Received 15 December 2021; revised 27 January 2022; accepted 20 February 2022. Date of publication 25 February 2022;
date of current version 8 March 2022.

Digital Object Identifier 10.1109/OJCOMS.2022.3154626

Dynamic Bayesian Network Based Security Analysis for Physical Layer Key Extraction

XUEQING HUANG[1] (Member, IEEE), NIRWAN ANSARI[2] (Fellow, IEEE), SIQI HUANG[3] (Student Member, IEEE), AND WENJIA LI[1] (Senior Member, IEEE)

(Invited Paper)

[1] Department of Computer Science, New York Institute of Technology, Old Westbury, NY 11568, USA

[2] Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ 07102, USA

[3] Department of Electrical and Computer Engineering, University of North Carolina at Charlotte, Charlotte, NC 28223, USA

CORRESPONDING AUTHOR: X. HUANG (e-mail: xhuang25@nyit.edu)

ABSTRACT Internet of Things (IoT) is envisioned to expand Internet connectivity of the physical world, and the mobile edge cloud can be leveraged to enhance the resource-constrained IoT devices. The performance of the cloud-enhanced IoT applications depends on various system-wide information, such as the wireless channel states between IoT devices and their corresponding serving edge cloud nodes. However, with the semi-trusted edge resources and the public nature of wireless channels, public sharing of system information should be avoided to better balance the tradeoff between performance and security. In this paper, the benefits of local information exchange is investigated, where the privately-owned physical layer channel information is leveraged to extract lightweight keys. For the point-to-point wireless communications links with multiple passive eavesdroppers, the security metric in terms of conditional min-entropy is evaluated via the proposed Dynamic Bayesian Model. The proposed model can flexibly incorporate various dynamic information flows in the system and quantify the information leakage caused by wireless broadcasting. The rigorously defined and derived security metrics for such a key generation pipeline has been verified via the real-world collected time-varying wireless channel data. The designed model can achieve previously inconceivable security properties.

INDEX TERMS Conditional min-entropy, dynamic Bayesian model, key extraction, physical layer security.

I. INTRODUCTION

INTERNET of Things (IoT) is envisioned to expand Internet connectivity and fuse the digital and physical world [1]–[5]. To enhance the resource-constrained IoT devices, the cloud vendors and the mobile operators are converging at the mobile edge computing environment [6]–[10]. The *performance* of IoT applications depends on the following four categories of system-wide information [11]–[19]: 1) static cloud resource configuration parameters (e.g., the storage capacity of edge cloud and communications resources of an IoT device) specified by the existing public standardization or protocols; 2) dynamic resource utilization status (e.g., the number of occupied resource blocks of a wireless access point and the buffer occupancy of a video streaming device); 3) status of the data pipeline, including the channel state of the wireless medium and the bandwidth measurement of the wired medium; 4) information of the data flow being transmitted in the pipeline.

With full knowledge of the system information, the quality of service (QoS) can be optimized by designing various resource provisioning schemes, such as deployment, association, allocation, and load balancing algorithms. The majority of the existing resource provisioning schemes rely on the trusted centralized controller to collect and exchange large scale system information, which is summarized as the input states (aforementioned static configurations or instantaneous parameters) in Fig. 1. This is a strong assumption for the cloud-enabled IoT infrastructure because

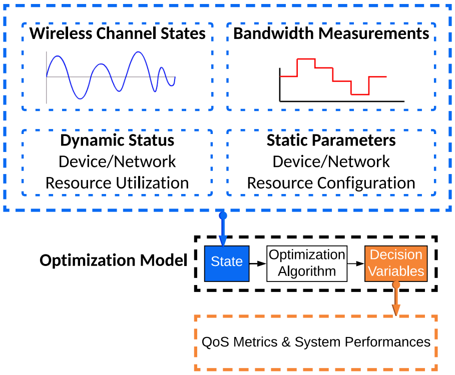


FIGURE 1. **Resource provisioning for performance optimization.**

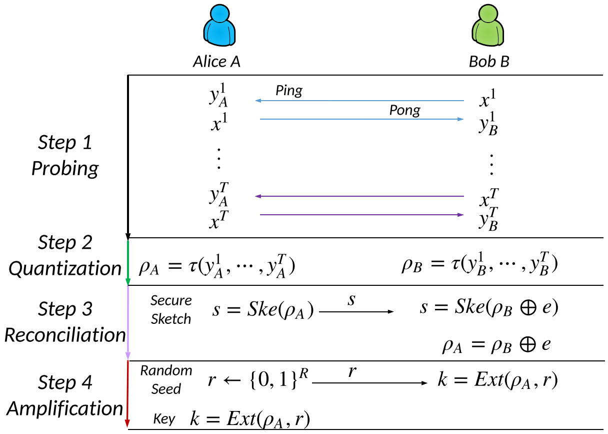


FIGURE 2. **Physical layer-based key extraction procedure.**

of the public nature of the cloud resources and wireless channels.

The intuition from the perspective of system *security*, however, is to avoid any public sharing of system information and remove the dependency on semi-trusted remote resources, i.e., guaranteeing the locality of both information and resource. Understanding the benefits brought by the local information exchange will not only better balance the tradeoff between performance and security, but also present opportunities to engineer the privately-owned information to extract lightweight keys and achieve previously inconceivable security properties.

This paper focuses on the following fundamental question about performance improvement and information leakage caused by information exchange in the cloud enhanced IoT networks. With limited information exchange, the physical (PHY) layer key generation pipeline can potentially enable two nodes to agree on a common key, such as extracting a common key based on the reciprocity of the wireless channel characteristics. **How can one rigorously define and derive the security metrics for such a key generation pipeline?**

Since system information exchange (channel state information) can in some cases cause security breach (physical layer key exposure), the information locality is engineered into the PHY layer key generation pipeline as illustrated in Fig. 2. Given the channel reciprocity of a point-to-point wireless link, there is a strong correlation between the probing signals received by Alice and Bob. Consequently, at the end of each key generation procedure, without publicly reporting or sharing channel information through the system, the two users can agree on a common key based on their respectively received time-varying signals. To theoretically measure the security performance of the generated key with multiple malicious users eavesdropping the broadcasted probing signals, we propose to leverage a probabilistic graphical model named Dynamic Bayesian Network (DBN).

By treating the channel statuses experienced by different users in the system as random variables, the DBN model structure becomes a directed acyclic graph that can encode the conditional dependence between any pair of variables as a directed edge, where the starting vertex of the edge is the parent variable, and the ending vertex is the child variable. To quantify the information leakage caused by passive eavesdropping during the probing period, all the users' received signals will be fed into the DBN model as time series. The correlation among consecutive signals can also be incorporated into the network structure of DBN, i.e., by flexibly adding the edges to encode the conditional dependence of variables in different time slots. The time series input and the conditional dependence edges allow DBN to adapt to the dynamic wireless channels.

With the proposed DBN model structure and the input data, we can 1) estimate the model parameters, i.e., the conditional probability of each random variable given its parents (such as the probability distribution of Alice's received signal given Bob's received signal), and 2) infer the corresponding conditional min-entropy to measure the uncertainty remaining in the key given the eavesdroppers' observations. As demonstrated by the data collected from the real-world experiments, extracting the common key from the local information can reduce the security vulnerabilities of the IoT applications deployed on cloud platforms.

As compared to the existing theoretical security analysis frameworks, the contributions of this work are summarized as follows. 1) The proposed probabilistic graphical model will scale well to multiple passive eavesdroppers and be much less demanding in terms of the probing sequence length needed to estimate the security measurement of the generated key. 2) With prior knowledge of the instantaneous channel fading distribution and white noise parameters, the information flow in DBN can be estimated with less complexity [20], [21]. 3) DBN also comprehensively includes all of the signals transmitted and received in the channel probing step, including information leakage caused by eavesdropping.

The remaining of the paper is organized as follows. Section II introduces the existing experimental and theoretical works on the physical layer key generation. Section III explains the proposed system model and the corresponding security metric. Section IV covers the detailed security quantification algorithms and the corresponding performance with the real world data, and Section VI presents the conclusion.

II. RELATED WORKS

The physical layer has been used for autonomous key generation by exploiting the physical communication channels or hardware-imperfection based attributes for the authentication process [22]–[26].

A. PHYSICAL LAYER-BASED KEY GENERATION PIPELINE

As illustrated in Fig. 2, current physical layer-based key generation schemes are largely based on the 4-step pipeline performed by two devices named Alice and Bob.

