Dense-Device-Enabled Cooperative Networks for Efficient and Secure Transmission

Shuai Han, Sai Xu, Weixiao Meng, and Cheng Li

Abstract

With the advancements in wireless networks, the number of user devices has increased dramatically, resulting in high device densities. Despite the resulting data traffic deluge, accompanied by severe security threats, wireless networks with high device densities are also breeding grounds for user cooperation. Considering various challenges and opportunities, this article attempts to enhance user cooperation utilizing big data generated from wireless networks toward achieving efficient and secure transmission. In particular, big data, viewed as a resource or tool, is employed to find potential connections among user devices, being followed by user cluster formation. Preliminary results demonstrate that big-data-driven user cooperation facilitates the utilization of wireless resources and reduces the secrecy loss originating from high device densities. Finally, this article identifies research topics for future studies on big-data-driven user cooperation and secure transmission in wireless networks.

Introduction

With the rapid development of wireless networks, the number of access devices and data volume will continue to grow exponentially due to the many new applications beyond personal communications [1]. Given this background, the core issue is the shortage of wireless resources [2]. In response to the pressure from mass quantity data transmission, wireless network architectures have been evolving. For instance, the potential techniques presented for the fifth generation (5G) of wireless communications have been exploited to achieve denser device configurations, higher communication rates, and lower latency [3]. On the other hand, data transmission may be subject to more serious security threats due to these high device densities [4]. Therefore, efficient and confidential data transmission is an important research issue.

In general, higher device densities result in more connections among user devices. This has recently begun to interest researchers for employing big data generated from various levels of wireless networks with dense devices to support both physical layer transmission and network optimization. These researchers believe that potential connections within wireless networks can be extracted from big data to help manage resources and enhance user cooperation to achieve higher utilization of wireless resources. The authors in [5] present a framework for efficiently analyzing massive data traffic from wireless networks to explore human movement behavior in densely populated areas. The authors in [6] propose a scheme of big-data-driven clustering, through which bandwidth can be efficiently allocated and used. The authors in [7] present the improvement in user quality of experience achieved by integrating big data analytics with network optimization. The authors in [8] show that big data can benefit the design and operation of software-defined networking (SDN).

On the other hand, connections among user devices also present certain disadvantages. To address security threats, PHY-security techniques, as an alternative to traditional high-complexity cryptography-based secrecy methods are often employed [4]. In PHY-security, the inherent randomness and differences in wireless channels are exploited to achieve confidentiality with proper coding and signal processing, which can ensure that confidential messages can only be decoded by authorized users. However, as device densities increase, the proximity of authorized users to unauthorized users may result in high correlations between their channels. In addition, the correlation enables unauthorized users to easily see and intercept more confidential information, which can produce a significant loss of secrecy [9]. Unfortunately, signal processing techniques at base stations, such as precoding/beamforming and embedding artificial noise (AN), are not useful enough or are completely ineffective.

Considering these coexisting opportunities and challenges, this article attempts to present a strategy that improves both the utilization of wireless resources and transmission confidentiality based on big data. Under this strategy, using big data, users satisfying a set condition can cluster. Through user cooperation in the formed clusters, not only are wireless resources utilized efficiently, but also secrecy loss due to high correlations can be effectively mitigated.

The remainder of this article is organized as follows. We first present a network model of user cooperation based on big data. Then, cluster formation and bandwidth allocation are discussed in detail. Next, a scheme using a cooperative jamming relay in clusters is given to reduce secrecy loss due to high correlations. Possible future research directions and conclusions are presented at the end of this article.

System Models

In wireless environments, higher device densities often result in more connections among user devices. By exploiting such connections, we can enhance the cooperation among users. On the

Digital Object Identifier:
10.1109/MNET.2018.1700292

Shuai Han , Sai Xu , and Weixiao Meng (corresponding Author) are with Harbin Institute of Technology; Cheng Li is with Memorial University of Newfoundland.

0890-8044/18/$25.00 © 2018 IEEE

other hand, mass quantities of data traffic can be generated, accompanied by more serious transmission security issues due to high correlations between channels. Faced with the opportunities and challenges brought about by dense devices, this section presents a network model of big-data-driven user cooperation, which can utilize wireless resources efficiently while reducing secrecy loss, as shown in Fig. 1.

User Clustering Based on Big Data

To enhance the utilization of wireless resources, clustering as a method of user cooperation is often considered in the conventional wisdom. In wireless environments, antenna deployments, the proximity of users, and scattering around users can cause correlative signal transmission paths between user devices, and thus high correlations between channels may occur when user devices are dense [10]. Under these circumstances, the similarity with respect to channel state information (CSI) among different user devices can be exploited to enable channel cluster formation [11]. With each formed cluster, messages from the base station are directly sent to cluster members by multicast, which can enhance resource utilization. However, the transmission rate in multicast is in general constrained by the minimum rate of all users receiving the same signals. With the bottlenecks that channel cluster schemes face, the authors in [6] employ big data as a new way to solve these problems. In particular, [6] adopt a similar idea to multicast but go further by clustering with the help of big data, therein effectively avoiding the disadvantage of multicast. However, [6] does not consider security issues.

