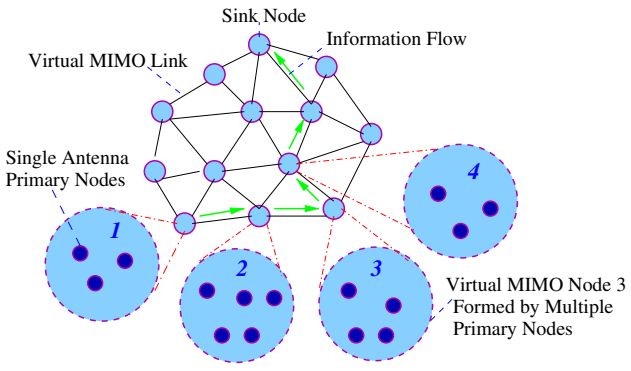


Fig. 1. System model.

Among the cooperative strategies, the amplify-and-forward and decode-and-forward are most widely used. In the amplify-and-forward strategy, the relay nodes simply boost the energy of the signal received from the sender and retransmit to the receiver. In the decode-and-forward strategy, the relay nodes will perform physical layer decoding (signal detection and demodulation) and then forward the decoded results. Although the amplify-and-forward relay has lower relay power consumption, it also amplifies the noise in the received signal and is not suitable for long-haul transmission. Moreover, decoding may be necessary when data aggregation and/or fusion is required at some local points such as cluster heads. Furthermore, considering that the decode-and-forward relay can be extended to combine with coding techniques and is easier to incorporate into network protocols [19], it will be considered in this paper.

Consider that the transmitting and receiving clusters have m_T and m_R nodes, respectively. The received signal at the virtual receiving MIMO node can be represented as [20]

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{s} + \boldsymbol{w} \tag{1}$$

where $\boldsymbol{y} = [y_1, y_2, \ldots, y_{m_R}]^T$ is a $m_R \times 1$ vector representing the received signals at the receiving cluster, $\boldsymbol{s} = [s_1, s_2, \ldots, s_{m_T}]^T$ is a $m_T \times 1$ vector representing the transmitted signal at the transmitting cluster, $\boldsymbol{H}$ is the a $m_R \times m_T$ matrix of channel coefficients, $\boldsymbol{w} = [w_1, w_2, \ldots, w_{m_R}]^T$ is a $m_R \times 1$ vector representing the additive Gaussian noise components, they are identically distributed and mutually statistically independent, each having zero mean and two-sided power spectral density $2N_0$.

Since this paper focuses on how to secure the communications in a cooperative MIMO network, we assume that the channel matrix $\boldsymbol{H}$ is known at the receiving cluster, but not at the transmitting cluster. This is feasible and can be achieved by using proper channel estimation method that is performed frequently enough to track the channel variations [21]. Usually the channel estimation is based on the known sequence of bits, which is unique for a certain transmitter and which is repeated in every transmission burst. Thus, the channel matrix $\boldsymbol{H}$ can be estimated for each burst separately by exploiting the known transmitted bits and the corresponding received samples.

2.2. Proposed cross-layer secured communication scheme

We propose a cross-layer secured communication scheme for cooperative MIMO wireless sensor networks, as shown in Fig. 2. Based on the security level set by the sink, each of the primary nodes in the receiving/detection process will determine whether it needs to perform compromised nodes detection during the sink defined time period. If detection is not needed, normal cooperative data transmission or relay will be conducted during this time period. Otherwise, compromised nodes detection will be performed. If the detection results indicate that there is no compromised node, normal cooperative data transmission or relay will be conducted for the rest of the time period. Otherwise, symbol recovery will be conducted to

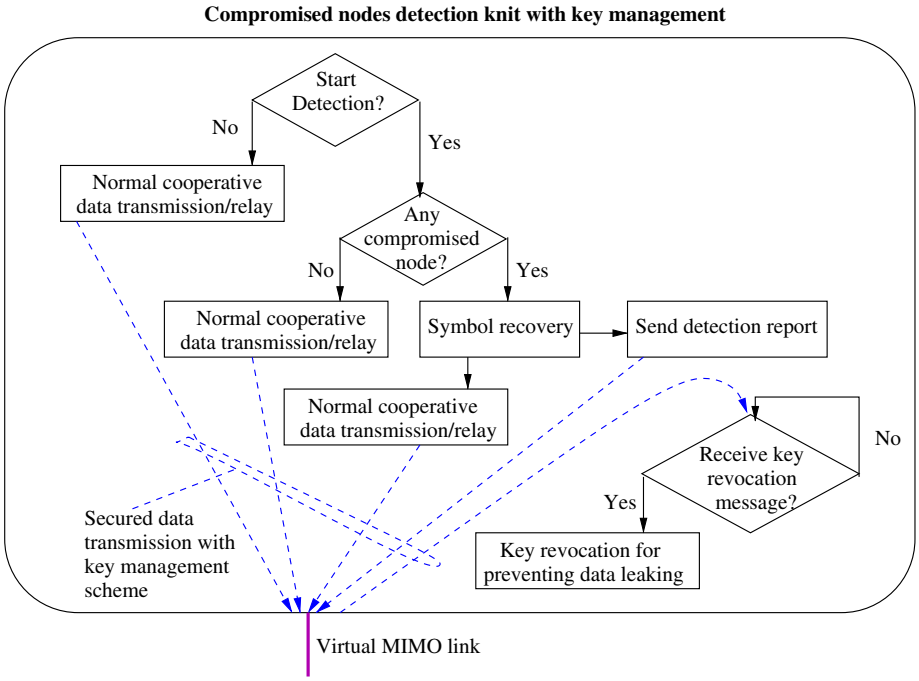


Fig. 2. Cross-layer secured communication scheme for cooperative MIMO networks.

eliminate the garbled symbol. The detection report will then be sent to the sink and normal cooperative data transmission or relay will be conducted for the rest of the time period. On the other hand, if the sink receives the report of compromised nodes, the key management system will invoke the key revocation to maintain the accuracy of the next detection and stop compromised nodes getting information from the network and reconfigure the cooperative MIMO network.

