Safeguarding 5G Wireless Communication Networks Using Physical Layer Security

Nan Yang, Lifeng Wang, Giovanni Geraci, Maged Elkashlan, Jinhong Yuan, and Marco Di Renzo

ABSTRACT

The fifth generation (5G) network will serve as a key enabler in meeting the continuously increasing demands for future wireless applications, including an ultra-high data rate, an ultra-wide radio coverage, an ultra-large number of devices, and an ultra-low latency. This article examines security, a pivotal issue in the 5G network where wireless transmissions are inherently vulnerable to security breaches. Specifically, we focus on physical layer security, which safeguards data confidentiality by exploiting the intrinsic randomness of the communications medium and reaping the benefits offered by the disruptive technologies to 5G. Among various technologies, the three most promising ones are discussed: heterogenous networks, massive multiple-input multiple-output, and millimeter wave. On the basis of the key principles of each technology, we identify the rich opportunities and the outstanding challenges that security designers must tackle. Such an identification is expected to decisively advance the understanding of future physical layer security.

Nan Yang is with the Australian National University.

Lifeng Wang and M. Elkashlan are with Queen Mary University of London.

Giovanni Geraci is with the Singapore University of Technology and Design.

Jinhong Yuan is with the University of New South Wales.

Marco Di Renzo is with Paris-Saclay University, Laboratory of Signals and Systems (UMR-8506), CNRS - CentraleSupelec - University Paris-Sud XI, 91192 Gif-sur-Yvette (Paris), France.

INTRODUCTION

Mobile wireless communication has experienced an unprecedented growth in data traffic in recent years, spurred by the popularity of various intelligent devices, the demand for exuberant multimedia content, and the rapid increase in the number of base stations (BSs). In particular, global mobile data traffic in 2013 was nearly 18 times the size of the entire global Internet in 2000, and monthly global mobile data traffic by 2018 will surpass 15 exabytes [1]. While the mature third generation network and the currently deploying fourth generation (4G) network may accommodate the data traffic surge for the next few years, they will not be able to support a very large number of devices with a huge network traffic demand in 2020 and beyond [2]. Against this backdrop, a number of disruptive trends and technologies shaping the fifth generation (5G) network are emerging worldwide through research and development. For example, academia is researching robust and efficient wireless transmission technologies for the 5G era, such as the heterogeneous network (HetNet), massive multiple-input multiple-output (MIMO), and millimeter wave (mmWave). At the same time, the industry is undertaking 5G standardization. Given the ubiquitousness and necessity of 5G connections in the near future, an enormous amount of sensitive and confidential information, e.g. financial data, electronic media, medical records, and customer files, will be transmitted via wireless channels. Thus, providing an unrivalled security service is one of the top priorities in the design and implementation of the 5G network.

Despite the current efforts from academia and industry, the security paradigms protecting the confidentiality of wireless communication in the 5G network remain elusive. Indeed, how to secure wireless data transmission is one of the core problems that any 5G network designer can face. Differing from the traditional approach which protects data security through cryptographic techniques, physical layer security is identified as a promising strategy that provides secure wireless transmissions by smartly exploiting the imperfections of the communications medium. Using this strategy, 5G network designers can effectively degrade the quality of signal reception at unauthorized receivers and devices, and therefore prevent them from acquiring confidential information from the received signal. With careful planning and execution, physical layer security will protect the communication phase of the network while cryptography will protect the processed data after the communication phase. As such, they will form a well-integrated security solution that efficiently safeguards sensitive and confidential data for the 5G era.

Notably, physical layer security offers two major advantages compared to cryptography, making it particularly suitable for the 5G network. First, physical layer security techniques do not depend on computational complexity, which implies that the achieved level of security will not be compromised even if the unautho-

0163-6804/15/$25.00 © 2015 IEEE

rized smart devices in the 5G network have powerful computational capabilities. This is in contrast to computation-based cryptography, which is based on the premise that the unauthorized devices have insufficient computational capabilities for hard mathematical problems. Second, physical layer security techniques have a high scalability. In the 5G network, devices are always connected to the nodes with different powers and computation capabilities at the different levels of the hierarchical architecture. Also, devices always join in or leave the network at random time instants, due to the decentralized nature of the network. As a consequence, cryptographic key distribution and management become very challenging. To cope with this, physical layer security can be used to either provide direct secure data communication or facilitate the distribution of cryptographic keys in the 5G network.

Given the potential of physical layer security for the 5G era, the goal of this article is to identify the opportunities and challenges offered by the disruptive technologies enabling 5G for achieving a high security level at the physical layer. Among the various technologies, we focus on the three most promising ones, which we now describe in detail.

- **HetNet**: The HetNet creates a multi-tier topology where multiple nodes are deployed with dissimilar characteristics such as transmit powers, coverage areas, and radio access technologies. Obviously, it offers a rather provocative departure from the conventional single-tier wireless network and creates a new trend to reduce the cost per bit of future wireless connections. In such a trend, the full exploitation of the opportunities offered by the multi-tier topology, such as spatial modeling of nodes, association of mobile users, and direct connection between devices, is a core component in the design of physical layer security. This exploitation is discussed later.
- **Massive MIMO**: By deploying a very large number of antennas (e.g. a few hundred) at BSs to serve many tens of users simultaneously, massive MIMO reaps all the benefits offered by conventional MIMO, but on a much larger scale. To leverage the advantages of massive MIMO in physical layer security, some challenges need to be resolved during the design process, such as pilot contamination, power management, channel reciprocity, and eavesdropper-targeted signal processing. Motivated by this, we argue for physical layer security along with these challenges later.
- **mmWave**: As an innovative solution to meet the 5G's requirement, mmWave communication systems use a huge swath of spectrum, from 30 GHz to 300 GHz, to shift wireless transmissions away from the nearly fully occupied spectral band for current wireless networks. Notably, mmWave technologies have been standardized for short-range transmission, e.g. IEEE 802.11ad, as well as deployed for small cell backhaul, e.g. Siklu's Etherhaul 1200T. Since secure mmWave transmission is a completely new and promising research frontier, we advocate the potential of mmWave communication for physical layer security in a later Section.

