Boosting Secret Key Generation for IRS-Assisted Symbiotic Radio Communications

Yang Liu[*][†][‡], Meng Wang[‡], Jing Xu[§], Shimin Gong[‡], Dinh Thai Hoang[¶], and Dusit Niyato[||]
[*]Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China
[†]University of Chinese Academy of Sciences, China
[‡]School of Intelligent Systems Engineering, Sun Yat-sen University, China
[§]School of Electronic Information and Communications, Huazhong University of Science and Technology, China
[¶]School of Electrical and Data Engineering, University of Technology Sydney, Australia
[||]School of Computer Science and Engineering, Nanyang Technological University, Singapore

Abstract—**Symbiotic radio (SR) has recently emerged as a promising technology to boost spectrum efficiency of wireless communications by allowing reflective communications underlying the active RF communications. In this paper, we leverage SR to boost physical layer security by using an array of passive reflecting elements constituting the intelligent reflecting surface (IRS), which is reconfigurable to induce diverse RF radiation patterns. In particular, by switching the IRS's phase shifting matrices, we can proactively create dynamic channel conditions, which can be exploited by the transceivers to extract common channel features and thus used to generate secret keys for encrypted data transmissions. As such, we firstly present the design principles for IRS-assisted key generation and verify a performance improvement in terms of the secret key generation rate (KGR). Our analysis reveals that the IRS's random phase shifting may result in a non-uniform channel distribution that limits the KGR. Therefore, to maximize the KGR, we propose a heuristic scheme and also a deep reinforcement learning (DRL) framework to control the switching of the IRS's phase shifting matrices. Simulation results show that the DRL approach for IRS-assisted key generation can significantly improve the KGR.**

I. Introduction

The concept of symbiotic radio (SR) emphases the dependence and mutual benefits among different radios, which have diverse transmission capabilities and resources constraints [1]. The design of SR systems aims to improve the overall network performance by exploiting the complement operations among different radios. An example of the SR system is the ambient backscatter communication networks, which allow passive reflecting devices to transmit private information along with the active RF communications [2]. The active radios can also benefit from the passive devices by reconfiguring the signal reflections in favor of the RF communications [3]. The benefit becomes more prevalent when the passive devices are equipped with a large array of reflecting elements, constituting an intelligent reflecting surface (IRS) [4], [5].

Currently, most of the studies on backscatter-aided or IRS-assisted SR systems focus on improving the energy- or spectrum-efficiency for wireless communications, e.g., [1]–[5]. In this paper, we intend to use the passive IRS to enhance physical-layer security of active RF communications in the SR system. Though the use of IRS for physical-layer security is not new, most of the existing works focus on an information-theoretic view by formulating a secrecy rate maximization problem, e.g., [6]–[8], which aims at improving the information rate between the legitimate transceivers while suppressing the information rate of the illegitimate eavesdroppers. This is often achieved by a joint optimization of the transmitter's beamforming matrix, the IRS's phase shifting matrix and the covariance matrix of artificial noise. However, the optimization framework typically relies on the channel information of both legitimate and illegitimate receivers. It also implies a fully cooperative IRS that is controllable by the legitimate transceivers or the base stations. Different from these works, we consider a more practical case in which the communications between legitimate transceivers are encrypted while the encryption process is vulnerable to the eavesdropping attack, which becomes our main problem in this paper.

Encrypted data transmission traditionally relies on a trustful infrastructure to distribute the secret keys to transceivers and to meet the security requirements, e.g., authenticity, confidentiality, integrity, and availability [9]. However, in a dynamic network, the exchange of secret keys between mobile devices becomes challenging as it is costly to maintain the trustful authority for key management. One promising way for decentralized SKG relies on extracting shared randomness in wireless channels between the transceivers. Due to channel reciprocity, the channel conditions sampled at two communicating transceivers are expected to be identical in one coherent time slot, while fluctuating randomly over different time slots [10], [11]. Therefore, ideally two transceivers will perceive the same sequence of channel samples. Most importantly, the sequence of channel samples can be viewed as location- or user-dependent RF fingerprint [12], which can be used to differentiate different receivers and thus to fight against the eavesdropping attack. By extracting the unique channel features at individual transceivers, we can encode the channel features into a bit stream and use it as the secret key for encrypted data transmission, e.g., [13]–[15].