To operate the proposed cross-layer secured communication scheme, there are three major tasks: (1) how to form the cooperative MIMO network with distributed primary nodes; (2) how to establish secret key and build secured communication and routing, and (3) when and how to detect the compromised nodes. The accomplishments of these three tasks are presented in the next three sections.

3. Cooperative network architecture and transmission scheme

This section discusses the cluster-based cooperative network architecture and the cooperative transmission scheme. It provides the fundamentals for the secured communication scheme described in Sections 4 and 5.

3.1. Cooperative MIMO networking architecture and formation

Let G be a network of single-antenna wireless nodes and V be the set of nodes in G. A d-clustering of V is a node-disjoint division of V, where the distance of two nodes in a d-cluster is not larger than d. Let A and B be two d-clusters with m_T ($m_T \geqslant 1$) and m_R ($m_R \geqslant 1$) nodes, respectively. If the distance of any node of A and any node of B does not exceed $D(D \gg d)$, a D-$m_T \times m_R$ cooperative MIMO transmission link can be defined between A and B, where node i in A uses its antenna as the ith antenna cooperating the transmission and node j in B uses its antenna as the jth antenna cooperating the reception. In order to avoid confusion, in this paper, a single-antenna wireless node in G is called as *primary node*, a d-cluster is called as *virtual MIMO node*, and a D-$m_T \times m_R$ cooperative MIMO transmission link is called as *virtual MIMO link*. Given d and D, a cooperative MIMO (CMIMO) radio network of G can be represented as an undirected graph $G_{CMIMO} = (V_{CMIMO}, E_{CMIMO})$, where V_{CMIMO} is the set of the d-clusters, and E_{CMIMO} is the set of edges. An edge $(A, B) \in E_{CMIMO}$ if and only if $A, B \in V_{CMIMO}$ and there is a D-$m_T \times m_R$ cooperative MIMO link between A and B. A cooperative MIMO network can be formed from the given G, d, and D as follows [8]:

1. the primary nodes in G self-form a cooperative MIMO radio network G_{CMIMO} by using a distributed clustering algorithm on G,
2. the virtual MIMO nodes (d-clusters) form a multi-hop backbone tree by using a distributed Spanning-Tree formation algorithm on G_{CMIMO}, and
3. the routing for data dissemination, data gathering and unicast is constructed by the paths of the backbone tree.

After the CMIMO network formation, the structures of the clusters and backbone tree are maintained. Each cluster A has a cluster ID. Each primary node in A retains the following information: ID of cluster A, IDs of all primary nodes in A, IDs and sizes of the clusters which are the neighbors of A in the backbone tree.

3.2. Cooperative transmission scheme

The multiple antennas in MIMO radio systems are used to provide diversity gain and multiplexing gain. In this pa-

per, we consider diversity gain only. There are two types of communication in a cooperative MIMO relay network: local/intra communication at virtual MIMO nodes and long-haul/inter communication between virtual MIMO nodes [3,22]. The following MIMO scheme cooperatively relays $k(k \geqslant 1)$ source data in a cluster A back to the destination:

MIMO Scheme for data relay between virtual nodes A and B

1. First hop between virtual nodes A and B:
   Fig. 3 shows the first hop cooperative transmission between virtual nodes A and B. It includes local transmission in virtual node A, and long-haul transmission between virtual nodes A and B.
   Step 1 (Local transmission at A): Each primary node i in A with source data I_i broadcasts its data to all other nodes using different timeslots. After this step, each node in A has source data sequence $I = I_1, I_2, \ldots, I_k$.
   Step 2 (long-haul transmission between A and B using multiple $m_T \times m_R$ MIMO link): Suppose there are $|A|$ and m_R cooperative nodes in transmission side A and in reception side B, respectively. m_T nodes in cluster A with smallest IDs will attend the data transmission, where
   $$m_T = \begin{cases} |A|, & \text{if } |A| \leqslant |D| \\ \text{round}((|D|+1)\cdot\frac{2}{3}), & \text{if } |A| > |D| \end{cases} \quad (2)$$
   round$(\cdot)$ stands for round to nearest integer, D is detection cluster, which is the larger cluster between cluster B and the cluster before A in the relay route, and $|D|$ is the number of cooperative nodes in D. Detailed derivation for Eq. (2) is given in Section 5. Since each node in A has the list of IDs of the primary nodes in A and knows the sizes of A's neighboring cluster, it can decide $|D|$, calculate m_T, and judge if it should attend the data transmission.
   After nodes's self-selection, each attending node with the ith smallest ID in A acts as the ith antenna and encodes the data sequence I using $m_T \times m_R$ MIMO coding. All m_T nodes in A broadcast encoded sequence to the m_R nodes in B at the same time. Each node of the m_R nodes in B receives combined m_T encoded sequences I.
   Step 3 (Local transmission at B): (i) Each primary node in B broadcasts the received data to all other primary nodes using different timeslots. (ii) After receiving the data from the other primary nodes, each primary node in B decodes the received data back to the original source data sequence I.
2. Other hops between virtual nodes B and C:
   The transmission in the hops other than the first hop consists the long-haul transmission between virtual nodes B and C that is similar to Step 2 in the first hop and the local transmission at C that is similar to Step 3 in the first hop. The only change is to replace A and B to B and C, respectively.