Step 1: Pairwise channel probing for $\mathbf{T} = \{1, \ldots, T\}$ time slots. With probing sequence $\{x^t | t \in \mathbf{T}\}$, Alice and Bob have their corresponding received signals $\{y_A^t | t \in \mathbf{T}\}$ and $\{y_B^t | t \in \mathbf{T}\}$, respectively.

Step 2: Quantization scheme τ to generate bit strings: ρ_A for Alice and ρ_B for Bob.

Step 3: Reconciliation to eliminate the discrepancy between two bit strings: Bob recovers ρ_A by using its own measurement ρ_B and error information e. Since Bob's string ρ_B can be treated as a noisy (fuzzy) version of Alice's string ρ_A, the error information e can be decoded based on s, which is generated by the secure sketch scheme *Ske* [27].

Step 4: Privacy amplification to increase the randomness of the generated common key k, where randomness extractor *Ext* can derive a uniformly random string k from ρ_A and an R-bit random seed r [23], [28].

Owing to the wireless channel reciprocity between Alice and Bob, channel impulse responses or received signal strengths derived from y_A^t and y_B^t will function as shared random sources to generate a common key [23], [29], and it was found that the above key probing sequences can be piggybacked with the data exchange sequences. Another field study [30] showed that when the probing sequences are only used for the key generation purpose, i.e., no piggybacking, the *energy consumption* and *time duration* can be more demanding, as compared with classical key establishment approaches. Although there is a discrepancy on the efficiency of the physical layer-based key generation procedure, which requires further investigation, traditional cryptography-based security solutions are becoming inadequate, due to the difficulty in providing physical protection of devices and initializing and securing key materials during the lifetime of an IoT system. Consequently, it is important to explore the physical layer-based key extraction pipeline for IoT devices.

B. SECURITY ANALYSIS FOR PHYSICAL LAYER KEY GENERATION

Empirical Security Analysis — In terms of security analysis for the generated key, the majority of the existing evaluations are performed using empirical methods. 1) For a *two-node system* with Alice and Bob, the following metrics have been measured in different phases of the pipeline: bit disagreement rate (BDR) before and after Step 3 [23], [30], secret mutual information shared via probing sequence in Step 1 [23], average min-entropy on each bit of the bit string in Step 2 to measure the lower bound of the randomness of the key [31], and secret bit rate to measure the average number of secret bits extracted from each reconciled channel response (Step 3) [32]. 2) For the *multi-node system* with passive/active attacker Eve, leaked information has been adopted to measure the mutual information among Alice, Bob, and Eve of the probing sequences in Step 1 [23].

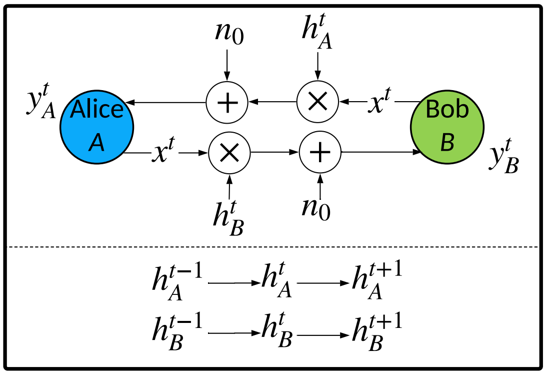

n_0
h_A^t
y_A^t
Alice
A
x^t
Bob
B
x^t
y_B^t
h_B^t
n_0
$h_A^{t-1} \rightarrow h_A^t \rightarrow h_A^{t+1}$
$h_B^{t-1} \rightarrow h_B^t \rightarrow h_B^{t+1}$

FIGURE 3. One pair of ping and pong probings between Alice and Bob.

Channel Modeling — The theoretical security performance of the physical layer-based key depends on randomness inherent in the wireless channel model. Channel modeling has laid the foundation for the first generation of wireless cellular technology over 40 years ago. As illustrated in Fig. 3, the relationship between the signal received by Alice y_A^t and the original probing signal x^t is determined by the physical wireless channel model: at time t, $y_A^t = h_A^t x^t + n_0$, where n_0 is the additive white Gaussian noise (AWGN) and h_A^t is the time-varying channel gain with destination indicated by the sub-index A. At any time t, the *instantaneous behavior* of the channel fading h_A^t can be statistically following Rayleigh, Rician, log-normal, or Nakagami probability distribution function [33]. The *time-variation* of the fading channel, i.e., the correlation between h_A^t and $h_A^{t'}$ can be modeled with Markov property [34], [35]. For example, with the first-order Markov model, future h_A^{t+1} and past h_A^{t-1} are independent given present h_A^t.

Theoretical Security Analysis — In the limited amount of theoretical frameworks, based on different wireless channel models, there are two directions on the security analysis.

1) With the assumption that h_A^t and h_A^{t+1} in Step 1 are independent and identically distributed (i.i.d.), i.e., no Markov property, the mutual information between Alice's channel estimation $(\tilde{h}_A)$ and Bob's channel approximation $(\tilde{h}_B)$ has been derived using an information theoretic approach [36]–[38], where the time index t is dropped

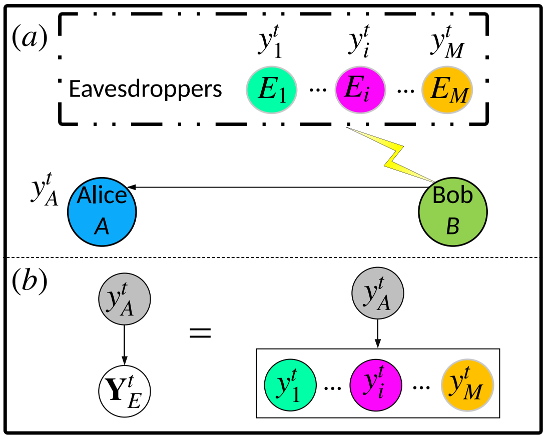


FIGURE 4. Eavesdropping: (a) wiretap channel model with M eavesdroppers $\{E_1, \ldots, E_i, \ldots, E_M\}$, and (b) HMM with multiple eavesdroppers' observations.

because of the stationary and memoryless assumption. It has been found that *the mutual information* $I(\tilde{h}_A, \tilde{h}_B)$ *will decrease with the key coherence time*, which can be approximated as the duration of one time slot because when the time interval between the ping signal and pong signal in one time slot is bigger than a certain threshold, the reciprocity between channels h_A and h_B cannot be guaranteed [39]. To obtain high secrecy capacity (mutual information over coherence time), the key generation system prefers shorter time slots. *The short time slot will invalidate the i.i.d. assumption for both* h_A^t *and* h_B^t*, and render the derived security performance measurement inaccurate.*

2) With channel fading h_A^t being kept in a black-box, Edman *et al.* [40] assumed the first order Markov property for Alice's received signal $\mathbf{Y}_A = \{y_A^t | t \in T\}$, and they subsequently adopted the hidden Markov model (HMM) to derive the *conditional min-entropy*, which measures the uncertainty remained in Alice's private sequential channel measurements $y_A^t \in \{s_0, \ldots, s_j, \ldots, s_{255}\}$ given Eve's passive sequential observation sequence, where 0-255 is the standard range of the received signal strength indicator (RSSI) [41]. Wang *et al.* [42] extended the scenario to the case with multiple passive eavesdroppers $\mathbf{M} = \{1, \ldots, M\}$. As illustrated in Fig. 4, at time t, the eavesdroppers' measurements are $\mathbf{Y}_E^t = \{y_i^t | i \in M\}$, with y_i^t being the ping signal transmitted by Bob and received by the i-th eavesdropper. However, it is found that HMM cannot scale well with multiple eavesdroppers because the state space (256^M possible states) of eavesdroppers' joint observation grows exponentially with M and the corresponding HMM parameters cannot be estimated accurately with the limited probing sequence given in Step 1.

III. SYSTEM MODEL

To answer the question "Is it possible to obtain the physical layer key with provable security?", we will design the security metrics for the common keys obtained from local channel state information. Given the fact that the keys are derived from the wireless channel between nodes, it is essential to prevent the system from revealing such information using the network. To theoretically quantify the information leakage occurred in the key generation pipeline, a probabilistic graphical model will be adopted to flexibly model the information flow caused by public channel probings and subsequent eavesdropping. The detailed signal model, channel model, and information flow model are given below.