As an evolved version of [6], we propose a big-data-driven cluster scheme under limitations imposed by security considerations. From Fig. 1, it is not difficult to find that our network model consists of a base station (Alice) and N authorized mobile users (Bobs) having security requirements $\mathcal{N} = \{1, 2, \cdots, N\}$. When Alice sends confidential messages to N Bobs, the transmissions are overheard by M ($M \ll N$ assumed) passive unauthorized users (Eves). We assume that Alice is equipped with N_A antennas, and Bob and Eve only have a single antenna each; all links are slowly fading Rayleigh channels. Each Bob can receive confidential messages from Alice directly or from other Bobs via short-range communication techniques such as WiFi-direct and device-to-device (D2D) techniques [12]. In our big-data-driven cluster scheme, big data from various levels of networks, as an additional resource, is applied to determine the relationships between Bobs, based on which clusters can be formed according to preset rules. In a cluster, a Bob satisfying the preset condition is selected as the "cluster head" to receive confidential messages directly from Alice; then, this Bob shares messages with other Bobs through short-range communication techniques in the cluster. For convenience, in a cluster, the cluster head and the other Bobs are denoted by Bob^h and Bob^{c}s, respectively.

Improving Secrecy by Cooperative Jamming

In the model, Alice is equipped with multiple antennas, with each Bob having only a single antenna. Thus, the transmission from Alice to Bob can be enhanced using a beamformer at Alice.

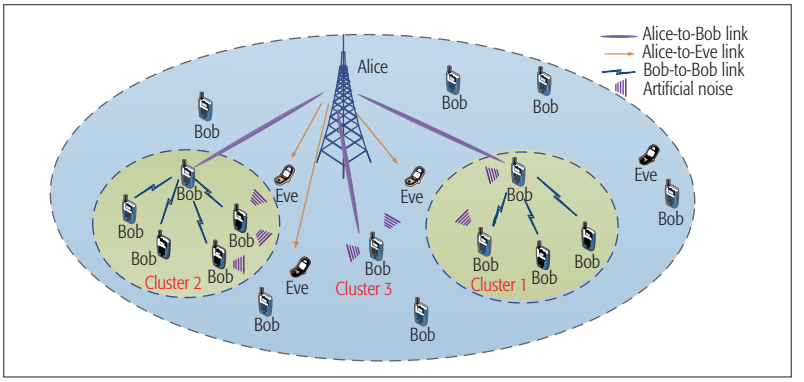

FIGURE 1. System model.

Since the CSI of the Alice-Eve link is not available, AN is often inserted into the transmitted signal to degrade the reception at Eve and consequently further increase the signal quality difference at Bob and Eve, aiming at maximizing the secrecy performance [13]. Unfortunately, such signal processing techniques at Alice are less effective in the scenario where main and wiretap channels (those are the channels from a transmitter to an authorized receiver and an unauthorized receiver, respectively) are highly correlated. This is mainly because signal processing techniques in PHY-security are inherently based on differences between wireless channels. With high device densities, there may exist at least one Eve nearby Bob^h, which often causes high correlations between Bob^h and Eve. As a result, it is easy for Eve to see and intercept more confidential information that Alice sends to Bob^h; thereby, the transmission from Alice to Bob^h suffers a significant secrecy loss.

To effectively mitigate the adverse effects of high correlations, we use Bob^{c}s in a cluster as cooperative jamming relays (Relays), which emit AN to help reduce secrecy loss. This scheme provides three benefits:

- The correlation between the Bob^c-Bob^h link and the Bob^c-Eve link may be smaller since there are greater differences between the two links.
- The distance from Bob^c to Eve is far less than that between Alice and Eve, which results in lower power consumption over travel paths.
- The equipment complexity at Alice is reduced significantly.

Obviously, high jamming efficiency can be guaranteed by selecting applicable Bob^{c}s as jamming relays. In addition to the cooperative jamming relay scheme, full-duplex-based jamming at Bob^h can also enhance secrecy despite self-interference [14], although this is beyond the scope of this article. Note that the spectrum occupied by the jamming signal is the same as that of the transmitted signal from Alice and differs from that in short-range communication for message sharing.

In summary, the given user cooperation strategy provides two advantages: i) improved utilization of wireless resources and ii) reduced secrecy loss due to high correlations between main and wiretap channels. In the later sections, we will discuss in detail big-data-driven cluster formation as well as bandwidth allocation and the operation of the cooperative jamming relay scheme.