The key generation rate (KGR) is limited by randomness in the channel conditions, which depends on the dynamics in the wireless environment. Therefore, it becomes impractical for stationary or slow-moving wireless devices. Though lots of existing works try to improve the KGR, they focus on the design of more complex transceiver structure or signal processing algorithms. The authors in [13] and [14] proposed to use multiple antennas to improve the KGR. The experimental results in [14] show that three-antenna transceivers can increase the KGR by more than four times over single-antenna systems. An adaptive SKG scheme is proposed in [15] to improve the KGR even with slow variations in the channel samples. Focusing on millimeter wave (mmWave) massive MIMO system, the authors in [16] allowed the directional mmWave beam to randomly rotate. This improves the channel randomness and therefore significantly improves the KGR for stationary receivers and co-located eavesdroppers. The authors in [17] proposed to generate secret keys by using reconfigurable aperture antennas. By randomly changing the antennas' load impedances, the signal reflections can create artificial fading and thus improve the KGR in static environment.

In this paper, we attempt to boost the KGR by using IRS to introduce artificial randomness into the wireless channels. The use of IRS for SKG firstly appears in [18], where a closed-form expression of the minimum secret key capacity is derived for an IRS-assisted wireless network. The authors in [18] also optimized the IRS's optimal phase shifting matrix to maximize the secret key capacity. A similar idea has been extended in a recent work [19], where a random switching scheme is proposed for IRS to disturb the wireless channels. Different from the random phase shifting schemes in [17] and [19], our analysis reveals that the SKG can be more efficient by devising a subtle phase control strategy. Our main contributions lie in the following aspects. Firstly we present the design principles for IRS-assisted SKG and verify an increased KGR by the random phase switching. Most importantly, to maximize the KGR, we further propose a heuristic scheme and also a deep reinforcement learning (DRL) framework to optimize the IRS's phase switching strategy, which have been verified via extensive simulation results.

II. System model

We consider encrypted data communications between legitimate transceiver in a vulnerable wireless environment. As shown in Fig. 1, the IRS is deployed in the wireless environment and can be exploited to prevent eavesdropping attack against illegitimate receivers, namely, the eavesdroppers, denoted as Eve or E, which can overhear the channels between the legitimate transceivers, denoted as Alice (or A) and Bob (or B) in Fig. 1, respectively. The secret keys for secure communications can be generated from the channel measurements at Alice and Bob. Assuming that Eve are separated from Alice and Bob at least one wave-length away, we can expect that Eve's channel measurements have a different pattern or distribution from that of Alice or Bob.

The unique channel features between Alice and Bob can be extracted and used for SKG, which can follow a similar time-slotted frame structure as that in [19]. Each time frame is divided into two parts, i.e., one part is for SKG and the other part is for encrypted data transmissions. The first part is further divided into multiple time slots allocated for exchanging the pilot packets between Alice and Bob. We require that the length of a time slot should be smaller than the channel coherent time. Based on channel reciprocity, the channel measurements of Alice and Bob within one time slot should be the same with high probability [10], [11]. This is the shared information between Alice and Bob while unknown to Eve. To avoid Eve's eavesdropping by, the channel measurements of Alice and Bob in each time slot should be fully randomized such that it becomes very unlikely for Eve to make a correct guess of the shared information.

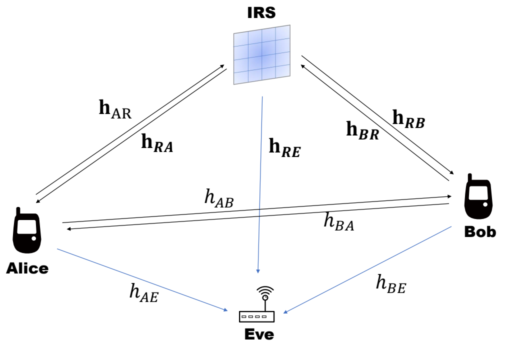


Fig. 1: Alice transmits encrypted data to Bob. The N-element IRS is deployed to prevent eavesdropping attack from Eves.

