This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2016.2569083, IEEE Transactions on Wireless Communications

United Against the Enemy: Anti-jamming Based on Cross-layer Cooperation in Wireless Networks

Liyang Zhang, Zhangyu Guan, *Member, IEEE*, and Tommaso Melodia, *Member, IEEE*

***Abstract*—Denial-of-service (DoS) attacks launched by malicious jammers can pose significant threats to infrastructure-less wireless networks without a centralized controller. While significant recent research efforts have dealt with such attacks and several possible countermeasures have been proposed, little attention has been paid to the idea of *cooperative anti-jamming*.**

Inspired by this observation, this paper proposes and studies a cooperative anti-jamming scheme designed to enhance the quality of links degraded by jammers. To achieve this objective, users are allowed to cooperate at two levels. First, they cooperate to optimally regulate their channel access probabilities so that jammed users gain a higher share of channel utilization. Second, users leverage multiple-input single-output cooperative communication techniques to enhance the throughput of jammed links. The problem of optimal cooperative anti-jamming is formulated as a distributed pricing-based optimization problem, and a best response algorithm is proposed to solve it in a distributed way. Simulations demonstrate that the proposed algorithm achieves considerable gains (compared to traditional non-cooperative anti-jamming) especially under heavy traffic or high jamming power. Furthermore, the proposed distributed algorithm is shown to achieve close-to-global optimality with moderate traffic load.

***Index Terms*—Anti-jamming, Cooperative Relay, Cross-layer Optimization, Distributed Algorithm.**

I. Introduction

WIRELESS networks are known to be vulnerable to denial-of-service (DoS) attacks, mostly as a consequence of their broadcast nature [2]–[4]. By radiating high-power radio-frequency signals, a malicious adversary can easily generate interference to degrade the perceived signal-to-interference-plus-noise ratio (SINR) and therefore the achievable link throughput of legitimate users. The situation may be exacerbated in infrastructure-less wireless networks with no centralized entity to coordinate the transmission strategies of different users.

Various techniques have been proposed to combat jamming attacks at different layers of the protocol stack. A common idea behind many of these techniques is to leverage additional levels of "diversity", for example, frequency hopping spread spectrum (FHSS) and direct sequence spread spectrum (DSSS) attempt to avert interference in the frequency or coding domains, respectively.

This material is based upon work supported in part by the National Science Foundation under Grant CNS-1218717 and the Air Force Research Laboratory under Contract FA8750-14-1-0074. A preliminary shorter version of this paper appeared in the Proceedings of IEEE International Conference on Computer Communications (INFOCOM), Toronto, Canada, April 2014.

L. Zhang, Z. Guan, and Tommaso Melodia are with the Department of Electrical and Computer Engineering, Northeastern University, Boston, MA 02115 USA (e-mail: {liyangzh, zguan, melodia}@ece.neu.edu).

FHSS has long been used to provide anti-jamming capabilities in wireless communications. By quickly shifting from one frequency carrier to another, FHSS allows legitimate users to actively avoid jamming attacks. However effective, FHSS has several shortcomings. First, it relies heavily on a pre-defined secret pattern. Therefore, it may not be suitable for ad hoc networks where it is difficult to share a common secret between transmitters and receivers; or for cognitive radio networks where the availability of spectrum "holes" may follow a random pattern [5]. Second, FHSS requires significantly more spectrum resources than single-carrier transmission strategies, and is not spectrally efficient [6]. Third, FHSS assumes that jammers can only jam one or a subset (but not all) of the available channels at the same time, while in some scenarios, it may be possible for a jammer to launch more powerful attacks by generating broadband interference on all the available channels. In fact, it has been shown that multi-channel jamming is feasible with cognitive radio technology [7].

In recent years, several *adaptive* frequency hopping strategies have been proposed to address the first shortcoming, including uncoordinated frequency hopping (UFH) [8] and message-driven frequency hopping (MDFH) [9]. The communication efficiency of UFH was analyzed theoretically in [10], and practical algorithms were proposed in [11], [12]; the anti-jamming properties of MDFH were analyzed in [13], [14]. However, in scenarios with scarce spectrum resources, anti-jamming techniques that can utilize the spectrum resource more flexibly and efficiently are needed.

DSSS provides an alternative way for spectrum spreading [15], [16]. Unlike FHSS, DSSS achieves this goal in the coding domain. Despite this difference, like FHSS, DSSS has limited spectral efficiency. Besides, DSSS relies on a specific physical layer transmission scheme (i.e., code-division multiple access, CDMA).

The rapid increase in spectrum requirements has motivated another family of anti-jamming techniques, suitable for multi-carrier networks. Unlike FHSS and DSSS, which achieve anti-jamming capabilities by expanding the transmitted waveform to a wider spectrum, these techniques try to achieve optimal use of the available spectrum resources. Thus, we refer to these techniques as *adaptive optimal resource allocation* approaches.

In optimal resource allocation techniques, a user, instead of trying to avoid a jammer altogether by using a different portion of the spectrum, relies on exploiting the current spectrum in the most efficient way to combat jamming. Optimal resource allocation techniques usually attempt to maximize

1536-1276 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2016.2569083, IEEE Transactions on Wireless Communications

the information-theoretic capacity of the link by allocating resources on several different channels at the same time. For example, in orthogonal frequency-division multiplexing (OFDM) systems iterative water-filling algorithms can be used to maximize the achievable capacity of legitimate users in the presence of jammers [17]. Since each user independently and selfishly selects its optimal transmission strategy, these approaches are often analyzed using tools from non-cooperative game theory. An extensive literature has emerged that relies on this or similar approaches [17]–[23].

It can be observed that all these techniques, along with those adopting FHSS, focus on how to effectively utilize different frequency channels against jamming. Therefore, they all take advantage of *frequency diversity*. In DSSS techniques, *coding diversity* is exploited. There are also techniques leveraging *space diversity* in various ways. Directional transmission is well-known to be able to enhance the throughput by creating highly directional links, and is thus widely used in anti-jamming [24] [25]. Multipath routing provides a way to circumvent the jammer [26] [27]. Even the mobility of the nodes can in certain scenarios be exploited to avoid jamming [28].

In this paper, we attempt to explore an additional degree of freedom, the *cooperative diversity* dimension, which has been underexplored in the context of anti-jamming techniques. Cooperative techniques can be jointly leveraged at the network, MAC, and physical layers to provide effective countermeasures against jamming. While in commercial networks it is natural for different terminals to operate selfishly and in a non-cooperative fashion, in sensor networks and tactical military networks, which are often managed by a single entity, cooperative behaviors can be more easily implemented. Interestingly, the idea of exploiting cooperation has also been proposed in different contexts. For example, to protect the secrecy of a wireless communication link, several studies have proposed to employ *cooperative jamming* (also known as "friendly" jamming) to prevent eavesdroppers from intercepting the transmitted information [29] [30] [31].

Inspired by this idea, we propose and study a cooperative anti-jamming scheme designed to maximize the *fairness-constrained network throughput* in the presence of jammers. The proposed algorithm jointly optimizes the channel access probabilities and cooperative relaying probabilities of legitimate users. Legitimate users cooperate at two levels. At the medium access control layer, a cooperative channel access scheme is proposed where the channel access probabilities of different users are optimally regulated so that users degraded severely by jammers have an increased share of air time. In this way, users with good links "trade" capacity with those with jammed links. The second step is to extend the cooperation from MAC to physical layer. It is well known that, by using cooperative relays, virtual multiple-input-single-output transmission links can be formed to increase the link capacity. In the proposed scheme, users able to enhance the link capacity of other users through cooperative transmissions cooperate as relays with a certain probability. A new distributed algorithm is proposed to jointly optimize these two levels of cooperations, with significant gains in terms of achievable network throughput.

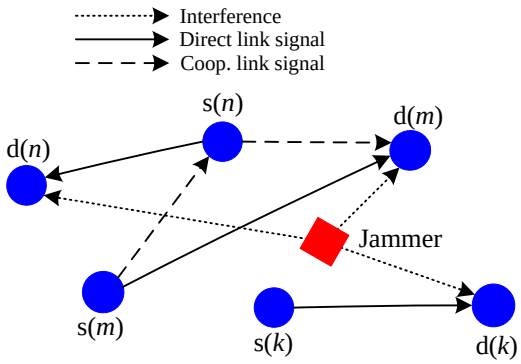


Fig. 1: An example topology with a jammer.

To summarize, we make the following contributions:

1) We propose the first *cooperative anti-jamming* scheme that jointly optimizes the cooperative behavior of nodes at the MAC and physical layers. To the best of our knowledge, this is also the first anti-jamming scheme based on a virtual multiple-input-single-output (MISO) variant of cooperative communications;
2) We formulate the *optimal cooperative anti-jamming* problem as an optimization problem with the objective of maximizing the fairness-constrained network throughput. We design a *distributed* solution algorithm based on dynamic pricing that is guaranteed to converge even if the socially optimal problem is not convex.
3) We design a *provably-optimal centralized algorithm* based on the branch and bound framework and convex relaxation techniques. The algorithm provides a performance benchmark for any distributed algorithm designed to solve similar problems.
4) We compare the performance of the cooperative distributed algorithm with a non-cooperative distributed algorithm and with the optimal centralized algorithm. We show that the cooperative algorithm achieves near-optimality under light and moderate traffic, and provides considerable gains compared to non-cooperative strategies.