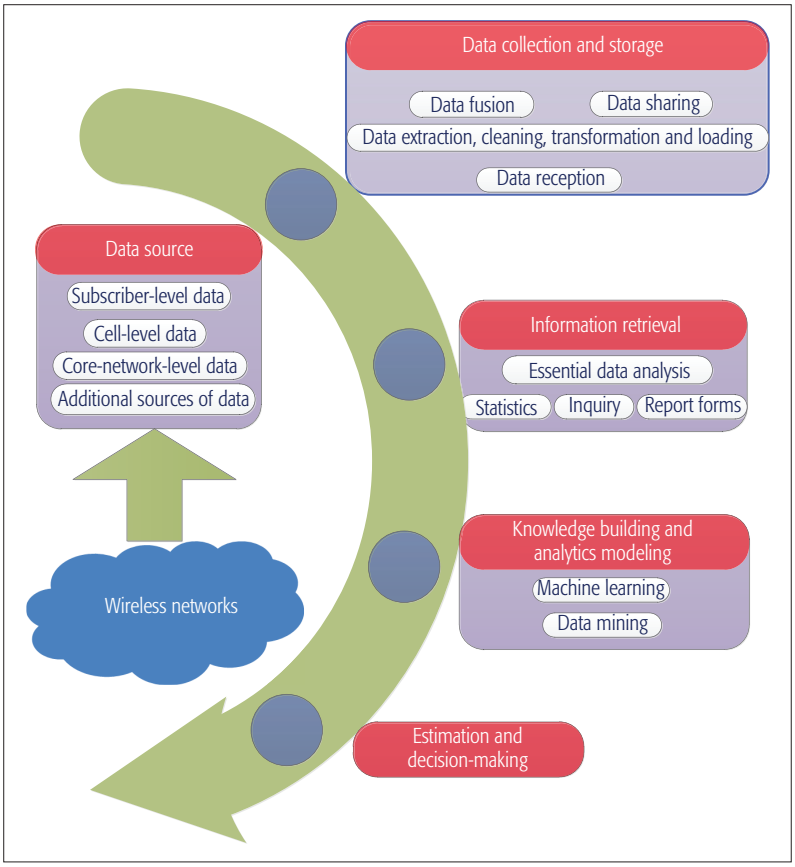

FIGURE 2. Big data processing procedure.

Data-Driven Cluster Formation

Clusters can be formed in two steps. First, according to certain rules, the cluster head Bob^h for each cluster is chosen, thereby forming an initial cluster consisting only of itself. Second, any other Bob, satisfying the preset condition, is added to the cluster in the form of Bob^c.

Cluster Head Selection

For each cluster, Bob^h must meet two requirements: Bob^h must ensure better confidentiality when communicating with Alice than Bob^c, and Bob^h must be able to share confidential messages securely and smoothly with as many Bob^{c}s as possible.

Secrecy Capacity: For the first requirement, we use the secrecy capacity to measure the secrecy performance when Bob$_n$ (n denotes Bob's index) communicates with Alice. The secrecy capacity, denoted by C_s^n, is defined as the maximum message transmission rate that confidential communications are able to be implemented with wiretap channel coding [4]. In particular, the secrecy capacity can often be written as the difference between the capacities of main and wiretap channels (when the difference is less than zero, it is zero as a matter of fact). We use C_m^n and C_w^n to denote the main and wiretap channel capacities, respectively. C_m^n is computed directly via the channel feedback. By contrast, C_w^n involves more information because of the passivity of Eve. Fortunately, the expectation of C_w^n can be estimated using various data collected from wireless networks such as statistical information about the channels, the degree of correlation between the channels, and information about the wireless environment. Then, C_s^n can be derived based on C_m^n and the estimation of C_w^n.

Ability for Sharing: For the second requirement, we use the number of Bobs (denoted by N_n) who can securely and smoothly obtain confidential messages shared by Bob$_n$ to measure the ability of Bob$_n$ for sharing. To determine N_n, the first task is to decide the conditions under which messages can be transmitted securely and smoothly between two Bobs. For this purpose, the *closeness* metric (denoted by $C_{ij} \in \{0, 1\}$) in [15] is used to describe whether confidential messages can be shared between two Bobs. In detail, for any two Bobs, denoted by Bob$_i$ and Bob$_j$, $C_{ij} = 1$ indicates that the confidential messages can be shared securely and smoothly between Bob$_i$ and Bob$_j$, while $C_{ij} = 0$ indicates otherwise. To obtain C_{ij}, data from wireless networks, such as the distance between two Bobs, mobility patterns, confidentiality, device type, storage, and radio measurements, are employed; these data, taken together, constitute a type of big data. By processing the big data, some useful knowledge can be extracted, which helps determine if a pair of Bobs can communicate confidentially with each other at a high transmission rate. With the evaluation, C_{ij} can be effectively estimated. The procedure of big data processing is summarized in Fig. 2. Since the procedure is complex and not our focus here, we directly assume that the necessary knowledge and models have been derived from the big data, followed by C_{ij} and N_n, without in-depth discussions about big data processing.

Utility Function: When the above two requirements are considered jointly, a utility function $Utility_n$ can be defined to evaluate the ability of Bob$_n$ to be a cluster head. In particular, we aim at maximizing the secrecy sum rate that Bob$_n$ can achieve for the whole cluster if it becomes the cluster head. Thus, $Utility_n$ is here defined as the sum of the secrecy rate each Bob acquires in a cluster if Bob$_n$ becomes the cluster head. Obviously, $Utility_n$ is the fusion of "secrecy capacity" and "sharing ability". In addition, $Utility_n$ is linear since it is an operation summing the secrecy rate for all cluster members. It is worth mentioning that $Utility_n$ is intuitively illustrated again with message frame structures in the later section for improved readability.

Cluster Formation

Once *Utility* for each Bob is obtained, we determine whether Bob can be the cluster head Bob^h. Then, other Bobs that have a nonzero *closeness* with Bob^h are added to the cluster of Bob^h as Bob^{c}s. Note that once a cluster is formed, all members of this cluster are removed from all Bobs who have not been in any cluster yet. Therefore, these removed Bobs will not participate in the next clustering, which ensures that no Bob can be added to multiple different clusters. The cluster formation procedure is summarized as follows:

1. Alice transmits a pilot signal, followed by the estimation and feedback of the CSI from Alice to each Bob.
2. Based on big data from wireless networks, the *closeness*, N_n, and C_s^n are deduced.
3. *Utility* for each Bob is calculated.