To increase the KGR, we leverage the N-element IRS (or R) to create a dynamic channel condition between Alice and Bob by switching the IRS's phase shifting matrix $\mathbf{\Theta} = \text{diag}(e^{j\theta_1}, e^{j\theta_2}, \ldots, e^{j\theta_N})$, where θ_n denotes the phase shift of individual reflecting element and $\text{diag}(\cdot)$ denotes a diagonal matrix. We assume that each reflecting element can be switched to B different states, where the constant B denotes the IRS's phase resolution. In each state $b \in \{1, 2, \ldots, B\}$, the phase shift is given by $\theta_n = \pi_b \in [0, 2\pi]$. Therefore, the IRS's phase shifting matrix $\mathbf{\Theta}$ can be chosen from a set $\mathbb{U}_{\mathbf{\Theta}}$ with the size $Q = B^N$. We have $Q = 2^N$ for a simple IRS implementation by using two-state diodes to control the phase shifts of individual reflecting elements.

Let h_{ab} (or $\mathbf{h}_{ab}$) denote the channel (or channel vector) from device-a to device-b where a, $b \in \{A, B, E, R\}$. We assume that each transceiver is equipped with a single antenna. In the t-th time slot, given the IRS's phase shifting matrix $\mathbf{\Theta}(t) \in \mathbb{U}_{\mathbf{\Theta}}$, the equivalent channels perceived by Alice and Bob are given as follows:

$$\begin{aligned} h_B(t) &= h_{AB} + \mathbf{h}_{RB}^H \mathbf{\Theta}(t) \mathbf{h}_{AR}, \\ h_A(t) &= h_{BA} + \mathbf{h}_{RA}^H \mathbf{\Theta}(t) \mathbf{h}_{BR}. \end{aligned} \tag{1}$$

Note that the pilot packets from Alice to Bob in the t-th time slot can also be overheard by Eve. Similar to (1), the perceived channels at Eve are given as follows:

$$\begin{aligned} h_{EA}(t) &= h_{AE} + \mathbf{h}_{RE}^H \mathbf{\Theta}(t) \mathbf{h}_{AR}, \\ h_{EB}(t) &= h_{BE} + \mathbf{h}_{RE}^H \mathbf{\Theta}(t) \mathbf{h}_{BR}. \end{aligned} \tag{2}$$

We denote $g_d(t) = ||h_d(t)||^2$ as the channel gain for $d \in \{A, B, EA, EB\}$, which can be easily sampled from the received signal strength. Note that $g_d(t)$ depends on the IRS's phase shifting matrix $\mathbf{\Theta}(t) \in \mathbb{U}_{\mathbf{\Theta}}$. By fast switching $\mathbf{\Theta}(t)$ in different time slots, we can create additional randomness in the channels perceived by Alice, Bob, and Eve. In the ideal case without measurement errors, the channel measurements at Alice and Bob are the same. Though Eve can overhear the pilot packets from Alice and Bob, the channel estimation $g_{EA}(t)$ (or $g_{EB}(t)$) will be very different from $g_B(t)$ (or $g_A(t)$) unless Eve is co-located with Alice or Bob.

III. Maximizing KGR for IRS-assisted Wireless Communications

In this part, we aim to maximize KGR by leveraging the IRS's phase reconfiguration to create shared randomness for the channels between Alice and Bob. We firstly present the general steps for SKG and then propose two phase configuration schemes to maximize the KGR.

A. General Steps for SKG

The secret key is a random bit stream shared between Alice and Bob and unknown to Eve, which is sampled, quantized, and encoded from the received signals at Alice and Bob. There are basically three steps to generate the shared secret keys based on the channel measurements [14].

1) Collecting Channel Information: A variety of channel measurements can be exploited as the unique feature for SKG, including channel impulse response, channel state information (CSI), and the received signal strength (RSS), i.e., $g_A(t)$ and $g_B(t)$, which can be easily extracted at the RF receiver. Note that the random variations in $g_A(t)$ and $g_B(t)$ will exhibit location- or user-dependent features. The deterministic components of $g_A(t)$ and $g_B(t)$, e.g., resulting from large-scale path loss, can be filtered out to extract the small-scale random variations. Without abuse of notations, we use $g_A(t)$ and $g_B(t)$ to denote the stochastic channel variations after filtering out the deterministic components. The same convention applies to Eve's channel measurements $g_{EA}(t)$ and $g_{EB}(t)$.

2) Quantizing Bit Streams: The channel reciprocity implies that the channel measurements $g_A(t)$ and $g_B(t)$ over different time slots should exhibit a similar dynamic behavior. To extract the shared features, we can quantize $g_A(t)$ and $g_B(t)$ into two sequences of bit streams. Due to fast channel fading, measurement or quantization errors, bit mismatch is usually unavoidable in practice, i.e., some bits generated by Alice are different from that generated by Bob. We expect a higher KGR for encrypted data communications if more matching bits can be found in two bit streams. To maximize the bit agreement ratio, we are required to design a proper quantization scheme that optimizes the number of quantization levels and the quantization intervals.