The rest of the paper is organized as follows. Section II presents the system model and problem statement, while Section III derives a model of the utility for each legitimate user. In Section IV, we present and analyze the distributed solution algorithm of cooperative anti-jamming. The centralized algorithm is proposed in Section V. Some practical issues are discussed in Section VI. In Section VII, we analyze the performance of the proposed algorithms. Finally, we draw conclusions in Section VIII.

II. System Model

As illustrated in Fig. 1, we consider a wireless ad hoc network composed of a set $\mathcal{N}$ of legitimate users each consisting of a source-destination pair of nodes. In this case, the network can be viewed as a set of concurrent node-to-node transmissions. Assuming that the transmitting nodes are always saturated, i.e., their queues are always backlogged, then we can also refer to each of the transmitter-receiver pairs as a

1536-1276 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2016.2569083, IEEE Transactions on Wireless Communications

session. We denote the source and destination nodes of each session $n \in \mathcal{N}$ as $\mathrm{s}(n)$ and $\mathrm{d}(n)$, respectively. There are a set $\mathcal{F}$ of frequency-orthogonal channels for the legitimate users to transmit on.

Jamming Model. We assume that there is one jammer node, denoted as j, constrained by a limited power budget that attempts to degrade the throughput of the legitimate users by generating interference on the available channels. The model can be easily extended to the case of multiple jammers. We assume that the jammer is able to emit wideband interference across all the available channels simultaneously. If we denote by $\boldsymbol{p}_{\mathrm{j}} = (p_{\mathrm{j}}^f)_{f\in\mathcal{F}}$ the jammer power allocation profile, where p_{j}^f is the power allocated on channel f, we have

$$\mathbf{1}^{\mathrm{T}}\boldsymbol{p}_{\mathrm{j}} \leq p_{\mathrm{j}}^{\max}, \tag{1}$$

where $p_{\mathrm{j}}^{\max}$ is the maximum power of the jammer, and $\mathbf{1}$ represents an $1 \times |\mathcal{N}|$ vector of '1' elements.

Because of the heterogeneity of different channels, the jammer must allocate its power budget in some way to achieve a good jamming effect. In previous literature [17]–[22], this is done by allocating the jamming power so that the achievable capacity of the jammed links is minimized. This strategy is theoretically optimal. However, it requires knowledge of many factors that are generally not available for the jammer, such as the channel gain of the jammed link and the power allocation of a jammed node. In this paper, we model the jammer in a similar way. Observing that it is often unrealistic to assume that the jammer is able to acquire the information mentioned above, we model two types of jammers with different capabilities in terms of information acquisition. The first is assumed to be only capable of sensing nearby transmissions by measuring the interference plus noise level at its own position; the second has perfect knowledge of the legitimate user strategies, as well as the channel gain to the jammed node, i.e., the same capabilities as the jammers in [17]–[22]. These two models set an upper and lower bound for a typical real jammer, which would have knowledge between these two ideal cases. We let the jammer update its own strategy once the jammed nodes have updated theirs. In this sense, the jammer is reactive.

It is worth noting that we have intentionally neglected some low-level details of the jamming pattern. For example, it is possible for a jammer to selectively jam only a small part of the transmitted packets; the jammer might also leverage location information to produce a highly-directive spatial jamming pattern. Our objective is not to propose a technique specialized against a particular jamming scheme. Rather, we seek to demonstrate the performance benefits of anti-jamming techniques based on leveraging cooperative diversity, which can be used by any anti-jamming schemes once appropriately specialized. The general model in the paper provides us with a way to showcase the proposed anti-jamming scheme; but the idea can be extended and specialized and is not limited to the simple jammer.

We do not consider the case of jamming of control messages. These messages typically incur small overhead and are transmitted with highly reliable modulations such as BPSK and with strong channel coding. If a jammer is capable of destructing the control messages, likely it is also able to completely disrupt the network. In such a case, techniques exploiting more degrees of diversity, including the space diversity, should be used, as discussed in Section I.

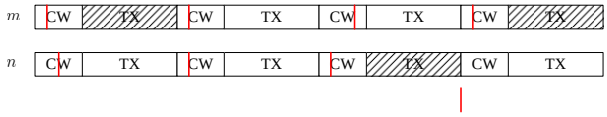


Fig. 2: Illustration of the multichannel slotted CSMA MAC.

MAC Layer Cooperation Model. The cooperative anti-jamming strategy is based on joint control of functionalities at the MAC and physical layers. At the MAC layer, users regulate their channel access probabilities to give higher opportunities to transmit to nodes that are being jammed. To achieve this objective, an opportunistic spectrum access scheme with tunable spectrum access probability is needed at the MAC layer. We consider one based on a state-of-the-art slotted multichannel CSMA. However, any other MAC protocols that enables stochastic channel access can be used, as long as the users are able to adjust the channel access probability. Details of the protocol are introduced below.

Similar to traditional CSMA-based MAC protocols, the considered protocol is based on contention. Transmission time is divided into a set of consecutive timeslots, and all nodes are assumed to be synchronized. At each timeslot, a user chooses (at most) one channel to transmit on, as in frequency hopping methods. However, the channel is chosen "randomly", i.e., a user chooses each channel with a certain probability. A node is also allowed to choose none and in this case it can serve as a relay for other users.

If a node chooses channel $f \in \mathcal{F}$, it first senses the channel at the beginning of the timeslot, to decide if the channel is available. Since there may be multiple users choosing the same channel, contention may happen. As in typical CSMA-based protocols, a competing node sets a random backoff and starts counting down, and the first one counting to zero wins the contention. To simplify the formulation, we assume that the contention window size is sufficiently large, so that the probability of collision is negligible. An example of the MAC protocol is illustrated in Fig. 2.

With such MAC protocol, a user can tune its channel access probability by simply adjusting the channel sensing probability on different channels. We let $q_n^f, f \in \mathcal{F}$ denote the channel sensing probabilities of user n on channel f. Since with non-zero probability node n may delay its own transmission and serve as relay for other nodes, we have $\sum_{f\in\mathcal{F}} q_n^f \leq 1$.

Physical Layer Cooperation Model. Physical-layer cooperation is obtained through relaying [32] [33]. Instead of transmitting its own traffic, a user can act as a relay and cooperatively transmit a packet on behalf of another user. Cooperative transmission is typically achieved by dividing the available transmission time into two phases: in the first phase, the transmitter broadcasts the message to both the destination and the relay; in the second phase, the relay forwards the received message to the destination, which then combines the two copies of the message and decodes. We

1536-1276 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2016.2569083, IEEE Transactions on Wireless Communications

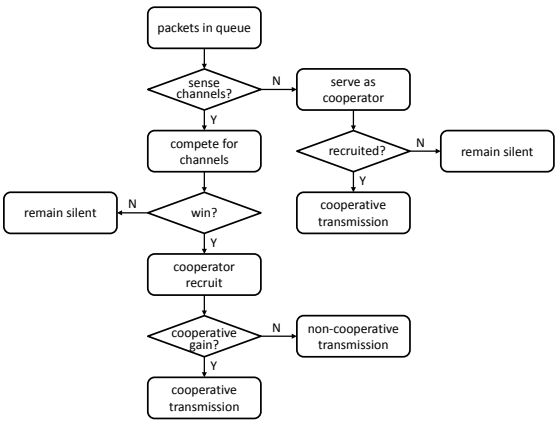


Fig. 3: Behavior of a legitimate user

focus on the decode-and-forward (DF) variant of cooperative communications, under which the relay node forwards the packet only when the information received from the source node can be successfully decoded. The analysis in this paper can be extended to other forwarding strategies, e.g., amplify-and-forward (AF).

To cope with the dynamic nature of the jammer, we consider a dynamic relay selection strategy to let the users form virtual MISO links. At each timeslot, a user that chooses not to sense any channels will act as a relay for another user if there is a positive cooperative gain[1].

Strategy of Legitimate Users. The behavior of a legitimate user can be illustrated through the flowchart in Fig. 3. When a user is backlogged, it selects its channel sensing probability for each channel. If it chooses not to sense any channel, it serves as a potential cooperator for other legitimate users. Intuitively, the cooperative relaying probability - which will be derived formally in Section III - is a function of the channel sensing probability q_n^f, of the strategy of the jammer $\boldsymbol{p}_{\rm j}$, and of the network topology. Users that choose to sense the same channel compete for channel access by setting random backoffs, and the winner has the privilege to transmit.

The factors determining cooperation at the two layers, i.e., channel access probability and cooperative relaying probability, are both functions of the channel access probability, for a given jammer strategy and network topology. Therefore, we can simply use the channel access probability as the strategy space of a legitimate user, and denote it as $\boldsymbol{q}_n = (q_n^f)_{f\in\widetilde{\mathcal{F}}}$ with $\widetilde{\mathcal{F}} \triangleq \mathcal{F}\cup\{0\}$, where q_n^f indicates the sensing probability of channel f, and q_n^0 denotes the probability that n does not sense any channel. Then, it needs be

$$q_n^f > 0,\ \forall n\in\mathcal{N},\ \forall f\in\widetilde{\mathcal{F}} \tag{2}$$

$$q_n^f \le 1,\ \forall n\in\mathcal{N},\ \forall f\in\widetilde{\mathcal{F}} \tag{3}$$

$$\mathbf{1}^{\rm T}\boldsymbol{q}_n = 1,\ \forall n\in\mathcal{N}. \tag{4}$$

We further denote by $\boldsymbol{q} = (\boldsymbol{q}_n)_{n\in\mathcal{N}}$ the sensing probability profile of all users in $\mathcal{N}$, and by $\boldsymbol{q}_{-n} = (\boldsymbol{q}_m)_{m\in\mathcal{N}/n}$ the profile of all users except for n.