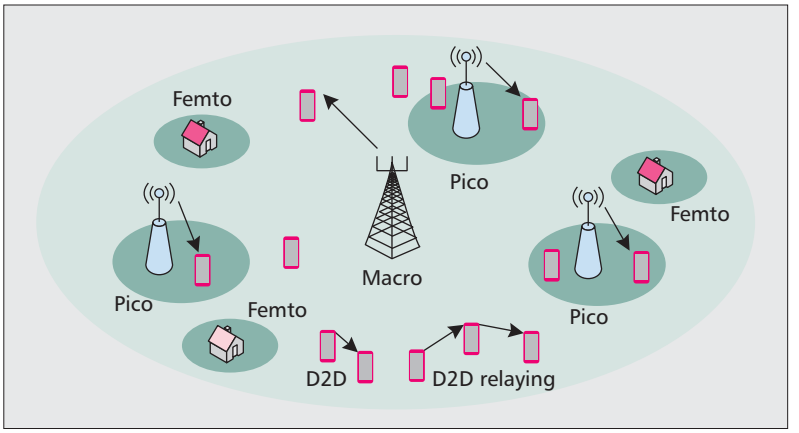

Figure 1. Heterogeneous network.

Physical Layer Security in Heterogeneous Networks

The HetNet is a promising network densification architecture in the 5G era. The aim of the HetNet is to provide a spectrum-efficient and energy-efficient solution that satisfies the dramatic growth in data demands of future wireless applications. In the HetNet, nodes with different transmit powers, coverage areas, and radio access technologies are deployed to form a multi-tier hierarchical architecture, as depicted in Fig. 1. Specifically, high-power nodes (HPNs) with large radio coverage areas are placed in *macro cell* tiers, while low-power nodes (LPNs) with small radio coverage areas are placed in *small cell* tiers. Small cells, such as pico cells and femto cells, are deployed under macro cell umbrellas to augment indoor coverage in highly populated buildings, and multi-tenant dwelling units, enterprises, and outdoor coverage in dense urban, suburban, or rural areas. In addition to the macro cell and small cell tiers that support HPN-to-device and LPN-to-device communications, respectively, the HetNet also involves a *device* tier that supports device-to-device (D2D) communications. The D2D communication allows geographically close devices to directly connect and interact with each other without using HPNs/LPNs, thus being a powerful enabler of low-latency and high-throughput data applications. Among multiple tiers, different radio access technologies such as WCDMA, LTE, WiMAX, and WLAN are adopted to provide various communication services. Therefore, the HetNet is a clearly different paradigm from conventional macrocell-only wireless networks. Obviously, the current physical layer security technologies for the direct single-user and multi-user transmissions [3, 4] and the relay-aided communications [5] in conventional networks cannot be readily applied in this paradigm. It follows that the compelling potential of the HetNet will trigger a new wave of innovation — in terms of spatial modeling, mobile association, and device connection — in securing multi-tier communications. These innovations are detailed in the following subsections.

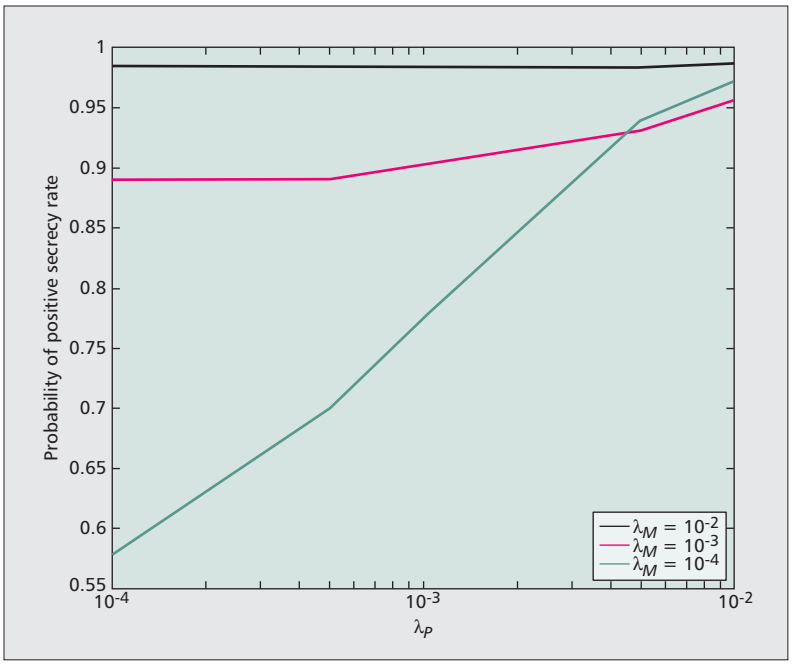

Figure 2. The probability of positive secrecy rate in a two-tier network where a macro cell tier is overlaid with a pico cell tier. The marco and pico cells are assumed to share the same frequency band. The locations of macro cell HPNs and pico cell LPNs follow independent homogeneous Poisson point processes (HPPPs) with densities λ_M and λ_P, respectively. The location of eavesdroppers also follows a HPPP with density $\lambda_E = 8 \times 10^{-4}$. User association is based on the maximal average received signal-to-noise ratio.