4. If all Bobs have *Utility* = 0, go to step 7; otherwise, let the Bob with the largest *Utility* be the cluster head and form an initial cluster $\mathcal{P}$.
5. Add all Bobs having a nonzero *closeness* with the cluster head to $\mathcal{P}$.
6. Set $\mathcal{N} = \mathcal{N} \backslash \mathcal{P}$ and go to step 3.
7. Cluster formation terminates.

In the above procedure, steps 3–6 form a loop. In each round, a cluster is formed; then, the members in this cluster are removed from all Bobs who have not clustered yet. The computational complexity of this procedure is low ($O(N^2)$), making it quick to execute. In addition, it is not difficult to deduce that the cluster formation scheme must be convergent. In the worst case, each cluster head forms a single cluster consisting of only itself. Actually, the cluster formation strongly depends on the *closeness* between two Bobs, which is closely related to big data.

Having accomplished clustering, Bob^h as a two-hop relay can forward the messages from Alice to each Bob^c, which has at least two advantages:

- When communicating with Alice, Bob^h has a higher secrecy capacity per unit bandwidth than Bob^c, which helps Bob^c obtain a higher secure transmission rate.
- Bob^{c}s requiring the same content only occupy a block of wireless resources.

Obviously, the advantages both facilitate the utilization of wireless resources and achieve a higher total secrecy capacity; thus, clustering can be used to guarantee secure transmission.

Note that the distance between Bob^h and Bob^{c}s in a cluster is far less than that between Alice and Bob^h. Considering that the power per unit area must decrease rapidly due to the path loss of a signal as a function of distance, the power consumed by short-range communication in a cluster is negligible. On the other hand, the big-data-driven cluster scheme is straightforward. To estimate the *closeness* metric more accurately, big data involving many practical factors must be effectively processed to model the relationship between Bobs. When analyzing big data to extract hidden information to optimize wireless communications, relevant artificial intelligence (AI) algorithms still need to be explored. Once the closeness is obtained, the utility function *Utility* related to multiple Bobs must be optimally defined. The definition may involve wireless resource allocation among Bobs, a rate limit for each content, cooperative methods, and so on.

Message Frame and Bandwidth Allocation

By clustering, we assume that K clusters are formed, although some clusters may only consist of Bob^h. With the formed clusters, Alice transfers confidential messages to Bob^h, and then Bob^h shares the messages with Bob^c. Bob^c decodes the shared messages to obtain the content that it requires. The transmission from Alice to Bob^h consumes certain wireless resources, such as frequency and time slots, and the allocation of wireless resources directly affects the secrecy sum capacity of the whole system. Without loss of generality, we investigate bandwidth allocation. It is assumed that the total bandwidth used to send confidential messages is B and that the bandwidth allocated to the k-th Bob^h is B_k, represented by a K-dimensional vector $\mathcal{B} = \{B_1, B_2, \cdots, B_K\}$.

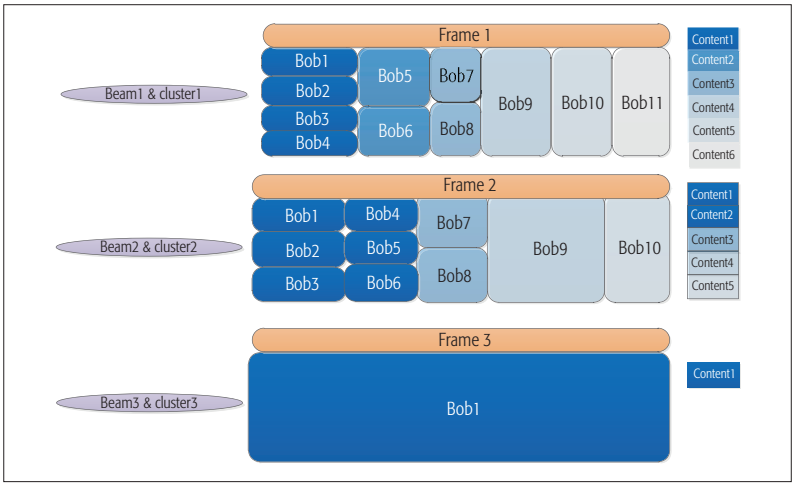

FIGURE 3. Message frame structures for clusters.

Messages from Alice are sent to Bob^h in the form of a frame, corresponding to *Utility*. As shown in Fig. 3, a message frame consists of several content blocks, which are distinguished by different shades of blue. A content block may be shared by multiple cluster members or be exclusive to one cluster member. In a frame, the number of content blocks and the transmission rate for each content can be determined jointly according to the predetermined protocols; this is related to the allocation of wireless resources based on practical services and targets.

Once the confidential messages are received, Bob^h shares them with Bob^c in the cluster. The optimization problem for the bandwidth allocation varies between different targets. When the objective is to maximize the total secrecy capacity of the network while restricting a cluster from obtaining an excessively large bandwidth, the bandwidth allocation is a low-computational-complexity linear programming (LP) problem that can be easily solved.