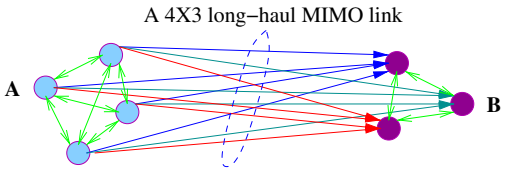


Fig. 3. Cooperative MIMO data transmission scheme.

Eq. (2) will also be used as the condition for compromised nodes detection in Section 5.

4. Security key management scheme

The key management system is based on shared/symmetric key cryptography. It only needs a small number of pre-loaded keys. Since localization itself is a very challenging problem, the key establishment in this work uses topology knowledge instead of the location knowledge which are used in existing work [23].

Types of keys: Fig. 4 shows the two types of keys used in the cooperative MIMO communications. They are:

- Shared keys, C-key(A), for local communication at each cluster.
- Shared keys, L-key(A, B), for long-haul communication at each link of two clusters A and B in the backbone tree. When the primary nodes in A and B cooperate on data transmission and reception, each node in A uses L-key(A, B) to encrypt the transmission data and each node in B uses the same key to decrypt the received data.

Key pre-distribution: For each primary node u in WSNs, a shared key, pre-key(b, u), is pre-distributed at the sink b and at the node u, respectively.

Key establishment: We proposed an key establishment algorithm as shown in Algorithm 1.

Algorithm 1. Key Establishment (Assume that the CMIMO network is already formed)

1: A special node u (e.g., the primary node with the smallest ID) at each cluster A sends a key request to the sink b with a plain message (u's ID, b's ID) and an encrypted message (u's ID, b's ID, u's member-list of the cluster, u's neighbor-list of the backbone) encrypted by using pre-key(u, b).
2: When b receives the key request from u, b decrypts the message by using pre-key(b, u). After b receives the key requests from all nodes, it has the topology of the whole cooperative MIMO network. Then, b generates a C-key(A) for each cluster A, and an L-key(A, B) for link AB of each cluster B in A's neighbor-list of the backbone. b disseminates the key response to each primary node x in cluster A as follows: a plain message (b's ID, x's ID) and an encrypted message (b's ID, x's ID, C-key(A), a list of L-key(A, B) for each cluster B in A's neighbor-list) encrypted by pre-key(b, x).
3: When primary node x receives a key response, x decrypts the message by using pre-key(x, b) to get C-key or L-keys.

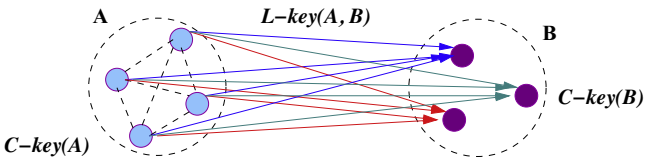


Fig. 4. Keys used in the proposed secured communication system.

Remark:

When b disseminates the key response to a primary node x, it delivers a package to x which includes the plain message (b's ID, x's ID), encrypted message (b's ID, x's ID, and a list of L-key(A, B)). The key responses for n primary nodes are distributed by performing depth-first travel on the backbone tree of the CMIMO network. Due to the cooperative communication, when a virtual MIMO node on the backbone tree receives a key response, all primary nodes in the virtual node receive the same key response. Therefore, the time required for key distribution is $O(n + t)$, where n is the number of primary nodes, and t is the size of the backbone tree which is the number of virtual MIMO nodes.

Secured Communication and Routing:

After the key establishment, the communication in each local virtual MIMO node A uses C-key(A) and in link AB at the backbone uses L-key(AB). Since the routing uses the paths on the backbone tree, cooperative data relay is secured.

In the proposed key management scheme, each primary node u needs only one pre-distribution key. After key establishment, each primary node u has one C-key and k L-keys, where k is the number of the neighbors that u's cluster has in the backbone tree. It is very small number of keys and affordable for small and inexpensive nodes. The total number of C-keys and L-keys in the whole network are n (the number of clusters), and $n - 1$ (the number of edges in the backbone tree), respectively. The proposed key management system is more efficient than other existing systems: it uses shared/symmetric key cryptography which requires small size of keys, it needs only a small number of keys at each primary nodes, and key establishment can be performed without location knowledge. Therefore, it is affordable for small and inexpensive nodes.

5. Compromised nodes detection with information and network recovery

In this Section, the algorithms for compromised nodes detection is presented first. After detection, the symbol recovery method is used to eliminate the impacts of the compromised nodes. The proposed detection and recovery can be applied to scenarios where the transmitting and detection clusters have different numbers of cooperative nodes. Finally, the key revocation algorithm is used to isolate the compromised nodes in the topology and the clusters are self-reconfigured. After key revocation, the compromised nodes are not able to affect the network or get information from the network.

5.1. Compromised nodes detection

In cooperative communications with multiple nodes and sophisticated relay rules, the security enforcement is a challenging and delicate task [11]. Here we present a physical layer identification approach that detects compromised nodes without increasing the transmission overhead. The system complexity will also be maintained in most cases.

Before we present the distributed algorithm for compromised nodes detection, we propose an algorithm shown in Algorithm 2 to determine at each cluster whether the compromised nodes detection is needed. In this algorithm, the sink broadcasts a time interval t_I and security level $0 \leqslant sl \leqslant 1$ at the beginning of the WSN deployment. It may broadcast the adjusted time interval t_I and security level sl during the operation of the WSN when necessary. The shorter the time interval or the lower the security level, the more compromised nodes detections will be performed.

Algorithm 2. Start detection at cluster D?