[1]Cooperative transmission does not always outperform direct transmission. The cooperative gain depends on the strategy of the jammer, the network topology and the instantaneous channel states - see Section III for details.

Legitimate Problem Statement. Our objective is to maximize the total utility of all legitimate users, which represents the fairness-constrained network throughput and will be defined formally in Section III, by choosing the optimal sensing probability profile for each user, for any given strategy of the jammer.

III. Problem Formulation

We consider the *expected capacity of a legitimate user* $n\in\mathcal{N}$, expressed as

$$C_n(\boldsymbol{q},\ \boldsymbol{p}_{\rm j}) = \sum_{f\in\mathcal{F}} q_n^f \rho_n^f(\boldsymbol{q},\ \boldsymbol{p}_{\rm j}) C_n^f(\boldsymbol{q},\ \boldsymbol{p}_{\rm j}), \tag{5}$$

where q_n^f is the probability that user n senses channel f, $\rho_n^f(\boldsymbol{q},\ \boldsymbol{p}_{\rm j})$ represents the probability that user n is able to successfully access the channel, and $C_n^f(\boldsymbol{q},\ \boldsymbol{p}_{\rm j})$ is the achievable capacity on that channel (through either direct transmission or by using a cooperative relay), for a given sensing probability profile $\boldsymbol{q}$ and jamming power profile $\boldsymbol{p}_{\rm j}$.

Channel Access Probability. MAC-layer cooperation is achieved through stochastic channel access. According to the slotted multichannel CSMA protocol described in Section II, user $n\in\mathcal{N}$ is able to successfully access channel $f\in\mathcal{N}$ if i) the channel is sensed to be idle at session n's source node $\mathrm{s}(n)$; and ii) session n wins the channel access competition. If we let $\widetilde{\rho}_n^f(\boldsymbol{p}_{\rm j})$ indicate the probability that channel f is idle and $\widehat{\rho}_n^f(\boldsymbol{q})$ the probability that session n wins the competition, the channel access probability $\rho_n^f(\boldsymbol{q},\ \boldsymbol{p}_{\rm j})$ in (5) can be expressed as

$$\rho_n^f(\boldsymbol{q},\ \boldsymbol{p}_{\rm j}) = \widetilde{\rho}_n^f(\boldsymbol{p}_{\rm j})\widehat{\rho}_n^f(\boldsymbol{q}). \tag{6}$$

If we let $p_{\rm th}$ represent the power threshold below which a channel is sensed idle, then $\widetilde{\rho}_n^f(\boldsymbol{p}_{\rm J})$ can be defined as

$$\widetilde{\rho}_n^f(\boldsymbol{p}_{\rm j}) \triangleq \mathrm{P}\left(p_{\rm j}^f H_{\mathrm{js}(n)}\cdot(h_{\mathrm{js}(n)}^f)^2 + (\sigma_{\mathrm{s}(n)}^f)^2 \le p_{\rm th}\right), \tag{7}$$

with $H_{\mathrm{js}(n)},\ h_{\mathrm{js}(n)}^f$ capturing path loss and fading of the link between the jammer and session n's source node $\mathrm{s}(n)$ on channel f, respectively, and $(\sigma_{\mathrm{s}(n)}^f)^2$ being the noise power. For $h_{\mathrm{js}(n)}^f$ Rayleigh distributed with fading factor $\Omega_{\mathrm{s}(n)}^f$, $\widetilde{\rho}_n^f(\boldsymbol{p}_{\rm j})$ in (6) can be written as

$$\widetilde{\rho}_n^f(\boldsymbol{p}_{\rm j}) = \int_0^{x_{\max}} 1 - e^{-x^2/\Omega_{\mathrm{s}(n)}^f}\ \mathrm{d}x, \tag{8}$$

with $x_{\max}$ calculated from (7) as

$$x_{\max} = \sqrt{(p_{\rm th} - (\sigma_{\mathrm{s}(n)}^f)^2)/(p_{\rm j}^f H_{\mathrm{js}(n)})}. \tag{9}$$

We now need to derive the probability that a user $n\in\mathcal{N}$ wins the medium access competition after sensing the channel $f\in\mathcal{F}$ to be idle. Denoting $\mathcal{N}_n^f\subset\mathcal{N}/n$ as the set of nodes competing with user n on channel f, the winning probability for user n can be written as $\frac{1}{1+|\mathcal{N}_n^f|}$, where $|\mathcal{N}_n^f|$ is the number of nodes in $\mathcal{N}_n^f$. Since each potential competing user $m\in\mathcal{N}/n$ joins the access competition with probability $q_m^f\widetilde{\rho}_m^f(\boldsymbol{p}_{\rm j})$,

1536-1276 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2016.2569083, IEEE Transactions on Wireless Communications

the cardinality of $\mathcal{N}_n^f$, i.e., $|\mathcal{N}_n^f|$, can be proven to be Poisson distributed with mean [34]

$$\mathrm{E}(|\mathcal{N}_n^f|) = \sum_{m \in \mathcal{N}/n} q_m^f \widetilde{\rho}_m^f(\boldsymbol{p}_\mathrm{j}). \tag{10}$$

Then, the overall probability of winning a medium access competition for user n, i.e, $\widehat{\rho}_n^f(\boldsymbol{q})$ in (6), can be expressed as

$$\widehat{\rho}_n^f(\boldsymbol{q}) = \sum_{k=0}^{|\mathcal{N}|-1} \frac{1}{1+k} \cdot \frac{\left(\mathrm{E}(|\mathcal{N}_n^f|)\right)^k e^{-\mathrm{E}(|\mathcal{N}_n^f|)}}{k!}. \tag{11}$$

Expected Capacity. Suppose that user $n \in \mathcal{N}$ has won the competition to access channel f. We can then derive the expected capacity achievable through either direct transmission or using a cooperative relay, i.e., $C_n^f(\boldsymbol{q},\ \boldsymbol{p}_\mathrm{j})$ in (5).

If direct transmission is used by n, denote the direct link capacity by $C_{n,f}^{\mathrm{dir}}(\boldsymbol{p}_\mathrm{j})$. Then, we have

$$C_{n,f}^{\mathrm{dir}}(\boldsymbol{p}_\mathrm{j}) = B \log\left(1 + \gamma_{n,f}^{\mathrm{s2d}}(\boldsymbol{p}_\mathrm{j})\right), \tag{12}$$

where B is the bandwidth of each channel, and

$$\gamma_{n,f}^{\mathrm{s2d}}(\boldsymbol{p}_\mathrm{j}) \triangleq \frac{p_n H_n \cdot (h_n^f)^2}{(\delta_{\mathrm{d}(n)}^f)^2 + p_\mathrm{j} H_{\mathrm{jd}(n)} \cdot (h_{\mathrm{jd}(n)})^2} \tag{13}$$

where p_n is the transmission power of user n; H_n and h_n^f are the path loss and fading, respectively; $(\delta_{\mathrm{d}(n)}^f)^2$ is the noise power at the destination of user n denoted by $\mathrm{d}(n)$ on channel f. The expected capacity achievable with a direct link, denoted by $\widehat{C}_{n,f}^{\mathrm{dir}}(\boldsymbol{p}_\mathrm{j})$, can be computed by averaging over all possible channel fading outcomes of the links between $\mathrm{s}(n)$ and $\mathrm{d}(n)$, and the jammer and $\mathrm{d}(n)$, i.e.,

$$\begin{aligned}\widehat{C}_{n,f}^{\mathrm{dir}}(\boldsymbol{p}_\mathrm{j}) = \int_0^\infty \int_0^\infty & C_{n,f}^{\mathrm{dir}}(\boldsymbol{p}_\mathrm{j}) \cdot \mathrm{P}(h_n^f = x_1) \\ & \cdot \mathrm{P}(h_{\mathrm{jd}(n)} = x_2) \mathrm{d}x_1\ \mathrm{d}x_2.\end{aligned} \tag{14}$$

As discussed in Section II, each source node $m \in \mathcal{N}/n$ serves as a potential relay with probability q_m^0. Therefore, with a certain probability, user n will receive cooperation assistance by one of the potential cooperators. Suppose user n chooses $\mathrm{s}(m)$ as the relay, then, the resulting cooperative capacity denoted by $C_{nm,f}^{\mathrm{cop}}(\boldsymbol{p}_\mathrm{j})$ can be expressed as [32]

$$C_{nm,f}^{\mathrm{cop}}(\boldsymbol{p}_\mathrm{j}) = \frac{B}{2} \log\left(1 + \min\left(\gamma_{nm,f}^{\mathrm{s2r}}, \gamma_{n,f}^{\mathrm{s2d}} + \gamma_{mn,f}^{\mathrm{r2d}}\right)\right), \tag{15}$$

where $\gamma_{nm,f}^{\mathrm{s2r}} = \gamma_{nm,f}^{\mathrm{s2r}}(\boldsymbol{p}_\mathrm{j})$ and $\gamma_{mn,f}^{\mathrm{r2d}} = \gamma_{mn,f}^{\mathrm{r2d}}(\boldsymbol{p}_\mathrm{j})$ represents the SINR (defined as in (13)) of the link from source to relay, and from relay to destination, respectively.