Improving Secrecy under High Correlation

In PHY-security, beamforming at Alice is often employed to enhance the signal quality at Bob while limiting the signal strength at Eve. By contrast, AN techniques are used to degrade the reception at Eve to increase the signal quality difference at Bob and Eve. To maximize the secrecy of the system, Alice simultaneously transmits an information-bearing signal and AN. However, when main and wiretap channels are highly correlated due to the proximity of Bob and Eve, Eve can see and intercept more confidential information, and thereby the transmission from Alice to Bob suffers a significant secrecy loss. Unfortunately, both signal processing techniques at Alice are hardly effective. Therefore, this section presents a cooperative jamming relay scheme, therein attempting to reduce the secrecy loss caused by highly correlated channels.

Here, we use the correlation coefficient between main and wiretap channel vectors to measure the degree of correlation. Since wireless environments often change very slowly, it is reasonable to assume that the average correlation coefficient can be estimated through multiple prior measurements of the environment and

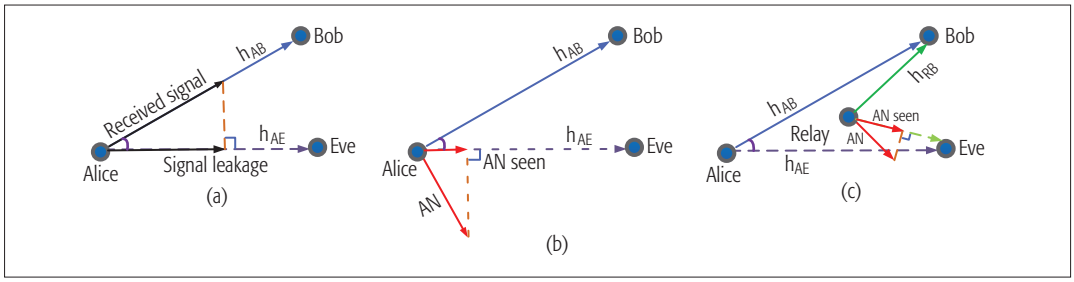

FIGURE 4. Illustrative diagrams of: a) beamforming; b) AN; c) cooperative jamming relay mechanisms.

A content block may be shared by multiple cluster members or be exclusive to one cluster member. In a frame, the number of content blocks and the transmission rate for each content can be determined jointly according to the predetermined protocols; this is related to the allocation of wireless resources based on practical services and targets.

feedback coming from other users. To intuitively describe the proposed cooperative jamming relay scheme, illustrative diagrams are given in Fig. 4, while the cosine of the angle between h_{AB} and h_{AE} corresponds to the correlation coefficient between main and wiretap channel vectors. Figures 4a and 4b are illustrative diagrams of beamforming and AN mechanisms, respectively. Since it is assumed that Eve receives the transmitted signal from Alice passively and that only statistical CSI of the Alice-Eve link is available to Alice, the signal vector is directed toward Bob via a beamformer, while the AN is scattered isotropically in the null space of the Alice-Bob channel. This approach can guarantee the maximum amount of information being received at Bob and prevent AN leakage into the Alice-Bob channel. In addition, some components of the AN may lie in the range space of the Alice-Eve link, which disrupts the reception at Eve. However, as the correlation between the main and wiretap channels increases, the angle between the channel vectors becomes increasingly smaller. As a result, more signal leaks into the Alice-Eve channel, while less AN interferes with the reception at Eve, which causes a significant secrecy loss. To overcome this disadvantage, a cooperative jamming relay is introduced, as shown in Fig. 4c. Generally speaking, the correlation between the Relay-Bob and Relay-Eve links is lower than that between the Alice-Bob and Alice-Eve links. Thus, more AN emitted by Relay is seen by Eve; this can effectively disrupt the reception at Eve and consequently increase the signal quality difference between Bob and Eve. Although there exists a possible situation whereby Eve can attempt to mitigate the interference signal from Relay using a directional antenna, the use of a directional antenna inevitably changes the Alice-Eve channel and amplifies the difference between the Alice-Bob and Alice-Eve links. In this case, the difficulty caused by the high correlation between the received signals at Bob and Eve disappears; therefore, this case is not considered.

In our network model, high device densities may produce a case whereby some Bob^hs may be surrounded by at least one Eve, which may result in high correlation between main and wiretap channels. Considering that the correlation between the Bob^c-Bob^h and Bob^c-Eve links may be smaller, some Bob^cs in a cluster are used as cooperative jamming relays to emit AN to help reduce the secrecy loss. Moreover, the power consumption over travel paths decreases.

To achieve optimal performance, the power allocation between an information-bearing signal transmitted by Alice and the AN from Bob^c must be given sufficient attention. Here, a simple power allocation between Alice and a cooperative jamming relay can be demonstrated as follows:

- Calculate the capacity of the Alice-Bob channel under an arbitrary power allocation ratio.
- Calculate the expectation of the capacity of the Alice-Eve channel with AN from Relay under an arbitrary power allocation ratio.
- Calculate the expectation of the secrecy capacity.
- Select the power allocation ratio maximizing the expectation of the secrecy capacity as the optimal power allocation ratio.