3) Extracting Shared Randomness: One purpose of this step is to remove the mismatched bits in two bit streams generated by Alice and Bob. This ensures that they share the same bit sequence and thus generate the same key for secure communications. The other purpose is to remove duplicated bits and ensure randomness in the shared bit sequence. When the channel coherence time is large, Alice and Bob may observe consecutive '0's or '1's in their bit sequences. By keeping the first '0' or '1' only, the shared bit sequence becomes randomized and unlikely to be guessed or estimated by the Eves. The authors in [14] achieved this purpose by allowing Alice and Bob to communicate and compare the start positions of consecutive '0's or '1's, namely, an excursion. Specifically, Alice firstly sends Bob a message containing the start positions of these excursions. Bob then checks its own bit stream at these positions. When observing the same excursions, Bob sends back Alice these positions to confirm the shared excursions. As such, Alice and Bob achieve a consensus on the excursions' start positions, which becomes the shared information for SKG.

B. IRS-Assisted SKG

The above procedures perform well for mobile users with dynamic channel conditions. However, the KGR becomes very limited for slow-moving or stationary users. This can be overcome by leveraging IRS to create artificial randomness into wireless channels. As shown in (1), by fast switching the IRS's phase shifting matrix $\mathbf{\Theta}(t)$, we can proactively create dynamic channel conditions and thus boost the KGR even in a stationary environment. To this end, the authors in [19] proposed the *random phase shifting* scheme for the IRS to operate independently from the legitimate transceivers. However, the random phase shifting scheme may not fully exploit the potential of IRS for SKG. The reasons can be explained from the following two aspects.

From an information-theoretic viewpoint, the maximum KGR is limited by the mutual information between Alice and Bob. Given the channel measurements $g_A(t)$ and $g_B(t)$, the mutual information can be numerically estimated by entropy following the procedures in [14]. To maximize the mutual information and therefore the estimated entropy, we need to exaggerate the randomness in the channel measurements. Observing from (1), the IRS-assisted channel gains are nonlinearly related to the IRS's phase shifting matrix $\mathbf{\Theta}(t)$, the direct and reflected channels. This implies that the IRS's random phase shifting may lead to non-uniform distributions of the channel measurements $g_A(t)$ and $g_B(t)$. This will result in an reduced entropy estimation and the KGR.

From the IRS's practical implementation, the feasible set of phase shifting matrices $\mathbb{U}_{\mathbf{\Theta}}$ can be very large, amounting to a huge number $Q = 2^N$ even for binary phase resolution. This implies ample redundancy in the IRS's phase reconfiguration. That is, lots of different phase shifting matrices may induce very similar channel gains. This implies that the random phase shifting scheme may lead to a small bit agreement ratio or *sample inefficiency*, i.e., we require more channel measurements to generate the secret key with a desirable size.

C. Phase Shifting Control Schemes

From the above observation, we conclude that the IRS's phase shifting should be able to introduce significant change to the channel gain each time when the IRS switches its phase shifting matrix. This motivates us to design new control algorithms for the IRS's phase shifting, instead of the uniformly random shifting scheme proposed in [17] and [19].

The most straightforward idea for phase control is to select a proper phase shifting matrix such that it can induce a significant change to the channel gains $g_A(t)$ and $g_B(t)$ in each time slot. To this end, we divide the key generation process into two phases. The first phase is for channel training, which is performed similarly to the random shifting scheme in [19]. This helps the IRS to build a mapping from its phase shifting matrices $\boldsymbol{\Theta}(t)$ to the channel gains $g_A(t)$ and $g_B(t)$. With this information, the IRS can optimize its phase shifting matrices in the second phase to improve the channel's randomness.

1) Channel Training: The channel training phase is further divided into T time slots. In each time slot, the IRS can randomly choose a different phase shifting matrix $\boldsymbol{\Theta}(t)$ similar to that in [19], and then the transceivers record the corresponding channel gains $g_A(t)$ and $g_B(t)$. After T time slots, the channel measurements are collected by the IRS controller to build a mapping from its phase shifting matrices to different channel gains. Note that we typically have $T \ll B^N$ and therefore the channel training does not necessarily enumerate all possible phase shifting matrices. Due to the IRS's phase shifting, the channel measurements can be highly fluctuating. Then, we can divide the channel measurements into $M \geq 2$ groups.