Note, from (12) and (15), that the cooperative capacity $C_{nm,f}^{\mathrm{cop}}(\boldsymbol{p}_\mathrm{j})$ can be higher or lower than the direct capacity (because of the $\frac{1}{2}$ coefficient in (15)). If we define the following indicator function

$$\mathrm{I}(x, y) \triangleq \begin{cases} 1, & \text{if } x > y \\ 0, & \text{otherwise,} \end{cases} \tag{16}$$

then, the expected capacity achievable through cooperative communication (assuming that cooperative transmission outperforms direct transmission) can be defined as

$$\widehat{C}_{nm,f}^{\mathrm{cop}}(\boldsymbol{p}_\mathrm{j}) \triangleq \mathrm{E}\left(C_{nm,f}^{\mathrm{cop}}(\boldsymbol{p}_\mathrm{j}) | \mathrm{I}\left(C_{nm,f}^{\mathrm{cop}}(\boldsymbol{p}_\mathrm{j}), C_{n,f}^{\mathrm{dir}}(\boldsymbol{p}_\mathrm{j})\right) = 1\right) \tag{17}$$

The expected capacity achieved through cooperative communication in (17) can be computed by averaging over all possible channel fading outcomes of the links, for each channel $f \in \mathcal{F}$:

1) $h_{\mathrm{jd}(n)}^f$: from jammer to $\mathrm{d}(n)$;
2) $h_{\mathrm{js}(m)}^f$: from jammer to $\mathrm{s}(n)$;
3) h_n^f: from $\mathrm{s}(n)$ to $\mathrm{d}(n)$;
4) h_{nm}^f: from $\mathrm{s}(n)$ to $\mathrm{s}(m)$;
5) $\hat{h}_{mn}^f$: from $\mathrm{s}(m)$ to $\mathrm{d}(n)$.

Therefore, we have

$$\begin{aligned}\widehat{C}_{nm,f}^{\mathrm{cop}}(\boldsymbol{p}_\mathrm{j}) = \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty & C_{nm,f}^{\mathrm{cop}}(\boldsymbol{p}_\mathrm{j}) \\ & \cdot \mathrm{I}\left(C_{nm,f}^{\mathrm{cop}}(\boldsymbol{p}_\mathrm{j}), C_{n,f}^{\mathrm{dir}}(\boldsymbol{p}_\mathrm{j})\right) \mathrm{P}(h_{\mathrm{jd}(n)}^f = x_1) \\ & \cdot \mathrm{P}(h_{\mathrm{js}(m)}^f = x_2) \mathrm{P}(h_n^f = x_3) \\ & \cdot \mathrm{P}(h_{nm}^f = x_4) \mathrm{P}(\hat{h}_{mn}^f = x_5) \\ & \mathrm{d}x_1\ \mathrm{d}x_2\ \mathrm{d}x_3\ \mathrm{d}x_4\ \mathrm{d}x_5.\end{aligned} \tag{18}$$

The resulting probability that user n achieves a capacity gain through cooperative relaying can then be represented as

$$\begin{aligned}\phi_{nm}^f(\boldsymbol{p}_\mathrm{j}) = q_m^0 \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty & \\ \mathrm{I}\left(C_{nm,f}^{\mathrm{cop}}(\boldsymbol{p}_\mathrm{j}), C_{n,f}^{\mathrm{dir}}(\boldsymbol{p}_\mathrm{j})\right) & \cdot \mathrm{P}(h_{\mathrm{jd}(n)}^f = x_1) \\ \cdot \mathrm{P}(h_{\mathrm{js}(m)}^f = x_2) & \mathrm{P}(h_n^f = x_3) \\ \cdot \mathrm{P}(h_{nm}^f = x_4) & \mathrm{P}(\hat{h}_{mn}^f = x_5) \\ \mathrm{d}x_1\ \mathrm{d}x_2\ \mathrm{d}x_3\ \mathrm{d}x_4\ \mathrm{d}x_5, &\end{aligned} \tag{19}$$

and the corresponding sum probability can be written as

$$\phi_n^f(\boldsymbol{p}_\mathrm{j}) = \sum_{m \in \mathcal{N}/n} \phi_{nm}^f(\boldsymbol{p}_\mathrm{j}). \tag{20}$$

Finally, the expected capacity achievable by user n over channel f can be expressed as

$$\begin{aligned}C_n^f(\boldsymbol{q},\ \boldsymbol{p}_\mathrm{j}) = & \sum_{m \in \mathcal{N}/n} q_m^0 \widehat{C}_{nm,f}^{\mathrm{cop}}(\boldsymbol{p}_\mathrm{j}) \\ & + \sum_{m \in \mathcal{N}/n} \left(1 - \phi_n^f(\boldsymbol{p}_\mathrm{j})\right) \widehat{C}_{n,f}^{\mathrm{dir}}(\boldsymbol{p}_\mathrm{j}).\end{aligned} \tag{21}$$

Note that (21) is exact when the probability that more than one cooperator participates in cooperative communication is very low. Otherwise, the capacity expression will be obtained as a sum of the expected cooperative capacities contributed by different cooperators. However, we can show through experimental results that in most cases this assumption is true. Readers are referred to Appendix A for details.

Social Problem Statement. So far, we have derived the expected capacity of each user $n \in \mathcal{N}$. If we consider a proportional fairness criterion, then the utility of each user can be defined as

$$U_n(\boldsymbol{q},\ \boldsymbol{p}_\mathrm{j}) \triangleq \log(C_n(\boldsymbol{q},\ \boldsymbol{p}_\mathrm{j})), \tag{22}$$

and the ideal objective is to maximize the sum utility of all users, i.e.,

$$\begin{array}{ll} \text{Given :} & \boldsymbol{p}_\mathrm{j} \\ \underset{\boldsymbol{q} \in (0,1]^{|\mathcal{N}|}}{\text{Maximize}} & U(\boldsymbol{q}, \boldsymbol{p}_\mathrm{j}) = \sum\limits_{n \in \mathcal{N}} U_n(\boldsymbol{q},\ \boldsymbol{p}_\mathrm{j}) \\ \text{Subject to :} & (2),\ (3),\ (4). \end{array} \tag{23}$$

1536-1276 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2016.2569083, IEEE Transactions on Wireless Communications

However, this objective is not easily achievable with distributed control. In fact, the optimization problem is non-convex and the utility expressions in (22) are rather complex. Moreover, the non-convexity also implies that only suboptimal solutions can be computed in polynomial time even with centralized algorithms. Since we would like to design distributed solutions with low complexity, we follow here a different approach and design a pricing-based distributed solution algorithm with provable convergence to a stationary point of the social problem.

We will introduce the distributed algorithm (which is our core contribution) in Section IV and also propose a centralized algorithm as a benchmark in Section V.

Jamming Strategy. Two types of reactive jammers are considered, differentiated based on the information available to them. The first type is only capable of measuring the interference plus noise level per channel, at its own location. Without any further information, all the jammer can do is to use this value as a rough estimation of the traffic on different channels. A reasonable strategy, under theses circumstances, is to allocate the power budget on the channels proportional to the estimated traffic on each of them. Denote the sensed interference plus noise level by j on channel f as D_j^f. Then the strategy is

$$p_{\mathrm{j}}^f = \frac{D_{\mathrm{j}}^f}{\sum_{f\in\mathcal{F}} D_{\mathrm{j}}^f} \cdot p_{\mathrm{j}}^{\max}, \forall f \in \mathcal{F}. \tag{24}$$

It is often the reality that the jammer has no further information other than this, so we refer to this type of jammer as the "standard" jammer.

We further model a more sophisticated jammer, which we refer to as the "optimal" jammer. The objective of an optimal jammer is to minimize

$$\sum_{n\in\mathcal{N}} \sum_{f\in\mathcal{F}} q_n^f \rho_n^f C_{n,f}^{dir}(\boldsymbol{p}_{\mathrm{j}}), \tag{25}$$

i.e., the sum of legitimate user capacity, assuming no physical layer cooperation. We do not consider physical layer cooperation here, since it is infeasible to solve the minimization of $\sum_{n\in\mathcal{N}} C_n^f(\boldsymbol{q},\ \boldsymbol{p}_{\mathrm{j}})$ [2]. The objective of the optimal jammer is, therefore, to minimize a lower bound of the sum capacity of the legitimate users. The gap between the result and the real sum capacity is introduced by physical layer cooperation. Apparently the "optimality" is built on the knowledge of some impractical assumptions (it is almost impossible for a jammer to know exactly how the user capacity changes with different jamming strategies), but we will use it as a benchmark to evaluate the anti-jamming scheme nonetheless.

The strategy of the optimal jammer is formally

$$\begin{array}{ll} \text{Given :} & \boldsymbol{q} \\ \underset{\boldsymbol{p}_j}{\text{Minimize}} & \sum\limits_{n\in\mathcal{N}} \sum\limits_{f\in\mathcal{F}} q_n^{\mathrm{f}} \rho_n^f C_{n,f}^{dir}(\boldsymbol{p}_{\mathrm{j}}) \\ \text{Subject to :} & \mathbf{1}^T \boldsymbol{p}_{\mathrm{j}} = p_{\mathrm{j}}^{max} \end{array} \tag{26}$$

The problem (26) can be solved using an approach similar to the water-filling algorithm, but in a "reverse" way. For details, see Appendix B.