Spatial Modeling

In the HetNet, different spatial modeling of HPNs' and LPNs' locations raises a natural question: "*How does the spatial modeling of nodes' locations affect and guide the physical layer security design?*" The rationale behind this question is that a HPN's location is currently modeled as a point at the center of a hexagonal grid, while LPNs' locations can be modeled as a uniform distribution, in particular a Poisson point process in the two-dimensional plane [6]. Evidently, the deterministic model for HPNs' locations provides no randomness, whereas the Poisson model for LPNs' locations corresponds to complete randomness. This necessitates different mathematical tools to be employed to accommodate the nature and properties of the two models into the design of physical layer security.

Due to the deterministic nature of the model, the impact of HPNs' locations on physical layer security can be evaluated by applying system-level simulations to approximate the distributions of signal-to-interference-plus-noise ratios (SINRs) and corresponding quality of service (QoS) parameters. When small cells are deployed as add-ons, however, the amount of interference grows dramatically. Accordingly, the complexity of simulations substantially increases, making performance evaluation and optimization more complex and time-intensive. More recently, the HPNs' locations have also been modeled as a PPP to gauge the degree of randomness [7, 8]. Against this background, network security designers need to develop effective methods rooted in probability processes and order statistics to characterize the SINR distributions and the QoS parameters from the theoretical perspective. In particular, the generalized methods characterizing channel dynamics such as signal fading, spatial correlation, practical path loss, random channel errors, and mobility-induced channel variations are of paramount importance since they reduce the simulation burden to the minimum. The development of the generalized methods will allow for accurate capture of the multi-tier structure of the HetNet.

The effect of HPNs' and LPNs' locations on physical layer security can be examined based on the knowledge from the fields of graph theory and stochastic geometry. This motivates network security designers to develop effective mathematical tools from these two areas such that the SINR distributions and the QoS parameters impacted by HPNs and LPNs can be characterized. Of course, the characterization needs to be as general as possible, which enables the unequivocal establishment of secure connectivity and the accurate assessment of secrecy capacity in the HetNet with arbitrary transmit powers and densities of HPNs and LPNs. Moreover, new solutions to the SINR distributions are required if an appropriate level of practical correlation is introduced into the placement of HPNs and LPNs. Such solutions will force a substantial advancement over the current studies relying on the assumption of independent placement of HPNs/LPNs.

Figure 2 evaluates the impact of HPNs' and LPNs' densities on the secrecy performance via simulations. In this figure we show the probability of positive secrecy rate in a two-tier network where a macro cell tier is overlaid with a pico cell tier. It is evident that the probability of positive secrecy rate[1] increases as the density of pico cell LPNs increases. Moreover, it is observed that if the density of pico cell LPNs increases beyond a critical point, a higher density of macro cell HPNs does not improve the secrecy performance any further. Therefore, Fig. 2 provides a guide for network security designers to decide the best density for the implementation of HPNs and LPNs in the HetNet. Of course, as previously discussed, the development of effective mathematical tools will enable us to undertake the evaluation involved in this figure in a computationally efficient manner.

Mobile Association

Associating mobile users with HPNs and LPNs leads to a challenging and promising question: "*What is the optimal strategy for users to select HPNs/LPNs under security constraints?*" In traditional macro-cell-only cellular networks, it is typically assumed that mobile users select the strongest HPN to connect such that the best channel quality with the highest SINR is obtained. Accordingly, the physical layer security technologies in the open literature are designed based on this assumption. However, in the multi-tier HetNet, such a selection causes a load balancing problem. This is due to the fact that the HPNs with high transmit power and large coverage areas are often fully loaded or even "over" loaded, whereas the LPNs with low transmit

[1] Here, the probability of positive secrecy rate reveals the probability that the secrecy rate is higher than zero, where the secrecy rate can be characterized by the difference between the capacity of the main channel and the capacity of the eavesdropper's channel.

power and small coverage areas are often very lightly loaded [6]. Such an unbalanced load is detrimental to the ubiquitous applications of real-time services with stringent delay constraints and high power consumption, e.g. streaming video and gaming. As such, the unbalanced load should be addressed in the design of physical layer security.

In order to secure transmission and overcome the unbalanced load problem, new security-oriented mobile association policies are required to monitor and balance the instantaneous load of HPNs and LPNs. In designing these policies, the optimization of secrecy performance, e.g. the secrecy rate and the secrecy outage probability, should be prioritized. Under this prioritization, some intelligent mobile association policies can be developed such that the mobile users are wisely assigned to some HPNs or LPNs based on the achievable secrecy performance, the instantaneous load, and other factors such as the transmit power, coverage area, and density of HPNs/LPNs. Considering that such intelligent and optimal policies would impose a high computational complexity, some simple yet suboptimal mobile association policies are required to achieve a complexity-quality tradeoff. Aided by these suboptimal policies, the near-optimal secrecy performance is guaranteed with a lower computational cost. In addition, the cooperation among HPNs and LPNs offers a feasible way to enhance the secrecy performance. To explore this feasibility, network security designers should develop new cooperative strategies to allow neighboring HPNs/LPNs to exchange the secure data for mobile users, the instantaneous load of themselves, and other factors of the network with each other for achieving close-to-maximum secrecy performance.