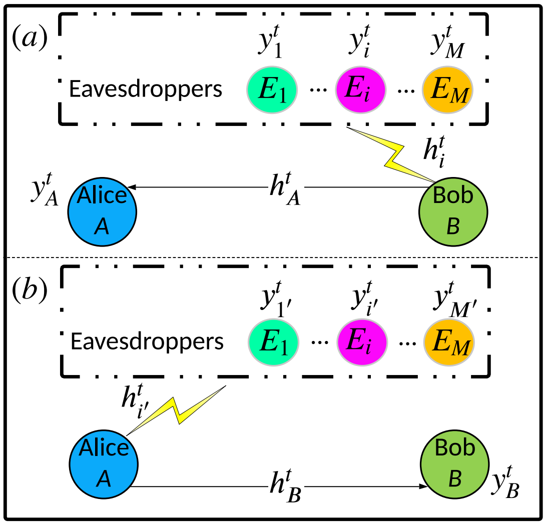


FIGURE 5. Signal probing with multiple eavesdroppers: (a) Ping signal received by Alice and M eavesdroppers. (b) Pong signal received by Bob and M eavesdroppers.

A. SIGNAL MODEL

As shown in Fig. 5, the wiretap channel model is adopted to evaluate the performance of the secret key extracted from the physical properties of the wireless channel between two legitimate wireless devices: Alice A and Bob B, with multiple distributed adversaries E_i, $i \in \mathbf{M} = \{1, \ldots, M\}$, eavesdropping the private data transmission. Alice and Bob will alternatively probe the pair-wise highly correlated wireless channel h_A^t and h_B^t, where h is the channel gain, subindex A indicates the destination node being Alice, and B for Bob. With the assumption of channel reciprocity, $h_A^t \approx h_B^t$, a common key will be derived based on a sequence of probing results over the time duration of $\mathbf{T} = \{1, \ldots, t, \ldots, T\}$.

In the t-th probing iteration, we denote $x^t \in \{0, 1\}$ as the probing symbol transmitted by Alice and Bob alternatively [23]. Then, the discrete-time model for the received signal is

$$\begin{cases} y_A^t = h_A^t x^t + n_0, \\ y_B^t = h_B^t x^t + n_0, \end{cases} \tag{1}$$

where the signals received by Alice and Bob are y_A^t and y_B^t, respectively. n_0 is the zero mean additive Gaussian noises with variance σ^2.

Each eavesdropper E_i, $i \in M$, will monitor the probing sequences sent by Alice and Bob. Let y_i^t be the i-th adversary's passive observation for the Bob→ Eavesdropper link, and $y_{i'}^t$ for the Alice→ Eavesdropper channel.

$$\begin{cases} y_i^t = h_i^t x^t + n_0, \\ y_{i'}^t = h_{i'}^t x^t + n_0, \end{cases} \tag{2}$$

where h_i^t and $h_{i'}^t$ are the gains for the channel from Bob/Alice to the Eavesdropper, respectively.

B. CHANNEL MODEL

We denote $\mathbf{H}_A = \{h_A^t | t \in \mathbf{T}\}$ as Bob's channel vector, which collects the fading coefficients on the links from Alice to Bob during the entire probing process. Similarly, Alice's channel vector is $\mathbf{H}_B = \{h_B^t | t \in \mathbf{T}\}$. Eve has two channel matrices $\mathbf{H}_E = \{h_i^t | t \in \mathbf{T}, i \in \mathbf{M}\}$ and $\mathbf{H}_{E'} = \{h_{i'}^t | t \in \mathbf{T}, i \in \mathbf{M}\}$. For each channel fading coefficient, we have the following two assumptions regarding the instantaneous behavior and the time variation behavior.

The instantaneous channel fading amplitude is assumed to be a Gaussian mixture model (GMM) with multiple components [43], [44]. With a sufficient number of mixing Gaussian components, it is mathematically convenient to adopt GMM as a universal approximator of densities. For instance, at time t, the channel from Bob to Alice is

$$h_A^t \sim \sum_{j=1}^{N_A} w_{A,j}^t \mathcal{N}\left(\mu_{A,j}^t, \left(\sigma_{A,j}^t\right)^2\right), \tag{3}$$

where N_A is the total number of Gaussian components, and the j-th component follows the Gaussian distribution with mean value μ_j and standard deviation $\sigma_{A,j}^t$. The weight of the j-th component is $w_{A,j}^t$ and $\sum_{j=1}^N w_{A,j}^t = 1$ [45].

The time variation of the wireless fading channels is assumed to satisfy the first order Markov property. Namely,

$$h_A^{t+1} \perp h_A^{t-1} | h_A^t, \tag{4}$$

where the future $(t+1)$ is conditionally independent of the past $(t-1)$ given the present t.

C. INFORMATION FLOW MODEL

To provision practical and provable security, fundamental methods are needed to bridge the gap between information theory and engineering solutions. We will adopt a general dynamic Bayesian network (DBN) model, which can represent the probabilistic relationships among interacting variables (nodes) using a parameterized directed acyclic graph. The directed edge encodes the conditional dependence of the child node on one or more parent nodes, where the starting vertex of the directed edge is the parent, and the ending vertex is the child.

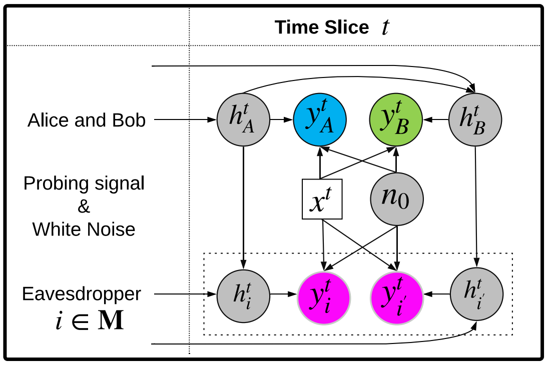


FIGURE 6. DBN = graph structure + local conditional distribution of each node given its parents. A colored node denotes a continuous observable variable, a shaded node indicates a latent continuous variable, and a square node means a discrete control variable.

As shown in Fig. 6, at each time slot t, the signals received by each node in the system is determined by the parent nodes (e.g., the signal received by the i-th eavesdropper from Alice, $y_{i'}^t$, depends on the probing signal x^t, the fading channel between Alice and eavesdropper $h_{i'}^t$, and white noise n_0). With the Markov property, the channel fading value at time slot t is correlated with the fading value at the previous time slot $t-1$. Note that the probing signal x^t is not a stochastic random variable, but an input variable externally imposed on the system. Meanwhile, owing to the channel reciprocity, one directed edge is added from h_A^t to h_B^t to indicate the correlation between these two channel measurements. With the proposed DBN network structure, since these two nodes (h_A^t and h_B^t) do not form a v-structure, i.e., they do not share a common child, the direction of the edge between h_A^t and h_B^t is not important and we choose the one to imply "Bob probes Alice first". Furthermore, Bob's probing signals will be received by both Alice and eavesdropper E_i, $i \in \mathbf{M}$, and thus one edge is added from h_A^t to h_i^t to indicate the information leakage caused by the correlation between these two channel measurements. Similarly, we have the $h_B^t \rightarrow h_{i'}^t$ edge.

The probabilistic model for the graph is a joint probability over all random variables. Since each variable is conditionally independent of all its non-descendants in the graph given the value of all its parents, joint distributions can be encoded as a product of local conditional distributions. The joint probability of the system is expressed as follows.

$$\begin{aligned}
&\Pr(\mathbf{Y}_A, \mathbf{H}_A, n_0, \mathbf{Y}_B, \mathbf{H}_B, \mathbf{Y}_E, \mathbf{Y}_{E'}, \mathbf{H}_E, \mathbf{H}_{E'}) \\
&\quad = \Pr\left[y_A^1 | x^1, h_A^1, n_0\right] \times \Pr\left[h_A^1\right] \times \Pr\left[n_0\right] \\
&\quad \times \Pr\left[y_B^1 | x^1, h_B^1, n_0\right] \times \Pr\left[h_B^1 | h_A^1\right] \\
&\quad \times \prod_{i=1}^{M} \left\{ \Pr\left[y_i^1 | x^1, h_i^1, n_0\right] \times \Pr\left[h_i^1 | h_A^1\right] \right\} \\
&\quad \times \prod_{i=1}^{M} \left\{ \Pr\left[y_{i'}^1 | x^1, h_{i'}^1, n_0\right] \times \Pr\left[h_{i'}^1 | h_B^1\right] \right\} \\
&\quad \times \left\{ \prod_{t=2}^{T} \left\{ \Pr\left[y_A^t | x^t, h_A^t, n_0\right] \times \Pr\left[h_A^t | h_A^{t-1}\right] \times \Pr\left[n_0\right] \right\} \right\} \\
&\quad \times \left\{ \prod_{t=2}^{T} \left\{ \Pr\left[y_B^t | x_t, h_B^t, n_0\right] \Pr\left[h_B^t | h_B^{t-1}\right] \Pr\left[h_B^t | h_A^t\right] \right\} \right\}
\end{aligned}$$

