Group Secret Key Generation via Received Signal Strength: Protocols, Achievable Rates, and Implementation

Hongbo Liu, Jie Yang, Yan Wang, Yingying Chen, C. Emre Koksal

Abstract—Secret key generation among wireless devices using physical layer information of radio channel has been an attractive alternative for ensuring security in mobile environments. Received Signal Strength (RSS) based secret key extraction gains much attention due to its easy accessibility in wireless infrastructure. However, the problem of using RSS to generate keys among multiple devices to ensure secure group communication in practice remains open. In this work, we propose a framework for collaborative key generation among multiple wireless devices leveraging RSS. To deal with mobile devices not within each other's communication range, we employ relay nodes to achieve reliable key extraction. To enable secure group communication, two protocols are developed to perform collaborative group key generation via star and chain topologies respectively. We further provide the theoretic analysis on the achievable secrecy rate for both star and chain topologies in the presence of an eavesdropper. Our prototype development using MICAz motes and extensive experiments using fading trend based key extraction demonstrate the feasibility of using RSS for group key generation in both indoor and outdoor environments, and concurrently achieving a lower bit mismatch rate compared to existing studies.

Index Terms: collaborative secret key extraction, received signal strength, group key extraction, mobile wireless network.

1 Introduction

The usage of wireless devices (e.g., PDAs, smartphones, and laptops) has become an inseparable part of our daily lives, which actively involves in information sharing and various data transactions in ways that previously were not possible. To ensure the successful deployment and adoption of these emerging applications, secure communication is crucial to support data transmission confidentiality, data integrity, and device authentication among multiple wireless devices. For example, police officers covering different street blocks need to share with each other the monitoring information along their daily patrol routes and the recording of the crime information by areas; soldiers carrying out a particular task need to share task plans and real-time monitoring results among themselves, but not to unauthorized parties. Another example is a group of travelers want to limit the sharing of travel plans, journals, pictures and video clips within the group through the peer-to-peer association.

There have been active researches in applying traditional cryptographic-based methods such as public key infrastructure (PKI) to wireless networks, these methods, however, may not be always applicable because of the limited resources on wireless devices (e.g., limited battery and computation power), and lacking of a fixed key management infrastructure due to highly dynamic mobile wireless environments (e.g., peer-to-peer association, neighborhood devices changing frequently). In addition, the openness of the wireless transmission medium makes the key establishment itself vulnerable to eavesdropping—adversaries within communication range of legitimate devices can monitor any information exchanges of key generation and renewal. In this study, we examine secure group communications among multiple wireless devices by exploiting physical layer information of the radio channel instead of using the traditional cryptographic-based methods.

The main advantage of the secret key generation utilizing physical layer information of the radio channel is that it allows any two wireless devices within transmission range of each other to extract a shared symmetric cryptographic key while does not require a fixed infrastructure or a secure communication channel [1]–[3]. Based on the *principle of channel reciprocity*, two wireless devices can extract identical secret bits independently by using the sampled sequence from the radio channel between them within the coherence time of the channel. Unlike existing key generation algorithms, such as Diffie-Hellman, which rely upon computational hardness of problems, secret key generation using channel randomness provided by the temporal and spatial variation of the radio channel can achieve information-theoretical secrecy [4].

Comparing to various physical layer information of the radio channel (such as channel phase [5], [6]), sampling Received Signal Strength (RSS) is an attractive approach to generate secret keys as the RSS readings are readily available in the existing wireless infrastructure and thus presents tremendous cost savings. However, previous studies on RSS based secret key generation mainly focus on improving the secret bit generation rate between a pair of wireless devices (by exploiting temporal and spatial variations of radio channel [7], [8], multiple antenna diversity [9], and multiple frequencies [10]). The problem of using RSS to practically perform key generation among multiple wireless devices to ensure secure group communication remains a challenge. Group secret key generation problem has been addressed mainly conceptually in [11]–[15].

In this work, we propose collaborative secret key ex-

- *Hongbo Liu is with Department of Computer Information and Technology, IUPUI, Indianapolis, IN, 46202. E-mail: hl45@iupui.edu. This work is conducted during his Ph.D. study at Stevens Institute of Technology.*
- *Yan Wang, Yingying Chen are with Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ, 07030. E-mail: {ywang48, yingying.chen}@stevens.edu*
- *Jie Yang is with Department of Computer Science and Engineering, Oakland University, Rochester MI, 48309. E-mail: yang@oakland.edu*
- *Can Emre Koksal is with Department of Electrical and Computer Engineering, Ohio State University, Columbus, OH, 43210. E-mail: koksal@osu.edu*

traction for a group of wireless devices using readily available RSS measurements, rather than relying on a key distribution infrastructure. The group of wireless devices involved in key generation may not be within each other's communication range. We address this issue by employing a relay node assisted approach and define a metric using difference of RSS to maintain secrecy among devices. To enable secure group communication, two protocols are developed in our framework via *star* and *chain* topologies respectively by exploiting RSS from multiple devices to perform group key generation collaboratively. Note that, it has been shown how nodes coordinate and group together to form the star and the chain topology in [16]. In particular, the collaborative key extraction via the star topology is designed for scenarios when multiple wireless devices are within each other's communication range (e.g., people traveling together), whereas the approach via the chain topology deals with scenarios when not all wireless devices under consideration are within each other's communication range, but they are interconnected (e.g., patrolling police officers and soldiers carrying out military tasks). We assume that all the nodes pass authentication before joining the group, and thus none of the group members act maliciously for both the star and the chain scenarios.

We analyze the reliability and scalability of the proposed collaborative secret key extraction framework by deriving the maximum achievable group key rate for our scheme under both star and chain topologies in the presence of eavesdroppers. Our secret key rate assumes full equivocation of the secret key bits at the eavesdropper, i.e., the *rate* of mutual information leakage to the eavesdroppers is 0. This implies that, regardless of the statistical method employed by the eavesdroppers, they cannot decode any of the key bits with high probability. We specify the amount of drop in key rate as a function of the size of the group. Additionally, to deal with various noises in real-world scenarios, we propose a secret key generation scheme exploiting the trend exhibited in RSS resulted from shadow fading to encode secret bits to work with our group key extraction framework. Our fading trend based key extraction aims to achieve a lower bit mismatch rate comparing to existing studies when maintaining a comparable bit generation rate. Compared with many traditional group key generation approaches, that perform pair-wise key generation first and then distribute to the rest of the nodes through secure links, our approach does not require the presence of links that are secured a priori. With a limited number of nodes in this group, the bit mismatch rate for group key extraction with our fading trend based method is acceptable, where using reconciliation technique could further recover the mismatched bits by using error correction code.

Furthermore, we build a system prototype using MICAz motes and conduct extensive experiments in both outdoor (e.g., park and street) and indoor (e.g., office building) environments to evaluate the effectiveness of our proposed collaborative key generation framework. Our experimental results confirm the feasibility of using RSS for group key generation among multiple wireless devices under various mobile scenarios. The results also demonstrate that our fading-trend assisted key extraction scheme can achieve a lower bit mismatch rate compared to existing studies when maintaining a comparable secret bit generation rate.

The rest of the paper is organized as follows: We place our work in the context of related research in secret key extraction in Section 2. We provide our framework overview and attack model in Section 3. We then describe the building block in our framework, relay node assisted collaborative key extraction, in Section 4. We next present our group key extraction protocols via the star topology in Section 5 and chain topology in Section 6 together with the corresponding theoretic analysis. We discuss the group key extraction under the hybrid topology in Section 7. To deal with various noises in practice, we show how to perform secret key extraction using RSS fading trend in Section 8. We present the prototype implementation and performance evaluation results in Section 9. Finally, we conclude our work in Section 10.

2 Related Work

There have been active theoretic studies on characterizing secrecy capacity using physical layer information. Wallace et al. [17], [18] present the mutual information secret bit rate bounds from theoretical channel models. Maurer et al. propose an information theoretic bound for secrecy rate between two nodes in the presence of an eavesdropper node [19]. And the theoretic basis for the feasibility of using channel state information in OFDM system for key generation is also explored [20].

Various radio channel features have been proposed for secret key extraction in literature. Phase difference is first proposed [21], in which differential phase of two-tone signal is measured and quantized to generate secret keys. Phase difference is further exploited [5], [6]. Sayeed et al. use random phase for secret key extraction in an OFDM system [5], whereas Wang et al. propose a scheme for efficient key establishment [6]. The impulse response of a wireless channel is used to generate a shared secret [4], [22]–[24]. Ultra-wideband radios are used to measure the impulse response [23], [24], while Mathur et al. and Ye et al. propose to estimate the impulse response from Wilar signals [4], [22], respectively. Statistics of the Angle-of-Arrival (AOA) is used as a signature for key generation [2], however, it requires an access point to have a programmable phased array antenn Received signal strength or channel gain is the most commonly used radio channel feature for secret key extraction due to it is readily available in existing wireless infrastructure, and thus it is easy to measure with little effort. Previous studies mainly focus on exploiting temporal and spatial variations of a radio channel [1], [3], [4], [7], [8], [25], [26], multiple antenna diversity [9], and multiple frequencies [10] for secret bit extraction between a pair of wireless devices. The change in the signal envelop during a transmission is used to encode and decode transmitted messages [25]. Li et al. use the universal software radio peripheral (USRP) and GNU radio to generate a 24-bit signature based on the measured channel gain [26]. Azimi et al. utilize the deep fading of channel gain that periodically occurs in mobile channels is proposed to extract secret bits [1]. Mathur et al. [4] generate secret bits using the RSS extracted from 802.11a packets with mobile devices. Patwari et al. focus on improving the secret bit generation rate in mobile wireless networks [7], [8], while Wilhelm et al. use multiple frequencies to generate secret keys in static wireless sensor networks [10]. Multiple-antenna diversity is also exploited to improve the bit generation rate [9]. However, none of these RSS based methods considers key generation for multiple wireless devices.

Different from the above studies, our group key generation method utilizing readily available RSS measurements is lightweight, and thus is a practical solution for different types of wireless networks. To show the practicality of our proposed method, we analyze the basic information theoretic limits of the achievable key rate (for the analysis of group secret key rates for the general source models, see [27] and [28]) and build a prototype using MICAs motes to evaluate it in both outdoor and indoor environments.

Also, there has been some theoretical investigations on group secret key generation [11]–[15]. The main approach there is to generate pairwise secret keys first, then to generate a group key, exploiting the pairwise keys. Particularly, Arazi. et.al. [11] introduce an ECC-based methodology for group key generation in ad hoc clusters of sensor nodes. Kim. et.al [12] propose to blend key trees with DiffieÐHellman key exchange for group key generation. Nitinawarat, et.al. [13], [14] develop a group key generation scheme that uses the existing practical Slepian-Wolf codes and Steiner tree packing in a multigraph for local pairwise and global key propagation respectively. Ye. et.al. [15] propose a method for secret key agreement in the kind of network based on well-established point-to-point techniques over a graphical representation of the network. The main objective is to achieve the group secret key capacity, derived in [28]. However, these capacity achieving schemes involve a separate information reconciliation phase with each user as well as separate transmissions of the group secret key. Consequently, the associated delay scales with the number of users, making them impractical for large systems. In our approach, we use some information broadcast, joint with the observation phase in such a way that the subsequent one-way public discussion involves merely a single broadcast for information reconciliation, hence reducing the delay at the expense of some sacrifice in the key rate. We also evaluate the achievable key rate of our scheme.

3 System Model

3.1 Framework Overview

Generating group secret key is essential to ensure secure communication among multiple wireless devices. Previous RSS-based key extraction schemes only work with pairwise devices within communication range of each other. In this framework, we focus on secret key extraction for a group of wireless devices by exploiting the RSS measurements from these devices collaboratively. There are a number of challenges arising from utilizing RSS measurements for group key generation. First, the RSS values obtained between a pair of devices cannot be securely passed to other devices, making it hard to reach key agreement among multiple devices without the availability of a fixed infrastructure. Second, due to the dynamics of mobile devices, the devices within the group that need to establish a secret key may not be within each other's communication range, making the existing RSS-based methods not applicable. To address these challenges, we define a metric called *DOSS* which represents the difference of signal strength measured at a particular wireless device from different radio channels. In our framework, instead of using RSS measurements directly, we utilize the DOSS values to facilitate key extraction.

Our framework consists of two protocols via either *star* and *chain* topologies to facilitate reliable secret key generation among multiple wireless devices. The collaborative key extraction protocol via the star topology is designed for the scenario when a group of wireless devices under consideration are within the communication range of each other. For example, a group of travelers are visiting the same scenic spot. In this case, a device in the group will be randomly picked to serve as the *virtual central node* by passing the DOSS values to other devices to perform key extraction collaboratively. Whereas under the scenario when not all the wireless devices in the group are within the communication range of each other, our collaborative key extraction protocol constructs a virtual chain topology where the devices in the group under consideration are connected with one another like a chain. Each device in the chain involves to pass the corresponding DOSS values to its neighbor device in the next step of the chain. The approach for chain topology may incur accumulated RSS noise across multiple devices. Our theoretic analysis discusses this issue in Section 6. Our framework is generic and can work on any secret key extraction methods using RSS measurements, such as [4] and [7]. Additionally, in Section 8, we also propose a fading trend based secret key extraction method to achieve a lower bit mismatch rate while maintaining the comparable bit generation rate when comparing to existing studies.