Device Connection

The introduction of D2D communication triggers a pertinent question on the security issue: "*How to protect data confidentiality between connected devices against data leakage?*" Doubtlessly, maintaining data security is an essential task in D2D communications since the transmitted data between connected devices may be overheard by all of the surrounding devices. This task becomes more arduous particularly given the fact that the connected devices may not be able to handle complex signal processing algorithms as HPNs and LPNs do. One possible solution to tackle this task is *closed access* [9], where the intended device has a list of "trusted" devices. In closed access, the non-listed devices can only communicate with the intended device by getting authenticated in the macro cell or the micro cell tier. Hence, the establishment of closed access safeguards the data exchange between the intended device and the "trusted" devices against eavesdropping.

It is worth noting that closed access may not always be implemented, due to the lack of authentication in the macro cell or the micro cell tiers. In this case, referred to as *open access*, not only surrounding devices but geographically close HPNs and LPNs may act as potential eavesdroppers for the connected devices, meaning that they benefit from listening to the transmitted data and pose an acute threat to data security. To address security issues in open access, network designers need to construct new secure data exchange strategies that fully consider the physical characteristics of unintended devices and malicious HPNs/LPNs, e.g. ambiguous location, uncertain mobility, and unknown configuration. In addition, the potential attacks and threats induced by unintended devices and malicious HPNs/LPNs need to be carefully analyzed and incorporated into the construction.

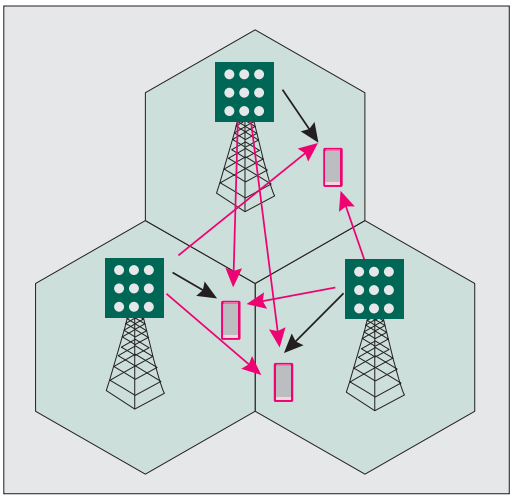

Figure 3. Cellular network with the deployment of massive MIMO.

Apart from direct D2D connections, another interesting paradigm in D2D communications is D2D relaying [10] where a device having better geometry to the transmitter device may act as a relay for the receiver device. The physical layer security design in this paradigm should exploit cooperative spatial diversity to maximize the secrecy performance. Despite the current relay-aided physical layer security techniques, such a design introduces new security problems to be solved. For example, the optimal selection of candidate relays need to be determined and the protection against untrusted relays needs to be investigated. Furthermore, if multiple devices are required for relaying the data between the connected devices, the impact of multi-hop coordination on the secrecy performance needs to be examined.

Physical Layer Security in Massive MIMO Systems

Massive MIMO systems are emerging as a new research field and have attracted significant interest from both scientists and industrialists. The benefits of the massive MIMO technique are realized by using very large antenna arrays (typically tens or even hundreds) at the transmitter and/or the receiver. In future cellular networks with massive MIMO, as depicted in Fig. 3, the number of antenna arrays at the BSs is much larger, e.g. 10 times, than the number of data streams served to all users in a cell [11]. Compared to the current counterpart, massive MIMO systems provide high power and spectrum effi-

Massive MIMO systems are emerging as a new research field and have attracted substantial interests from both scientists and industrialists. The benefits of the massive MIMO technique are realized by using very large antenna arrays (typically tens or even hundreds) at the transmitter and/or the receiver.

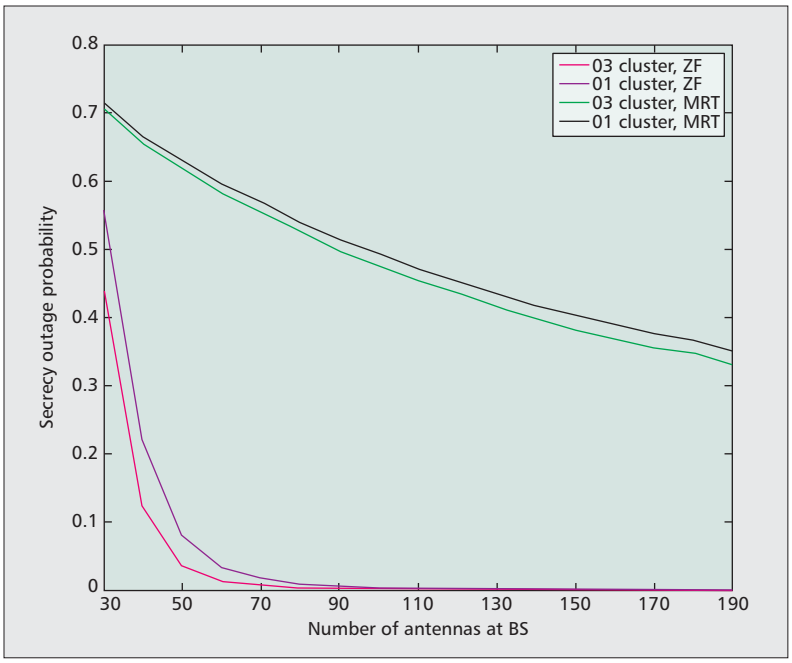

Figure 4. The secrecy outage probability of downlink with MRT and ZF precoding techniques at the BS. We consider three hexagonal cells without sectorization, where the radius of each cell is 350 meters. In each cell, seven single-antenna users are uniformly distributed and one of them is an eavesdropper. In this figure "01 cluster" means that the frequency reuse pattern is 1 while "03 cluster" means that the frequency reuse pattern is 3.

ciencies by exploiting the large arrays gain offered by low-complexity transmission designs. Moreover, random impairments such as small-scale fading and noise can be averaged out when a large number of antennas are deployed at the BS [12]. Furthermore, the interference, channel estimation errors, and hardware impairments [13] vanish when the number of antennas grows large, leaving only pilot contamination as the performance limit [14].