Case Study for Proposed Strategy

This section presents two cases for evaluating the performance of the proposed strategy, in which user cooperation based on big data is applied to improve the utilization of wireless resources and transmission confidentiality. The two given cases are used to validate the effectiveness of the proposed big-data-driven cluster and cooperative jamming relay schemes. Note that in these cases, simulations are performed using random data generated by a computer instead of real-world data.

First Case

This case shows how the big-data-driven cluster scheme affects the overall secrecy performance of networks. We consider the simplified scenario whereby each message frame only consists of one content block and all cluster members share the same content. Certain parameters are set as follows: There exist $N = 100$ Bobs, with $M = 2$ passive Eves; the antenna number at Alice is $N_A = 4$; the total bandwidth $B = 600$ kHz; the variances of the Alice-Bob and Alice-Eve channels are set as $\sigma^2_{AB} = \sigma^2_{AE} = 1/2$; the SINR at Bob is a random variable between −3 dB and 10 dB; and considering that all Bobs have equal status and that it is difficult to obtain the locations of passive Eves, the estimated value of the expected wiretap *SINR* is set as 0.7 times that of the corresponding Bob due to the high correlation between main and wiretap channels. In addition, there is no upper limit on the threshold of the data rate r for a cluster; this setting attempts to maximize the total secrecy capacity of the network without restricting a cluster from obtaining an excessively large bandwidth.

For simplicity, the *closeness* metric of the clus-

ter formation scheme is modeled as the Euclidean distance between two Bobs. In particular, Alice is located at the center of a cell with a radius R = 100 m, and the locations of Bobs, following a uniform distribution, are randomly generated. When the Euclidean distance between two Bobs is less than $0.1R$ = 10 m, confidential messages can be shared securely and smoothly by them.

In Fig. 5, we plot the average total secrecy capacity vs. the number of Bobs with M = 2 Eves. It can be seen that the curve of the big-data-driven cluster scheme perfectly agrees with that of the traditional scheme without clustering for one to four Bobs. This is because when there are fewer Bobs, multiuser diversity gain plays a main role in improving the total secrecy capacity of the system with less cluster gain. However, as the number of Bobs continues increasing, the cluster gain becomes increasingly obvious, and thus, the average total secrecy capacity for the cluster scheme still rapidly increases. By comparison, the growth of the curve of the traditional scheme tends to be flat since the multiuser diversity gain remains unchanged in the case of many Bobs. In short, more users indirectly share the bandwidth of the cluster head as a result of clustering; therefore, the big-data-driven clustering scheme is attractive.

Second Case

This case is given to evaluate the secrecy performance gain achieved by the proposed cooperative jamming relay scheme. In the simulations, we consider a scenario involving only one Bob and one Eve. Certain parameters are set as follows. The number of antennas at Alice is N_A = 2, 4 or 8; the number of antennas at Relay is N_B = 2; the total transmitted power P is 3 dBW; the additive noise at Bob and Eve is assumed to follow a normal distribution; the variances of the Alice-Bob and Alice-Eve channels are set as $\sigma^2_{AB} = \sigma^2_{AE} = 1/2$; and the variances of the Relay-Bob and Relay-Eve channels are $\sigma^2_{RB} = \sigma^2_{RE} = 1$. Note that because we use the normalized noise power, the total transmitted power can be measured in dBW.

Figure 6 plots the relationship between the average secrecy capacity $\overline{C}_s$(bps/Hz) and the square of the correlation coefficient between the main and wiretap channel vectors ρ^2. Note that the average secrecy capacity here only represents the statistics consisting of large numbers of possible instantaneous results, rather than the ergodic secrecy capacity, which is characterized as the average of the capacities for all possible channel realizations without strict delay restrictions. From Fig. 6, we observe that with increasing ρ^2, $\overline{C}_s$ continuously decreases. Compared to the traditional scheme employing beamforming and AN at Alice simultaneously, the proposed scheme using a cooperative jamming relay has a higher $\overline{C}_s$ for the same ρ^2, which indicates that secrecy loss due to high correlation can be significantly reduced using a cooperative jamming relay. On the other hand, since the power of the AN is equally divided into different dimensions of the null space of the Alice-Bob channel, the influence that AN has on secrecy is greater with fewer antennas. This result is the opposite of the beamforming method, where the latter improves secrecy more when Alice is equipped with more antennas. Thus, it is reasonable that the curves

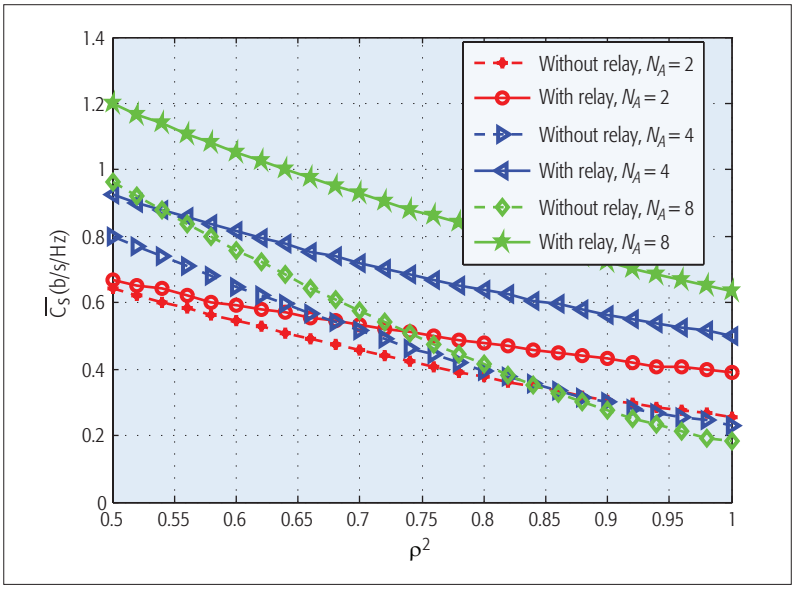

FIGURE 5. The relationship between the average secrecy capacity $\overline{C}_S$(b/s/Hz) and the square of the correlation coefficient between the main and wiretap channel vectors ρ^2.