1: At the beginning of each time interval t_I, the random number generator at the node h with the smallest ID in the cluster D generates a uniformly distributed random number l between 0 and 1.
2: h compares l with the sink defined security level, sl.
3: **if** $l > sl$ **then**
4: Node h broadcasts the detection message to other nodes in the cluster.
5: Each node in the detection cluster perform compromised nodes detection in this time interval before data transmission and relay.
6: **else**
7: No detection. The clusters perform normal cooperative data transmission or relay operation.
8: **end if**

After the cluster decides that the compromised nodes detection is needed, the cluster will use symbols of time span t_d for detection, where $t_d \ll t_I$. The starting point of t_d is uniformly distributed in the sink defined time interval t_I.

Remarks:

1. In this proposed scheme for cooperative compromised nodes detection, the detection cluster is always the receiving cluster, either in the receiving side of a transmission-receiving pair or through listening. On the other hand, the detection is conducted at random times, the compromised nodes in a transmitting cluster do not know when to pretend to be trustworthy nodes. If the compromised nodes always pretend to be trustworthy nodes by transmitting correct data, they are not considered to be compromised nodes because all the data they transmitted are correct.
2. Detecting the compromised nodes at random times has two advantages. First, in the fixed-time detection scheme [24], the compromised nodes can pretend to be trustworthy nodes by sending the correct data only

at the detection time. However, in this proposed scheme, the compromised nodes cannot pretend to be trustworthy nodes since they do not know when the detection process will be performed. Second, the security level is adapted according to the detected number of compromised nodes, so that the energy and communications of the whole WSN can be saved.

In multi-hop WSNs, except the sink cluster, any one of the clusters will serve as the cluster to-be-detected when it is selected as relay cluster by the routing scheme. Consider two consecutive detection pairs as shown in Fig. 5, where cluster B is the cluster for detecting compromised nodes in cluster A and cluster C is the detection cluster for cluster B. If there are compromised nodes in B, B may detect the compromised nodes in A with higher error rate. However, these compromised nodes will be detected by C and removed from B with key revocation in the (B, C) detection pair. After this, B will not have compromised nodes and can detect compromised nodes in A with high accuracy. Therefore, without loss of generality, in the following algorithm for compromised nodes detection, we assume that the detection cluster does not have compromised node.

Fig. 6 shows the model for compromised nodes detection, where A is the transmitting cluster and D is the detection cluster.

In the following description of the compromised nodes detection algorithm, the diversity gain is obtained by letting all transmission nodes in cluster A transmit the same data stream. It can be easily extended when the space–time code is used as explained in remarks. As shown in Fig. 6, the m_T primary nodes in A transmit $s_1 = s_2 = \cdots = s_{m_T} = (I = I_1, I_2, \ldots, I_k)$, where $I_i, i = 1, 2, \ldots, k$ are data symbols. The primary nodes in D collect the data for detection. When the detection is needed, the node with the smallest ID in D requests all the nodes in its cluster to broadcast the received symbols at different time slots. Due to the small distance between the nodes in the same cluster, it is reasonable to assume that there is no error during this broadcast process when appropriate error correction coding scheme is used. After local broadcast, each primary node in D has complete received data sequence $\boldsymbol{y}$ and will perform distributed detection to identify compromised nodes in cluster A. Algorithm 3 describes the compromised nodes detection algorithm at each primary node in D.

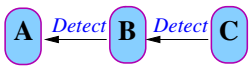


Fig. 5. Example of consecutive detection pairs.

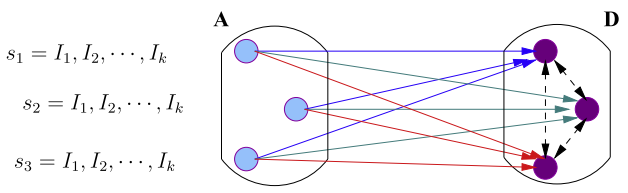


Fig. 6. Model for compromised nodes detection.

Algorithm 3. Compromised nodes detection

1: After receive complete data sequence $\boldsymbol{y}$, each primary node in the detection cluster D performs Inverse Channel Detection [20] to estimate the transmitted symbols s. The inverse channel detector multiplies a weighting matrix that is inverse or pseudo-inverse of the channel matrix $\boldsymbol{H}$ with the received symbols to estimate the transmitted symbols s, that is, $\hat{\boldsymbol{s}} = \boldsymbol{W}^H \boldsymbol{y}$, where $\boldsymbol{W}$ is an $|D| \times m_T$ weighting matrix and $|D|$ is the number of nodes in D, $(\cdot)^H$ represents Hermitian transpose. According to the system model in Section 2, the matrix of channel coefficients $\boldsymbol{H}$ is assumed known to the detection cluster. Moreover, since $m_T \leqslant |D|$ as shown in Eq. (2), $\boldsymbol{W}$ can be determined by

$$\boldsymbol{W} = \begin{cases} \boldsymbol{H}^{-1}, & \text{if } m_T = |D| \\ (\boldsymbol{H}^H \boldsymbol{H})^{-1} \boldsymbol{H}^H, & \text{if } |D| > m_T \end{cases} \tag{3}$$

2: Based on the assumption that all data streams s_i $(1 \leqslant i \leqslant m_T)$ transmitted by the non-compromised nodes should be the same, the detecting primary node can identify the compromised nodes x_i (if any) and record their IDs by checking whether x_i transmitted the same symbols as the majority nodes. To check that, the recovered data streams from different transmitting nodes will be sorted into groups, where nodes are assigned to the same group if they contain identical symbols. The group with largest number of nodes is assumed to be trustworthy nodes. All the other nodes are classified as compromised nodes.