Given the fact that massive MIMO will serve as an essential enabling technology for the 5G wireless network, we next discuss the design of physical layer security based on the tremendous potential of massive MIMO systems. Needless to say, such a design opens a new and promising research avenue, extending current research efforts in conventional MIMO systems to a new area.

Low Power Consumption

In massive MIMO systems, the secrecy performance can be remarkably enhanced by adopting a reduced power consumption. The enhancement is attributed to two major causes. First, since the transmit power level is cut, the receive signal-to-noise ratios (SNRs) at the eavesdroppers are highly reduced. This leads to a significant decrease in the eavesdroppers' channel capacities. Second, given the transmit power and the expected secrecy rate at the transmitter, the secrecy outage probability can be arbitrarily small when the number of antennas grows unbounded.[2] In Fig. 4 we show the secrecy outage probability for a rate threshold of 2 bits/s/Hz in the downlink of a three-hexagonal-cell network. We consider two commonly-used precoding methods: maximal ratio transmission (MRT) and zero-forcing (ZF) [11]. It is seen from this figure that ZF outperforms MRT because the intra-cell interference can be cancelled through ZF. It is also seen that the secrecy outage probability profoundly declines when the number of antennas at the BS grows large. As such, the minimum power consumption achieving the target secrecy performance level needs to be determined. In this determination, the development of new and powerful mathematical tools, e.g. random matrix theory, will eliminate the burden of performance evaluation incurred by time-consuming simulations.

[2] Eavesdroppers are typically passive to hide their existence. As such, the secrecy outage is addressed as a principal concern of security.

Time Division Duplex Operation

Massive MIMO systems are recommended to operate in a time division duplex (TDD) mode [11], which is different from conventional MIMO systems that generally operate in a frequency division duplex (FDD) mode. This is due to the fact that the channel training overhead in the FDD mode scales linearly with the number of transmit antennas, which in turn imposes a severe limit on the number of antennas. In the TDD mode, the training burden is independent of the number of BS antennas and channel reciprocity is exploited. In TDD massive MIMO systems, eavesdroppers may experience particular difficulties for wiretapping, because downlink pilot signals from the BS to the users are not required in the TDD mode. Specifically, the BS with massive antenna arrays obtains the uplink channel state information (CSI) via uplink pilot signals from the users. It then obtains the downlink CSI relying on the reciprocity between the uplink and downlink. As such, it becomes difficult for eavesdroppers to know the CSI between themselves and the BS, as well as the CSI from other users to the BS. Therefore, how to design secure transmission under the assumption of imperfect (or no) CSI at the eavesdroppers is of practical importance in massive MIMO systems. Moreover, pilot contamination occurs in the TDD mode if the pilot signals employed in different cells are not orthogonal. As such, the effect of an inaccurate channel estimate caused by pilot contamination on the secrecy performance should be understood and counteracted. In addition, the TDD operation requires reciprocity calibration [11]. In practical systems, the hardware chains at the BS and users may not be reciprocal between the uplink and the downlink. This motivates the examination of the impact of improper calibration on the secrecy performance.

Artificial Noise

The deliberate deterioration of the eavesdropper's channel quality in massive MIMO systems is a fruitful avenue to explore. In conventional MIMO systems, the artificial noise (AN)-based transmission has been identified as an effective method to cause interference to the eavesdroppers and degrade their received signals. In massive MIMO systems, new challenges are opened for AN-based transmission. For example, transmitting AN signals in a spatial null space may not be practical since the computation complexity of the null space is extremely high for the

large-dimensional channel matrix. Moreover, random and independent AN is averaged out given the availability of a large number of antennas. Therefore, new AN-based transmission schemes need to be developed. Correspondingly, the optimal power allocation between information signals and AN signals needs to be determined and the achievable secrecy performance needs to be evaluated.

Antenna Correlation

Antenna correlation is a practical challenge underlying the deployment of massive MIMO systems. Specifically, a significant amount of correlation may exist between large antenna arrays, due to either the limited aperture of the antenna array or a lack of scattering. For the uplink transmission, for example, the antenna correlation is experienced in different diversity branches at the BS, due to the non-isotropic antennas with reduced separation. Although the impact of antenna correlation on the secrecy performance of conventional MIMO systems has been revealed, e.g. [15], very little detailed work has specifically been carried out to analyze the effect of antenna correlation on the secrecy performance of massive MIMO systems. The research efforts in this area are of enormous value since they enable us to decide how to compensate for antenna correlation in the uplink and downlink massive MIMO systems.