2) Uniform Multi-level Quantization: Theoretically, the number M of quantization levels is limited by the mutual information between $g_A(t)$ and $g_B(t)$, which can be approximated by the estimated entropy [14], calculated as $\ell = \sum_{g_A} p(g_A) \log_2 p(g_A)$, where $p(g_A)$ represents the empirical distribution of the channel measurement $g_A(t)$. Then, the maximum quantization level M will be bounded by $M \leq 2^{\ell}$. To maximize the mutual information, we expect that the occurrence frequencies in each quantization interval are comparable by controlling the IRS's phase shifting matrices. Considering the uniform M-level quantization, the i-th quantization interval is given by $\mathbb{I}_i = [g_{\min} + (i-1)\Delta_g, g_{\min} + i\Delta_g)$ for $i \in \{1, 2, \ldots, M\}$, where $\Delta_g = (g_{\max} - g_{\min})/M$, $g_{\min}$ and $g_{\max}$ denote the minimum and maximum channel gains, respectively. Then we can quantize each channel sample to a certain quantization level and encode it by a bit string.

3) Heuristic Phase Shifting Control: The number of quantization levels is limited by the mutual information while the channel training sequence can be very large. This implies a many-to-one mapping from different phase shifting matrices to a certain quantization level. That is, for the i-th quantization level, the IRS can determine a subset $\mathbb{U}_i$ of phase shifting matrices $\boldsymbol{\Theta}(t)$ that leads to similar channel measurements falling in the same quantization interval $\mathbb{I}_i$, i.e.,

$$\mathbb{U}_i = \{\boldsymbol{\Theta}(t) \in \mathbb{U}_{\boldsymbol{\Theta}} : g(t) \in \mathbb{I}_i\}, \quad \forall\, i = \{1, 2, \ldots, M\}.$$

Given the above multi-level quantization and the grouping of phase shifting matrices, we can devise the heuristic phase control algorithm following a two-step procedure. In the first step, the IRS randomly chooses a subset $\mathbb{U}_i$ of phase shifting matrices following a uniform distribution. This ensures that the channel measurements have equal occurrence frequencies in different quantization intervals. As such, we can maximize the mutual information by using less channel measurements, and thus increase the sample efficiency or the bit agreement ratio. In the second step, the IRS can randomly choose the phase shifting matrix within the set $\mathbb{U}_i$ to disturb the two-way channels between Alice and Bob. The KGR is expected to increase comparing to the random phase shifting scheme in [19]. For a special case with binary quantization, we can simply divide the channel measurements into two groups, e.g., a high and low RSS groups. Correspondingly, the IRS's phase shifting matrices can also be divided into two groups.

4) DRL for Phase Shifting Control: Besides the heuristic phase control, we also design a DRL algorithm to improve the sample efficiency by finding a minimal subset of phase shifting matrices to stir the wireless environment. DRL algorithm allows the IRS to observe the channel conditions, measure the channel randomness, and then make decisions on the selection of phase shifting matrices. This can be formulated into Markov Decision Process (MDP), denoted by a tuple $(\mathcal{S}, \mathcal{A}, \mathcal{R})$.

- The system state $\mathbf{s}_t \in \mathcal{S}$ includes a subset of phase shifting matrices $\mathbb{U}_{\boldsymbol{\Theta}}(t)$ currently in use, the corresponding channel measurements and the estimated entropy.
- The action $\mathbf{a}_t = [a_t(1), a_t(2), \ldots, a_t(M)] \in \mathcal{A}$ is a binary variable, indicating whether a phase shifting matrix can be used or not by the IRS in the next decision epoch. If $a_t(m) = 1$ (or 0), the m-th phase shifting matrix will be selected (or removed) in the next decision epoch. Here M denotes the maximum number of phase shifting matrices that can be used by the IRS. To ensure stable learning, we practically have $M \ll 2^N$ for binary phase resolution. In this case, we can uniformly sample M phase matrices from the feasible set $\mathbb{U}_{\boldsymbol{\Theta}}$. That is, the size of $\mathbb{U}_{\boldsymbol{\Theta}}(t)$ in the system state will be limited by M. As such, we can reduce the sizes of action and state spaces. We can also assume continuous phase control at the IRS. This can be optimized directly by the deep deterministic gradient policy (DDPG) algorithm [20].
- The immediate reward $r_t \in \mathcal{R}$ is evaluated by based on $\mathbf{s}_t$ and $\mathbf{a}_t$. It is proportional to the maximum mutual information that can be achieved by switching the IRS's phase shifting matrix $\boldsymbol{\Theta} \in \mathbb{U}_{\boldsymbol{\Theta}}(t)$. In particular, we can randomly generate a large sequence of phase shifting matrices from $\mathbb{U}_{\boldsymbol{\Theta}}(t)$ and record the corresponding channel measurements. The reward r_t is then numerically evaluated by the entropy of the channel measurements [14].