3.2 Attack Model

We consider a passive adversary, an *Eavesdropper*, who follows the legitimate mobile devices involving in group key extraction. The eavesdropper's channel gain observation is independent of the channel gain observations of every other legitimate mobile device. It overhears all the public discussion during key generation and can obtain the secret key extraction algorithm and corresponding parameters for key generation. The Eavesdropper is assumed to be located at least $\lambda/2$ away from legitimate devices. Over distances of half a wavelength, wireless channel gains decorrelate in multipath fading environments, hence leading to independent observations at the eavesdropper and the legitimate nodes. This eliminates the possibility of the eavesdropper to extract any information on the gains of the legitimate channels, merely based on the observations over its own channel [29]. However, by accumulating the channel information broadcasted during public discussion phase from multiple wireless devices, the eavesdropper may be able to derive part or all of the group secret key as the number of users increase. To counter that, a subsequent privacy amplification phase should be used by the legitimate users.

4 Relay Node Assisted Collaborative Key Extraction

To achieve group key extraction, one fundamental issue needs to be addressed is when a pair of wireless devices are not within each other's communication range. We propose an approach using relay nodes for key extraction when two wireless devices cannot communicate directly. In particular, we design a collaborative key extraction scheme under the assistance of relay nodes. We assume that all the relay nodes will be authenticated before deployed in our collaborative key extraction scheme. Since there is

no common radio channel that two devices (e.g., Alice and Bob) can measure directly when they are not within each other's communication range, we propose to use the collaborative efforts from one or more relay nodes, who connect between these two devices, to assist in secret key generation between them. However, due to the open nature of wireless medium, any information forwarded by relay nodes will be eavesdropped, which makes it infeasible to pass RSS measurements directly to either Alice or Bob for secret key generation. To solve this problem, we define a metric called *DOSS*, which represents the *difference of signal strength* measured at each relay node from two different radio channels that the relay nodes connected to other devices. Instead of passing the RSS readings, the DOSS values will be passed to other devices to facilitate key extraction. Without obtaining the exact RSS measurements, an adversary cannot regenerate the same secret key between Alice and Bob.

Our relay node assisted collaborative key extraction scheme can work with many existing studies leveraging RSS measurements for secret key extraction. Particularly, Mathur et.al [4] uses single threshold to quantize the RSS measurements for secret key generation. Multi-quantization technique is applied on RSS measurements to extract secret keys in wireless networks [3], [7]. Zhu et.al [30] extracts secret keys by improving the level-crossing technique in a noisy vehicular environment. We also propose a new key extraction method utilizing fading trend, aiming to achieve a lower bit mismatch rate. The details of this method is presented in Section 8.

4.1 Basic Protocol

We use three mobile devices, including Alice, Bob and Ryan, to illustrate the basic idea of the relay node assisted secret key extraction scheme. We denote Ryan's observation on channel (A,R) between Alice and Ryan in slot t with:

$$\hat{Y}^R_{A,R}(t) = Y_{A,R}(t) + W^R_{A,R}(t),$$

where $Y_{A,R}(t)$ is the actual channel gain in slot t and $W^R_{A,R}(t)$ is the observation noise, assumed to be i.i.d. for all A and R.

Step 1: Alice, Bob and Ryan consist of a one-hop network, where Alice and Bob communicate via the relay node, Ryan.

Step 2: Any two neighboring devices among Alice, Bob and Ryan exchange the probe packets for extracting channel measurements. The RSS measured at Ryan from its neighboring devices Alice and Bob are $\hat{Y}^R_{A,R}(t)$ and $\hat{Y}^R_{B,R}(t)$, respectively. Alice and Bob obtain the RSS measurements $\hat{Y}^B_{R,A}(t)$ and $\hat{Y}^B_{R,B}(t)$ from Ryan, respectively.

Step 3: Ryan calculates the DOSS values based on the radio channels it uses to communicate with Alice and Bob, $\delta_R(t) = \hat{Y}^R_{B,R}(t) - \hat{Y}^R_{A,R}(t)$, and then forwards it to Bob.

Step 4: Once the DOSS values from Ryan arrives at Bob, Bob is able to estimate the radio channel between Alice and Ryan: $Y^A_{R,A}(t) = \hat{Y}^A_{R,B}(t) + \delta_R(t)$. Since Alice can directly measure the radio channel between Ryan and Alice: $\hat{Y}^A_{R,A}(t)$, both Alice and Bob have obtained the common channel information of radio channel between Alice and Ryan. Thus, secret keys can be generated secretly between Alice and Bob by using the key extraction algorithm.

One alternative is to utilize all the channel information along the path between Alice and Bob by letting the relay node, Ryan, send the DOSS values to both Alice and Bob. However, we find that the generated key presents the same secrecy as this simple approach, which only uses the channel information between Alice and Ryan. Figure 1 illustrates our proposed protocol by employing 1 relay node, Ryan. The protocol can be easily extended to the case with multiple relay nodes, which is further discussed in Section 6. We note that this protocol is generic to any key extraction algorithms using RSS. In this work, we apply the fading trend based scheme introduced in Section 8 to our group key extraction framework and compare its performance with exiting methods using RSS.

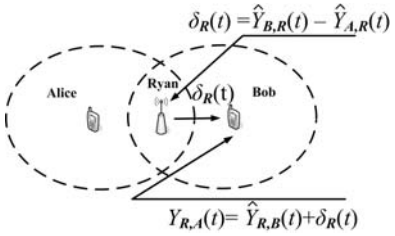


Fig. 1. Illustration of relay node assisted collaborative key extraction.

5 Group Key Extraction via the Star Topology

We examine two typical scenarios in mobile wireless networks when performing group key extraction for multiple devices. The first one is when all wireless devices inside the group under consideration are within each other's communication range, which means any two devices are directly connected. For example, a group of travelers are visiting different places and would like to establish secure communication among themselves. In this scenario, we randomly choose one device as the *virtual central node* and the rest of the devices in the group forms a star topology. The virtual central node facilitates the group key extraction by passing the DOSS values to other nodes and perform key extraction collaboratively. When not all wireless devices within the group under consideration are within each other's communication range, they are interconnected with either group or non-group members. We form the devices within the group to a virtual chain topology, where nodes are sequentially connected. In this section, we focus our attention on presenting the group key extraction protocol via the star topology and defer the discussion on the group key extraction protocol via the chain topology in Section 6.

5.1 Protocol Design

There are four steps in the protocol via the star topology. We assume there are n nodes in the group. Each group member is represented as j, where $j = c, 1, 2, ..., n-1$.

Step 1: First, the group will randomly select a group member, say c, serving as the virtual central node. The secret key will be extracted based on the radio channel between c and another randomly selected device, say node 1. Figure 2 illustrates the topology with the central node and the other members of the group. In each time slot $t, 1 \leq t \leq T$, the group repeats Steps 2-3:

Step 2: Each group member $j, j = 1, \cdots, n-1$ obtains channel measurement $\hat{Y}^j_{c,j}(t)$ by exchanging probe packets with c and quantizes the observation with an accuracy level Δ. The quantizer maps $\hat{Y}^j_{c,j}(t)$ to the closest integer multiple

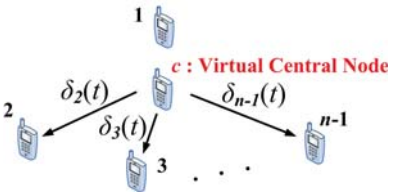


Fig. 2. Illustration of the group key extraction protocol via the star topology.

of Δ to obtain the quantized observation, $\hat{Y}^{\Delta,j}_{c,j}(t)$. In the meanwhile, c obtains the channel gain, $\hat{Y}^{c}_{j,c}(t)$ from all j's.

Step 3: Next, c calculates the DOSS value:

$$\delta_j(t) = \hat{Y}^{c}_{j,c}(t) - \hat{Y}^{c}_{1,c}(t), \tag{1}$$

for $j, j = 2, \cdots, n-1$. Then, it broadcasts the quantized DOSS values, $\delta^{\Delta}_{j}(t)$.

Step 4: Finally, the central node initiates a one-way public discussion and initial key bits are extracted via information reconciliation. All the nodes share the same initial key after information reconciliation, but at the same time some information is revealed to the attacker. Nodes then go through privacy amplification to generate the secret key, which is independent of all observations (including those overheard during public discussion) of the attacker. We elaborate on the details of these two steps in the following section.

The topology and the broadcasts are illustrated in Figure 2.

5.2 Achievable Group Secret Key Rate with the Star Topology

In this section, we analyze the achievable key rate of our scheme under the network via the star topology in the presence of an eavesdropper. We also provide the details of the public discussion and privacy amplification phases in the sequel.

The key generation problem we are considering lies under source-type models for secret-key agreement [31]. Nodes make noisy observations (i.e., the source) of the gains of particular channels in each time slot $t \in \{1, \ldots, T\}$. We denote the observation of node $j \in \{c, 1, \ldots, n-1\}$ of channel (j,i) in slot t with:

$$\hat{Y}^{j}_{i,j}(t) = Y_{i,j}(t) + W^{j}_{i,j}(t),$$

where $W^{j}_{i,j}(t)$ is i.i.d., $Y_{i,j}(t) = Y_{j,i}(t)$ for all $i, j \in \{c, 1, \ldots, n-1\}$, and $Y_{i,j}(t)$ is independent (but not necessarily identically distributed) over different channels, (i,j) and i.i.d. across time t. We consider a favorable scenario for the eavesdropper, in particular, the eavesdropper has noiseless observation of $Y_{i,e}(t)$ for all i's and t. In vector notation, we denote any random sequence $Z(t)$ with boldface $\mathbf{Z} \triangleq [Z(1)\ Z(2)\ \cdots\ Z(T)]$. Also, let the entire sequence of observations of e be matrix $\mathscr{Y}^{e} \triangleq [\mathbf{Y}_{1,e}\ \mathbf{Y}_{2,e}\ \cdots\ \mathbf{Y}_{(n-1),e}]$. The virtual central node c chooses one of the users and uses the observed channel gain as reference to generate the group key. In the following derivations, we take Node 1 as the reference node. With the reference channel gain represented with $Y_{c,1}(t)$, node c calculates [1] $\mathbf{S}_{\text{pub}}(t) \triangleq [\delta_2(t), \delta_3(t), \cdots, \delta_{n-1}(t)]$ and broadcasts the quantized version, $\mathbf{S}^{\Delta}_{\text{pub}}(t) = [\delta^{\Delta}_2(t), \delta^{\Delta}_3(t), \cdots, \delta^{\Delta}_{n-1}(t)]$, overheard by all nodes, including the eavesdropper. We start the analysis by stating the amount of common information between node j and node c before privacy amplification. Here we use notation $\mathscr{S}_{\text{pub}} \triangleq [\mathbf{S}_{\text{pub}}(1)\ \cdots\ \mathbf{S}_{\text{pub}}(T)]$ and the quantized version $\mathscr{S}^{\Delta}_{\text{pub}} \triangleq [\mathbf{S}^{\Delta}_{\text{pub}}(1)\ \cdots\ \mathbf{S}^{\Delta}_{\text{pub}}(T)]$. Next, we derive the maximum achievable group key rate between the nodes as $T \to \infty$ for the star topology.