Confidential Broadcasting

In massive MIMO systems, each BS simultaneously communicates with a large number of users. One challenge to multiuser security is achieving confidential broadcasting in the downlink. In particular, each message needs to be kept confidential from all the users other than the intended one, i.e. each user can be treated as an eavesdropper for all messages other than its own. In order to preserve this confidentiality, a precoder needs to be associated with each data stream not only to limit the interference at other users, but also to limit the information leakage. Designing the optimal precoder often involves optimization problems that can only be solved numerically. More practical and near-optimal precoders are thus required. Therefore, it is pivotal to provide design guidelines and to quantify the optimal achievable secrecy performance of linear precoders that guarantee confidential broadcasting in massive MIMO systems.

Hardware Impairments

In contrast with conventional MIMO systems with ideal hardware, the inexpensive hardware components used by massive MIMO systems may give rise to hardware impairments [13]. Although hardware impairments deteriorate the legitimate receivers' channels, the impact of hardware impairments vanishes asymptotically when large-scale arrays are deployed. We note that the presence of hardware impairments also deteriorates the eavesdroppers' channels, which appears to be beneficial for security enhancement. It is therefore worth investigating physical layer security in massive MIMO systems with non-ideal hardware.

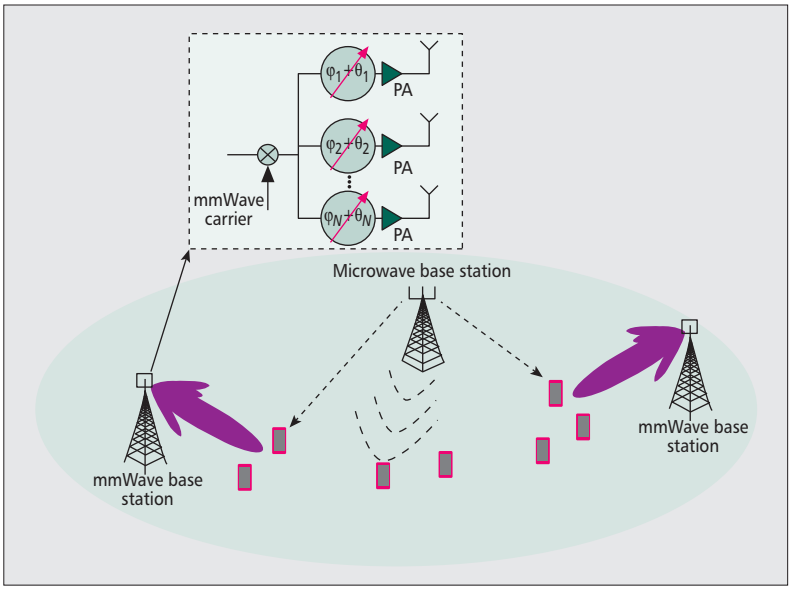

Figure 5. Deployment of mmWave BSs.

Physical Layer Security in Millimeter Wave Communication

Almost all mobile communication systems today restrict their operation to the spectrum in the range of 300 MHz–3GHz. Unfortunately, this spectral band has now become nearly fully occupied. In the 5G network, mmWave communication systems, operating in the frequency range of 30–300 GHz, have been recognized as a promising solution to remove the restriction and meet a thousand-fold capacity increase [16]. As depicted in Fig. 5, mmWave BSs can be deployed with microwave BSs to ensure reliable and fast data transmissions.

Although some efforts need to be made to render the GHz frequency bands available on mobile cellular networks, a series of research initiatives have been undertaken to explore the potential of mmWave communication technologies. Needless to say, security and privacy issues need to be addressed in the implementation of mmWave communication systems. We believe that the investigation of physical layer security in mmWave communication systems is a very promising and highly rewarding area, due to the following factors.

Large Bandwidth: Current maximum aggregated bandwidth in 4G LTE is 20+20 MHz by using carrier aggregation. However, mmWave communication systems provide GHz bandwidths. Therefore, the secrecy outage probability in the passive eavesdropping scenario is remarkably reduced if the transmitter sets a lower transmit secrecy rate in mmWave communications. Also, high secrecy throughput can be obtained with large mmWave bandwidths.

Short-Range Transmission: Compared to the current microwave communication systems, mmWave signals in the higher frequencies experience an increase in free-space path loss by several orders of magnitude. Therefore, only geographically neighboring eavesdroppers are