$$\times\left\{\prod_{t=2}^{T}\prod_{i=1}^{M}\left\{\Pr\left[y_i^t|x^t, h_i^t, n_0\right]\Pr\left[h_i^t|h_i^{t-1}\right]\Pr\left[h_i^t|h_A^t\right]\right\}\right\}$$
$$\times\left\{\prod_{t=2}^{T}\prod_{i=1}^{M}\left\{\Pr\left[y_{i'}^t|x^t, h_{i'}^t, n_0\right]\Pr\left[h_{i'}^t|h_{i'}^{t-1}\right]\Pr\left[h_{i'}^t|h_B^t\right]\right\}\right\} \tag{5}$$

where $\mathbf{Y}_A = \{y_A^t | t \in \mathbf{T}\}$ are the signals received by Alice over T time slots. Similarly, Bob's signal vector is $\mathbf{Y}_B = \{y_B^t | t \in \mathbf{T}\}$. Eve has two channel matrices $\mathbf{Y}_E = \{y_i^t | t \in \mathbf{T}, i \in \mathbf{M}\}$ and $\mathbf{Y}_{E'} = \{y_{i'}^t | t \in \mathbf{T}, i \in \mathbf{M}\}$.

D. SECURITY EVALUATION FOR THE PROBING PHASE OF THE KEY GENERATION PIPELINE

The security strength of a pre-shared key between Alice and Bob can be measured as the number of bits (Shannon entropy) in the key, that is, to break any "L-bits security strength" key k, an attacker is required to perform around 2^L operations . Meanwhile, the concept of min-entropy provides a worst case estimation of the randomness of the key, where the lower bound of uncertainty is given below [31].

$$H_\infty(K) = \min_{k\in\mathcal{K}} -\log\Pr\left[K=k\right]$$
$$= -\log\left(\max_{k\in\mathcal{K}}\Pr\left[K=k\right]\right), \tag{6}$$

where the random variable key is K and $\mathcal{K}$ is the set of all possible L-bits keys. $\Pr[K = k]$ is the probability of generating the pre-shared key $k \in \mathcal{K}$. The success probability of an adversary guessing the key on the first try using an optimal guessing scheme is $2^{-H_\infty(K)}$.

Suppose the adversary can eavesdrop some information O about the key K. The conditional min-entropy, $H_\infty(K|O)$, is adopted to indicate the remaining uncertainty . Given an observation o, the probability of adversaries successfully obtaining the key using the maximum likelihood decoder on the first try is [46]:

$$2^{-H_\infty(K|O=o)} = \max_{k\in\mathcal{K}}\Pr\left[K=k|O=o\right]. \tag{7}$$

Then, the *average performance* of Eve using a maximum likelihood decoder to obtain the correct key will be $2^{-H_\infty(K|O)}$, and the corresponding conditional min-entropy is formally defined as follows.

$$H_\infty(K|O) = -\log\left(\mathop{\mathbb{E}}_{o\leftarrow O}\left[2^{-H_\infty(K|O=o)}\right]\right)$$
$$-\log\left(\sum_{o\in O}\Pr\left[O=o\right]\left(\max_{k\in\mathcal{K}}\Pr\left[K=k|O=o\right]\right)\right), \tag{8}$$

where ← refers to sampling of a random variable.

As shown in Fig. 2, during the key generation procedure, adversaries can eavesdrop the channel probing signal, the secure sketch, and the random seed. In this paper, we primarily focus on the information leakage incurred during the probing phase. That is, the observation O includes two sets of signals $\mathbf{Y}_E$ and $\mathbf{Y}_{E'}$. Since both Alice and Bob's key K will be extracted from the commonly agreed $\mathbf{Y}_A$, the conditional min-entropy during the probing phase will be $H_\infty(\mathbf{Y}_A|\mathbf{Y}_E, \mathbf{Y}_{E'})$, with $2^{-H_\infty(\mathbf{Y}_A|\mathbf{Y}_E,\mathbf{Y}_{E'})}$ being given in Eq. (9). The objective of a key generation system is to maximize the conditional min-entropy, i.e., minimizing Eq. (9).

$$2^{-H_\infty(\mathbf{Y}_A|\mathbf{Y}_E,\mathbf{Y}_{E'})}$$
$$= \sum_{\left\{y_{(1:M)}^{(1:T)}, y_{(1':M')}^{(1:T)}\right\}}\left\{\Pr\left[\mathbf{Y}_E = y_{(1:M)}^{(1:T)}, \mathbf{Y}_{E'} = y_{(1':M')}^{(1:T)}\right]\right.$$
$$\left.\times\max_{y_A^{(1:T)}}\Pr\left[\mathbf{Y}_A = y_A^{(1:T)}|\mathbf{Y}_E = y_{(1:M)}^{(1:T)}, \mathbf{Y}_{E'} = y_{(1':M')}^{(1:T)}\right]\right\}. \tag{9}$$

IV. CONDITIONAL MIN-ENTROPY QUANTIFICATION

The objective of a key generation system is to maximize the conditional min-entropy, i.e., minimizing $2^{-H_\infty(\mathbf{Y}_A|\mathbf{Y}_E,\mathbf{Y}_{E'})}$ in Eq. (9). To evaluate the key generation performance during the probing phase, we mainly have to solve Eq. (9) by performing the following two categories of sub-tasks: learning and inference, over the designed dynamic Bayesian network.

Task 1) DBN learning learns the parameters based on the known structure and partially observed data in Fig. 6. In particular, based on the collected signals $\mathbf{Y}_A, \mathbf{Y}_B, \mathbf{Y}_E, \mathbf{Y}_{E'}$, the learning algorithm will estimate all the distributions in Eq. (5), including the probability distribution of each root node (node has no children) and the conditional probability distributions of the remaining nodes in the graph.

Task 2) Marginal inference calculates the marginal probability of eavesdroppers' observation.

$$\Pr\left[\mathbf{Y}_E = y_{(1:M)}^{(1:T)}, \mathbf{Y}_{E'} = y_{(1':M')}^{(1:T)}\right]. \tag{10}$$

Task 3) Maximum a Posteriori (MAP) inference decodes the most likely signal received by Alice given Eve's observations, and returns the corresponding posteriori probability.

$$\max_{y_A^{(1:T)}}\Pr\left[\mathbf{Y}_A = y_A^{(1:T)}|\mathbf{Y}_E = y_{(1:M)}^{(1:T)}, \mathbf{Y}_{E'} = y_{(1':M')}^{(1:T)}\right]$$
$$= \frac{\max_{y_A^{(1:T)}}\Pr\left[\mathbf{Y}_A = y_A^{(1:T)}, \mathbf{Y}_E = y_{(1:M)}^{(1:T)}, \mathbf{Y}_{E'} = y_{(1':M')}^{(1:T)}\right]}{\Pr\left[\mathbf{Y}_E = y_{(1:M)}^{(1:T)}, \mathbf{Y}_{E'} = y_{(1':M')}^{(1:T)}\right]}, \tag{11}$$

where the denominator is irrelevant of $y_A^{(1:T)}$ and can be ignored.

A. MODEL STRUCTURE SIMPLIFICATION

DBN learning aims to learn the parameters based on the known structure and partially observed data in Fig. 6. In particular, based on the collected signals $\mathbf{Y}_A, \mathbf{Y}_B, \mathbf{Y}_E, \mathbf{Y}_{E'}$, the learning algorithm will be designed to estimate the distributions of each node, including the probability distribution of each root node (node with no children) and the conditional probability distributions of the remaining nodes given their parents in the graph.

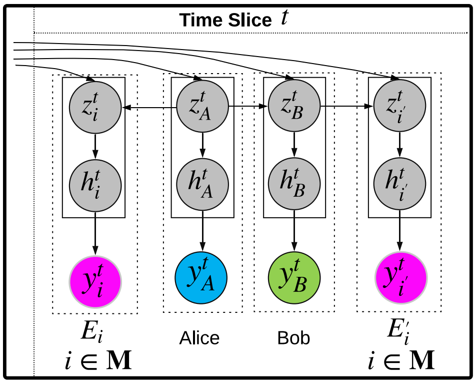


FIGURE 7. Simplified DBN.