Based on a system state $\mathbf{s}_t \in \mathcal{S}$, the DRL agent refines its action $a_t \in \mathcal{A}$ for the IRS's phase control, leading to an immediate reward $r_t \in \mathcal{R}$.

IV. Simulation Results

In this part, we present simulation results to verify the capability of using IRS for secret key generation in wireless networks. We compare our algorithm with three benchmarks: (1) SKG without IRS, (2) IRS with fixed phase matrix, (3) IRS with random phase matrix as that in [19]. We consider

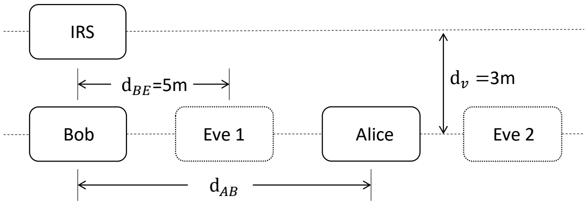
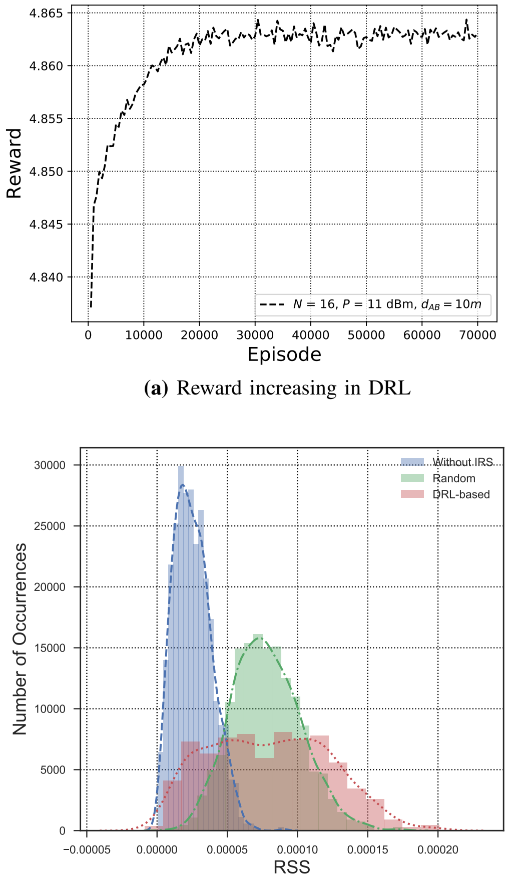


Fig. 2: Locations of legitimate transceivers and eves.

Uniform Linear Array (ULA) at IRS with $N = 16$ elements in the 2D plane as shown in Fig. 2. The vertical distance between Bob and IRS is set to $d_v = 3$ meters. The distance between Alice and Bob in simulation is set as $d_{AB} = 10$ meters. Two eavesdroppers appear on the horizontal line from Alice to Bob, denoted as Eve1 and Eve2. The path loss is given by $L = (L_0) + 10\alpha \log_{10}(d)$, where $L_0 = 30$ dB is the pass loss one meter away from the transmitter, and the path loss exponent is $\alpha = 5$. The small-scale fading of all involved channels follows a Rayleigh fading model. The noise power is set to $\delta_d^2 = -70$ dBm. We firstly verify the performance gain of the DRL algorithm, and then compare it with three benchmarks. The impacts of IRS's size and the its deploying location on the KGR are studied in the simulation.

To verify the DRL algorithm for KGR, we evaluate how the channel measurements at Alice and Bob change with the IRS's phase shifting control. We draw the reward value of the DRL algorithm in Fig. 3 (a), and show the distribution of channel measurements by using different SKG schemes in Fig. 3 (b). Fig. 3 (b) shows that the DRL algorithm can maximize the KGR by constantly trying to find the optimal IRS switching strategy. The RSS without using IRS follows the Rayleigh distribution. The random phase shifting scheme can approximate the RSS into Gaussian shape distribution, which leads to an improved KGR. Instead, by using the DRL-based phase control algorithm, the RSS approximately follow the uniform distribution, which implies an increased entropy estimation and therefore a higher KGR.