3: When compromised node j in A is detected by primary node u in D, the encrypted detection report with a plain message (u's ID, the sink's ID b) and an encrypted message (u's ID, the sink's ID b, j's ID) encrypted by pre-key(u, b) will be sent by each detecting primary node. This message will then be relay to the sink b by the cooperative MIMO scheme for data relay. The sink will use majority rule to determine whether a reported node is really compromised or not, that is, if more than half of the primary nodes in the detection cluster claimed that node j is compromised, the sink will classify node j as compromised node.

Remarks:

1. Comparing with [11], no tracing symbols are needed in the proposed compromised nodes detection algorithm, therefore, the proposed algorithm does not have the overhead and the system complexity is lower without the need of tracing symbol transmission and extraction.
2. The proposed algorithm for compromised nodes detection will also work well when the space–time code is used. In this case, the symbol-by-symbol comparison will be replaced by the pattern comparison, where each

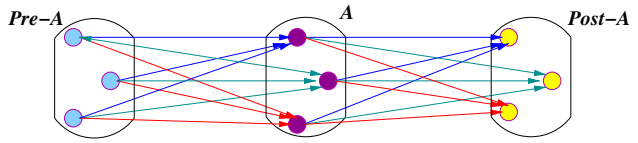


Fig. 7. Selection of cluster for detection.

pattern that includes several symbols is determined by the selected space–time code. Therefore, the full MIMO benefits can still be maintained with the proposed algorithm.

3. Simple sorting algorithm can be used for grouping in the second step when detecting the compromised nodes in transmitting cluster A. The total number of comparisons in the algorithm is $O(m_T \log(m_T))$, where m_T is the number of primary nodes in cluster A.

In above proposed algorithm for compromised nodes detection, in order to illustrate the selection of detection cluster and the maximum number of compromised nodes that can be identified, let's consider a two-hop data relay path as show in Fig. 7, where *Pre_A* cluster relays data to cluster A, then A relays data to cluster *Post_A*. Let $|A|, |Pre\_A|$, and $|Post\_A|$ denote the number of nodes in clusters A, Pre_A, and *Post_A*, respectively. We consider three cases to cover all the scenarios for selecting the detection cluster to identify the compromised nodes in cluster A and determining the maximum number of compromised nodes that can be identified.

Select the detection cluster to identify the compromised nodes in A

Case 1. $|A| \leqslant |Post\_A|$ (A is cluster 1 in Fig. 1) In this case, the compromised nodes in cluster A are detected by cluster *Post_A*. The maximum number of compromised nodes that can be identified is $|A|/2 - 1$.

Case 2. $|A| > |Post\_A|$ and $|A| \leqslant |Pre\_A|$ (A is cluster 3 in Fig. 1) In this case, the compromised nodes in cluster A are detected by cluster *Pre_A* when A relays data to cluster *Post_A*. The *Pre_A* cluster will listen A's transmission. The maximum number of compromised nodes that can be identified is $|A|/2 - 1$.

Case 3. $|A| > |Post\_A|$ and $|A| > |Pre\_A|$ (A is cluster 2 in Fig. 1).
In this case, the compromised nodes in cluster A are detected by the larger cluster of *Pre_A* and *Post_A*. According to the CMIMO Network transmission scheme, m_T nodes are self-selected to transmit or relay data and the rest keep idle. The maximum number of detectable compromised nodes in A and the value of m_T can be determined using the following equations:

$$\begin{aligned} N_{max} + m_T &= |D| \\ N_{max} &= \frac{m_T}{2} - 1 \end{aligned} \tag{4}$$

where N_{max} is the maximum number of compromised nodes in the transmitting cluster A and $|D|$ is the number of nodes in the detection cluster. The first part of Eq. (4) guarantees that the number of transmitted nodes is not exceed that the number of nodes in detection cluster which is required by Eq. (3), even when the compromised nodes keep sending garbled data when they are required to keep idle. The second part of Eq. (4) guarantees that the number of the compromised nodes is less than half of the transmitting nodes, so that by comparing the transmitted data in different nodes, the compromised nodes could be identified. By solving Eq. (4), we have

$$\begin{aligned} N_{max} &= (|D| + 1) \cdot \frac{1}{3} - 1 \\ m_T &= (|D| + 1) \cdot \frac{2}{3} \end{aligned} \tag{5}$$

If the solutions of N_{max} and/or m_T in Eq. (5) are not integer, they will be rounded to the closest integer. For example, when $|D|$ is 4, 5, or 6, we can detect one compromised node by setting m_T to 3, 4, or 5. When $|D|$ is 7, we can detect two compromised node by setting m_T to 5.

Based on the above approach for selecting detection cluster, we add which cluster D detects in the algorithm *Start detection at cluster D* as follows:

Since the nodes of D know the size of *Pre_D* and *Post_D*, where the cluster *Pre_D* is the cluster that relays data to D, and the cluster *Post_D* is the cluster that D relays data to.***

1. if $|D| \geqslant |Pre\_D|$ and/or $|D| \geqslant |Post\_D|$
   D detects *Pre_D* and/or *Post_D*.
2. else if m_T of *Pre_D* and/or *Post_D* is $(|D| + 1) \cdot \frac{2}{3}$
   D detects *Pre_D* and/or *Post_D*.

If the cluster D needs to detect the compromised nodes in both clusters *Pre_D* and *Post_D*, it first checks cluster *Pre_D* then cluster *Post_D*.

Summarizing the above three cases, it is clear that the detection algorithm can identify all the compromised nodes if they are less than half in the cooperative transmission cluster and the number of all nodes in this cluster is not larger than that of all its neighboring clusters. If the number of nodes in this cluster is larger than that of its all neighboring clusters, the proposed detection approach can detect all the compromised nodes if they are less than one third of the number of all nodes in the neighboring cluster that has the larger number of nodes.