One *challenge* for conducting parameter learning is designing a general framework applicable to the various probability distribution for the instantaneous wireless channel behaviors. Another *issue* is that the mixed template model has discrete input variable x^t, continuous hidden states (e.g., h_A^t), and continuous observed states (e.g., y_A^t), thus rendering the problem intractable. For example, the conditional probability of h_B^t given h_A^t may be very complex. As shown in Eq. (3), to simplify the mathematical derivation process without losing the parameter estimation accuracy, Gaussian mixture model is adopted to model the instantaneous channel fading amplitude. Moreover, to avoid complex conditional distribution caused by varying x^t, we assume that the probing sequence consists of all 1s, i.e., $x^t = 1, \forall t \in \mathbf{T}$; this assumption will provide the lower bound of the randomness in the key.

With the above two assumptions, a latent variable is introduced to indicate which Gaussian component is chosen, and the DBN in Fig. 6 can be simplified to Fig. 7. For instance, GMM for h_A^t is represented in a hierarchical fashion, where z_A^t is a multi-nominal label for the mixture component. h_A^t is conditionally Gaussian given $z_A^t = j$. Meanwhile, y_A^t is a linear combination of two independent Gaussian distributions: the j-th component that follows the Gaussian distribution and the white noise. To our best knowledge, this is the first general theoretical framework that can effectively consider all of the information leakage caused by the broadcasting nature of wireless channels.

$$\begin{cases} \Pr[z_A^t = j] = w_{A,j}^t, j \in \{1, \ldots, N_A\}, \\ \Pr[h_A^t] = \sum_{j=1}^{N_A} \Pr[h_A^t | z_A^t = j] \Pr[z_A^t = j] \\ \qquad := \sum_{j=1}^{N_A} w_{A,j}^t \mathcal{N}\Big(\mu_{A,j}^t,\ (\sigma_{A,j}^t)^2\Big), \\ \Pr[y_A^t] = \Pr[h_A^t + n_0] \\ \qquad := \sum_{j=1}^{N_A} w_{A,j}^t \mathcal{N}\Big(\mu_{A,j}^t,\ (\sigma_{A,j}^t)^2 + \sigma^2\Big), \end{cases} \tag{12}$$

where z_A^t is a multi-nominal label for the mixture component, h_A^t is conditionally Gaussian given $z_A^t = j$, and y_A^t is a linear combination of two independent Gaussian distributions.

B. ALGORITHMS FOR TASKS 1-3

Parameter Learning: To learn the network parameters (conditional probability distribution of each variable), the expectation–maximization (EM) algorithm is adopted to iteratively improve the maximum likelihood of the latent variables [47]. 1) Assume the latent variable follows the multinomial distributions. 2) Expectation (E) step creates a function for the expectation of the log-likelihood evaluated using the current estimate for the parameters. 3) The maximization (M) step computes parameters that maximize the expected log-likelihood found on the E step, and these parameter-estimates are then used to determine the distribution of the latent variables in the next E step. The EM iteration alternates between performing the E step and M step until convergence.

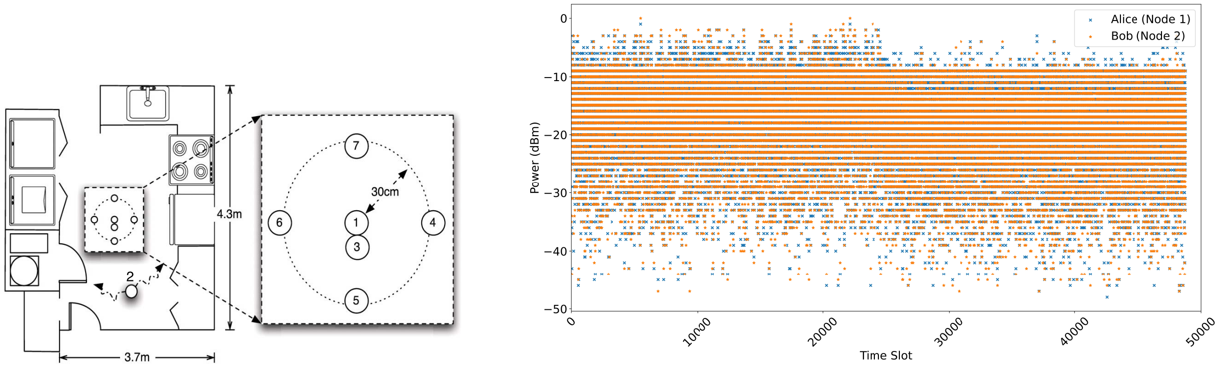
Marginal Inference: The junction tree algorithm will make probabilistic inference. The intuition is to use independence properties of the graph to decompose a global calculation on a joint probability into a linked set of local computations. The inference data structure named junction tree will elicit the relationship between locality and probabilistic inference. This algorithm will work with any probabilistic graphical model by generalizing variable elimination.

MAP Inference: The Annealed MAP algorithm [48] can find the most likely configuration of values of a set of nodes given observations of another subset of nodes. The Annealed MAP algorithm approximates the most probable sequence of states by simulated annealing [49]. While the solution is approximate, it performs well in practice and it gives an idea of the order of magnitude of the true maximum. The Annealed MAP algorithm drastically extends the class of MAP problems that can be solved.

V. SECURITY PERFORMANCE EVALUATION

For the point-to-point system with $M = 5$ eavesdroppers, RSSI has been collected by the sensor deployment as illustrated in Fig. 8 (a) for a duration of 48805 time slots. Nodes 1 and 2 are Alice and Bob, respectively. The remaining nodes 2-7 are eavesdroppers. After removing the missing values, the clean dataset contains 48779 consecutive RSSI values from 7 nodes. The RSSI value reported by each device has been quantized to a limited number of values, such as 48 unique dBm values in Fig. 8 (b).

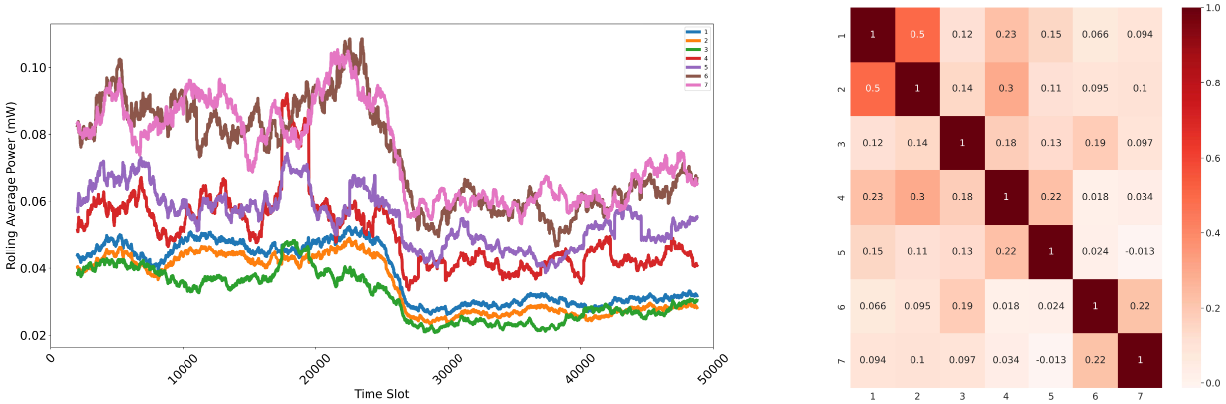
To better visualize each node's RSSI, Fig. 9 (a) presents the average of the RSSI points on either side of time t with a window size of 2000 time slots. Note that for the rolling average, RSSI in dBm has been converted to mW. The pairwise Pearson correlation coefficient in Fig. 9 (b) shows the linear relationship between Alice and Bob (Nodes 1 and 2) is the strongest with 0.5 correlation coefficient, which verifies the preliminary reciprocity of the signals received by Alice and Bob.



(a) Data Collection Procedure [40].

(b) Alice and Bob's RSSI.

FIGURE 8. Real-world Data.



(a) Rolling Average of RSSI (mW).

(b) Pair-wise Correlation.

FIGURE 9. RSSI of seven nodes in the system.

A. ASSUMPTION VERIFICATION OF THE REAL-WORLD DATA

To quantify the security performance of the physical layer key generated form the real-world RSSI sequences, the following properties have been verified so that they can fit the proposed DBN model.

Assumption 1 (Stationary distribution): Since RSSI collected by each node can be treated as a time series, the Augmented Dickey-Fuller (ADF) test has been conducted to determine how strongly a time series is defined by a trend. There are a number of unit root tests and ADF is one of the more widely used. It uses an autoregressive model and optimizes an information criterion across multiple different lag values. The results show that each node's RSSI is a stationary process.