The objective of the one-way public discussion initiated by c after the observation phase is to achieve information reconciliation. At the end of public discussion, all nodes should obtain $[\hat{\mathbf{Y}}^{\Delta,c}_{1,c}, \hat{\mathbf{Y}}^{\Delta,c}_{2,c}, \cdots, \hat{\mathbf{Y}}^{\Delta,c}_{n-1,c}]$, which will be the common information among the group. This problem can be viewed as an instance of Slepian-Wolf coding ([32], Chapter 14): node c constructs a random binning structure with $2^{TR_{\text{key}}}$ codewords, where $R_{\text{key}}(T) = \frac{1}{T} H\left(\hat{\mathbf{Y}}^{\Delta,c}_{1,c}, \ldots, \hat{\mathbf{Y}}^{\Delta,c}_{n-1,c}\right)$. For the entire group to be able to decode the common information, the number of bins should be no less than $2^{TR_{\text{bin}}(T)}$, where

$$R_{\text{bin}}(T) = \max_{1 \le j \le n-1} \frac{1}{T} H\left(\hat{\mathbf{Y}}^{\Delta,c}_{1,c}, \ldots, \hat{\mathbf{Y}}^{\Delta,c}_{n-1,c} \mid \mathscr{S}^{\Delta}_{\text{pub}}, \hat{\mathbf{Y}}^{\Delta,j}_{c,j}\right).$$

Thus, $R_{\text{bin}}(T)$ corresponds to the rate of information node c needs to provide to the group member with the "worst" observation for the generation of the initial key. With the Slepian-Wolf source encoding with the above random binning structure, the key mismatch probability goes to 0 as $T \to \infty$. Let us define the asymptotic group information rate as:

$$\begin{aligned} R_{\text{star}} &\triangleq \min_{1 \le j \le n-1} \lim_{T \to \infty} \frac{1}{T} I([\hat{\mathbf{Y}}^{\Delta,c}_{1,c}, \ldots, \hat{\mathbf{Y}}^{\Delta,c}_{n-1,c}]; [\mathscr{S}^{\Delta}_{\text{pub}}, \hat{\mathbf{Y}}^{\Delta,j}_{c,j}]) \\ &= \min_{1 \le j \le n-1} \lim_{T \to \infty} \frac{1}{T} I([\hat{\mathbf{Y}}^{c}_{1,c}, \ldots, \hat{\mathbf{Y}}^{c}_{n-1,c}]; [\mathscr{S}_{\text{pub}}, \hat{\mathbf{Y}}^{j}_{c,j}]) \qquad (2) \\ &= \lim_{T \to \infty} [R_{\text{key}}(T) - R_{\text{bin}}(T)], \qquad (3) \end{aligned}$$

where (2) follows from Eq.(9.52) in [32]. With the broadcasts that node c makes and based on its own observations, the eavesdropper has side information $[\mathscr{S}^{\Delta}_{\text{pub}}, \mathscr{Y}^{e}]$ and let

$$\begin{aligned} R_e &\triangleq \lim_{T \to \infty} \frac{1}{T} I([\hat{\mathbf{Y}}^{\Delta,c}_{1,c}, \ldots, \hat{\mathbf{Y}}^{\Delta,c}_{n-1,c}]; [\mathscr{S}^{\Delta}_{\text{pub}}, \mathscr{Y}^{e}]) \\ &= \lim_{T \to \infty} \frac{1}{T} I([\hat{\mathbf{Y}}^{c}_{1,c}, \ldots, \hat{\mathbf{Y}}^{c}_{n-1,c}]; [\mathscr{S}_{\text{pub}}, \mathscr{Y}^{e}]), \end{aligned}$$

where the second equality follows similarly from Eq.(9.52) in [32] and that $\mathscr{Y}^{e}$ is independent of all other channel gains. Now, we can write the achievable secret key rate as:

$$R^{\text{sec}}_{\text{star}} = R_{\text{star}} - R_e, \tag{4}$$

which follows from the corollary subsequent to Theorem 1 in [33], since $[\hat{\mathbf{Y}}^{\Delta,c}_{1,c}, \ldots, \hat{\mathbf{Y}}^{\Delta,c}_{n-1,c}] \leftrightarrow [\mathscr{S}^{\Delta}_{\text{pub}}, \hat{\mathbf{Y}}^{\Delta,j}_{c,j}] \leftrightarrow [\mathscr{S}^{\Delta}_{\text{pub}}, \mathscr{Y}^{e}]$ form a Markov chain for all j, $1 \le j \le n-1$. This rate can be achieved after privacy amplification via universal hashing [34]. Note that, this rate corresponds to the group secret key rate provided in Theorem 1 of [35].

In what follows we will evaluate $R^{\text{sec}}_{\text{star}}$ for the case in which $Y_{ji}(t)$ and $W^{j}_{ij}(t)$ are i.i.d. processes (i.e., symmetric channels) for all channels $i, j \in \{c, 1, \ldots, n-1\}$. Since the observation of each node is identically distributed, R_{star} can be found by analyzing the secret key rate for an arbitrary

1. Note that this is a sequence of vectors the associated set of observations can be transmitted in time sequentially, as the observations are made, rather than all at once.

user other than node 1, e.g., node 2. The secret key rate as given in Eq. (4) can be written after dropping the time indices for simplicity as:

$$R_{\text{star}}^{\text{sec}} = I([\hat{Y}_{1,c}^c,\ldots,\hat{Y}_{(n-1),c}^c];[\mathbf{S}_{\text{pub}},\hat{Y}_{c,2}^2]) - I([\hat{Y}_{1,c}^c,\ldots,\hat{Y}_{(n-1),c}^c];[\mathbf{S}_{\text{pub}},Y_{c,e},Y_{1,e},\ldots,Y_{(n-1),e}]) \tag{5}$$

$$= I([\hat{Y}_{1,c}^c,\ldots,\hat{Y}_{(n-1),c}^c];\hat{Y}_{c,2}^2 \mid \mathbf{S}_{\text{pub}}) + I([\hat{Y}_{1,c}^c,\ldots,\hat{Y}_{(n-1),c}^c];\mathbf{S}_{\text{pub}}) - I([\hat{Y}_{1,c}^c,\ldots,\hat{Y}_{(n-1),c}^c];Y_{c,e},Y_{1,e},\ldots,Y_{(n-1),e} \mid \mathbf{S}_{\text{pub}}) - I([\hat{Y}_{1,c}^c,\ldots,\hat{Y}_{(n-1),c}^c];\mathbf{S}_{\text{pub}}) \tag{6}$$

$$= I([\hat{Y}_{1,c}^c,\ldots,\hat{Y}_{(n-1),c}^c];\hat{Y}_{c,2}^2 \mid \mathbf{S}_{\text{pub}}) \tag{7}$$

$$= I(\hat{Y}_{1,c}^c;\hat{Y}_{c,2}^2 \mid \mathbf{S}_{\text{pub}}) + I([\hat{Y}_{2,c}^c,\ldots,\hat{Y}_{(n-1),c}^c];\hat{Y}_{c,2}^2 \mid \hat{Y}_{1,c}^c,\mathbf{S}_{\text{pub}}) \tag{8}$$

$$= I(\hat{Y}_{1,c}^c;\hat{Y}_{c,2}^2 \mid \mathbf{S}_{\text{pub}}) \tag{9}$$

$$= h(\hat{Y}_{1,c}^c \mid \mathbf{S}_{\text{pub}}) - h(\hat{Y}_{1,c}^c \mid \mathbf{S}_{\text{pub}},\hat{Y}_{c,2}^2)$$
$$= h(\mathbf{S}_{\text{pub}} \mid \hat{Y}_{1,c}^c) + h(\hat{Y}_{1,c}^c) - h(\mathbf{S}_{\text{pub}}) - \left[h(\mathbf{S}_{\text{pub}},\hat{Y}_{c,2}^2 \mid \hat{Y}_{1,c}^c) + h(\hat{Y}_{1,c}^c) - h(\mathbf{S}_{\text{pub}},\hat{Y}_{c,2}^2)\right] \tag{10}$$

$$= h(\hat{Y}_{1,c}^c - \hat{Y}_{2,c}^c,\ldots,\hat{Y}_{1,c}^c - \hat{Y}_{(n-1),c}^c \mid \hat{Y}_{1,c}^c) - h(\hat{Y}_{1,c}^c - \hat{Y}_{2,c}^c,\ldots,\hat{Y}_{1,c}^c - \hat{Y}_{(n-1),c}^c) - h(\hat{Y}_{1,c}^c - \hat{Y}_{2,c}^c,\ldots,\hat{Y}_{1,c}^c - \hat{Y}_{(n-1),c}^c,\hat{Y}_{c,2}^2) + h(\hat{Y}_{1,c}^c - \hat{Y}_{2,c}^c,\ldots,\hat{Y}_{1,c}^c - \hat{Y}_{(n-1),c}^c,\hat{Y}_{c,2}^2)$$
$$= h(-\hat{Y}_{2,c}^c,\ldots,-\hat{Y}_{(n-1),c}^c) - h(\hat{Y}_{1,c}^c - \hat{Y}_{2,c}^c,\ldots,\hat{Y}_{1,c}^c - \hat{Y}_{(n-1),c}^c) - h(-\hat{Y}_{2,c}^c,\ldots,-\hat{Y}_{(n-1),c}^c,\hat{Y}_{c,2}^2) + h(\hat{Y}_{1,c}^c - \hat{Y}_{2,c}^c,\ldots,\hat{Y}_{1,c}^c - \hat{Y}_{(n-1),c}^c,\hat{Y}_{c,2}^2) \tag{11}$$

$$= h(-\hat{Y}_{2,c}^c,\ldots,-\hat{Y}_{(n-1),c}^c) - h(\hat{Y}_{1,c}^c - \hat{Y}_{2,c}^c,\ldots,\hat{Y}_{1,c}^c - \hat{Y}_{(n-1),c}^c) - h(-\hat{Y}_{3,c}^c,\ldots,-\hat{Y}_{(n-1),c}^c) - h(-\hat{Y}_{2,c}^2,\hat{Y}_{2,c}^2) + h(\hat{Y}_{1,c}^c - \hat{Y}_{2,c}^c,\ldots,\hat{Y}_{1,c}^c - \hat{Y}_{(n-1),c}^c,\hat{Y}_{c,2}^2), \tag{12}$$

where Eq. (5) follows by dropping the vector notation, since observations are i.i.d, Eq.(6) follows by the application of chain rule on both terms of the right side of Eq.(5), Eq.(7) follows since $\hat{Y}_{j,c}^c$ is independent of $(Y_{c,e},Y_{1,e},\ldots,Y_{(n-1),e})$ for all j's, Eq.(8) follows by chain rule, and Eq.(9) follows since, given $\hat{Y}_{1,c}^c$ and $\mathbf{S}_{\text{pub}}$, one can determine $\hat{Y}_{j,c}^c$ for all j's with probability 1, Eq.(10) follows from the chain rule for entropies, Eq.(11) follows since, given $\hat{Y}_{1,c}^c$, all the uncertainty in $\hat{Y}_{1,c}^c - \hat{Y}_{j,c}^c$ is in $\hat{Y}_{j,c}^c$ and that $\hat{Y}_{1,c}^c$ and $\hat{Y}_{2,c}^2$ are independent, and Eq.(12) is by the chain rule.

Now, we evaluate the key rate for the scenario in which the channels are i.i.d. Rayleigh fading. Thus, $Y_{j,i}(t)$ is 0-mean circularly symmetric complex Gaussian with an identical variance σ_Y^2 per dimension for all channels $i,j \in \{c,e,1,\ldots,n-1\}$. Also, in our evaluations, we assume $W_{i,j}^j(t)$ to be i.i.d., 0-mean circularly symmetric complex Gaussian with a variance σ_W^2 per dimension for all $i,j \in \{c,1,\ldots,n-1\}$. Finally, let $\gamma_m \triangleq \frac{\sigma_Y^2}{\sigma_W^2}$ be the measurement SNR, where m is the number of quantization levels. With these assumptions, all five differential entropies in Eq. (12) are those of Gaussian random vectors. In particular, let $\mathbf{1}_{n\times n}$ and $\mathbf{I}_{n\times n}$ denote respectively, the matrix of all 1's and the identity matrix of size $n\times n$. Then, we can write the following for these vectors in Eq. (12).