5.2. Symbol recovery method that eliminates the garbled symbol

When the compromised nodes are detected, the sink will forward the ID of the compromised node to the receiving cluster. The receiving cluster then decodes the message by simply setting the columns in channel matrix that corresponds to the compromised nodes to zero. This will eliminate the use of the malicious data by ignoring the symbols transmitted from the compromised nodes.

5.3. Key revocation and network recovering

If the sink determines that there is compromised node in cluster A, it will start the key revocation and network recovering process. This approach is used to prevent compromised nodes from getting information in the network and sending false reports.

In key revocation, the sink b takes the following actions for key revocation and network recovery:

1. The sink b sends all nodes v in cluster A other than x a key revocation information with a plain message (b's ID, v's ID) and an encrypted message (b's ID, v's ID, new C-key(A), and ID list of the compromised nodes) encrypted by pre-key(b, v).
2. For each A's neighbor B in the backbone tree, b sends each node v in A and in B other than x a key revocation information with a plain message (b's ID, v's ID) and an encrypted message (b's ID, v's ID, new L-key(A, B)) encrypted by pre-key(b, v).
3. When node v in A and B receives a key revocation information from the above steps 1) and 2), it decrypts the message by pre-key(v, b) and gets a new C-key or L-key. In this way, the C-key for local communication in virtual node A and the L-key for long-haul communication between A and each of its neighboring clusters in the backbone tree are revoked. The compromised node x does not have the new keys and will be not able to get information from the network.

Remark:

When a compromised node in a cluster A is detected by A's parent, a report is sent back to the sink b from A's parent to A's grandparent, and then from A's grandparent to A's great-grandparent, etc., on the backbone tree. If each node keeps a record which indicates from whom it receives the report, b can send the key revocation packages to A and its neighbors via the path that reverses the path that the report traveled. Therefore, the time for key revocation is $O(h + k)$, where h is the height of the backbone tree and k is the number of A's neighbors.

6. Simulation results

In this section we investigate the performance of the proposed compromised nodes detection algorithm and the cooperative secured communication system through computer simulations. Since this paper deals with secured communication scheme, MATLAB, a commonly-used simulation tool for communications research, is selected. In the simulations, multiple single-antenna sensor nodes are physically grouped together to form a MIMO system. The active compromised nodes attack is considered. Instead of relaying the received information, the compromised nodes transmit randomly generated symbols. Similar to the existing works presented in [11], the multi-path scattered environment is considered. The channels are block Rayleigh fading channels, i.e., the channel coefficient matrix $\boldsymbol{H}$ is constant during the transmission of one symbol, but is randomly changing between symbols. Different channels are identically distributed and statistically independent. Binary phase shift keying (BPSK) is chosen as the modulation scheme. 100 received symbols are used in the proposed algorithms for compromised nodes identification. The maximum likelihood detector is used for symbol demodulation.

Since compromised nodes detection is performed between one transmitting virtual MIMO node and one receiving virtual MIMO node, the performance evaluation only evaluates one hop as shown in Fig. 6 to demonstrate the effectiveness of the proposed algorithm. Fig. 8 shows the accuracy of the proposed algorithm for compromised nodes detection when there is one compromised node in the transmitting cluster. The accuracy is defined as the ratio of correctly identified compromised nodes and normal nodes to all nodes. Since all the compromised nodes detection cases require $|D| \geqslant m_T$, where $|D|$ is the number of nodes in the detection cluster and m_T is the self-selected nodes in transmission, in this simulation, the transmitting cluster has 4 nodes and the detection cluster has 5 nodes. It is clear that the proposed algorithm has high identification accuracy, since when the SNR is larger than -4 dB, the proposed algorithm can identify the compromised node close to 100%.

Fig. 9 compares the performance of the proposed cooperative communication system with the conventional system that does not detect compromised nodes in terms of bit error rate (BER). There is one compromised node in the transmitting cluster, and the transmitting cluster has 4 nodes and the detection cluster has 5 nodes. The BER performance of the system when there is no compromised nodes is also presented with the dashed line as a reference of the optimum performance. The dotted line is for the conventional system where the receiving cluster does not detect compromised nodes and uses the garbled data from the compromised nodes in symbol demodulation. The solid line is for the proposed system. Comparing with the conventional system, it is clear that the proposed system significantly improves the reliability of the communication when SNR is higher than -8 dB. Comparing with the system without compromised node, the performance loss of the proposed system

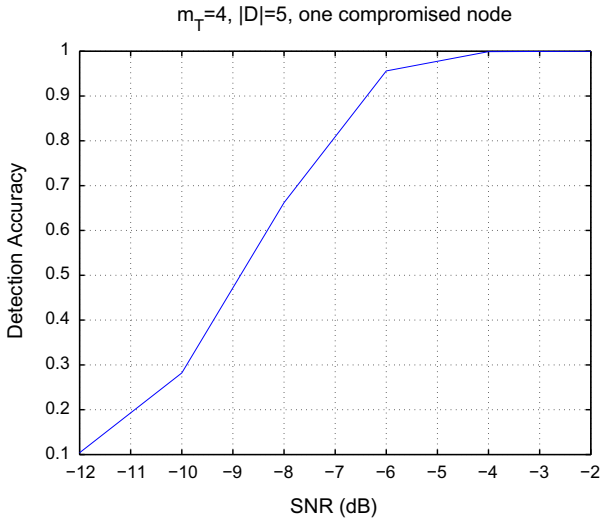


Fig. 8. Accuracy of the proposed compromised nodes detector.