[2]Notice that, the binary indicator $\mathrm{I}\left(C_{nm,f}^{\mathrm{cop}}(\boldsymbol{p}_{\mathrm{j}}), C_{n,f}^{\mathrm{dir}}\right)$ implies that (18) is not a differentiable function in $\boldsymbol{p}_j$.

IV. Distributed Solution Algorithm

The distributed solution algorithm is designed based on the recent framework results in [35], with the objective to achieve a stationary solution point of the social problem (23). Specifically, we design an iterative best-response algorithm based on a pricing mechanism. At each iteration, each session n maximizes its own utility minus a pricing term that acts as a penalty imposed to each session for being too aggressive in choosing its own strategy and thus "hurting" other sessions. The challenge in applying the framework is to design the pricing term of each iteration so that the distributed algorithm converges to a "good" stationary point (if more than one exists) of the social problem in (23) [3]. Since we are designing an algorithm for legitimate users for which the strategy of the jammer, i.e., $\boldsymbol{p}_{\mathrm{j}}$ is a given parameter, we will neglect it from the utility function for simplicity in this section and Section V.

We denote by $\boldsymbol{q}^\nu$ the sensing probability profile of iteration ν (with $\nu = 1, 2, \dots$). The pricing term for session $n \in \mathcal{N}$, denoted as $\Gamma_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n}^\nu)$, can be written as

$$\Gamma_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n}^\nu) \triangleq (\boldsymbol{q}_n)^{\mathrm{T}} \left(\Gamma_n^f(q_n^f, \boldsymbol{q}_{-n}^\nu)\right)_{f\in\widetilde{\mathcal{F}}} - \frac{\tau_n}{2}\|\boldsymbol{q}_n - \boldsymbol{q}_n^\nu\|^2, \tag{27}$$

where

$$\Gamma_n^f(q_n^f, \boldsymbol{q}_{-n}^\nu) \triangleq \sum_{m\in\mathcal{N}/n} \frac{\partial U_m(\boldsymbol{q}^\nu)}{\partial q_n^f} \tag{28}$$

represents the marginal decrease of the sum-utility of the other sessions due to a variation of session n's sensing probability associated with channel f. Here $-\frac{\tau_n}{2}\|\boldsymbol{q}_n - \boldsymbol{q}_n^\nu\|^2$ is a proximal regulation with constant τ_n, whose value needs to be chosen properly to guarantee strong concavity of the resulting penalized utility function, and at the same time to prevent each session n from being too conservative in changing its sensing probability profile. To discuss the convergence, we first introduce Lemma 1.

Lemma 1: Given the sensing probability profiles of all other users $\boldsymbol{q}_{-n}$, the utility function $U_n(\boldsymbol{q}_n,\ \boldsymbol{q}_{-n})$ defined in (22) is strongly concave with respect to $\boldsymbol{q}_n$.

Proof: Since $C_n(\boldsymbol{q})$ is a linear function of $\boldsymbol{q}_n$ given $\boldsymbol{q}_{-n}$ and $C_n(\boldsymbol{q}) > 0$, $U_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n})$ in (22) is concave over $\boldsymbol{q}_n$. Therefore all we need to prove is that the second derivative $\nabla^2_{\boldsymbol{q}_n} U_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n})$ is bounded for $\forall \boldsymbol{q}_{-n} \in \Phi_{-n} \triangleq (\Phi_m)_{m\in\mathcal{N}/n}$ with

$$\Phi_m \triangleq \{\boldsymbol{q}_m | \text{constraints} : (2), (3), (4)\}. \tag{29}$$

The second derivative of $U(\boldsymbol{q}_n, \boldsymbol{q}_{-n})$ with respect to $\boldsymbol{q}_n$ can be written as

$$\begin{aligned} \nabla^2_{\boldsymbol{q}_n} U(\boldsymbol{q}_n, \boldsymbol{q}_{-n}) = &\frac{1}{C_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n})} \nabla^2_{\boldsymbol{q}_n} C_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n}) \\ &- \frac{1}{C_n^2(\boldsymbol{q}_n, \boldsymbol{q}_{-n})} \nabla_{\boldsymbol{q}_n} C_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n}). \end{aligned} \tag{30}$$

It can be verified that both $\nabla_{\boldsymbol{q}_n} C_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n})$ and $\nabla^2_{\boldsymbol{q}_n} C_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n})$ are bounded for closed Φ_n. Since we

[3]Note that problem (23) is non-convex. Therefore, it is computationally infeasible in general to find the optimal solution. The proposed algorithm is "good" in the sense that: (i) it is theoretically guaranteed to converge to a non-local-minimum stationary point; (ii) in practice, it achieves close-to-globally-optimal performance.

1536-1276 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2016.2569083, IEEE Transactions on Wireless Communications

let $q_n^f > 0$ for $\forall n \in \mathcal{N}, f \in \mathcal{F}$, $\frac{1}{C_n(\boldsymbol{q}_n,\boldsymbol{q}_{-n})}$ and $\frac{1}{C_n^2(\boldsymbol{q}_n,\boldsymbol{q}_{-n})}$ are also bounded. Hence, $\nabla^2_{\boldsymbol{q}_n} U(\boldsymbol{q}_n, \boldsymbol{q}_{-n})$ is bounded. ■

Because Lemma 1 guarantees strong concavicity, we can set $\tau_n = 0$. The formal description of the algorithm is given in Algorithm 1, where the penalized version of utility function $U_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n})$, denoted by $\widetilde{U}_n(\boldsymbol{q}_n,\ \boldsymbol{q}_{-n})$, is defined as

$$\widetilde{U}_n(\boldsymbol{q}_n,\ \boldsymbol{q}_{-n}) \triangleq U_n(\boldsymbol{q}_n,\ \boldsymbol{q}_{-n}) + \Gamma_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n}^{\nu}, 0), \tag{31}$$

with $\Gamma_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n}^{\nu}, 0)$ defined in (27). The convergence properties of Algorithm 1 are given in Theorem 1 below, where ζ is a parameter to guarantee the convergence of the algorithm.

Algorithm 1: Pricing Jacobi Algorithm

Data : $\{\zeta^{\nu}\} > 0$; Set $\nu = 0$.
(S.1) : If $\boldsymbol{q}^{\nu}$ satisfies a suitable termination criterion: STOP;
(S.2) : For all $n \in \mathcal{N}$, compute

$$\widehat{\boldsymbol{q}}_n(\boldsymbol{q}^{\nu}) \triangleq \underset{\boldsymbol{q}_n \in \Phi_n}{\arg\max}\ \widetilde{U}_n(\boldsymbol{q}_n,\ \boldsymbol{q}_{-n}) \tag{32}$$

(S.3) : Set $\boldsymbol{q}_n^{\nu+1} = \boldsymbol{q}_n^{\nu} + \zeta^{\nu}(\widehat{\boldsymbol{q}}_n(\boldsymbol{q}^{\nu}) - \boldsymbol{q}_n^{\nu})$.
(S.4) : $\nu \leftarrow \nu + 1$ and go to (S.1).

Theorem 1 (Convergence Condition): Given the social problem (23), suppose that $\{\zeta^{\nu}\}$ is chosen so that

$$\zeta^{\nu} \in (0, 1],\ \zeta^{\nu} \to 0,\ \text{and}\ \sum_{\nu} \zeta^{\nu} = +\infty. \tag{33}$$

Then, either Algorithm 1 converges in a finite number of iterations to a stationary solution of (23), or every limit point of the sequence $\{\boldsymbol{\zeta}^{\nu}\}$ (at least one of such points exists) is a stationary solution of (23). Moreover, no such point is a local minimum of the social function.

Proof: Based on the Descent Lemma in [36], it can be proven that the algorithm always converges to a feasible solution point of the social problem in (23). Then, together with Lemma 1, it can be further proven that each such solution point is also stationary for the social problem. An example of sequence η^{ν} satisfying conditions (33) in Theorem 1 is [35]:

$$\zeta^{\nu} = \frac{\zeta^{\nu-1} + \alpha(\nu)}{1 + \beta(\nu)},\ \nu = 1, \ldots, \tag{34}$$

where $\alpha(\nu) = \alpha$ and $\beta(\nu) = \nu\beta$ with $\alpha, \beta \in (0, 1)$ and $\alpha \leq \beta$. ■

V. Centralized Solution Algorithm

We now present a centralized solution algorithm to solve the social problem (23) to provide a performance benchmark for the distributed solution algorithm proposed in Section IV.

Objective. Denote U^* as the global optimum of the social problem, and $\varepsilon \in (0, 1]$ as a predefined optimality precision. Then the objective of the algorithm is to obtain an ε-optimal solution $\boldsymbol{q}$ satisfying

$$U(\boldsymbol{q}) \geq \varepsilon U^*. \tag{35}$$

Here, the optimality precision ε can be set as close to 1 as we wish at the price of computational complexity.

Denote $\mathrm{UP}_{\mathrm{glb}}$ as a global upper bound, and $\mathrm{LR}_{\mathrm{glb}}$ as a global lower bound on the sum-utility $U(\boldsymbol{q})$ in (23). Then it must be

$$\mathrm{LR}_{\mathrm{glb}} \leq U^* \leq \mathrm{UP}_{\mathrm{glb}}. \tag{36}$$

Then, the algorithm searches for the ε-optimal solution by iteratively updating $\mathrm{UP}_{\mathrm{glb}}$ and $\mathrm{LR}_{\mathrm{glb}}$ so that, the two bounds get closer and closer to each other, until

$$\mathrm{LR}_{\mathrm{glb}} \geq \varepsilon \cdot \mathrm{UP}_{\mathrm{glb}}. \tag{37}$$

We implement the above iteration based on a combination of the *branch-and-bound* framework and convex relaxations [37].