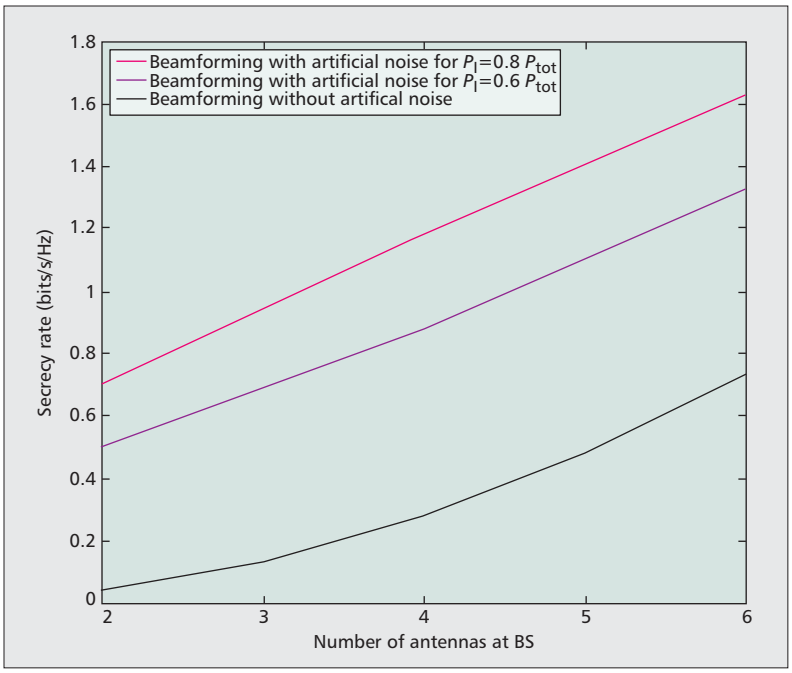

Figure 6. Secure mmWave downlink transmission with AN: An N-antenna mmWave BS transmits the confidential messages to a single-antenna user in the presence of a single-antenna eavesdropper. The BS uses analog beamforming with AN. The total transmit power is P_{tot} = 43 dBm, the power allocated to the information signal is P_I, and the power allocated to the AN signal is $P_{tot} - P_I$.

able to overhear the signals, whereas geographically remote users cannot capture the data transmission.

Directionality: In mmWave systems, highly directional communication with narrow beams is employed for suppressing the interference from neighbors. Therefore, the receive SNRs at the eavesdroppers may be extremely low such that the eavesdroppers are not able to recover information signals from the overheard messages.

Large Antenna Arrays: Large antenna arrays provide high beamforming gains to mitigate the propagation attenuation and save transmit power. In light of the array aperture constraint, current cellular systems[3] in the microwave frequency bands are expected to implement large antenna arrays in a two-dimensional (2-D) or three-dimensional (3-D) array structure. However, 2-D or 3-D arrays increase the coupling effects due to the increase in the number of adjacent antennas [17]. For a fixed array aperture, the shorter wavelengths at the mmWave frequencies enable the mmWave BSs to pack more antennas. Therefore, mmWave systems with large antenna arrays offer a wealth of opportunities at the physical layer security to secure mmWave communication.

Based on the aforementioned factors, the aim of physical layer security design in mmWave communication systems is to fully exploit the potentials of these factors. In this design, several challenging tasks need to be solved. First, the propagation characteristics at higher frequencies need to be precisely modeled. Indeed, an accurate and comprehensive quantification of the impact of path loss, blocking, penetration, and rain absorption on mmWave transmission enables network security designers to theoretically capture the properties of mmWave channels and address these properties in their design. Second, new secure transmission schemes need to be developed. It has been shown that beamforming is a key enabler of mmWave mobile broadband service [18]. Since digital beamforming with a large number of radio frequency (RF) chains incurs a very high implementation cost and power consumption, secure mmWave transmission needs to be designed based on analog beamforming and RF beamforming with a small number of RF chains. Against this background, the transmission of AN signals becomes promising in mmWave communication systems. With the aid of analog beamforming with phase shifters, the beam pattern of AN signals can be easily restricted to the orthogonal direction to the beam pattern of information signals. As depicted in Fig. 6, the secrecy rate is profoundly improved by incorporating AN signals into secure transmission. Moreover, the power allocated to the information signal plays a pivotal role in determining the secrecy performance. For example, beamforming with AN transmission with $P_I = 0.8P_{tot}$ yields a higher secrecy rate than with $P_I = 0.6P_{tot}$ in the system considered in Fig. 6. Motivated by this, an interesting question that needs to be explored is how to optimally allocate the transmit power between the information signal and AN signal in mmWave communication systems. Apart from the AN-based transmission, other secure schemes such as hybrid beamforming that mixes analog and digital signal processing techniques can be devised to secure mmWave transmission. In addition, a secure backhaul link between mmWave BSs in the mmWave cellular networks needs to be established.

[3] In current cellular systems, only a small number of antennas at the BS are used. For example, LTE allows for up to 8 antenna ports.

Conclusions

With the introduction of small cell deployments and D2D connections, the use of a very large number of antennas, and the exploration of the underutilized mmWave frequency spectrum, we believe that the 5G network is well positioned to meet the ever-increasing demand on data-centric applications over the next decade. The path to 5G is essentially irreversible, and will impose a profound impact on the design of physical layer security. In this article we have identified the scientific opportunities and discussed the technical challenges driven by the HetNet, massive MIMO, and mmWave communication. The novel solutions we have developed can take data confidentiality to a whole new level, inaugurating a brand new security paradigm that is truly worthy of the 5G designation.

References

[1] Cisco, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast update, 2013–2018, Cisco White Paper, Feb. 2014.
[2] W. Roh *et al.*, "Millimeter-Wave Beamforming as an Enabling Technology for 5G Cellular Communications: Theoretical Feasibility and Prototype Results," *IEEE Commun. Mag.*, vol. 52, no. 2, Feb. 2014, pp. 106–13.
[3] Y.-W. Hong, P.-C. Lan, and C.-C. Kuo, "Enhancing Physical-Layer Secrecy in Multiantenna Wireless Systems: An Overview of Signal Processing Approaches," *IEEE Signal Proc. Mag.*, vol. 30, no. 5, Sept. 2013, pp. 29–40.