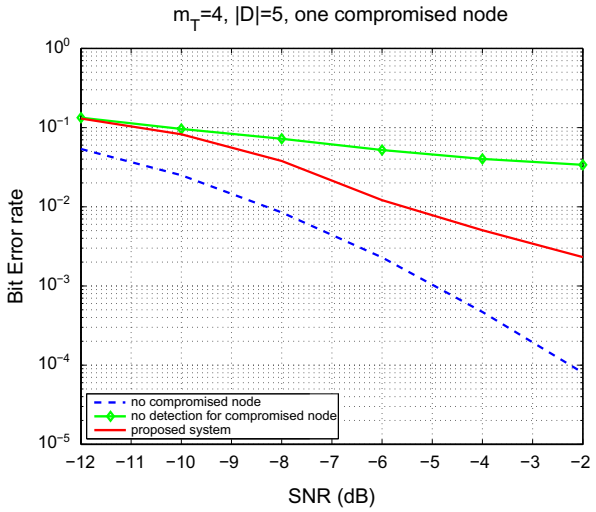


Fig. 9. Performance comparison.

Table 1
Performance evaluation for secure cooperative MIMO networks.

System number	$\|D\|$	m_T	Number of compromised node	BER	Detection accuracy
1	4	3	0	0.01955	NA
2	4	3	0	0.02115	NA
3	4	3	0	0.01916	1.000
1	4	3	1	0.01981	NA
2	4	3	1	0.13356	NA
3	4	3	1	0.05659	0.973
1	5	5	1	0.00313	NA
2	5	5	1	0.03794	NA
3	5	5	1	0.00991	0.963
1	7	5	2	0.00054	NA
2	7	5	2	0.08657	NA
3	7	5	2	0.01324	0.923

is because the diversity gain is smaller when the data from the compromised nodes are eliminated.

Table 1 gives more simulation results in terms of detection accuracy and the BER for different $m_T, |D|$, and the number of compromised nodes. The SNR is -6 dB. The system number 1, 2 and 3 represents the system free of compromised nodes, the system that does not detect compromised nodes and uses the garbled data from the compromised nodes in symbol demodulation, and the proposed system, respectively. For the first two types of systems, there is no compromised nodes detection and therefore the detection accuracy is not applicable (NA) to them. Simulations have been conducted for all the cases with $4 \leqslant |D| \leqslant 7, 4 \leqslant m_T \leqslant 7$, and $-12 \leqslant SNR \leqslant -2$. The rest results are available upon request. It is clear that the proposed schemes significantly improved the accuracy of the data transmission of the cooperative communication system when comparing with the system without compromised node detection.

7. Conclusions

This paper proposed a secure cooperative MIMO communication system under active compromised nodes, where some of the relay nodes are compromised and try to corrupt the communications by sending garbled signals. To combat the compromised nodes, we propose a cross-layer secured communication scheme for cooperative MIMO communication in wireless sensor networks to overcome the external and active compromised nodes attacks. The scheme combines cryptographic technique implemented in higher layers with data assurance analysis at the physical layer to provide better communication security. The cryptography provides secured data transmission between authorized nodes and it also secures key revocation and network recovery. An information theory based algorithm for compromised nodes detection is proposed. It can identify all the compromised nodes in the latest configured cooperative node groups, provided that the compromised nodes are less than half in the cooperative cluster and the number of all nodes in this cluster is not larger than that of its all neighbors. If the number of all nodes in this cluster is larger than that of its all neighbors, the proposed detection approach could detect all the compromised nodes if they are less than one third of the number of all nodes in the neighboring cluster that has the largest number of nodes. The effectiveness and efficiency of the proposed algorithm for compromised nodes detection is demonstrated through computer simulations. The simulation results also show the significant improvement in the accuracy of received information.

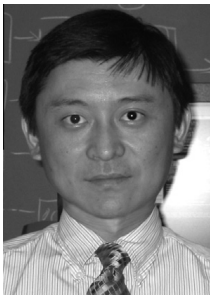
Acknowledgment

This work is supported in part by U.S. AFRL/RY Contract FA8650-05-D-1912, 2010–2011.

References

[1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, A survey on sensor networks, IEEE Commun. Mag. 40 (2002) 102–114.
[2] J. Mietzner, R. Schober, L. Lampe, W.H. Gerstacker, P.A. Hoeher, Multiple-antenna techniques for wireless communications – a comprehensive literature survey, IEEE Commun. Surveys Tuts. 11 (2009) 87–105.
[3] S. Hussain, A. Azim, J.H. Park, Energy efficient virtual MIMO communication for wireless sensor networks, Telecommun. Syst. 42 (2009) 139–140.
[4] S.K. Jayaweera, Virtual MIMO-based cooperative communication for energy-constrained wireless sensor networks, IEEE Trans. Wireless Commun. 5 (2006) 984–989.
[5] W. Chen, Y. Yuan, C. Xu, K. Liu, Z. Yang, Virtual MIMO protocol based on clustering for wireless sensor network, in: Proc. IEEE Symp. Computer and Communications, Cartagena, Murcia, Spain, 2005, pp. 335–340.
[6] J. Liu, Energy-efficient cross-layer design of cooperative MIMO multi-hop wireless sensor networks using column generation, Wireless Personal Commun. J. 66 (2012) 185–205.
[7] M.R. Islam, J. Kim, Energy efficient cooperative MIMO in wireless sensor network, in: Proc. IEEE International Conference on Intelligient Sensors, Sensor Networks and InformationProcessing (ISSNIP), Sydney, Australia, 2008, pp. 505–510.
[8] W. Chen, H. Miao, L. Hong, J. Savage, H. Adas, Cross layer design of heterogeneous virtual MIMO radio networks with multi-optimization, in: Proc. IEEE International Parallel and Distributed Processing Symposium on Advances in Parallel and Distributed Computing Models, Atlanta, GA, USA, 2010, pp. 1–8.
[9] Y. Zhou, Y. Fang, Y. Zhang, Securing wireless sensor networks, a survey, IEEE Commun. Surveys Tuts. 10 (2008) 6–28.
[10] C. Karlof, D. Wagner, Secure routing in wireless sensor networks: attacks and countermeasures, in: Proc. 1st IEEE International Workshop on Sensor Network Protocols and Applications, Anchorage, AL, USA, 2003, pp. 113–127.