Algorithmic Framework. We solve a series of subproblems of the original social problem (23), obtained by partitioning its domain into a set of subdomains. Denote $\Phi_{\mathcal{N}} = \prod_{n\in\mathcal{N}} \Phi_n$ as the joint domain of all the users in $\mathcal{N}$ with Φ_n defined in (29); and $\Phi = \{\Phi_{\mathcal{N}}^i, i = 0, 1, 2, \ldots\}$ as the set of subdomains, with i denoting the subdomain index, $\Phi_{\mathcal{N}}^i = \Phi_{\mathcal{N}}$ for $i = 0$, and $\Phi_{\mathcal{N}}^i \subset \Phi_{\mathcal{N}}$ for the others. For each subproblem $\Phi_{\mathcal{N}}^i$, denote the local upper and lower bounds on sum-utility $U(\boldsymbol{q})$ by $\mathrm{UP}(\Phi_{\mathcal{N}}^i)$ and $\mathrm{LR}(\Phi_{\mathcal{N}}^i)$, respectively. Then, the global upper bound $\mathrm{UP}_{\mathrm{glb}}$ and lower bound $\mathrm{LR}_{\mathrm{glb}}$ are updated as follows.

$$\mathrm{UP}_{\mathrm{glb}} = \max_{i=0,1,\ldots} \{\mathrm{UP}(\Phi_{\mathcal{N}}^i)\} \tag{38}$$

$$\mathrm{LR}_{\mathrm{glb}} = \max_{i=0,1,\ldots} \{\mathrm{LR}(\Phi_{\mathcal{N}}^i)\}. \tag{39}$$

The algorithm then checks how close the obtained global bounds are to each other. If the termination criterion (37) is satisfied, the algorithm terminates and sets the ε-optimal solution as $U(\boldsymbol{q}) = \mathrm{LR}_{\mathrm{glb}}$, and sets $\boldsymbol{q}$ accordingly; otherwise, the algorithm chooses one subdomain from Φ, partitions it into two smaller subdomains, then calculates the local upper and lower bounds for them each, and again updates $\mathrm{UP}_{\mathrm{glb}}$ and $\mathrm{LR}_{\mathrm{glb}}$. The above procedure is repeated until the gap between $\mathrm{UP}_{\mathrm{glb}}$ and $\mathrm{LR}_{\mathrm{glb}}$ converges to 0, and hence [according to (36)] converges to the global optimum U^*.

Convex Relaxation. In the above iterations, for a given $\Phi_{\mathcal{N}}^i$, the corresponding local upper bound $\mathrm{UP}(\Phi_{\mathcal{N}}^i)$ needs to be easy to compute. To this end, we rely on convex relaxation, i.e., we relax the original nonlinear nonconvex problem into a convex one that is easy to solve using standard convex programming techniques. We call the solution obtained by solving the relaxed optimization problem *relaxed solution*. Since the relaxed solution is also a feasible solution, we compute the sum throughput based on (22), and set the local lower bound $\mathrm{LR}(\Phi_{\mathcal{N}}^i)$ to the resulting solve.

To relax the objective function in (23) to be convex, we only need to relax the individual utility function of each user. Different approaches can be used (see [37] for details of possible relaxation techniques). Here, we adopt a simple but effective relaxation method based on the observations that $U_n(\boldsymbol{q})$ is a monotonically decreasing function with respect to q_m^f for any $f \in \mathcal{F}$. For given $\Phi_{\mathcal{N}}^i$, denoting the range of q_m^f as $[q_{m,f}^{\mathrm{L}}\ q_{m,f}^{\mathrm{U}}]$, $U_n\left(\boldsymbol{q}_n, (q_{m,f}^{L})_{m\in\mathcal{N}/n}^{f\in\mathcal{F}}\right)$ provides an upper bound on $U_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n})$. By deriving the first and second derivatives of $U_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n})$ with respect to $\boldsymbol{q}_n$, it can be seen that $U_n(\boldsymbol{q}_n)$

1536-1276 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2016.2569083, IEEE Transactions on Wireless Communications

is a concave function whose global optimum can be easily computed, e.g., by using standard interior-point methods [38].

Variable Partition. We select the subdomain $\Phi^i_{\mathcal{N}}$ with the highest local upper bound from Φ for partition, i.e.,

$$i = \arg\max_i \mathrm{UP}(\Phi^i_{\mathcal{N}}). \tag{40}$$

The selected subproblem is then partitioned into two new subproblems by partitioning one of its variables, i.e., $\{q^f_n,\ n \in \mathcal{N}, f \in \mathcal{F}\}$. We select the variable with the largest range and partition it from the half, i.e., selecting q_{n^*,f^*} that satisfies

$$\{n^*, f^*\} = \arg\max_{n\in\mathcal{N}, f\in\mathcal{F}} (q^{\mathrm{U}}_{n,f} - q^{\mathrm{L}}_{n,f}) \tag{41}$$

and partition it as

$$q^{\mathrm{M}}_{n^*,f^*} = \frac{q^{\mathrm{U}}_{n^*,f^*} + q^{\mathrm{L}}_{n^*,f^*}}{2}, \tag{42}$$

which results in two new subproblems with domains of $[q^{\mathrm{L}}_{n^*,f^*}\ q^{\mathrm{M}}_{n^*,f^*}]$ and $[q^{\mathrm{M}}_{n^*,f^*}\ q^{\mathrm{U}}_{n^*,f^*}]$, respectively.

VI. Practical Considerations

Information Exchange. Player strategies (for both legitimate users and jammer) are coupled. A user needs to have some knowledge on strategies of other users to update its own. $q^f_m, \forall m \neq n, \forall f \neq 0$ are not needed by n, because it only affects ρ^f_n. And ρ^f_n can be inferred from previous channel contention results. $q^0_m \forall m \neq n$ cannot be inferred. Therefore, each node needs to periodically broadcast its own q^0, i.e., its probability to serve as a relay, to the nodes within contention range. Legitimate users can exchange information through a control channel.

The strategy of a legitimate user is also affected by the jammer's strategy, which can be estimated over a period of observation. Note, the interference at $d(n)$ is $p_{\mathrm{j}} H_{\mathrm{j}d(n)} \cdot (h_{\mathrm{j}d(n)})^2$, with $h_{\mathrm{j}d(n)})^2$ being a random variable with known distribution. As long as there are sufficient observations, $p_{\mathrm{j}} H_{\mathrm{j}d(N)}$ can be estimated. The exact value of p_{j} is not necessary since $H_{\mathrm{j}} d(n)$ is considered fixed in our scenario.

The standard jammer, only has to measure the interference plus noise level at its own position. It is assumed that all "necessary" information, such as the legitimate user strategies and the channel gain to the receivers, is available at the optimal jammer. Clearly, it is not feasible in practice. However, it must be pointed out that we are not trying to design "practical" jamming strategies in this paper; instead, we are trying to analyze the performance of the proposed anti-jamming scheme. The optimal jamming strategy provides an extreme scenario and the performance with it provides a benchmark in evaluating the anti-jamming scheme.

Cooperator Recruitment. A critical component of the proposed scheme is cooperator recruitment. The process of cooperator recruitment is performed after channel access competition. To reserve the channel and select the transmission scheme (i.e., rate and modulation), the transmitter and receiver exchange messages as follows. The transmitter sends an RTS to the receiver, and the receiver replies with a CTS message. During the message exchange, the receiver can estimate the channel coefficient from transmitter to receiver, and send it back to the transmitter. Meanwhile, the neighbors that have elected to serve as potential relays can also listen to the message exchange and estimate the channel coefficient from transmitter and receiver to themselves. After the channel has been reserved, the transmitter sends out a relay recruitment message to its neighbors. Potential relays reply with their estimated channel coefficients. In this way, the transmitter can calculate the capacity of the direct link as well as the cooperative throughput with each relay candidate, and therefore decide whether to use a relay and which relay to use. If there are multiple candidates, the transmitter will choose the one with the highest cooperative gain. Compared to traditional CSMA/CA, the only extra overhead is the message exchange between transmitter and relay candidates.

It is worth noting that, the expected cooperative capacity in (21) is obtained by summing up the contributions of every potential relay, so an important prerequisite for this equation to approximate well the real cooperative capacity is that the probability that two or more relays with cooperative gain exist simultaneously is negligible.

Fortunately, in the scenario we are focusing on, i.e., a network with moderate or heavy traffic load and a powerful jammer, it is reasonable to assume that most of the legitimate users are affected by the jammer simultaneously, and thus the probability that relays exist with positive cooperative gain is not high in most cases (see Appendix A). Besides, even if a relay is able to enhance another user's link through cooperation, it will only act as relay with a certain, typically small, probability (only when it chooses not to sense any channel). In most cases, then, the probability that multiple potential relays exist for a user in the same timeslot is very low, and the approximation in (21) is quite accurate.

For scenarios in which many cooperation opportunities exist, the cooperative capacity in (21) becomes an upper bound on the real value because of the overlap in cooperation probabilities.