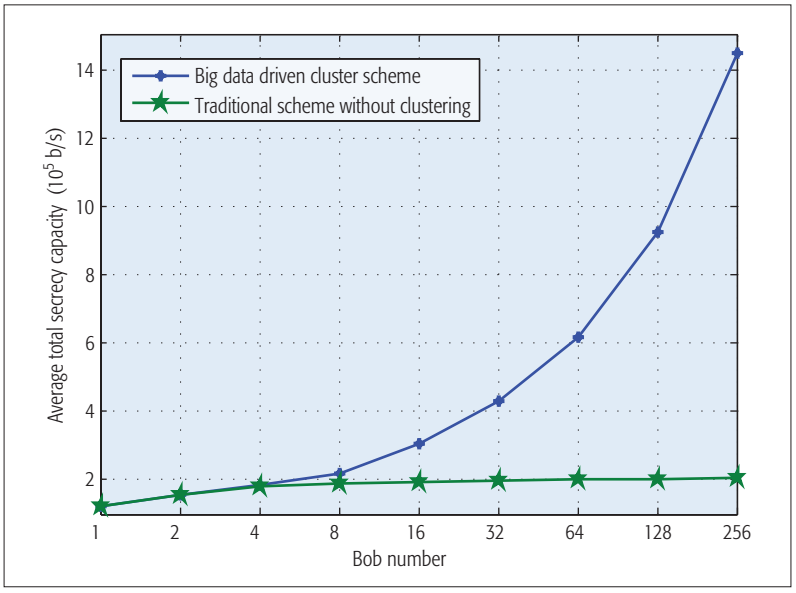

FIGURE 6. Average total secrecy capacity vs. the number of Bobs with M = 2 Eves.

for the traditional scheme without the jamming relay intersect, therein jointly considering the two factors of beamforming and AN.

Future Research Work

With the coexisting opportunities and challenges brought about by dense devices and the corresponding data deluge, big-data-driven user cooperation in wireless networks can effectively facilitate the utilization of wireless resources and increase transmission confidentiality. However, as a promising research area, the following open issues remain:

- To maximize the utilization of wireless resources, protocols related to message frame structure and resource allocation among cluster heads need to be strictly designed based on the given situation. Moreover, optimal or suboptimal power allocation

As an emerging and promising method, big data is explored to strengthen user cooperation. Using big data, users satisfying an established condition can cluster. Through these formed clusters, not only are wireless resources utilized efficiently, but also the secrecy loss due to high correlations can be effectively mitigated with AN produced by cluster members.

among a transmitter and cooperative jamming relays can effectively reduce secrecy loss due to high correlation; therefore, it is necessary to solve the involved optimization issues.

- Big-data-driven clustering and cooperative jamming relay schemes can improve the total secrecy rate of a system. To achieve better performance, it is suggested to combine both techniques. Joint usage of both techniques involves complex mathematical processing; thus, the quantification of performance indicators is a challenging task.
- To obtain more convincing results, one must employ real big data involving various practical factors instead of randomly generated data. Analyzing big data to extract hidden information can optimize wireless communications, and certain involved techniques, such as deep learning and mining raw data, must be explored.
- The big-data-driven clustering and cooperative jamming relay schemes in this article are not limited to mobile communication networks and may be extended to Internet of Things (IoT) as well as vehicle networks. On the other hand, the combination of the proposed schemes and potential 5G techniques is also a promising research direction, with no contradiction between the two.

Conclusions

Increasing the number of devices in wireless networks can result in high device density, therein presenting both opportunities and challenges simultaneously. As an emerging and promising method, big data is explored to strengthen user cooperation. Using big data, users satisfying an established condition can cluster. Through these formed clusters, not only are wireless resources utilized efficiently, but also the secrecy loss due to high correlations can be effectively mitigated with AN produced by cluster members.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (No. 91438205), the Provincial Natural Science Foundation of Heilongjiang (No. ZD2017013), and the National Natural Science Foundation of China (No. 61771169).