In Fig. 4, we reveal that an increasing transmit power of the pilot signal will increase the KGR. Besides, the DRL-based phase control is much better than the other schemes. We set the size of IRS as $N = 16$ and increase transmit power of Alice and Bob gradually. As shown in Fig. 4, the KGR increases with the transmit power and converges to a limiting high value. It is worth noticing that the KGR obtained by the DRL method is much higher than that of the other methods. This is because that the KGR is not only increased by a higher transmit power but also enhanced by the uniform RSS distribution due to the IRS's more intelligent phase control. This implies that the DRL method can fully exploit the advantages of the IRS's phase reconfigurability. We also noticed that the mutual information of Eve is very close to zero, which can be explained by the channel conditions being independent from that of Alice and Bob.



(a) Reward increasing in DRL

(b) Increased entropy estimation

Fig. 3: Comparing DRL-based scheme with other SKG schemes.

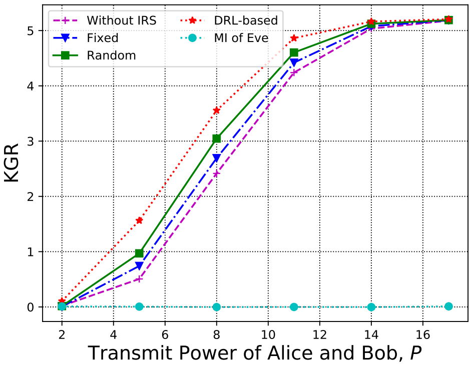


Fig. 4: IRS-assisted KGR with different transmit power.

In Fig. 5, we evaluate how the size of IRS effects the SKG. We expect that a larger size of IRS has better flexibility to stir the wireless environment and thus result in a higher KGR. To verify this, we gradually increase the size of IRS and record the KGR in Fig. 5. We can achieve a high mutual information by increasing the number of elements, however, the mutual

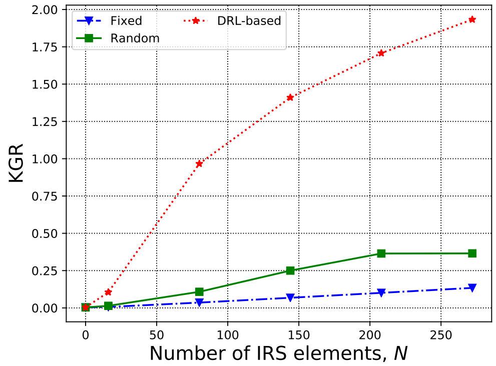


Fig. 5: IRS-assisted KGR with different size of IRS.

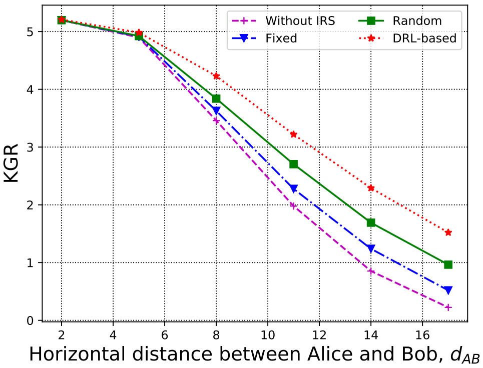


Fig. 6: IRS-assisted KGR with different communication distances.

information increases initially and then becomes saturated at a high level even the size of IRS still increases. This implies that the optimal size of IRS is limited. In Fig. 6, we evaluate the influence of the direct link quality on IRS-assisted KGR. Specifically, we vary the horizontal distance between Alice and Bob (i.e., d_{AB}) from 2 to 20 meters, and fix the transmit power of Alice and Bob as $P = 8$ dBm. Generally, the quality of the direct link becomes worse off as the distance between Alice and Bob increases. A worse off SNR at the receiver implies more mismatch bit streams in SKG, resulting a reduced KGR. However, for the DRL-based algorithm, we SKG process can be more robust in non-preferable channel conditions by exploiting the IRS's reconfigurability.