Assumption 2 (Mixture Gaussian distribution): To verify the mixture Gaussian distribution of RSSI, the EM algorithm has been applied to Alice's received signal. The result has shown a one-dimensional Gaussian mixture model with seven components. The first panel of Fig. 10 shows the model selection criteria, Bayesian information criterion (BIC), as a function of the number of components. The BIC curve has an elbow point at the 7-component model. The second panel shows a histogram of the data, along with the best-fit model for a mixture with seven components. The third panel shows the probability that a given point is drawn from each class (component) as a function of its position, and the last panel shows the weight of each Gaussian component. Similar Gaussian mixture model can be applied to other nodes' RSSI measurements.

Assumption 3 (Markov property): To test the Markov assumption of each user's received RSSI sequence, the conditional characteristic function (CCF) of the current measurements given those taken in the past should be estimated. By dividing the Alice's RSSI sequence into multiple chunks, each chunk consists of 100 measurements. Using these chunks as learning data, the CCF test [50] conclude that Alice's RSSI measurement follows the 4-th order Markov property. For probabilistic graphical models, when the Markov assumption is not satisfied, the foundation of algorithms for the three tasks may be violated, thus leading to deterioration of their performance to different degrees. Consequently, in the following section when evaluating the minimum conditional entropy, the order of Markov property will be considered.

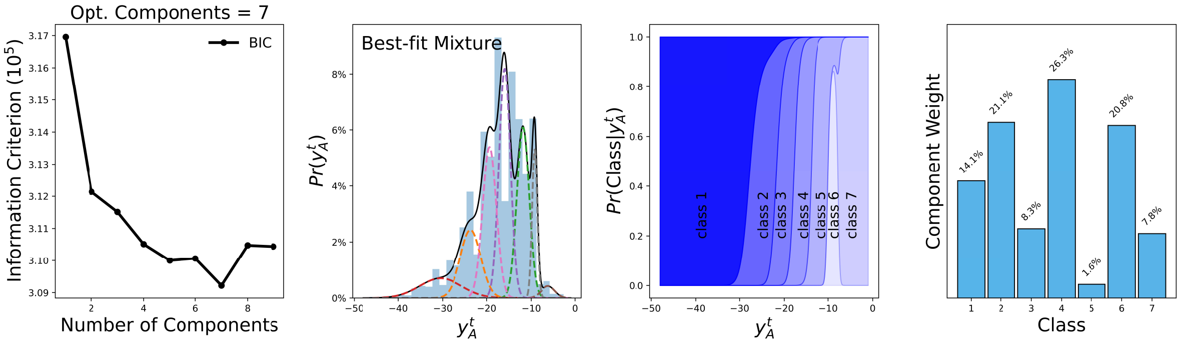
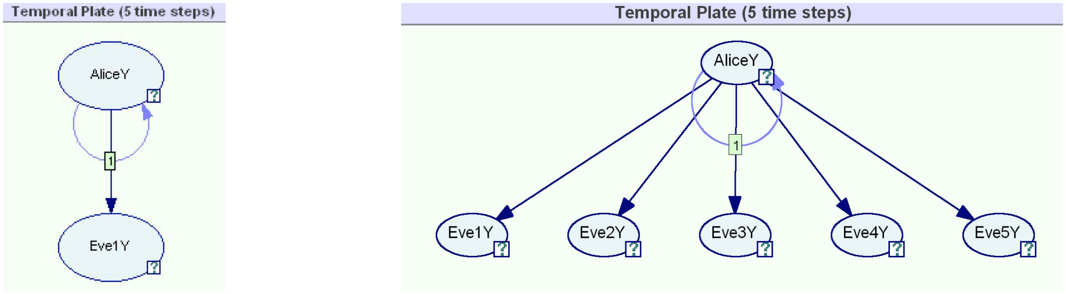


FIGURE 10. The mixture Gaussian decomposition of Alice's RSSI sequence (dBm).



(a) Alice and one eavesdropper [40].

(b) Alice and five eavesdroppers [42].

FIGURE 11. HMM based models (where 1 in the arc means first order Markov).

TABLE 1. Parameter learning performance.

Model	$M = 1$		$M = 5$	
	Validation	Test	Validation	Test
HMM	0.11	0.14	0.09	0.21
DBN	0.46	0.48	0.46	0.48

B. MODEL COMPARISONS

The three tasks for the conditional min-entropy estimation are conducted using the GeNIe Modeler and SMILE Engine,[1] which provides artificial intelligence modeling and machine learning software based on Bayesian networks. The performance of the proposed DBN model is compared with two existing models illustrated in Fig. 11: 1) the Hidden Markov model with Alice and one eavesdropper [40]; 2) the Hidden Markov model with Alice and five eavesdroppers [42].

Parameter Learning Accuracy: Suppose each user's RSSI sequence has been divided into chunks, each with $T = 5$ consecutive measurements. For the user-specific 9755 chunks, the parameter learning algorithm, i.e., EM algorithm, uses 90% as the training set with 10-fold cross validation, and the remaining 10% is used as the test set. The model training/validation procedure set y_A^t as target, and then learn the model parameters such that the target can be estimated based on the observed y_i^t.

The validation and training accuracy for y_A^0 is given in Table 1. For the HMM model with $M = 1$ eavesdropper, the validation and test performances are 0.11 and 0.14, respectively. That is, for the chunks in the validation dataset, the model can only accurately predicted 11% of the Alice's RSSI value at time $t = 0$, and the remaining 89% of chunks are predicted with error. The low accuracy for the test dataset demonstrates that the physical layer key generation pipeline can be very secure.

For the HMM model, the learning accuracy decreases with the number of eavesdroppers M because the state space for $y_A^t \times y_i^t$ increases from 48^2 to more than 48^6. Note that although Fig. 8 shows 48 distinct RSSI measurement values for Alice, some eavesdroppers have up to 54 distinct RSSI values (not included in the figure because of the space limit). With the same number of data records (9775), the parameter learning accuracy decreases when the state space increases.

As we can see, the proposed DBN model outperforms the two existing HMM based models, because the test accuracy is up to 48%. This shows that the physical layer-key generation pipeline is not as secure as though by previous works. When M increases, the DBN model accuracy degradation is much less obvious because the state space for z_A^t is much smaller. Instead of 48 states, the number of possible Gaussian components for Alice is 7 (Fig. 10) and the numbers of components for eavesdroppers are all less than 10 (not demonstrated in the figure because of the space limit). This concludes the DBN model can better fit the collected real-world data.

Minimum Conditional Entropy Estimation: For the system with Alice and one eavesdropper, among the collected

[1] BayesFusion LLC, http://www.bayesfusion.com/.

TABLE 2. Minimum conditional entropy estimates.

Model	HMM	DBN			
Evidence	Collected y_1^t		Collected z_1^t	All z_1^t	
Markov Order	1				4
$H_\infty(K\|O)$	14.94	14.39	10.72	10.25	8.56

9755 user-specific RSSI chunks, each unique eavesdropper's sequence of length $T = 5$ has been fed into the learned HMM model and DBN model as evidence. For each evidence chunk $y_1^t, t \in \mathcal{T}$, the marginal probability of the eavesdropper's RSSI observations and the maximum posterior probability of Alice's RSSI measurements have been calculated via the junction tree and annealed MAP algorithm. The summation of the product of marginal probability and maximum posterior probability are used to estimate the minimum conditional entropy, and the results are 14.94 and 14.39 for the HMM model and the proposed DBN model, respectively. The reduction in the entropy estimates proves that the DBN model can better evaluate the information leakage. Note that estimates with collected y_1^t in Table 2 are not the exact minimum conditional entropy because not all the possible evidence chunks are included in the 9755 chunks. When all of the possible evidence chunks are included, the product summation will increase and the entropy will drop for both HMM and DBN models.

For the DBN model, instead of using RSSI measurement as evidence, $z_1^t, t \in \mathcal{T}$ has been used as evidence as well. Since the state space of z_1^t is smaller, the entropy is reduced to 10.72. Similarly, when all of the possible z_1^t is considered, the minimum conditional entropy of DBN model with first order Markov property is reduced further to 10.25. Furthermore, when the Markov order is increased to 4, the entropy is reduced to 8.56. This concludes that the proposed DBN model can better quantify the information leakage and obtain more accurate security performance measurement for the physical layer key generation pipeline.