[4] A. Mukherjee *et al.*, "Principles of Physical Layer Security in Multiuser Wireless Networks: A Survey," *IEEE Commun. Surveys Tutorials*, vol. 16, no. 3, 3rd Quarter 2014, pp. 1550–73.
[5] R. Bassily *et al.*, "Cooperative Security at the Physical Layer: A Summary of Recent Advances," *IEEE Signal Proc. Mag.*, vol. 30, no. 5, Sept. 2013, pp. 16–28.
[6] J. Andrews, "Seven Ways that HetNets are a Cellular Paradigm Shift," *IEEE Commun. Mag.*, vol. 51, no. 3, Mar. 2013, pp. 136–44.
[7] H. Wang, X. Zhou, and M. C. Reed, "Physical Layer Security in Cellular Networks: A Stochastic Geometry Approach," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, June 2013, pp. 2776–87.
[8] G. Geraci *et al.*, "Physical Layer Security in Downlink Multi-Antenna Cellular Networks," *IEEE Trans. Commun.*, vol. 62, no. 6, June 2014, pp. 2006–21.
[9] M. Tehrani, M. Uysal, and H. Yanikomeroglu, "Device-to-Device Communication in 5G Cellular Networks: Challenges, Solutions, and Future Directions," *IEEE Commun. Mag.*, vol. 52, no. 5, May 2014, pp. 86–92.
[10] N. Bhushan *et al.*, "Network Densification: The Dominant Theme for Wireless Evolution into 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, Feb. 2014, pp. 82–89.
[11] E. Larsson *et al.*, "Massive MIMO for Next Generation Wireless Systems," *IEEE Commun. Mag.*, vol. 52, no. 2, Feb. 2014, pp. 186–95.
[12] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and Spectral Efficiency of Very Large Multiuser MIMO Systems," *IEEE Trans. Commun.*, vol. 61, no. 4, Apr. 2013, pp. 1436–49.
[13] E. Björnson *et al.*, "Massive MIMO Systems with Non-Ideal Hardware: Energy Efficiency, Estimation, and Capacity Limits," http://arxiv.org/pdf/1307.2584v1.pdf.
[14] J. Zhu, R. Schober, and V. Bhargava, "Secure Transmission in Multicell Massive MIMO Systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, Sept. 2014, pp. 4766–81.
[15] N. Yang *et al.*, "Physical Layer Security of TAS/MRC with Antenna Correlation," *IEEE Trans. Info. Forensics Security*, vol. 8, no. 1, Jan 2013, pp. 254–59.
[16] T. Rappaport et al., "Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!," *IEEE Access*, vol. 1, May 2013, pp. 335–49.
[17] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling Up MIMO: Opportunities and Challenges with Very Large Arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, Jan. 2013, pp. 40–60.
[18] Z. Pi and F. Khan, "An Introduction to Millimeter-Wave Mobile Broadband Systems," *IEEE Commun. Mag.*, vol. 49, no. 6, June 2011, pp. 101–07.

Biographies

Nan Yang (nan.yang@anu.edu.au) received the Ph.D. degree in electronic engineering from the Beijing Institute of Technology, Beijing, China, in 2011. He is currently a future engineering research leadership fellow and lecturer at the Australian National University. He received the IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award in 2014, the Exemplary Reviewer Certificate of *IEEE Communications Letters* in 2012 and 2013, and the Best Paper Award at the IEEE VTC in 2013. His research interests include collaborative networks, network security, massive MIMO, millimeter wave, and molecular communications.

Lifeng Wang (lifeng.wang@qmul.ac.uk) received the M.S. degree in electronic engineering from the University of Electronic Science and Technology of China, Sichuan, China, in 2012. He is currently working toward the Ph.D. degree in electronic engineering at Queen Mary University of London, London, U.K.. His research interests include MIMO, cooperative communications, cognitive radio, and physical-layer security.

Giovanni Geraci (giovanni_geraci@sutd.edu.sg) received the Ph.D. in electrical engineering from the University of New South Wales, Sydney, Australia, in 2014. He is currently a postdoctoral research fellow at the Singapore University of Technology and Design, Singapore. His research interests include wireless communications, signal processing, applied mathematics, and information technology.

Maged Elkashlan (maged.elkashlan@qmul.ac.uk) received the Ph.D. degree in electrical engineering from the University of British Columbia, Vancouver, BC, Canada, in 2006. Since 2011 he has been with the School of Electronic Engineering and Computer Science, Queen Mary University of London, London, U.K., as an assistant professor. He currently serves as an editor for *IEEE Transactions on Wireless Communications*, *IEEE Transactions on Vehicular Technology*, and *IEEE Communications Letters*. His research interests fall into the broad areas of communication theory, wireless communications, and statistical signal processing for distributed data processing, millimeter-wave communications, cognitive radio, and wireless security.

Jinhong Yuan (j.yuan@unsw.edu.au) received the Ph.D. degree in electronics engineering from the Beijing Institute of Technology, Beijing, China, in 1997. In 2000 he joined the School of Electrical Engineering and Telecommunications, University of New South Wales, Sydney, Australia, where he is currently a telecommunications professor. He has co-authored three Best Paper Awards and one Best Poster Award. He currently serves as an associate editor for *IEEE Transactions on Communications*. His current research interests include error control coding and information theory, communication theory, and wireless communications.

Marco Di Renzo (marco.direnzo@lss.supelec.fr) received the Ph.D. degree in electrical and information engineering from the University of L'Aquila, Italy, in 2007. Since January 2010 he has been a tenured researcher with the French National Center for Scientific Research and a faculty member at the Laboratory of Signals and Systems. He currently serves as an editor for *IEEE Transactions on Communications* and *IEEE Communications Letters*. His main research interests are in the area of wireless communications theory.

With the introduction of small cell deployments and D2D connections, the use of a very large number of antennas, and the exploration of the underutilized mmWave frequency spectrum, we believe that the 5G network is well positioned to meet the ever-increasing demand on data-centric applications over the next decade.