V. Conclusions

In this paper, we propose a DRL approach to boost the secret key generation problem in an IRS-assisted system. Different from the conventional random shifting scheme, we have designed a DRL-based phase control algorithm, which aims to reshape the RSS distribution at the receivers. The numerical results reveal that the RSS resulting from the DRL-based control algorithm approximately follows the uniform distribution, which can significantly increase the KGR.

References

[1] R. Long, Y. Liang, H. Guo, G. Yang, and R. Zhang, "Symbiotic radio: A new communication paradigm for passive internet of things," *IEEE Internet of Things Journal*, vol. 7, no. 2, pp. 1350–1363, 2020.

[2] Y. C. Liang, Q. Zhang, E. G. Larsson, and G. Y. Li, "Symbiotic radio: Cognitive backscattering communications for future wireless networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 4, pp. 1242–1255, 2020.

[3] S. Gong, Y. Zou, D. T. Hoang, J. Xu, W. Cheng, and D. Niyato, "Capitalizing backscatter-aided hybrid relay communications with wireless energy harvesting," *IEEE Internet of Things Journal*, vol. 7, no. 9, pp. 8709–8721, 2020.

[4] S. Gong, X. Lu, D. Hoang, D. Niyato, L. Shu, D. Kim, and Y. Liang, "Toward Smart Wireless Communications via Intelligent Reflecting Surfaces: A Contemporary Survey," *IEEE Commun. Surv. Tut.*, vol. 22, no. 4, pp. 2283–2314, 2020.

[5] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface aided wireless communications: A tutorial," *arXiv preprint arXiv:2007.02759*, 2020.

[6] X. Yu, D. Xu, and R. Schober, "Enabling secure wireless communications via intelligent reflecting surfaces," in *proc. IEEE GLOBECOM*, Dec. 2019.

[7] M. Cui, G. Zhang, and R. Zhang, "Secure wireless communication via intelligent reflecting surface," *IEEE Wireless Commun. Lett.*, vol. 8, pp. 1410–1414, Oct. 2019.

[8] X. Guan, Q. Wu, and R. Zhang, "Intelligent reflecting surface assisted secrecy communication: is artificial noise helpful or not?" *IEEE Wireless Commun. Lett.*, vol. 9, no. 6, pp. 778–782, Jun. 2020.

[9] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, "A survey on wireless security: Technical challenges, recent advances, and future trends," *Proc IEEE Inst. Electr. Electron. Eng.*, vol. 104, no. 9, pp. 1727–1765, 2016.

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

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

[12] D. Kreiser, Z. Dyka, S. Kornemann, C. Wittke, L. Kabin, O. Stecklina, and P. Langendörfer, "On wireless channel parameters for key generation in industrial environments," *IEEE Access*, vol. 6, pp. 79 010–79 025, 2017.

[13] J. Wallace and R. Sharma, "Automatic secret keys from reciprocal MIMO wireless channels: Measurement and analysis," *IEEE Trans. Inf. Forensics Secur.*, vol. 5, no. 3, pp. 381–392, 2010.

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

[15] S. Premnath, S. Jana, J. Croft, P. Gowda, M. Clark, S. Kasera, N. Patwari, and S. Krishnamurthy, "Secret key extraction from wireless signal strength in real environments," *IEEE Trans. Mob. Comput.*, vol. 12, no. 5, pp. 917–930, 2012.

[16] L. Jiao, N. Wang, and K. Zeng, "Secret beam: Robust secret key agreement for mmwave massive mimo 5g communication," in *proc. IEEE GLOBECOM*. Abu Dhabi, United Arab Emirates: IEEE, Dec. 2018, pp. 1–6.

[17] R. Mehmood and J. Wallace, "Wireless security enhancement using parasitic reconfigurable aperture antennas," in *proc. European Conf. Ant. Propag. (EUCAP)*. Rome, Italy: IEEE, Apr. 2011, pp. 2761–2765.

[18] Z. Ji, P. Yeoh, D. Zhang, G. Chen, Y. Zhang, Z. He, H. Yin, and Y. Li, "Secret key generation for intelligent reflecting surface assisted wireless communication networks," *arXiv preprint arXiv:2008.06304*, 2020.

[19] Z. Ji, P. L. Yeoh, G. Chen, C. Pan, Y. Zhang, Z. He, H. Yin, and Y. Li, "Random shifting intelligent reflecting surface for otp encrypted data transmission," *arXiv preprint arXiv:2010.14268*, 2020.

[20] T. Lillicrap, J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.