References

[1] S. Bi *et al.*, "Wireless Communications in the Era of Big Data," *IEEE Commun. Mag.*, vol. 53, no. 10, Oct. 2015, pp. 190–99.
[2] H. Lee, S. Vahid, and K. Moessner, "A Survey of Radio Resource Management for Spectrum Aggregation in LTE-Advanced," *IEEE Commun. Surveys and Tutorials*, vol. 16, no. 2, Nov. 2014, pp. 745–60.
[3] G. J. Andrews *et al.*, "What Will 5G Be?," *IEEE JSAC*, vol. 32, no. 6, 2014, pp. 1065–82.
[4] Y. Liu, H. H. Chen, and L. Wang, "Physical Layer Security for Next Generation Wireless Networks: Theories, Technologies, and Challenges," *IEEE Commun. Surveys and Tutorials*, Aug. 2016.
[5] Y. Qiao *et al.*, "A Mobility Analytical Framework for Big Mobile Data in Densely Populated Area," *IEEE Trans. Vehicular Tech*, vol. 66, no. 2, Feb. 2017, pp. 1443–55.
[6] B. Fan, S. Leng, and K. Yang, "A Dynamic Bandwidth Allocation Algorithm in Mobile Networks with Big Data of Users and Networks," *IEEE Netw.*, vol. 30, no. 1, 2016, pp. 6–10.
[7] K. Zheng *et al.*, "Big Data-Driven Optimization for Mobile Networks toward 5G," *IEEE Netw.*, vol. 30, no. 1, 2016, pp. 44–51.
[8] L. Cui, F. R. Yu, and Q. Yan, "When Big Data Meets Software-Defined Networking: SDN for Big Data and Big Data for SDN," *IEEE Netw.*, vol. 30, no. 1, 2016, pp. 58–65.
[9] H. Jeon *et al.*, "Bounds on Secrecy Capacity over Correlated Ergodic Fading Channels at High SNR," *IEEE Trans. Inf. Theory*, vol. 75, no. 4, 2011, pp. 1975–83.
[10] D.-S. Shiu *et al.*, "Fading Correlation and Its Effect on the Capacity of Multielement Antenna Systems," *IEEE Trans. Commun.*, vol. 48, Mar. 2000, pp. 502–13.
[11] Q. Zhang *et al.*, "User Clustered Opportunistic Beamforming for Stratospheric Communications," *IEEE Commun. Let.*, vol. 20, no. 9, 2016, pp. 1832–35.
[12] L. Jiajia *et al.*, "Device-to-Device Communication in LTE-Advanced Networks: A Survey," *IEEE Commun. Surveys and Tutorials*, vol. 17, no. 4, Dec. 2015, pp. 1923–40.
[13] Y. Liu, H. H. Chen, and L. Wang, "Secrecy Capacity Aanalysis of Artificial Noisy MIMO Channels — An Approach Based on Ordered Eigenvalues of Wishart Matrices," *IEEE Trans Inf. Forensics and Security*, vol. 99, 2016, pp. 1-1.
[14] A. Sabharwal *et al.*, "In-Band Full-Duplex Wireless: Challenges and Opportunities," *IEEE JSAC*, vol. 32, no. 9, 2014, pp. 1637–52.
[15] S. Han *et al.*, "An Agile Confidential Transmission Strategy Combining Big Data Driven Cluster and OBF," *IEEE Trans. Vehicular Tech.*, vol. 99, 2017, pp. 1-1.

Biographies

Shuai Han [S'11, M'12, SM'17] (hanshuai@hit.edu.cn) is an associate professor in the Department of Electronics and Communication Engineering, Harbin Institute of Technology. He is also a Vice Chair of the IEEE Harbin ComSoc Chapter and Vice Chair of the IEEE Harbin VTS Chapter. He received his B.S., M.E. and Ph.D. degrees from Harbin Institute of Technology in 2004, 2007 and 2011, respectively. His research interests include wireless sensor networks, wireless communications, the global navigation satellite system and indoor location.

Sai Xu [S'16] (fenicexusai@163.com) received his B.S. degree in physics from Hebei Normal University, Shijiazhuang, China, in 2012, and his M.E. degree in electronic science and technology from Harbin Institute of Technology (HIT), Harbin, China, in 2015. Currently, he is a Ph.D. candidate at HIT, as well as a joint Ph.D. student in the Electrical Engineering Department, University of California, Los Angeles, USA. His current research interests include artificial intelligence and big data for network optimization and security.

Wei-Xiao Meng [M'04] (wxmeng@hit.edu.cn) received the B.Eng., M.Eng., and Ph.D. degrees from Harbin Institute of Technology (HIT), Harbin, China, in 1990, 1995, and 2000, respectively. From 1998 to 1999, he worked at NTT DoCoMo on adaptive array antennas and dynamic resource allocation for beyond 3G as a senior visiting researcher. He is now a full professor and the vice dean of the School of Electronics and Information Engineering of HIT. His research interests include broadband wireless communications and networking, MIMO, GNSS receivers and wireless localization technologies. He has published three books and over 220 papers in journals and international conferences. He is the chair of the IEEE Communications Society Harbin Chapter, a Fellow of the China Institute of Electronics, and a senior member of the IEEE and the China Institute of Communication.}

Cheng Li [M'97] (licheng@mun.ca) received the B.Eng. and M.Eng. degrees from Harbin Institute of Technology, Harbin, P. R. China, in 1992 and 1995, respectively, and the Ph.D. degree in electrical and computer engineering from Memorial University, St. Johns, NL, Canada, in 2004. He is currently a full professor in the Department of Electrical and Computer Engineering, Faculty of Engineering and Applied Science, Memorial University, St. Johns, NL, Canada. His research interests include mobile ad hoc and wireless sensor networks, wireless communications and mobile computing, switching and routing, and broadband communication networks.