C. THE DBN-BASED SECURITY QUANTIFICATION FOR IOT NETWORKS

For IoT networks, the security performance of the physical layer-based key depends on the randomness inherent in the wireless channel. If the key generation pipeline is deployed in the stationary wireless sensor network, the slowly changing wireless channel characteristics will increase the order of the Markov property, because the consecutive probing signals received by each user, such as Alice, can be highly correlated, even identical. In this case, the proposed DBN model can still measure the security performance of the generated key because an edge between correlated vertices (from z_A^t to z_A^{t+1}) can be flexibly added into the model structure. However, the security performance will drop, as shown in Table 2 with the Markov order being 4, i.e., there are 4 directed edges from z_A^t to z_A^{t+1}, z_A^{t+2}, z_A^{t+3}, and z_A^{t+4}, respectively.

For the IoT system with high order Markov property, to guarantee the conditional min-entropy, one approach is downsampling the probing sequence [40], which can increase the time interval of time series being fed into the DBN model. The second approach relies on the privacy amplification scheme designed in Step 4 of Fig. 2. The first approach will take longer to generate one key, and the second approach requires the computing resources of the IoT devices. Nonetheless, the proposed model can guide whether the security enhancement schemes have reached the desired performance level.

VI. CONCLUSION

For IoT applications, there has been essentially very little understanding of the security level in the physical layer attribute-based keys, and yet they appear to be a promising authentication tool for security-sensitive applications with resource-constrained IoT devices. The fundamental to engineering/deploying IoT applications with security requirements may lie in understanding the security level along the key generation pipeline. In this paper, by rigorously quantifying the security metric of the physical layer key, we have taken various information leakage into consideration. As compared with the existing HMM based security performance evaluation models, the proposed DBN model can better quantify the information leakage with more flexibility and less data.

ACKNOWLEDGMENT

The authors thank Dr. Qiang Tang (The University of Sydney) for providing the real-world data and for useful discussions.

REFERENCES

[1] X. Sun and N. Ansari, "EdgeIoT: Mobile edge computing for the Internet of Things," *IEEE Commun. Mag.*, vol. 54, no. 12, pp. 22–29, Dec. 2016.

[2] N. Neshenko, E. Bou-Harb, J. Crichigno, G. Kaddoum, and N. Ghani, "Demystifying IoT security: An exhaustive survey on IoT vulnerabilities and a first empirical look on Internet-scale IoT exploitations," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2702–2733, 3rd Quart., 2019.

[3] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, and W. Zhao, "A survey on Internet of Things: Architecture, enabling technologies, security and privacy, and applications," *IEEE Internet Things J.*, vol. 4, no. 5, pp. 1125–1142, Oct. 2017.

[4] T. Saarikko, U. H. Westergren, and T. Blomquist, "The Internet of Things: Are you ready for what's coming?" *Bus. Horizons*, vol. 60, no. 5, pp. 667–676, 2017.

[5] L. Coetzee and J. Eksteen, "The Internet of Things—Promise for the future? An introduction," in *Proc. IST Africa Conf.*, Gaborone, Botswana, May 2011, pp. 1–9.

[6] "AWS Wavelength Project." Amazon. [Online]. Available: https://aws.amazon.com/wavelength/ (accessed Mar. 2020).

[7] X. Sun and N. Ansari, "Adaptive avatar handoff in the cloudlet network," *IEEE Trans. Cloud Comput.*, vol. 7, no. 3, pp. 664–676, Jul.–Sep. 2019.

[8] S. Wang, R. Urgaonkar, T. He, M. Zafer, K. Chan, and K. K. Leung, "Mobility-induced service migration in mobile micro-clouds," in *Proc. IEEE Mil. Commun. Conf.*, Baltimore, MD, USA, Oct. 2014, pp. 835–840.

[9] S. Vissicchio and L. Cittadini, "Safe, efficient, and robust SDN updates by combining rule replacements and additions," *IEEE/ACM Trans. Netw.*, vol. 25, no. 5, pp. 3102–3115, Oct. 2017.

[10] A. Manzalini, R. Minerva, F. Callegati, W. Cerroni, and A. Campi, "Clouds of virtual machines in edge networks," *IEEE Commun. Mag.*, vol. 51, no. 7, pp. 63–70, Jul. 2013.

[11] A. Kiani, N. Ansari, and A. Khreishah, "Hierarchical capacity provisioning for fog computing," *IEEE/ACM Trans. Netw.*, vol. 27, no. 3, pp. 962–971, Jun. 2019.

[12] Q. Fan and N. Ansari, "Workload allocation in hierarchical cloudlet networks," *IEEE Commun. Lett.*, vol. 22, no. 4, pp. 820–823, Apr. 2018.

[13] A. Kiani and N. Ansari, "Optimal code partitioning over time and hierarchical cloudlets," *IEEE Commun. Lett.*, vol. 22, no. 1, pp. 181–184, Jan. 2018.

[14] A. Kiani and N. Ansari, "Toward hierarchical mobile edge computing: An auction-based profit maximization approach," *IEEE Internet Things J.*, vol. 4, no. 6, pp. 2082–2091, Dec. 2017.

[15] O. Galinina, A. Pyattaev, S. Andreev, M. Dohler, and Y. Koucheryavy, "5G multi-RAT LTE-WiFi ultra-dense small cells: Performance dynamics, architecture, and trends," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 6, pp. 1224–1240, Jun. 2015.

[16] "Security in the Internet of Things lessons from the past for the connected future, white paper," Wind River Inc., Alameda, CA, USA, Rep. Accessed Jul. 2020. [Online]. Available: https://www.semanticscholar.org/paper/SECURITY-IN-THE-INTERNET-OF-THINGS-Lessons-from-the-Shipley/a15d5029f6901bfa36c3bd9e726385d01c8c28b5

[17] F. Meneghello, M. Calore, D. Zucchetto, M. Polese, and A. Zanella, "IoT: Internet of threats? A survey of practical security vulnerabilities in real IoT devices," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8182–8201, Oct. 2019.

[18] X. Huang and N. Ansari, "Content caching and distribution at wireless mobile edge," *IEEE Trans. Cloud Comput.*, early access, May 18, 2020, [Online]. Available: https://ieeexplore.ieee.org/document/9095382, doi: 10.1109/TCC.2020.2995403.

[19] S. Huang, X. Huang, and N. Ansari, "Budget-aware video crowdsourcing at the cloud-enhanced mobile edge," *IEEE Trans. Netw. Service Manag.*, vol. 18, no. 2, pp. 2123–2137, Jun. 2021.

[20] V. Mihajlović and M. Petkovic, "Dynamic Bayesian networks: A state of the art," CTIT, New York, NY, USA, Rep. TR-CTIT-34, vols. 1–34, Oct. 2001.

[21] S. Ito and T. Sagawa, *Information Flow and Entropy Production on Bayesian Networks*. Weinheim, Germany: Wiley, 2016, pp. 63–99, ch. 3.

[22] X. Wang, P. Hao, and L. Hanzo, "Physical-layer authentication for wireless security enhancement: Current challenges and future developments," *IEEE Commun. Mag.*, vol. 54, no. 6, pp. 152–158, Jun. 2016.

[23] K. Zeng, "Physical layer key generation in wireless networks: Challenges and opportunities," *IEEE Commun. Mag.*, vol. 53, no. 6, pp. 33–39, Jun. 2015.

[24] W. Xi *et al.*, "Instant and robust authentication and key agreement among mobile devices," in *Proc. ACM SIGSAC Conf. Comput. Commun. Security*, 2016, pp. 616–627.

[25] J. E. Hershey, A. A. Hassan, and R. Yarlagadda, "Unconventional cryptographic keying variable management," *IEEE Trans. Commun.*, vol. 43, no. 1, pp. 3–6, Jan. 1995.

[26] A. Badawy, T. Elfouly, T. Khattab, A. Mohamed, and M. Guizani, "Unleashing the secure potential of the wireless physical layer: Secret key generation methods," *Phys. Commun.*, vol. 19, pp. 1–10, Jun. 2016.

[27] Y. Dodis, L. Reyzin, and A. Smith, "Fuzzy extractors: How to generate strong keys from biometrics and other noisy data," *SIAM J. Comput.*, vol. 38, no. 1, pp. 97–139, Mar. 2008.

[28] L. Jiao, N. Wang, P. Wang, A. Alipour-Fanid, J. Tang, and K. Zeng, "Physical layer key generation in 5G wireless networks," *IEEE Wireless Commun.*, vol. 26, no. 5, pp. 48–54, Oct. 2019.

[29] J. Zhang, T. Q. Duong, A. Marshall, and R. Woods, "Key generation from wireless channels: A review," *IEEE Access*, vol. 4, pp. 614–626, 2016.

[30] D. Kreiser *et al.*, "On wireless channel parameters for key generation in industrial environments," *IEEE Access*, vol. 6, pp. 79010–79025, 2018.