Computational Cost. A major component in the computational complexity of the proposed algorithm lies in the computation of both expected cooperative capacity and probability with cooperative gain, shown in (18) and (19), respectively. The integrations seem impractical. However, as we will show through rigorous analysis, only moderate computational cost is incurred.

Before the formal analysis, several remarks on the integrations are listed,

1) By using reasonable approximations, the integrations can be dramatically simplified, so that the computational cost is significantly reduced.
2) Since the legitimate users take mixed strategies, it requires a jammer to observe a sufficiently long period for the legitimate strategies before a strategy update (see details in following discussions). Thus, the computational cost is diluted over time.
3) For any value of $\mathbf{p}_{\mathrm{j}}$, both (18) and (19) only need to be computed once. Therefore a legitimate user can maintain a lookup table for possible $\mathbf{p}_{\mathrm{j}}$ values over a finite set.

We will elaborate these remarks in detail.

1536-1276 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2016.2569083, IEEE Transactions on Wireless Communications

Experiment	1	2	3	4	5	6	7	8	9	10
Results (ms)	387.212	391.213	380.824	380.795	383.794	411.635	380.777	380.758	380.745	392.714

TABLE I: CPU time consumed to compute a loop of 10^6 logarithms

First of all, the integrations in (18) and (19) can be approximated with much simpler forms. In fact, the integrations are computed to account for every possible combinations of the channel gains for $I(C^{\rm cp}_{nm,f}(\mathbf{p}_{\rm j}, C^{\rm dir}_{n,f}(\mathbf{p}_{\rm j}))=1$. Recalling (15), we can use the SINR values instead of channel gains as the integrated variables. For example, (18) can be rewritten as

$$\widehat{C}^{\rm cop} = \int_{\gamma^{\rm s2r}} \int_{\gamma^{\rm s2d}+\gamma^{\rm r2d}} \frac{B}{2} \log\left(1+\min\left(\gamma^{\rm s2r}, \gamma^{\rm s2d}+\gamma^{\rm r2d}\right)\right) \cdot I(C^{\rm cop}, C^{\rm dir}) P(\gamma^{\rm r2d}) P(\gamma^{\rm s2d}+\gamma^{\rm r2d}) \cdot {\rm d}\gamma^{\rm r2d} {\rm d}(\gamma^{\rm s2d}+\gamma^{\rm r2d}), \quad (43)$$

with subscripts omitted. In this way, the integrations involve only 2 integrals, and the computational complexity is significantly reduced.

The distributions of the new integrated variables are not well defined as the channel gains (which are assumed to be Rayleigh distributed), but in practice, integrations are often approximated with summation over a number of small subintervals. In this case, the probability for each subinterval can be computed and stored in advance. Suppose the number of subintervals for each of the integrated variables is $N_{\rm int}$, then the time complexity for computing (18) and (19) is $O(N^2_{\rm int})$.

Second, in this paper we assume that the jammer tries to minimize the expected sum rate of the jammed nodes. Recall that the legitimate nodes follow mixed strategies, with a certain probability of taking an action on each of the channels. At the same time, the jammer, which is not aware of the exact distributions,[4] has to learn them based on observations. This implies that, in order to obtain a good estimate, the jammer must listen to the legitimate transmissions for a period long enough for the accumulated observations to be statistically significant. Because at least $|\mathcal{N}||\mathcal{F}|$ transmission periods are required for each node to try each channel at least once, this period is at least at the order of $O(|\mathcal{N}||\mathcal{F}|)$[5].

Third, for the same value of $\mathbf{p}_{\rm j}$, the corresponding (18) and (19) only need to be computed once. Therefore, a node can maintain a lookup table for the possible values of $\mathbf{p}_{\rm j}$ over a finite set. It updates the corresponding entry once a new $\mathbf{p}_{\rm j}$ value is encountered, and refers to the table once an old $\mathbf{p}_{\rm j}$ is used by the jammer.

Based on the above analysis, we design an on-demand look-up table based integration computation algorithm. For every updated jammer strategy that is not already in the look-up table, the computation for the corresponding item is conducted. The integrations are computed $\forall m \in \mathcal{V}_n, \forall f \in \mathcal{F}$ [6]. Therefore, the time complexity for one strategy update is $O(|\mathcal{V}_n||\mathcal{F}|N^2_{\rm int})$. However, as already discussed, the strategy update period also scales at an order (usually much) higher than $O(|\mathcal{N}||\mathcal{F}|)$. Therefore, the effective computational complexity is actually at a scale smaller than $O(N^2_{\rm int})$, with a controllable $N_{\rm int}$ affecting the precision of the integration.

Algorithm 2: On-demand Look-up Table Based Integration Computation Algorithm for Legitimate Users n

Data: The distances from $s(n)$ to $r(n)$ and the neighboring nodes.
(S.1) : Initialize look-up-table $T_m, \forall m \in \mathcal{V}_m$.
(S.2) : If jammer strategy $\mathbf{p}_{\rm j}$ is updated:
(S.3) : For $m \in \mathcal{V}_n, f \in \mathcal{F}$:
(S.4) : If the item in T_m corresponding to $\mathrm{p_j}$ is already computed: read it from T_m;
(S.5) : Else: compute (18) and (19) for $m \in \mathcal{V}_n$ with $\mathrm{p_j}$ and store it to T_m;

We conclude the above analysis with a case study. There are 10 legitimate users and 5 channels. The strategy update period is set to $10|\mathcal{N}||\mathcal{F}|$, so it takes 5×10^2 transmission periods to update the strategies. If $N_{\rm int}$ is set to 100, then there are 5×10^5 steps in all the loops, assuming that all nodes are mutually connected. Therefore, in every transmission period, 1000 computations of the integrated variables in (18) and (19) are needed.

Assuming the throughput is $10\,{\rm Mbits/s}$, a packet with 1000 Bytes takes around $800\,\mu$s to be transmitted. So the task is to compute 1000 logarithm computations (because the bottleneck to compute the variables inside the integrations is clearly the logarithm computation) within $800\,\mu$s.

To show these computations are feasible, we conducted experiments on a Raspberry Pi 2 Model B, with a quad-core ARM Cortex-A7 CPU clocked at 900 MHz [39]. To be specific, we run a C++ program with a loop composed of 10^6 steps. At each step, a logarithm function is computed. We use the clock() function to calculate the CPU time used. The experiment is conducted 10 times, and the results are shown in Table. I. In average, it takes $348.872\,{\rm ms}$ for a loop with 10^6 logarithm computations, or equivalently, $348.872\,\mu$s for 1000 steps, i.e., less than half of the transmission period.

VII. Performance Evaluation

System Setup. The topology is generated randomly. Specifically, all nodes are located in a $500\,{\rm m}\times500\,{\rm m}$ area, with the distance between the transmitter-receiver pairs generated uniformly between $250\,{\rm m}$ and $350\,{\rm m}$. The location of the jammer is fixed to the position $(250, 250)$. In most experiments, we assume that there are 2 different channels. The Rayleigh fading coefficients of the channels are set to different values generated uniformly from $[\frac{1}{2}, \frac{2}{3}]$. We set the path loss factor to 4.

The power of legitimate users is set to 1 W, while the average noise power is set to 1×10^{-10} W. The power of the jammer is set to different values in different experiments, but generally it is much higher than that of the users to better highlight the anti-jamming performance. We use in the

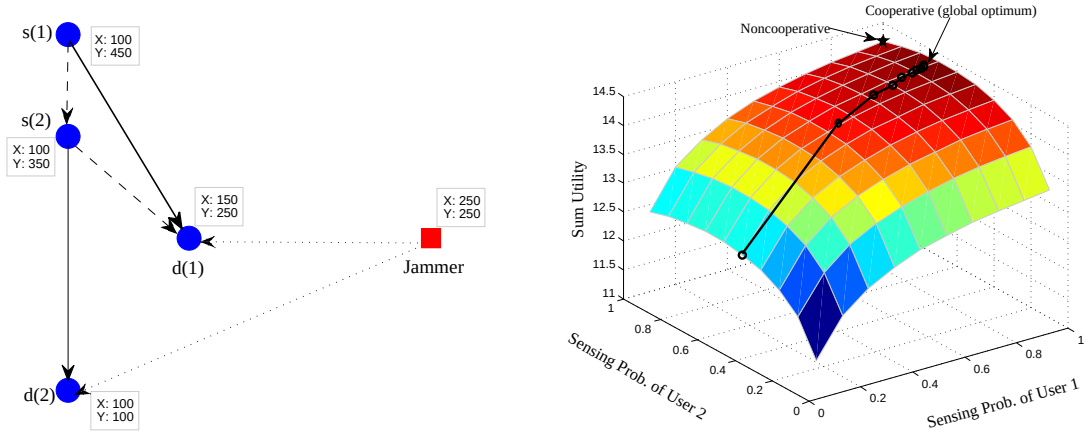
[4] we assume the 2nd type of jammer has perfect knowledge on the legitimate user strategy (only for worst-case study).

[5] It is usually much larger, but we will not delve into the analysis. We will show that even the order of $O(|\mathcal{N}||\mathcal{F}|)$ is sufficient for our argument to hold.

[6] There are 2 integrations for each (m, f). However, since (18) and (19) differ only on the integrated parts, they can be done within the same loop.

1536-1276 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2016.2569083, IEEE Transactions on Wireless Communications



(a) (b)

Fig. 4: (a) Toy topology. X, Y: x and y coordinates, respectively; (b) Convergence of the distributed algorithm (to the global optimum in this example).

simulations the standard jammers, unless otherwise specified. The bandwidth of each channel is set to 20 kHz.