1. The vector consisting of the negative RSS measurements between the virtual central node c and node $i, i = 2,\cdots,n-1$, i.e., $\left[-\hat{Y}_{2,c}^c,\ldots,-\hat{Y}_{(n-1),c}^c\right]$ is following the distribution:

$$\left[-\hat{Y}_{2,c}^c,\ldots,-\hat{Y}_{(n-1),c}^c\right] \sim \mathscr{N}\left(\mathbf{0},\sigma_Y^2\left(1+\gamma_m^{-1}\right)\mathbf{I}_{(n-2)\times(n-2)}\right) \tag{13}$$

2. The vector consisting of the DOSS value between the node 1 and node $i, i = 2,\cdots,n-1$, $\left[\hat{Y}_{1,c}^c - \hat{Y}_{2,c}^c,\ldots,\hat{Y}_{1,c}^c - \hat{Y}_{(n-1),c}^c\right]$ follows the distribution as:

$$\left[\hat{Y}_{1,c}^c - \hat{Y}_{2,c}^c,\ldots,\hat{Y}_{1,c}^c - \hat{Y}_{(n-1),c}^c\right] \sim \mathscr{N}\left(\mathbf{0},\sigma_Y^2\left(1+\gamma_m^{-1}\right)\left(\mathbf{I}_{(n-2)\times(n-2)} + \mathbf{1}_{(n-2)\times(n-2)}\right)\right) \tag{14}$$

3. We have the distribution for the vector consisting of the negative RSS measurements between the virtual central node c and node $i, i = 3,\cdots,n-1$, $\left[-\hat{Y}_{3,c}^c,\ldots,-\hat{Y}_{(n-1),c}^c\right]$, shown as:

$$\left[-\hat{Y}_{3,c}^c,\ldots,-\hat{Y}_{(n-1),c}^c\right] \sim \mathscr{N}\left(\mathbf{0},\sigma_Y^2\left(1+\gamma_m^{-1}\right)\mathbf{I}_{(n-3)\times(n-3)}\right) \tag{15}$$

4. The distribution for $\left[-\hat{Y}_{2,c}^c,\hat{Y}_{c,2}^2\right]$ is shown as:

$$\left[-\hat{Y}_{2,c}^c,\hat{Y}_{c,2}^2\right] \sim \mathscr{N}\left(\mathbf{0},\sigma_Y^2\begin{bmatrix} 1+\gamma_m^{-1} & -1 \\ -1 & 1+\gamma_m^{-1}\end{bmatrix}\right) \tag{16}$$

5. For $\left[\hat{Y}_{1,c}^c - \hat{Y}_{2,c}^c,\ldots,\hat{Y}_{1,c}^c - \hat{Y}_{(n-1),c}^c,\hat{Y}_{c,2}^2\right]$, its distribution follows:

$$\left[\hat{Y}_{1,c}^c - \hat{Y}_{2,c}^c,\ldots,\hat{Y}_{1,c}^c - \hat{Y}_{(n-1),c}^c,\hat{Y}_{c,2}^2\right] \sim \mathscr{N}(\mathbf{0},K_{\mathbf{Z}}) \tag{17}$$

where the entries of covariance matrix $K_{\mathbf{Z}}$ in the top left portion with coordinates $[1,n-2]\times[1,n-2]$ are identical to $\sigma_Y^2\left(1+\gamma_m^{-1}\right)\left(\mathbf{I}_{(n-2)\times(n-2)} + \mathbf{1}_{(n-2)\times(n-2)}\right)$; the entries in the $(n-1)$st column and the $(n-1)$st row are all 0, except for $K_{\mathbf{Z}}(2,n-1) = K_{\mathbf{Z}}(n-1,2) = -\sigma_Y^2$ and $K_{\mathbf{Z}}(n-1,n-1) = \sigma_Y^2\left(1+\gamma_m^{-1}\right)$.

Noting that $\det(\mathbf{1}_{n\times n} + \mathbf{I}_{n\times n}) = n+1$ and $\det(K_{\mathbf{Z}})$ can be computed[2] and equals to $\sigma_Y^{2(n-1)}\left[n-1-\frac{n-2}{\left(1+\gamma_m^{-1}\right)^2}\right]$. Thus, we can obtain the achievable group key rate in star topology as:

$$\begin{aligned} R_{\text{star}}^{\text{sec}} &= \log\left[\left(2\pi e\sigma_Y^2\left(1+\gamma_m^{-1}\right)\right)^{n-2}\right] \\ &\quad - \log\left[\left(2\pi e\sigma_Y^2\left(1+\gamma_m^{-1}\right)\right)^{n-2}\cdot(n-1)\right] \\ &\quad - \log\left[\left(2\pi e\sigma_Y^2\left(1+\gamma_m^{-1}\right)\right)^{n-3}\right] \\ &\quad - \log\left[\left(2\pi e\sigma_Y^2\right)^2\left[\left(1+\gamma_m^{-1}\right)^2-1\right]\right] \\ &\quad + \log\left[\left(2\pi e\sigma_Y^2\left(1+\gamma_m^{-1}\right)\right)^{n-1}\left(n-1-\frac{n-2}{\left(1+\gamma_m^{-1}\right)^2}\right)\right] \\ &= \log\left(1+\frac{1/(n-1)}{\left(1+\gamma_m^{-1}\right)^2-1}\right). \end{aligned}$$

From the equation above, we observe that the achievable group key rate $R_{\text{star}}^{\text{sec}}$ in star topology only depends on the group size n and SNR γ_m on wireless channel. As the group size increases, the achievable group key rate decreases, indicating that a larger group is more vulnerable to eavesdropper attacks. This is because more channel statistical information, embedded in the DOSS information, is exposed to the attacker as node c broadcasts all DOSS values in each block. Furthermore, if SNR on radio channel increases, the achievable group key rate becomes higher. Higher SNR would result in less ambiguity on group key extraction, which benefits the key agreement among group members and is independent from the potential key extraction of the attacker.

Discussion: Secret key generation algorithms typically go through a subsequent public discussion and privacy amplification phases at the end of the entire observation phase

2. $\det(K_{\mathbf{Z}})$ can be found by (1) multiplying the last row and the last column by -1, (2) shifting the rows by 1 so that the jth row becomes $(j+1)$st row and $(n-1)$st row becomes the first row and likewise for the columns, (3) a cofactor expansion.

(e.g., see [36]). While maximizing the length of the secret key generated, this leads to a high delay, since the amount of exchange for reconciliation increases with the length of the observation. For instance, in [35], simple schemes are proposed for group key generation: First, the center node generates pairwise secret keys, separately with each node. Then, it broadcasts one of the keys (shortest one) xor'ed with the key associated with each user, so each user can reconstruct the shortest key. This way, the group secret key capacity derived in [35] is achieved. However, these capacity achieving schemes involve a separate information reconciliation phase with each user as well as a separate transmission (of the secret key). Consequently, the delay subsequent to the observation phase scales with the number of nodes.

On the other hand, our scheme broadcasts some information during the observation phase ($\mathbf{S}_{\text{pub}}(t)$ in each time slot t) in an on-line manner[3] The subsequent one-way public discussion involves a single broadcast for information reconciliation, as opposed to a separate broadcast for each pairwise key as in the secret key capacity achieving scheme. Clearly, this reduction of the delay and the public discussion overhead is significant in practice. Note however that, this comes at the expense of some sacrifice in the key rate.

6 Group Key Extraction via the Chain Topology

Under the scenarios when not all the wireless devices in the group are within the communication range of each other, our collaborative key extraction protocol via the chain topology constructs a virtual topology where the devices in the group under consideration are connected with one another like a chain as depicted in Figure 3. Each device in the chain involves to pass the corresponding DOSS values to its neighbor device in the next step of the chain. We note that the virtual chain topology is a special case of the tree topology, i.e., hybrid of star and chain topologies, and represents the worst case scenario in terms of accumulated noise during group key extraction using RSS.

6.1 Protocol Design

We assume there are n wireless devices in the group. A chain topology is formed with c and $n-1$ as the head and tail node respectively, and the radio channel between c and 1 is chosen as the channel for secret bit extraction for all the members. There are four steps in group key extraction protocol via the chain topology.

Step 1: First, the group will select a pair, say c and 1, which are within the range of each other. The reference radio channel to be used by all group members is the one between c and 1. In each time slot $t, 1 \le t \le T$, the group repeats Steps 2-3:

Step 2: Each group member observes the channel gain between its neighboring nodes. Each group member $j \in \{1,\ldots,n-2\}$ has two neighbors and collects two channel gain measurements $\hat{Y}^{j}_{j-1,j}$ and $\hat{Y}^{j}_{j+1,j}$. Nodes c and $n-1$ have only one neighbor, $\hat{Y}^{c}_{1,c}(t)$ and $\hat{Y}^{n-1}_{n-2,n-1}(t)$ respectively. Then, as in the star scenario, the observations are quantized with an accuracy level Δ. The quantized observations at node j are represented with, $\hat{Y}^{\Delta,j}_{\cdot,j}(t)$.

3. This can be fully integrated with the observation phase, as the nodes can use the signal broadcast by node c to observe the channel gains, as opposed to using separate beacons.

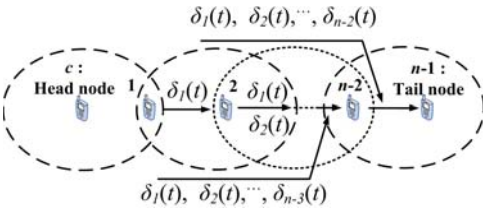


Fig. 3. Illustration of chain-based group key extraction protocol.

Step 3: The DOSS value:

$$\delta_j(t) = \hat{Y}^{j}_{j+1,j}(t) - \hat{Y}^{j}_{j-1,j}(t) \tag{18}$$

is measured by $j, j = 1,\cdots,n-2$. Then $\delta_j(t)$ is forwarded by traversing j's subsequent nodes on the chain until it reaches $n-1$ as shown in Figure 3. Nodes keep the quantized version, $\delta^{\Delta}_j(t)$ of RSS values.

Step 4: Finally, similar to the star scenario, the head node initiates a one-way public discussion and initial key bits are extracted via information reconciliation. Subsequently, nodes go through privacy amplification to generate the secret key. We elaborate on the details of these two steps in the following section.

Note that, another possible solution is to utilize all the channel gains in the chain to generate the key, rather than the gain between c and 1. However, we find that the secret key rate does not change in this alternate approach.

6.2 Achievable Group Secret Key Rate with the Chain Topology

With the chain topology, all nodes are ordered from the center node, c to the tail node, $n-1$ toward the tail of the chain. In each time slot, each node observes the gain of its channel to its immediate neighbors. The model for the observations between any pair of nodes is identical to the model given under the star topology. Here, we use the channel gain, $Y_{c,1}(t)$, of the first hop to generate the group key. The eavesdropper is assumed to have noiseless observation of $Y_{j,e}(t)$ for all j's and t. At the end of each slot t, node 1 broadcasts quantized difference $\delta^{\Delta}_1(t)$, where $\delta^{\Delta}_1(t) = \hat{Y}^{1}_{2,1}(t) - \hat{Y}^{1}_{c,1}(t)$. Then node 2 relays this information as well as $\delta^{\Delta}_2(t)$ to node 3, where $\delta^{\Delta}_2(t) = \hat{Y}^{2}_{3,2}(t) - \hat{Y}^{2}_{1,2}(t)$. Each node j along the chain broadcasts all the information it receives, as well as $\delta^{\Delta}_j(t)$. The exchange at time t ends when node $n-1$ receives $\mathbf{S}^{\Delta}_{\text{pub}}(t) = [\delta^{\Delta}_1(t), \delta^{\Delta}_2(t), \ldots, \delta^{\Delta}_{n-2}(t)]$. Similar to the scenario with the star topology, we use notation $\mathscr{S}_{\text{pub}} = [\mathbf{S}_{\text{pub}}(1) \ \cdots \ \mathbf{S}_{\text{pub}}(T)]$ and the quantized version $\mathscr{S}^{\Delta}_{\text{pub}} = [\mathbf{S}^{\Delta}_{\text{pub}}(1) \ \cdots \ \mathbf{S}^{\Delta}_{\text{pub}}(T)]$. Note that the subsequent steps are parallel to the derivation we have for the star topology.

Since $Y_{c,1}(t)$ is used to generate the group key, here $R_{\text{key}}(T) = \frac{1}{T} H\left(\hat{\mathbf{Y}}^{\Delta,c}_{1,c}\right)$. The number of bins used for a Slepian-Wolf code needs to be $2^{TR_{\text{pub}}(T)}$, where

$$R_{\text{pub}}(T) = \frac{1}{T} H\left(\hat{\mathbf{Y}}^{\Delta,c}_{1,c} \mid \mathscr{S}^{\Delta}_{\text{pub}}, \hat{\mathbf{Y}}^{\Delta,n-1}_{(n-2),(n-1)}\right).$$

One can realize that the above rate is the rate that node $n-1$ requires to reconstruct $\hat{\mathbf{Y}}^{\Delta,c}_{1,c}$. This is also the group information rate, since the worst node happens to be node $n-1$, which is at the farthest point of the network from node c. With the Slepian-Wolf source encoding with the

above random binning structure, the key mismatch probability goes to 0 as $T \to \infty$. Similar to (2),(3), derived for the star topology, the asymptotic group information rate is:

$$\begin{aligned} R_{\text{chain}} &\triangleq \lim_{T\to\infty} \frac{1}{T} I(\hat{\mathbf{Y}}_{1,c}^{\Delta,c}; \mathscr{S}_{\text{pub}}^{\Delta}, \hat{\mathbf{Y}}_{(n-2),(n-1)}^{\Delta,n-1}) \\ &= \lim_{T\to\infty} \frac{1}{T} I(\hat{\mathbf{Y}}_{1,c}^{c}; \mathscr{S}_{\text{pub}}, \hat{\mathbf{Y}}_{(n-2),(n-1)}^{n-1}) \qquad (19) \\ &= \lim_{T\to\infty} [R_{\text{key}}(T) - R_{\text{bin}}(T)], \qquad (20) \end{aligned}$$