[31] J. Wan, A. Lopez, and M. A. A. Faruque, "Physical layer key generation: Securing wireless communication in automotive cyber-physical systems," *ACM Trans. Cyber-Phys. Syst.*, vol. 3, no. 2, pp. 1–26, Oct. 2018.

[32] J. Zhang, S. K. Kasera, and N. Patwari, "Mobility assisted secret key generation using wireless link signatures," in *Proc. IEEE INFOCOM*, San Diego, CA, USA, Mar. 2010, pp. 1–5.

[33] O. Georgiou, C. P. Dettmann, and J. P. Coon, "Network connectivity: Stochastic vs. deterministic wireless channels," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Sydney, NSW, Australia, Jun. 2014, pp. 77–82.

[34] P. Sadeghi, R. A. Kennedy, P. B. Rapajic, and R. Shams, "Finite-state Markov modeling of fading channels—A survey of principles and applications," *IEEE Signal Process. Mag.*, vol. 25, no. 5, pp. 57–80, Sep. 2008.

[35] Z. Yang, L. Zhou, G. Zhao, and S. Zhou, "Blockage modeling for inter-layer UAVs communications in urban environments," in *Proc. 25th Int. Conf. Telecommun. (ICT)*, Saint-Malo, France, Jun. 2018, pp. 307–311.

[36] L. Lai, Y. Liang, H. V. Poor, and W. Du, *Key Generation From Wireless Channels*. Boca Raton, FL, USA: CRC Press, Nov. 2013, pp. 47–68.

[37] R. F. Schaefer, A. Khisti, and H. V. Poor, "Secure broadcasting using independent secret keys," *IEEE Trans. Commun.*, vol. 66, no. 2, pp. 644–661, Feb. 2018.

[38] R. F. Schaefer, A. Khisti, and H. V. Poor, "How to use independent secret keys for secure broadcasting of common messages," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Hong Kong, Jun. 2015, pp. 1971–1975.

[39] H. V. Poor and R. F. Schaefer, "Wireless physical layer security," *Proc. Nat. Acad. Sci.*, vol. 114, no. 1, pp. 19–26, 2017.

[40] M. Edman, A. Kiayias, Q. Tang, and B. Yener, "On the security of key extraction from measuring physical quantities," *IEEE Trans. Inf. Forensics Security*, vol. 11, pp. 1796–1806, 2016.

[41] X. Guo, N. Ansari, F. Hu, Y. Shao, N. R. Elikplim, and L. Li, "A survey on fusion-based indoor positioning," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 1, pp. 566–594, 1st Quart., 2020.

[42] X. Wang, Y. Hou, X. Huang, D. Li, X. Tao, and J. Xu, "Security analysis of key extraction from physical measurements with multiple adversaries," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Kansas City, MO, USA, May 2018, pp. 1–6.

[43] P. S. Rossi, D. Ciuonzo, K. Kansanen, and T. Ekman, "Performance analysis of energy detection for MIMO decision fusion in wireless sensor networks over arbitrary fading channels," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7794–7806, Nov. 2016.

[44] O. Alhussein, B. Selim, T. Assaf, S. Muhaidat, J. Liang, and G. K. Karagiannidis, "A generalized mixture of Gaussians for fading channels," in *Proc. IEEE 81st Veh. Technol. Conf. (VTC Spring)*, Glasgow, U.K., 2015, pp. 1–6.

[45] B. Selim, O. Alhussein, S. Muhaidat, G. K. Karagiannidis, and J. Liang, "Modeling and analysis of wireless channels via the mixture of Gaussian distribution," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 8309–8321, Oct. 2016.

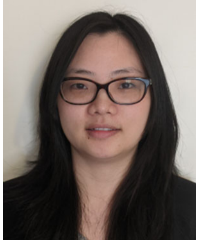
[46] J. B. Perazzone, P. L. Yu, B. M. Sadler, and R. S. Blum, "Physical layer authentication via fingerprint embedding: Min-entropy analysis : Invited presentation," in *Proc. 53rd Annu. Conf. Inf. Sci. Syst. (CISS)*, Baltimore, MD, USA, Mar. 2019, pp. 1–6.

[47] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *J. Roy. Stat. Soc. Ser. B, Methodol.*, vol. 39, no. 1, pp. 1–22, 1977.

[48] C. Yuan, T.-C. Lu, and M. J. Druzdzel, "Annealed map," in *Proc. 20th Conf. Uncertainty Artif. Intell.*, 2004, pp. 628–635.

[49] N. Ansari and E. Hou, *Computational Intelligence for Optimization*. Boston, MA, USA: Springer Publ. Company, Incorp., 1997.

[50] C. Shi, R. Wan, R. Song, W. Lu, and L. Leng, "Does the Markov decision process fit the data: Testing for the Markov property in sequential decision making," in *Proc. 37th Int. Conf. Mach. Learn.*, 2020, pp. 1–11.



XUEQING HUANG (Member, IEEE) received the B.E. degree in communications engineering from the Hefei University of Technology, Hefei, the M.E. degree in information and communication engineering from the Beijing University of Posts and Telecommunications, Beijing, and the Ph.D. degree in electrical engineering from the New Jersey Institute of Technology, Newark, NJ, USA. She is an Assistant Professor of Computer Science with the New York Institute of Technology. Her research interests include mobile edge computing, with current emphases on resources allocation, and security schemes for IoT applications.

NIRWAN ANSARI (Fellow, IEEE) received the B.S.E.E. degree (*summa cum laude* with a perfect GPA) from the New Jersey Institute of Technology (NJIT), the M.S.E.E. degree from the University of Michigan, and the Ph.D. degree from Purdue University.

He is a Distinguished Professor of Electrical and Computer Engineering with NJIT. He has authored *Green Mobile Networks: A Networking Perspective* (Wiley-IEEE, 2017) with T. Han, and coauthored two other books. He has also (co)authored more than 600 technical publications. He has also been granted more than 40 U.S. patents. His current research focuses on green communications and networking, cloud computing, drone-assisted networking, and various aspects of broadband networks.

Dr. Ansari received several excellence in teaching awards, a few best paper awards, the NCE Excellence in Research Award, several ComSoc TC technical recognition awards, the NJ Inventors Hall of Fame Inventor of the Year Award, the Thomas the Alva Edison Patent Award, the Purdue University Outstanding Electrical and Computer Engineering Award, NCE 100 Medal, the NJIT Excellence in Research Prize and Medal, and designation as a COMSOC Distinguished Lecturer. He has guest-edited a number of special issues covering various emerging topics in communications and networking. He has served on the editorial/advisory board of over ten journals, including as an Associate Editor-in-Chief of *IEEE Wireless Communications Magazine*. He was elected to serve in the IEEE Communications Society (ComSoc) Board of Governors as a member-at-large, has chaired some ComSoc technical and steering committees, is currently the Director of ComSoc Educational Services Board, has been serving in many committees, such as the IEEE Fellow Committee and has been actively organizing numerous IEEE International Conferences/Symposia/Workshops. He is frequently invited to deliver keynote addresses, distinguished lectures, tutorials, and invited talks. He is also a Fellow of National Academy of Inventors.

SIQI HUANG (Student Member, IEEE) received the B.E. degree in software engineering from Sun Yat-sen University in 2015. He is currently pursuing the Ph.D. degree with the University of North Carolina at Charlotte. His research interests include mobile edge computing and mobile augmented/virtual reality (AR/VR).

WENJIA LI (Senior Member, IEEE) received the Ph.D. degree in computer science from the University of Maryland Baltimore County, Baltimore, MD, USA, in 2011.

In 2014, he joined the Department of Computer Science, New York Institute of Technology (NYIT), New York, NY, USA, as a Tenure-Track Assistant Professor and has been a Tenured Associate Professor since September 2020. Prior to joining NYIT, he was an Assistant Professor of Computer Science with Georgia Southern University, Statesboro, GA, USA, from 2011 to 2014. He has authored or coauthored over 80 peer-reviewed publications in various journals and conference proceedings. His research has been supported by the National Institute of Health and the U.S. Department of Transportation Region 2 University Transportation Research Center. His current research interest include cyber security, mobile computing, and wireless networking, particularly security, trust, and policy issues for wireless networks, cyber–physical systems, Internet of Things, and intelligent transportation systems. He was a recipient of the 2019 IEEE Region 1 Technological Innovation (Academic) Award. He has served in the Organizing Committee of many international conferences, such as ACM WiSec, IEEE MDM, IEEE IPCCC, and IEEE Sarnoff, and also served as a Reviewer for many prestigious journals, such as IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING, and IEEE INTERNET OF THINGS JOURNAL.