[11] Y. Mao, M. Wu, Tracing malicious relays in cooperative wireless communications, IEEE Trans. Inform. Forensics Sec. 2 (2007) 198–212.
[12] X. Chen, K. Makiki, K. Yen, N. Pissinou, Sensor network security: a survey, IEEE Commun. Surveys Tuts. 11 (2009) 52–73.
[13] A. Jain, K. Kant, M.R. Tripathy, Security solutions for wireless sensor networks, in: Proc. IEEE Second Intl. Conf. Advanced Computing and Communication Technologies, Rohtak, Haryana, India, 2012, pp. 430–433.
[14] V.C. Sekhar, M. Sarvabhatla, Security in wireless sensor networks with public key techniques, in: Proc. IEEE Intl. Conf. Computer Communication and Informatics, Coimbatore, India, 2012, pp. 1–16.
[15] H.K.D. Sarma, A. Kar, R. Mall, Secure routing protocol for mobile wireless sensor network, in: Proc. IEEE Sensors Applications Symposium, San Antonio, TX, USA, 2011, pp. 93–99.
[16] X. Li, J. Hwu, Using antenna array redundancy and channel diversity for secure wireless transmissions, J. Commun. 2 (2007) 24–32.
[17] H. Kim, J.D. Villasenor, Secure MIMO communications in a system with equal numbers of transmit and receive antennas, IEEE Commun. Lett. 12 (2008) 386–388.
[18] H. Wen, G. Gong, A MIMO based Cross-layer Approach to Augment the Security of Wireless Networks, Tech. Rep. CACR 2008-21, University of Waterloo, 2008.
[19] J.N. Leneman, D.N.C. Tse, G.W. Wornell, Cooperative diversity in wireless networks: efficient protocols and outage behavior, IEEE Trans. Inform. Theory 50 (2004) 3062–3080.
[20] J.G. Proakis, Digital Communications, ffith ed., McGraw-Hill, New York, NY, 2008.
[21] S.M. Alamouti, A simple transmit diversity technique for wireless communications, IEEE J. Sel. Areas Commun. 16 (1998) 1451–1458.
[22] S. Cui, A.J. Goldsmith, A. Bahai, Energy-efficiency of mmio and cooperative mimo techniques in sensor networks, IEEE J. Sel. Areas Commun. 22 (2004) 1089–1098.
[23] X. Du, M. Guizani, Y. Xiao, H.-H. Chen, A routing-driven elliptic curve cryptography based key management scheme for heterogeneous sensor networks, IEEE Trans. Wireless Commun. 8 (2009) 1223–1229.
[24] W. Chen, M. McNeal, L. Hong, Cross-layered design of security scheme for cooperative MIMO networks, in: Proc. IEEE International Conference on Wireless Information Technology and Systems, Honolulu, HI, USA, 2010, pp. 1–4.



Dr. Liang Hong received the B.S. and the M.S. degrees in Electrical Engineering from Southeast University, Nanjing, China in 1994 and 1997, respectively, and the Ph.D. degree in Electrical Engineering from University of Missouri, Columbia, Missouri in 2002. He was a Research Associate at Siemens Corporate Research, Princeton, New Jersey from February 2003 to July 2003. Since August 2003, he has been with the Department of Electrical & Computer Engineering at Tennessee State University where he is now Associate Professor. His research interests include cognitive radio, modulation classification, error control coding, noise and interference reduction, and wireless multimedia communications and networks. He has authored and coauthored over 30 technical refereed papers and served as the technical committee member for many international conferences. He also served as session chairs for several academic conferences. Throughout his academic career, he has served as Co-PI of NSF funded research projects and involved in several federal and industry funded projects.

Dr. Wei Chen received the B.A. degree in Mathematics in Shanghai Maritime University, China in 1982, and the M.S. and the Ph.D. degrees in Computer and Information Science in Osaka University, Japan in 1991 and 1994, respectively. She was an Assistant Professor and Associate Professor in Nagoya Institute of Technology, Japan during 1994–1998 and 1999–2000, respectively. She was an Associate Professor in Nanzan University, Japan in 2001. She has been a Professor in Tennessee State University since 2002. Her research interests include cognitive radio networks, wireless sensor, ad hoc, and mobile networks, network security, parallel/distributed computing, and bioinformatics. Dr. Chen has been the principal investigator in Sensor Directorate at MLP program (Air Force) since 2006. She and her team have conducted a number of innovative research projects such as "Cross-Layered Design of MIMO Radio Networks" and "Cooperative and Networked Methods for Spectrum Sharing and Interference Reduction". She has also led and engaged in the research of wireless sensor networks, network security, bioinformatics and some other areas funded by ARO, DTRA, NASA and NSF. She received IBM Faculty Award in 2010, Research Awards in Tennessee State University (TSU) in 2006 and Most Outstanding Faculty Award in College of Engineering at TSU in 2012. She has over 80 peer-reviewed journal/conference publications, and has served as editor/associated-editor, committee chair/member for a number journals and international conferences.