From the broadcasts during the observation phase and based on its own observations, the eavesdropper has side information $[\mathscr{S}_{\text{pub}}, \mathscr{Y}^e]$ and let

$$R_e \triangleq \lim_{T\to\infty} \frac{1}{T} I(\hat{\mathbf{Y}}_{1,c}^{\Delta,c}; \mathscr{S}_{\text{pub}}^{\Delta}, \mathscr{Y}^e) = \lim_{T\to\infty} \frac{1}{T} I(\hat{\mathbf{Y}}_{1,c}^{c}; \mathscr{S}_{\text{pub}}, \mathscr{Y}^e),$$

where the second equality follows similarly from Eq.(9.52) in [32]. Now, we can write the achievable secret key rate as:

$$R_{\text{chain}}^{\text{sec}} = R_{\text{chain}} - R_e, \qquad (21)$$

identical to the analysis for the star topology. Likewise, for i.i.d. $Y_{ji}(t)$ and $W_{ij}^{j}(t)$, the secret key rate as given in Eq. (21) can be simply written after dropping the time indices as:

$$\begin{aligned} R_{\text{chain}}^{\text{sec}} =& I(\hat{Y}_{1,c}^{c}; \mathbf{S}_{\text{pub}}, \hat{Y}_{(n-2),(n-1)}^{n-1}) \qquad (22) \\ & - I(\hat{Y}_{1,c}^{c}; \mathbf{S}_{\text{pub}}, Y_{c,e}, Y_{1,e}, \ldots, Y_{(n-1),e}) \\ =& I(\hat{Y}_{1,c}^{c}; \hat{Y}_{(n-2),(n-1)}^{n-1} \mid \mathbf{S}_{\text{pub}}) + I(\hat{Y}_{1,c}^{c}; \mathbf{S}_{\text{pub}}) \\ & - I(\hat{Y}_{1,c}^{c}; Y_{c,e}, Y_{1,e}, \ldots, Y_{(n-1),e} \mid \mathbf{S}_{\text{pub}}) - I(\hat{Y}_{1,c}^{c}; \mathbf{S}_{\text{pub}}) \qquad (23) \\ =& I(\hat{Y}_{1,c}^{c}; \hat{Y}_{(n-2),(n-1)}^{n-1}) \mid \mathbf{S}_{\text{pub}}) \qquad (24) \\ =& h(\hat{Y}_{1,c}^{c} \mid \mathbf{S}_{\text{pub}}) - h(\hat{Y}_{1,c}^{c} \mid \mathbf{S}_{\text{pub}}, \hat{Y}_{(n-2),(n-1)}^{n-1}) \\ =& -h(\mathbf{S}_{\text{pub}}) + h(\hat{Y}_{1,c}^{c}, \mathbf{S}_{\text{pub}}) + h(\mathbf{S}_{\text{pub}}, \hat{Y}_{(n-2),(n-1)}^{n-1}) \qquad (25) \\ & - h(\hat{Y}_{1,c}^{c}, \mathbf{S}_{\text{pub}}, \hat{Y}_{(n-2),(n-1)}^{n-1}) \end{aligned}$$

where (22) follows by dropping the vector notation, since observations are i.i.d, (23) follows by the application of chain rule on both terms of the right side of (22), (24) follows since $\hat{Y}_{1,c}^{c}$ is independent of $[Y_{c,e}, Y_{1,e}, \ldots, Y_{(n-1),e}]$, and (25) follows from the chain rule for entropies. Identical to the star topology, now, we evaluate the key rate for the scenario in which the channels are i.i.d. Rayleigh fading with the same parameters given in Section 5.2. Hence, all four differential entropies in Eq. (25) are those of Gaussian random vectors as characterized in what follows:

1. For the DOSS vector between any two neighboring user $i = 1, \cdots, n-2$, $\mathbf{S}_{\text{pub}}$, we have:

$$\mathbf{S}_{\text{pub}} = \left[\hat{Y}_{2,1}^{1} - \hat{Y}_{c,1}^{1},\ \hat{Y}_{3,2}^{2} - \hat{Y}_{1,2}^{2}, \ldots, \hat{Y}_{(n-1),(n-2)}^{n-2} - \hat{Y}_{(n-3),(n-2)}^{n-2}\right] \sim \mathscr{N}\left(\mathbf{0}, K_{\mathbf{S}_{\text{pub}}}\right) \qquad (26)$$

where $K_{\mathbf{S}_{\text{pub}}}$ is an $(n-2)\times(n-2)$ covariance matrix with diagonal entries. Particularly, for all $j \in \{1, \ldots, n-3\}$:

$$\begin{aligned} K_{\mathbf{S}_{\text{pub}}}(j,j) &= 2\sigma_Y^2(1+\gamma^{-1}) \\ K_{\mathbf{S}_{\text{pub}}}(j,j+1) &= K_{\mathbf{S}_{\text{pub}}}(j+1,j) = -\sigma_Y^2 \end{aligned} \qquad (27)$$

and all other entries of $K_{\mathbf{S}_{\text{pub}}}$ are identical to 0.

2. For $\left[\hat{Y}_{1,c}^{c}, \mathbf{S}_{\text{pub}}\right]$, we have:

$$[\hat{Y}_{1,c}, \mathbf{S}_{\text{pub}}] \sim \mathscr{N}\left(\mathbf{0}, K_{\hat{Y}_{1,c}, \mathbf{S}_{\text{pub}}}\right) \qquad (28)$$

where $K_{\hat{Y}_{1,c}^{c}, \mathbf{S}_{\text{pub}}}$ is an $(n-1)\times(n-1)$ covariance matrix, where the bottom right portion with entries $[2, n-1]\times[2, n-1]$ is identical to $K_{\mathbf{S}_{\text{pub}}}$, and the first row and the first column are as follows:

$$\begin{aligned} K_{\hat{Y}_{1,c}^{c}, \mathbf{S}_{\text{pub}}}(1,1) &= \sigma_Y^2(1+\gamma^{-1}) \\ K_{\hat{Y}_{1,c}^{c}, \mathbf{S}_{\text{pub}}}(2,1) &= K_{\hat{Y}_{1,c}^{c}, \mathbf{S}_{\text{pub}}}(1,2) = -\sigma_Y^2 \end{aligned} \qquad (29)$$

and the rest of the first column and the first row are all 0's.

3. For $\left[\mathbf{S}_{\text{pub}}, \hat{Y}_{(n-2),(n-1)}^{n-1}\right]$, we have:

$$\left[\mathbf{S}_{\text{pub}}, \hat{Y}_{(n-2),(n-1)}^{n-1}\right] \sim \mathscr{N}\left(\mathbf{0}, K_{\mathbf{S}_{\text{pub}}, \hat{Y}_{(n-2),(n-1)}^{n-1}}\right) \qquad (30)$$

where $K_{\mathbf{S}_{\text{pub}}, \hat{Y}_{(n-2),(n-1)}^{n-1}}$ is an $(n-1)\times(n-1)$ covariance matrix with the top left portion of entries $[1, n-2]\times[1, n-2]$ is identical to $K_{\mathbf{S}_{\text{pub}}}$, and the last row and the last column are as follows:

$$\begin{aligned} K_{\mathbf{S}_{\text{pub}}, \hat{Y}_{(n-2),(n-1)}^{n-1}}(n-1, n-1) &= \sigma_Y^2(1+\gamma^{-1}) \\ K_{\mathbf{S}_{\text{pub}}, \hat{Y}_{(n-2),(n-1)}^{n-1}}(n-2, n-1) &= K_{\mathbf{S}_{\text{pub}}, \hat{Y}_{(n-2),(n-1)}^{n-1}}(n-1, n-2) = \sigma_Y^2 \end{aligned} \qquad (31)$$

and the rest of the last column and the last row are all 0's.

4. For $\left[\hat{Y}_{1,c}^{c}, \mathbf{S}_{\text{pub}}, \hat{Y}_{(n-2),(n-1)}^{n-1}\right]$, we have:

$$\left[\hat{Y}_{1,c}^{c}, \mathbf{S}_{\text{pub}}, \hat{Y}_{(n-2),(n-1)}^{n-1}\right] \sim \mathscr{N}\left(\mathbf{0}, K_{\hat{Y}_{1,c}^{c}, \mathbf{S}_{\text{pub}}, \hat{Y}_{(n-2),(n-1)}^{n-1}}\right) \qquad (32)$$

where $K_{\hat{Y}_{1,c}^{c}, \mathbf{S}_{\text{pub}}, \hat{Y}_{(n-2),(n-1)}^{n-1}}$ is an $n\times n$ covariance matrix with the top left portion of entries $[1, n-1]\times[1, n-1]$ is identical to $K_{\hat{Y}_{1,c}^{c}, \mathbf{S}_{\text{pub}}}$,and the last row and the last column are as follows:

$$\begin{aligned} K_{\hat{Y}_{1,c}^{c}, \mathbf{S}_{\text{pub}}, \hat{Y}_{(n-2),(n-1)}^{n-1}}(n, n) &= \sigma_Y^2(1+\gamma^{-1}) \\ K_{\hat{Y}_{1,c}^{c}, \mathbf{S}_{\text{pub}}, \hat{Y}_{(n-2),(n-1)}^{n-1}}(n-1, n) &= K_{\hat{Y}_{1,c}^{c}, \mathbf{S}_{\text{pub}}, \hat{Y}_{(n-2),(n-1)}^{n-1}}(n, n-1) = \sigma_Y^2 \end{aligned} \qquad (33)$$

and the rest of the last column and the last row are all 0's.

The determinants of the above covariance matrices can be calculated recursively as follows. Firstly, let us define $d_n^{(1)} \triangleq \det\left(\frac{1}{\sigma_Y^2} K_{\mathbf{S}_{\text{pub}}}\right)$ for the n users. One can observe from the cofactor expansion of $\frac{1}{\sigma_Y^2} K_{\mathbf{S}_{\text{pub}}}$ that,

$$d_n^{(1)} = 2(1+\gamma_m^{-1}) d_{n-1}^{(1)} - d_{n-2}^{(1)} \qquad (34)$$

With the initial conditions $d_2^{(1)} = 4(1+\gamma_m^{-1})^2 - 1$ and $d_1^{(1)} = 2(1+\gamma_m^{-1})$, we can evaluate $\det\left(K_{\mathbf{S}_{\text{pub}}}\right) = \sigma_Y^{2(n-2)} d_n^{(1)}$ for any given $n > 2$, recursively.

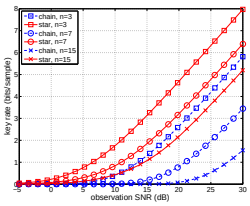
Similarly, let us define $d_n^{(2)} \triangleq \det\left(\frac{1}{\sigma_Y^2} K_{\hat{Y}_{1,c}^{c}, \mathbf{S}_{\text{pub}}}\right)$ (see Item 2) and expand $\frac{1}{\sigma_Y^2} K_{\hat{Y}_{1,c}^{c}, \mathbf{S}_{\text{pub}}}$ via a cofactor expansion, we can find the recursive relation

$$d_n^{(2)} = (1+\gamma_m^{-1}) d_n^{(1)} - d_{n-1}^{(1)} \qquad (35)$$

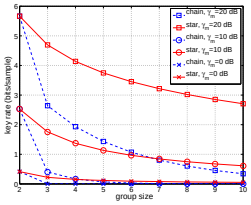
It is not difficult to see for $d_n^{(3)} \triangleq \det\left(\frac{1}{\sigma_Y^2} K_{\mathbf{S}_{\text{pub}}, \hat{Y}_{(n-2),(n-1)}^{n-1}}\right)$ that $d_n^{(3)} = d_n^{(2)}$ (by multiplying the final row and column of $d_n^{(3)}$ by -1, we can obtain a symmetric version of $d_n^{(2)}$).

Lastly, using a similar cofactor expansion and utilizing the above observations, one can also deduce that, for $d_n^{(4)} \triangleq \det\left(\frac{1}{\sigma_Y^2} K_{\hat{Y}_{1,c}^{c}, \mathbf{S}_{\text{pub}}, \hat{Y}_{(n-2),(n-1)}^{n-1}}\right)$, the following recursive relationship holds:

$$d_n^{(4)} = (1+\gamma_m^{-1}) d_n^{(2)} - d_{n-1}^{(2)} \qquad (36)$$



(a) Group key rate vs. observation SNR



(b) Group key rate vs. group size

Fig. 4. Analytical results of the achievable group key rate vs. observation SNR and the group size.

Thus, all the determinants can be calculated recursively. Combining all of the above, we derive the achievable group key generation rate as:

$$\begin{aligned} R_{\text{chain}}^{\text{sec}} =& 2\log\left[\left(2\pi e\sigma_Y^2\right)^{n-1}\left((1+\gamma_m^{-1})d_n^{(1)}-d_{n-1}^{(1)}\right)\right] \\ &-\log\left[\left(2\pi e\sigma_Y^2\right)^{n-2}d_n^{(1)}\right] \\ &-\log\left[\left(2\pi e\sigma_Y^2\right)^{n}\left((1+\gamma_m^{-1})d_n^{(2)}-d_{n-1}^{(2)}\right)\right] \\ =&\log\left(\frac{\left[(1+\gamma_m^{-1})d_n^{(1)}-d_{n-1}^{(1)}\right]^2}{d_n^{(1)}\left[(1+\gamma_m^{-1})d_n^{(2)}-d_{n-1}^{(2)}\right]}\right), \end{aligned} \tag{37}$$

Examining the $R_{\text{chain}}^{\text{sec}}$, similar trend on the achievable group key rate via the chain topology can be observed as in the star topology when varying the group size n and the SNR γ_m. In addition, when the DOSS value propagates along the chain topology, the noisy measurements will be accumulated. Thus, the observation of SNR keeps decreasing as the group size increases. Whereas group key extraction via the star topology does not involve such noise accumulation, since it has the maximum hop size equal to one. Therefore, the decreasing SNR makes the achievable group key rate in the chain topology decrease much faster than that in the star topology. We illustrate this in the next section.

7 Discussion on Group Key Extraction via the Hybrid Topology

In Fig. 4, we present the analytical results for the achievable group key rates, $R_{\text{star}}^{\text{sec}}$ and $R_{\text{chain}}^{\text{sec}}$ as a function of the observation SNR, γ_r, and the group size, n. As expected, the group key rate decreases with the group size and increases with the observation SNR[4]. The important observation is that, the key rate decreases much faster with the chain topology as shown in Figure 4. This is due to the fact that the observation SNR keeps decreasing as the chain size increases. Since the noise accumulates each time the channel gain difference is passed on over the relay. Indeed, the use of chain topology leads to a power penalty between 7-12 dB, compared to using the star topology, as the number of users vary between 3-15. This implies that, given a network, users should form as large a star topology as possible. However, as the network size grows, chains are become necessary to connect far-away users, since the observation SNR decreases significantly with the increased distances, due to path losses.

With the decreased observation SNR, the performance of pure star topology degrades as illustrated in Fig. 4. Thus, in an extended network, one should use a hybrid topology

4. Note that, observation SNR can be increased by using a higher-powered pilot signals to measure the channel gains.

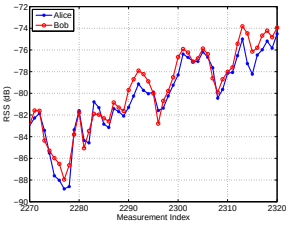


Fig. 5. Segments of RSS measurements from a pair of mobile devices in park.

in which nodes in a close vicinity connect to form star topologies and such clusters are connected to each other via chains. Using our results, one can find the correct balance between the size of the stars and beyond what distances to start forming chains. When a portion of the group members are isolated from the rest of the group, non-group device members could be employed to connect the sub-groups. Intra-group communication should form a hybrid topology following the guidance as we discussed and inter-group communication uses the relay node assisted collaborative key extraction. We leave this analysis as a future study.

8 Fading Trend Based Secret Key Extraction

Our proposed collaborative secret key extraction can work with any secret key extraction method leveraging RSS measurements. To cope with the high bit mismatch rate in previous studies while maintaining a similar key generation rate, we propose a fading trend based secret key extraction algorithm that helps to better capture the similarity presented by channel reciprocity as opposed to using the RSS measurements directly. Using RSS measurements directly during key extraction may suffer various noises in real mobile environments and leads to a higher bit mismatch rate. We take the view point that there should be similar fading trend presented in the RSS measurements between a pair of wireless devices according to the channel reciprocity.

8.1 Algorithm

Given the RSS measurements from the same radio channel, the RSS readings measured by a pair of wireless devices, e.g., Alice and Bob, within the coherence time should be identical based on the principle of wireless channel reciprocity. In practice, there will be mismatch due to the half-duplex operating mode of standard transceivers (e.g., one device cannot send and receive packets at the same time) and the measurement errors. However, we find that the fading exhibited in RSS measurements over time for a pair of mobile devices follows similar increasing or decreasing trend despite of the mismatch of absolute values, as shown in Figure 5. This observation inspires us to utilize the fading trend to reduce the secret bit mismatch rate when extracting secret bits from RSS measurements.

The proposed fading trend based secret key extraction algorithm includes three components: *interpolation*, *fading trend estimation* and *thresholding*. Two variants are proposed in the thresholding step: basic RSS fading trend and median thresholding (**RTM**) and extended RSS fading trend and quantization (**RTQ**). The algorithm flow is displayed in Algorithm 1.

We use the following standard notations: (a) "$\wedge$" is the AND operation; (b) "$\vee$" stands for OR operation; (c) $\hat{Y}(t)$

denotes RSS measurement extracted from the probe packet at time t.

Interpolation: Due to the half-duplex operating mode of standard transceivers, the probe packet transmitted by Alice and Bob has a short delay, which results in the channel measurements asymmetry. And this becomes one of the sources causing RSS reading mismatch. To address this issue, we use the cubic Farrow filter based interpolation technique on top of the measurement RSS readings so that Alice and Bob are able to estimate the RSS measurements at common time instants [37].

Fading trend estimation: The objective of this step is to extract one secret bit at RSS each measurement that exhibits *fading trend*. To determine the fading trend on one particular RSS measurement $\hat{Y}(t)$, we examine its previous sample $\hat{Y}(t-1)$ and the second following sample $\hat{Y}(t+2)$. Here we define $\Phi^1 = \hat{Y}(t) - \hat{Y}(t-1)$ and $\Phi^2 = \hat{Y}(t+2) - \hat{Y}(t)$. If the set of RSS measurements $\{\hat{Y}(t-1), \hat{Y}(t), \hat{Y}(t+2)\}$ consist of a monotone sequence, i.e., Φ^1 and Φ^2 has the same positive or negative relationship, a fading trend is determined. Using this approach, for the fading trend estimation at each measurement, there is only one overlapped RSS sample. Thus, the possible correlation caused by fading trend estimation is minimized. The secret bit, $b_t(1)$, encoded at $\hat{Y}(t)$ is determined as 1 or 0 which corresponds to increasing or decreasing fading trend, as computed in equation 39 displayed in Algorithm 1.

Thresholding: Two variants of secret bits extraction are proposed at this step.

RTM: This basic version of our proposed scheme uses the median value of all RSS measurements, θ, as the single threshold to extract another secret bit for each RSS measurement. The bit, $b_t(2)$ is encoded as 1 or 0 depending on whether $\hat{Y}(t)$ is larger than θ or not, as described in equation 40 of Algorithm 1.

RTQ: The extended version of our key generation scheme extracts multiple bits per RSS measurement in addition to the trend based quantization at the previous step. Instead of using single threshold, we are inspired by the idea of quantization in signal processing to extract secret bits via multiple thresholds. In order to extract $m-1$ bits per measurement, the RSS measurements $\hat{Y}(t)$ is quantized into $2^{(m-1)}$ equally-likely levels. Let $F(\hat{Y}(t))$ be the cumulative distribution function of $\hat{Y}(t)$. The thresholds used for extracting secret bits are determined by the inverse of $F(\hat{Y}(t))$,

$$\rho_k = F^{-1}(\frac{k}{2^w}), k = 1, \cdots, 2^{m-1} - 1 \tag{38}$$

In addition, $\rho_0 = min(\hat{Y}(t))$ and $\rho_{2^{m-1}} = max(\hat{Y}(t))$. When $\hat{Y}(t)$ falls between any neighboring thresholds, Gray coding [38] are employed for extracting $m-1$ bits, $b_t(i), i = 2, \cdots, m$, from $\hat{Y}(t)$.

By examining through the measurements, all the RSS readings exhibiting the fading trend can be found. Alice and Bob will exchange their own set of index that includes all the measurements have the fading trend. The measurements at the common indexes are then encoded to secret bits by using our proposed fading trend estimation and thresholding. The remaining set of measurements without the fading trend will be quantized to secret bits by using existing multi-level quantization method [3]. Therefore, in our method, both the fading trend estimation and multiple thresholding are used in secret key extraction. One of the encouraging observations from our various experimental scenarios is that we found over 75% of RSS measurements exhibit a fading trend.

Algorithm 1 Algorithm flow for fading trend based secret bit extraction per RSS measurement.

Require: **INPUT:**
$\hat{Y}(t-1), \hat{Y}(t), \hat{Y}(t+2)$: the RSS readings measured from the probe packet with time index $t-1, t, t+2$;
OUTPUT:
$[b_t(1), b_t(2), \cdots, b_t(m)]$: m-bit secret bit sequence extracted from RSS measurement $\hat{Y}(t)$;
PROCEDURES:

1: **Interpolation:**
Using cubic Farrow filter based interpolation technique.

2: **Fading trend estimation:**
For a set of RSS measurements $\{\hat{Y}(t-1), \hat{Y}(t), \hat{Y}(t+2)\}$,

$$b_t(1) = \begin{cases} 0 & \Phi^1 < 0 \wedge \Phi^2 < 0 \\ 1 & \Phi^1 > 0 \wedge \Phi^2 > 0 \end{cases} \tag{39}$$

3: **Thresholding:**
RTM:

$$b_t(2) = \begin{cases} 0 & \hat{Y}(t) < \theta \\ 1 & \hat{Y}(t) \geq \theta \end{cases} \tag{40}$$

RTQ:
$b_t(i), i = 2, \cdots, m$: Using quantization via multiple thresholds.

8.2 Bit Mismatch Probability Analysis

We next provide a theoretic analysis of the probability of bit disagreement when using the fading trend for secret bit encoding. $\hat{Y}^A_{B,A}(t)$ and $\hat{Y}^B_{A,B}(t)$ are measured RSS readings at Alice and Bob respectively,

$$\hat{Y}^A_{B,A}(t) = Y^A_{B,A}(t) + W^A_{B,A}(t)$$
$$\hat{Y}^B_{A,B}(t) = Y^B_{A,B}(t) + W^B_{A,B}(t),$$

where $t = \tau - 1, \tau, \tau + 2$. The RSS measurements are determined by the radio channel and noise $W(t)$ at different time instants. $W(t)$ is assumed as i.i.d Gaussian noise, following $N(0, \sigma^2)$. According to the reciprocity principle, for each time instant t, $Y^A_{B,A}(t)$ should be equal to $Y^B_{A,B}(t)$. Assuming each RSS measurement is independent, both Φ^i_A and $\Phi^i_B, i = 1, 2$, also follow Gaussian distribution with variance $2\sigma^2$, where Φ^i_A and $\Phi^i_B, i = 1, 2$, has the same definition as Φ_i for Alice and Bob respectively. The following conditions need to be fulfilled if there is a bit disagreement:

$$\begin{aligned} &\{\Phi^1_A > 0 \wedge \Phi^1_B < 0 \wedge \Phi^2_A > 0 \wedge \Phi^2_B < 0\} \\ \vee \quad &\{\Phi^1_A < 0 \wedge \Phi^1_B > 0 \wedge \Phi^2_A < 0 \wedge \Phi^2_B > 0\} \end{aligned} \tag{41}$$

where

$$\Phi^1_A = \hat{r}_A(k) - \hat{r}_A(k-1) \sim N(r_A(k) - r_A(k-1), 2\sigma^2)$$
$$\Phi^2_A = \hat{r}_A(k+2) - \hat{r}_A(k) \sim N(r_A(k+2) - r_A(k), 2\sigma^2)$$

Then the probability for bit disagreement can be derived as:

$$\begin{aligned} &Pr(err) = Pr(\Phi^1_A > 0 \wedge \Phi^1_B < 0 \wedge \Phi^2_A > 0 \wedge \Phi^2_B < 0) \\ &+ Pr(\Phi^1_A < 0 \wedge \Phi^1_B > 0 \wedge \Phi^2_A < 0 \wedge \Phi^2_B > 0) \\ &= (1 - F(\Phi^1_A = 0))(1 - F(\Phi^2_A = 0))F(\Phi^1_B = 0)F(\Phi^2_B = 0) + \\ &F(\Phi^1_A = 0)F(\Phi^2_A = 0)(1 - F(\Phi^1_B = 0))(1 - F(\Phi^2_B = 0)), \end{aligned} \tag{42}$$

where $F()$ is the cumulative distribution function for Gaussian distribution.

To illustrate, Figure 6 depicts the probability density function of Φ^1_A. If the mean value of Φ^1_A has a large deviation from 0, which means the signal strength changes sharply from $Y^A_{B,A}(\tau - 1)$ to $Y^A_{B,A}(\tau)$ due to the fading effects,

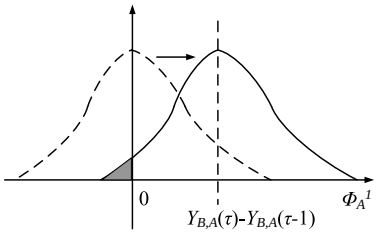


Fig. 6. Illustration of the bit disagreement probability analysis.

the probability that $\Phi_A^1 < 0$ shown as the shaded area will be extremely small. Due to the reciprocity principle, $Y_{B,A}^A(\tau) - Y_{B,A}^A(\tau-1)$ equals to $Y_{A,B}^B(\tau) - Y_{A,B}^B(\tau-1)$, which implies that Φ_A^i and $\Phi_B^i, i = 1,2$, have the same mean value, and it results in the probability of $\Phi_B^1 < 0$ to be also small. Therefore, the first term of equation (42) should be a small value, which indicates a small bit disagreement probability. Similar analysis can be applied for the second term in equation (42) as well.

In this work, we utilize the fading trend based key extraction method as the basis for our two group key extraction protocols (via star and chain topologies) in our framework. We compare the performance of our method with existing ones in next section.

8.3 Secret Key Reconciliation and Privacy Amplification

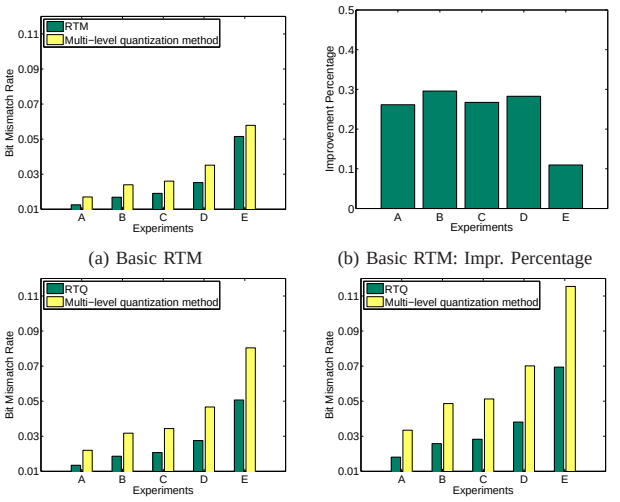
Information Reconciliation: After the fading trend assisted bit extraction, Alice and Bob end up with a bit sequences, K_A and K_B, respectively. According to the reciprocity property, the two bit sequence should theoretically identical, but due to the estimation error resulted from noise or interference etc., there are some bit mismatches existing between K_A and K_B. To reconcile the bit discrepancies, existing information reconciliation techniques, such as error correction codes ((12, 23) Golay code), is deployed [39]. An (n,m) error correction code C includes a one-to-one encoding function $f_{enc}()$ mapping from m-bit sequence to n-bit sequence $(n > m)$, and a many-to-one decoding function $f_{dec}()$ mapping any n-bit sequence to one of 2^k n-bit sequence which is called codewords of C. Alice first computes the closest codeword to K_A in C through the decoding function $f_{dec}(K_A)$, and then calculates the offset $P = K_A - f_{dec}(K_A)$, which is sent to Bob. Upon receiving P, Bob can decode K_A by the following operation: $P + f_{dec}(K_B - P)$. At the end of this step, Both Alice and Bob can have the common K_A with high probability due to the error correcting ability of C.

Privacy Amplification: Since the information during the reconciliation stage can also be heard by Eve in the public channel, partial information about the secret key between Alice and Bob may be exposed to Eve. To ensure the shared secret key completed unknown to Eve, the technique of privacy amplification can be used to solve this problem. The way to realize privacy amplification is to use the encoding function $f_{enc}()$ to obtain the k-bit pre-image of the n-bit codeword $f_{dec}(K_A)$.

9 System Prototype and Experimental Evaluation

9.1 Prototype Implementation

We build a group key extraction system prototype, in which one *initial node* and several *participant nodes* generate a group key collaboratively. The initial node is used to start the procedure of extracting the group key via both star and chain topologies. It is responsible for 1) actively sending probe packets to other participating nodes to collect RSS measurements; 2) being the central node in the star topology and calculating the DOSS value for participating nodes. To ensure the reciprocity of wireless channels, each participating node sends out probe packets once receiving probe packets from any other node. To cooperate the secret key extraction via the chain topology, the participating node is designed to calculate the DOSS value, and inserts it to the probe packet that will be essentially relayed to the end of the chain.



(a) Basic RTM (b) Basic RTM: Impr. Percentage

(c) Extended RTQ: 3-bit (d) Extended RTQ: 4-bit

Fig. 7. Bit mismatch rate under various experimental scenarios.

Our prototype uses Crossbow MICAz motes, which support 2.4GHz IEEE 802.15.4 communication at a high speed of 250kbps. We implement a mobile wireless network using 6 MICAz motes: one acts as the initial node and the other five are participating nodes. One additional mote is connected to a laptop acting as sink. Probe packets are broadcasted at the rate of 20pkt/sec. The probe packet includes the sending node ID and the packet sequence number so that the sink node can distinguish different probe packets. When a node receives a probe packet, it extracts the sending node ID and packet sequence number, and calculates the DOSS for its two neighboring channels. Next, all these information are inserted into the probe packet that the node will send out. After the sink node receives the probe packet from other nodes, it extracts the related information and stores it in the database. Finally, the sink can calculate all the RSS measurements on the channel between any pair of nodes based on the DOSS information, and feed them as the input of the fading trend based key extraction algorithm for generating secret bits.

Experimental Setup and Scenarios: We conduct experiments by running our mobile wireless network to collect RSS measurements in both outdoor and indoor environments. Our outdoor environments include *park* and *street*. The park is covered with tall trees, multiple small roads and fountains. Our street environment is from Hoboken train station to Stevens Institute of Technology spanning over 10 street blocks. During our experiments, we measure RSS under two different conditions: one is having pedestrians passing through our mobile wireless network,

and the other is not having pedestrians passing through. Thus, for outdoor environments we have four experimental scenarios numbered as: *A (park, with pedestrian)*, *B (park, without pedestrian)*, *C (street, with pedestrian)*, and *D (street, without pedestrian)*. In our indoor environment, the RSS measurements are collected in classrooms, stairs and hallways, indicated as *E (building)*. The outdoor experiments are performed under the presence of dynamic environmental movements (including people walking, kids running, and cars driving around) and all the motes involved in secret key generation are constantly moving. There are total 25 data sets, each lasts for about 5 minutes.

Metrics: To evaluate the performance of our framework, we use the following metrics:

Bit mismatch rate (BMR): For key extraction between a pair of wireless devices, the bit mismatch rate is defined as the number of bits that do not match between two devices divided by the total number of secret bits extracted. For group key extraction, it is defined as the averaged bit mismatch rate from all pairs of devices in the group.

Bit generation rate (BGR): The bit generation rate represents as the number of secret bits extracted per RSS measurement.

Randomness: The standard NIST test suite is employed to measure the randomness of the generated secret bit string.

9.2 Evaluation of Fading Trend based Key Extraction Algorithm

9.2.1 Bit Mismatch Rate (BMR)

We compare our fading trend based key extraction scheme with the representative previous studies [7], which uses multi-level quantization.

Basic RTM scheme: Figure 7(a) shows the bit mismatch rate versus different experimental scenarios from *A* to *E* for both our method and the multi-level quantization method when maintaining the same secret bit generation rate at 2 bits per measurement. We observe that our method outperforms the multi-level quantization approach by over 26% for outdoor environments, particularly, 26%, 30%, 27%, 28% for scenarios A∼D respectively, and around 11% for indoor environment as shown in figure 7(b). In addition, the scenarios with pedestrians passing between mobile devices achieve lower bit mismatch rate, indicating the presence of larger fading, which benefits our proposed method.

Extended RTQ scheme: Figure 7(c) and (d) presents the bit mismatch rate for RTQ scheme and the multi-level quantization method when generating 3 bits and 4 bits from one RSS measurement. By comparing Figure 7(c) and figure 7(d), we observe that as the number of secret bits extracted per RSS measurement increases, the bit mismatch rate also increases for both methods. However, our proposed method outperforms the multi-level quantization method for more than 40% under each scenario, and the performance improvement becomes more significant as the number of encoded bits increases. The increased bit mismatch rate for both methods is caused by the increasing number of thresholds for quantizing RSS measurements. However, due to the fading trend employed, the bit mismatch rate of our method does not increase as much as the multi-level quantization method when the number of encoded bits increases.

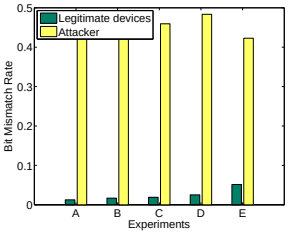


Fig. 8. Bit mismatch rate for legitimate devices and attacker under different scenarios.

Comparison of BMR between legitimate user and eavesdropper: Figure 8 presents our experimental results of a pair of MICAz nodes with the presence of eavesdropper (using an additional MICAz mote placed at 30 cm away) for 2 bits per measurement. We find that the bit mismatch rate incurred by eavesdropper is much higher than that of between the pair of legitimate devices under different scenarios identified as *A*, *B*, *C*, *D*, and *E* in Section 9.1. This observation validates the high security of using channel measurements for secret key extraction.

9.2.2 Randomness

To ensure that the secret key generated is substantially random, the standard randomness test suite from NIST [40] is employed to verify the effectiveness of the secret bits extracted after secret key reconciliation and before privacy amplification [6]. Since the bit length generated from our experiments should meet the recommended size of the NIST tests, we run 8 NIST tests and calculate their p-values. The test results for 5 different experimental scenarios are listed in Table 1. All the cases pass the test with the *p*-value much larger than 0.01, which is the threshold to pass the test. The NIST results show that privacy amplification increases the randomness of the secret key extracted, so as the effectiveness of the secret key extracted.

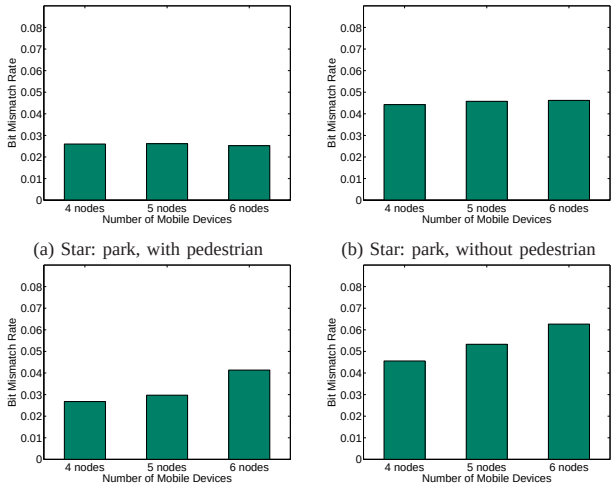
9.3 Group Key Extraction via the Star Topology

We next study how the number of nodes in the group affects the BMR for group key extraction via the star topology in Figure 9 (a) and (b). We observe that the bit mismatch rate is stable when the group size increases under both scenarios A and B when maintaining the bit generation rate at 2 bits per measurement.

The results are consistent with our theoretical analysis in Section 5.2. A slight difference exists on the bit mismatch rate under each scenario among different group sizes due to the noise does not strictly follow identical Gaussian distribution in practice. Furthermore, we found that the performance of our protocol is better under scenario A with

Test	A	B	C	D	E
Freq.	0.55	0.42	0.23	0.55	0.55
Block Freq.	0.86	0.87	0.81	0.87	0.96
Cum. sums (Fwd)	0.72	0.54	0.22	0.72	0.96
Cum. sums (Rev)	0.81	0.81	0.39	0.81	0.96
Runs	0.84	0.50	0.51	0.84	0.69
Longest run of 1s	0.76	0.42	0.51	0.84	0.83
FFT	0.65	0.65	0.65	0.65	0.17
Approx. Entropy	0.92	0.65	0.39	0.92	0.92
Serial	0.50	0.50	0.50	0.50	0.50
	0.50	0.50	0.50	0.50	0.97

TABLE 1
NIST statistical test suite results



(a) Star: park, with pedestrian (b) Star: park, without pedestrian

(c) Chain: park, with pedestrian (d) Chain: park, without pedestrian

Fig. 9. Performance of group key extraction.

pedestrians, thus confirms the effectiveness of our fading trend based key extraction scheme.

9.4 Group Key Extraction via the Chain Topology

Figure 9(c) and (d) present the BMR for group key extraction via the chain topology. We observe that as the group size increases, the bit mismatch rate under both scenarios A and B increases when maintaining the bit generation rate at 2 bits per measurement. For scenario A, the bit mismatch rate increases from 0.034 to 0.058 when the number of group members changes from 4 to 6, whereas scenario B has the bit mismatch rate increasing from 0.056 to 0.073. This is due to the increasing noise variance when DOSS values are accumulated along the chain. According to the analysis in Section 4, the bit mismatch rates are still within the error tolerance range of Golay code.

9.5 Communication Cost

Besides the probe exchange between two neighboring nodes for both star and chain topologies, the virtual central node needs to disseminate the DOSS values to the group members for star topology, whereas the group members need to forward the DOSS value by traversing their subsequent nodes on the chain for chain topology. The communication cost for both topologies is similar, and is linear to the size of the group. For example, every 1 or 2 bits in the group key would incur $2(n-1)+(n-2)=3n-4$ packets communication cost among the whole group, where n is the number of nodes defined in section 5 and 6.

10 Conclusions

In this paper, we address the problem of group key extraction by exploiting physical layer information of radio channel. In particular, the group key is extracted when multiple wireless devices work collaboratively with the readily available Received Signal Strength (RSS) in radio channels, without relying on a fixed infrastructure. We propose a relay node assisted mechanism that solves the issue when mobile devices are not within each other's communication range. Our relay node assisted mechanism uses difference of signal strength to ensure the security of the key extraction, and achieves a lower bit mismatch rate comparing to existing studies while maintaining a similar key generation rate when employing key extraction based on fading trend. To enable secure group communication, two protocols via star and chain topologies are developed in our framework by exploiting RSS from multiple devices to perform group key generation collaboratively. The collaborative key extraction protocol via the star topology is designed for scenarios when the group of wireless devices under consideration is within the communication range of each other, while the protocol via the chain topology involves handling the scenarios when not all wireless devices inside the group are within the communication range of each other. We derive the maximum achievable group key rate by our approach. Our analysis provides important insights on the amount of drop in key rate as the group size grows and enables us to find the best network topology the group can form in order to achieve a high key rate. Our prototype using a mobile wireless network with multiple MICAz motes confirms the feasibility of leveraging RSS for group key generation among multiple wireless devices. The effectiveness of group key extraction via star and chain topologies built on top of fading-trend based key extraction and relay node assisted mechanism is demonstrated through extensive experimental study in both outdoor (e.g., park and street) and indoor (e.g., office building) environments.

11 Acknowledgment

The preliminary results of this project have been published in International Conference on Computer Communications (INFOCOM) 2012 [41]. This work was supported in part by the National Science Foundation under grant numbers CNS-0954020, CCF-1018270, CNS-1318751, CNS-1318748, CNS-1054738, CCF-0916664 and Army Research Office W911NF-13-1-0288. This work was done while H. Liu was at Stevens Institute of Technology.

References

[1] B. Azimi-Sadjadi, A. Kiayias, A. Mercado, and B. Yener, "Robust key generation from signal envelopes in wireless networks," in *ACM CCS*, 2007, pp. 401–410.

[2] T. Aono, K. Higuchi, T. Ohira, B. Komiyama, and H. Sasaoka, "Wireless secret key generation exploiting reactance-domain scalar response of multipath fading channels," *IEEE Transactions on Antennas and Propagation*, vol. 53, no. 11, pp. 3776–3784, 2005.

[3] S. Jana, S. Premnath, M. Clark, S. Kasera, N. Patwari, and S. Krishnamurthy, "On the effectiveness of secret key extraction from wireless signal strength in real environments," in *ACM MobiCom*, 2009, pp. 321–332.

[4] S. Mathur, W. Trappe, N. Mandayam, C. Ye, and A. Reznik, "Radio-telepathy: extracting a secret key from an unauthenticated wireless channel," in *ACM MobiCom*, 2008, pp. 128–139.

[5] A. Sayeed and A. Perrig, "Secure wireless communications: Secret keys through multipath," in *IEEE ICASSP*, 2008, pp. 3013–3016.

[6] Q. Wang, H. Su, K. Ren, and K. Kim, "Fast and scalable secret key generation exploiting channel phase randomness in wireless networks," in *IEEE INFOCOM*, 2011.

[7] N. Patwari, J. Croft, S. Jana, and S. Kasera, "High-rate uncorrelated bit extraction for shared secret key generation from channel measurements," *IEEE Transactions on Mobile Computing*, pp. 17–30, 2009.

[8] J. Croft, N. Patwari, and S. Kasera, "Robust uncorrelated bit extraction methodologies for wireless sensors," in *ACM/IEEE ICNP*, 2010, pp. 70–81.

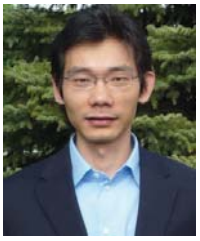
[9] K. Zeng, D. Wu, A. Chan, and P. Mohapatra, "Exploiting multiple-antenna diversity for shared secret key generation in wireless networks," in *IEEE INFOCOM*, 2010, pp. 1–9.

[10] M. Wilhelm, I. Martinovic, and J. Schmitt, "Secret keys from entangled sensor motes: implementation and analysis," in *ACM Wisec*, 2010, pp. 139–144.

[11] O. Arazi and H. Qi, "Self-certified group key generation for ad hoc clusters in wireless sensor networks," in *IEEE INFOCOM*, 2005.

[12] Y. Kim, A. Perrig, and G. Tsudik, "Tree-based group key agreement," *TISSEC*, 2004.
[13] S. Nitinawarat, C. Ye, A. Barg, P. Narayan, and A. Reznik, "Secret key generation for a pairwise independent network model," *IEEE Transaction on Information Theory*, 2010.
[14] S. Nitinawarat and P. Narayan, "Perfect omniscience, perfect secrecy, and steiner tree packing," *IEEE Transactions on Information Theory*, 2010.
[15] C. Ye and A. Reznik, "Group secret key generation algorithms," in *IEEE ISIT*, 2007.
[16] F. Miller, A. Vandome, and J. McBrewster, *Network Topology*. Alphascript Publishing, 2009.
[17] J. Wallace, "Secure physical layer key generation schemes: Performance and information theoretic limits," in *IEEE ICC*, 2009, pp. 1–5.
[18] J. Wallace, C. Chen, and M. Jensen, "Key generation exploiting mimo channel evolution: Algorithms and theoretical limits," in *EuCAP*, 2009, pp. 1499–1503.
[19] U. Maurer and S. Wolf, "Unconditionally secure key agreement and the intrinsic conditional information," *IEEE Transactions on Information Theory*, vol. 45, no. 2, pp. 499–514, 1999.
[20] S. Draper, A. Sayeed, and T. Chou, "Minimum energy per bit for secret key acquisition over multipath wireless channels," in *IEEE ISIT*, 2009, pp. 2296–2300.
[21] J. Hershey, A. Hassan, and R. Yarlagadda, "Unconventional cryptographic keying variable management," *IEEE Transactions on Communications*, vol. 43, no. 1, pp. 3–6, 1995.
[22] C. Ye, A. Reznik, G. Sternberg, and Y. Shah, "On the secrecy capabilities of itu channels," in *IEEE VTC-2007 Fall*, 2007.
[23] M. Madiseh, M. McGuire, S. Neville, and A. Shirazi, "Secret key extraction in ultra wideband channels for unsynchronized radios," in *CNSR*, 2008, pp. 88–95.
[24] R. Wilson, D. Tse, and R. Scholtz, "Channel identification: Secret sharing using reciprocity in ultrawideband channels," *IEEE TIFS*, vol. 2, no. 3, pp. 364–375, 2007.
[25] M. Tope and J. McEachen, "Unconditionally secure communications over fading channels," in *IEEE MILCOM*, 2001.
[26] Z. Li, W. Xu, R. Miller, and W. Trappe, "Securing wireless systems via lower layer enforcements," in *Proceedings of the 5th ACM workshop on Wireless security*, 2006, pp. 33–42.
[27] I. Csiszar and P. Narayan, "Common randomness and secret key generation with a helper," *IEEE Transactions on Information Theory*, vol. 46, pp. 344–366, 2000.
[28] I. Csiszar and R. Ahlswede, "Secrecy capacities for multiple terminals," *IEEE Transactions on Information Theory*, vol. 50, pp. 3046–3061, 2004.
[29] A. Goldsmith, *Wireless Communications*. New York, NY, USA: Cambridge University Press, 2005.
[30] X. Zhu, F. Xu, E. Novak, C. C. Tan, Q. Li, and G. Chen, "Extracting secret key from wireless link dynamics in vehicular environments," in *IEEE INFOCOM*, 2013.
[31] Y. Liang, H. V. Poor, and S. S. (Shitz), *Information Theoretic Security*, ser. Foundations and Trends in Communications and Information Theory. Hanover, MA: now Publishers, Inc., 2008, vol. 5.
[32] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
[33] R. Ahlswede and I. Csiszar, "Common randomness in information theory and cryptography, part i: Secret sharing," *IEEE Transactions on Information Theory*, vol. 39, pp. 1121–1132, 1993.
[34] U. Maurer and S. Wolf, "Secret key agreement over a non-authenticated channel part iii," *IEEE Transactions on Information Theory*, vol. 49, pp. 822–851, 2003.
[35] I. Csiszar and P. Narayan, "Secrecy capacities for multiple terminals," *IEEE Transactions on Information Theory*, vol. 50, pp. 3047–3061, 2004.
[36] U. Maurer, "Secret key agreement by public discussion from common information," *IEEE Transactions on Information Theory*, vol. 39, pp. 733–742, 1993.
[37] C. W. Farrow, "A continuously variable digital delay element," in *IEEE International Symposium on Circuits and Systems*, 1988.
[38] C. Ye, A. Reznik, and Y. Shah, "Extracting secrecy from jointly gaussian random variables," in *IEEE ISIT*, 2006, pp. 2593–2597.
[39] M. Stojanovic, *Wiley Encyclopedia of Telecommunications*. John Wiley & Sons, Inc., 2003.
[40] A. Rukhin and et al., "A statistical test suite for random and pseudorandom number generators for cryptographic applications," in *NIST Special Publication 800-22 Revision 1*, 2001.
[41] H. Liu, J. Yang, Y. Wang, and Y. Chen, "Collaborative secret key extraction leveraging received signal strength in mobile wireless networks," in *IEEE INFOCOM*, 2012, pp. 927–935.

Hongbo Liu joins IUPUI as an assistant professor in Department of Electrical and Computer Engineering since Aug. 2013. He received his Ph.D. degree in Electrical Engineering from Stevens Institute of Technology. His research interests include mobile and pervasive computing, cyber security and privacy and smart grid. He is the recipient of the Best Paper Award from ACM MobiCom 2011 and Best Paper Runner-up Award of IEEE CNS 2013.



Jie Yang received the PhD degree in computer engineering from Stevens Institute of Technology in 2011. He is currently an assistant professor in the Department of Computer Science and Engineering at Oakland University. His research interests include cyber security and privacy, mobile and pervasive computing, wireless localization systems, and mobile social networks. His research is supported by NSF and ARO. He is the recipient of the Best Paper Award from ACM International Conference on Mobile Computing and Networking (MobiCom) 2011 and the Outstanding Research Award in 2009 from Stevens Institute of Technology. His research has received wide press coverage including MIT Technology Review, The Wall Street Journal, CNET News, and Yahoo News. He is a member of the IEEE.

Yan Wang is a Ph.D. candidate of the Electrical and Computer Engineering Department at Stevens Institute of Technology. His research interests include mobile computing, information security and privacy and pervasive computing . He is currently working in the Data Analysis and Information SecuritY (DAISY) Lab with Prof. Yingying Chen.

Yingying (Jennifer) Chen is an Associate Professor in the Department of Electrical and Computer Engineering at Stevens Institute of Technology. Her research interests include cyber security and privacy, mobile computing, and mobile healthcare. She has published over 80 journals and referred conference papers in these areas. She received her Ph.D. degree in Computer Science from Rutgers University. Prior to joining Stevens, she was with Alcatel-Lucent. She is the recipient of the NSF CAREER Award and Google Research Award. She also received NJ Inventors Hall of Fame Innovator Award. She is the recipient of the Best Paper Award from ACM MobiCom 2011. She also received the IEEE Outstanding Contribution Award from IEEE New Jersey Coast Section each year 2005-2009. Her research has been reported in numerous media outlets. She is on the editorial boards of IEEE Transactions on Mobile Computing (IEEE TMC), IEEE Transactions on Wireless Communications (IEEE TWireless), and IEEE Network Magazine.

Can Emre Koksal (S'96 M'03 SM'13) received the B.S. degree in Electrical Engineering from the Middle East Technical University in 1996, and the S.M. and Ph.D. degrees from MIT in 1998 and 2002, respectively, in Electrical Engineering and Computer Science. He was a Postdoctoral Fellow in the Computer Science and Artificial Intelligence Laboratory at MIT until 2004, and a Senior Researcher at EPFL until 2006. Since then, he has been with the Electrical and Computer Engineering Department at Ohio State University, currently as an Associate Professor. His general areas of interest are wireless communication, communication networks, information theory, stochastic processes, and financial economics. He is the recipient of the National Science Foundation CAREER Award in 2011, the OSU College of Engineering Lumley Research Award in 2011, and the co-recipient of an HP Labs - Innovation Research Award in 2011. The paper he co-authored was a best student paper candidate in MOBICOM 2005. Since 2013, he has been an Associate Editor for IEEE Transactions on Wireless Communications and Elsevier Computer Networks.