Case Study. We first show a toy example to gain insights on the convergence and optimality of the algorithm. We consider a 2-user-1-channel scenario. We vary the sensing probability of both users from 0 to 1 and calculate the expected utility for every possible tuple. We set the Rayleigh fading factors to 0.5. The locations of the nodes are as shown in Fig. 4(a).

We observe from Fig. 4 that user 2 is able to cooperate with user 1. We verify that the algorithm converges to $(0.95, 0.71)$ (the black line in the figure shows the convergence path). We can also easily verify that the convergence point is the global optimum. The non-cooperative optimum (each user senses the channel with probability 1 since there is only 1 channel in this case) is compared with our algorithm. Specifically, the total utility of the non-cooperative algorithm is 14.18 , while the total utility of the distributed algorithm is 14.53. The gain is moderate because the traffic in this case is fairly light (2 users competing for 1 channel).

Convergence Analysis. We now evaluate the convergence speed of the proposed algorithm. We set the number of users to 10, and the number of channels to 2. The power of the jammer is 10 W. The result is shown in Fig. 5(a). We only plot the strategy updates of users 1 to 5 on channel 1 for readability. We observe that the strategies converge quite quickly, i.e., within 20 iterations.

The average convergence speed of our algorithm is also shown in Fig. 5(b). We vary the number of users from 2 to 10 in steps of 2, with the same settings for other parameters. For a fixed number of users, we randomly generate 20 topologies and calculate the average convergence speed. The algorithm is considered to have converged if the element-wise absolute difference of two consecutive iterations is no larger than 0.005. We observe that the distributed algorithm converges within about 30 iterations in all cases.

Utility Comparison. Now we compare the utility of our distributed algorithm vs. the frequency hopping algorithm and the centralized algorithm. In the frequency hopping algorithm, users select the best instantaneously available channel. Since we aim at maximizing the sum-log capacity, to make the comparison fair, we consider a frequency hopping algorithm designed to maximize the same objective function. The scenario is the same as described above. We vary the number of users between 2 and 18 in steps of 2, as shown in Fig. 6.

In all considered scenarios there are gains for our algorithm, up to 19.6%. We observe that, since the utility represents the logarithm of capacity, a gain of 19.6% is considerable. Compared with the centralized algorithm, when the number of users is small, our algorithm achieves utility very close to the upper bound of the centralized algorithm. To be specific, when the number of users is below 12, we achieve more than 80% of the upper bound of the centralized optimal value. In these cases, the lower bounds and the upper bounds of the centralized algorithm converge to one point, so the upper bound is actually the optimal point and our distributed algorithm has near-optimal utility performance. When the number of users is large, there are gaps between our algorithm and the upper bound. However, in these cases, the branch-and-bound-based algorithm fails to converge within the maximum number of iterations we set.[7] Therefore, the upper bound does not necessarily represent the actual optimal value. So, for these two cases, we are unable to make any conclusive statement about the global optimality.

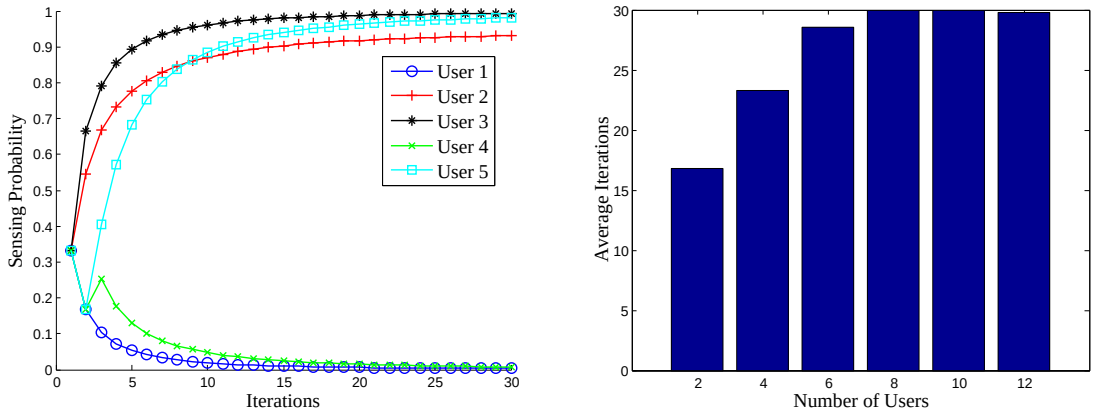
Besides the density of users, the jamming power is also an important factor affecting the performance of the anti-jamming algorithm. To analyze this, we fix the number of users to 6. We vary the power of the jammer from 1 to 20 W. To better illustrate the impact of the jamming power, we fix the topology and the Rayleigh fading factors for each jamming power.

The results are shown in Fig. 7, where the power of the jammer is normalized by the power of the legitimate nodes. It can be observed that as the power of the jammer increases, the utility of both algorithms decreases. Our algorithm always outperforms the non-cooperative algorithm, with a gain up to 34.9%. An important observation is that the gain increases

[7]Although theoretically the algorithm will eventually converge, there is no guarantee of convergence speed. In our experiment, the maximum number of iterations is set to 30000.

1536-1276 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2016.2569083, IEEE Transactions on Wireless Communications



(a) (b)

Fig. 5: (a) Example of convergence (b) Average number of iterations for convergence vs number of users.

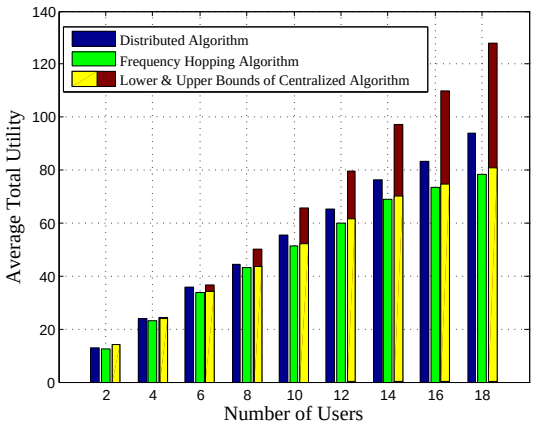


Fig. 6: Utility vs number of users.

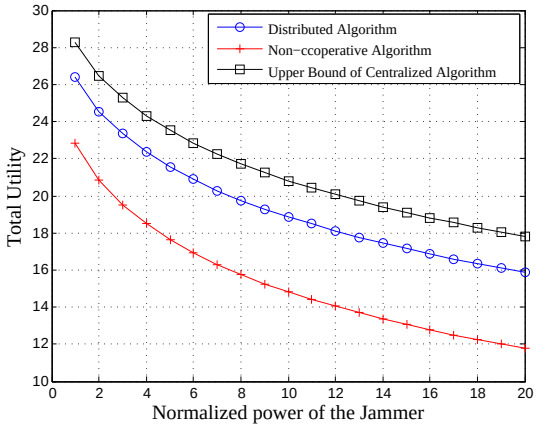


Fig. 7: Utility vs jamming power.

when the power of the jammer increases. This implies that, when the jamming attack is not very severe, i.e., at low power, and there is enough room for the frequency hopping scheme to counteract the effect of jamming, the advantage of cooperation is not so evident. Therefore, cooperative diversity is particularly beneficial under severe jamming attacks.

Performance Against Different Jamming Strategies. We now show the performance of the proposed anti-jamming scheme against different jamming strategies. To this end, we run simulations in which both standard and optimal jammers are considered for the same topology. We consider both utility and sum rate (since it is the objective of the optimal jammer) for different topologies and show the results in Fig. 8. For all the simulated topologies, the number of nodes is set to 5 and number of channels set to 2.

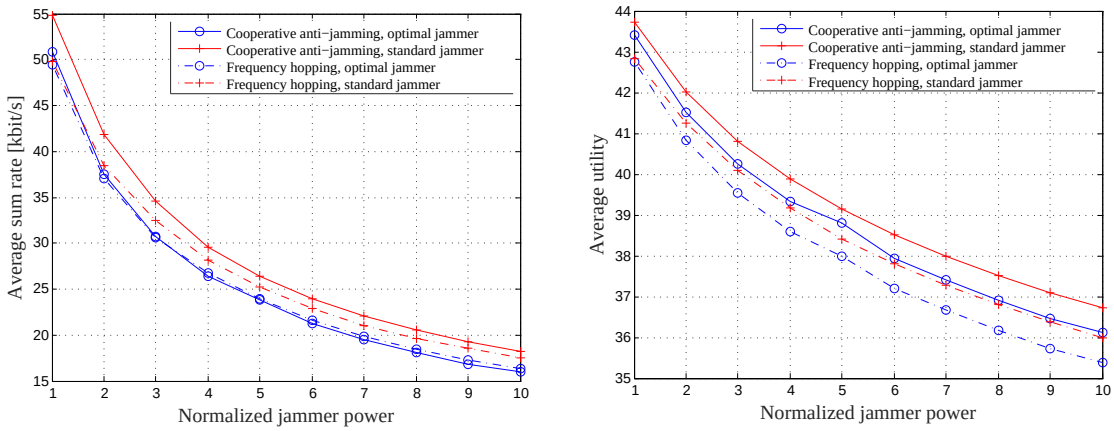
The superiority of the proposed cooperative anti-jamming scheme in terms of utility is confirmed in Fig. 8(a), where cooperative anti-jamming is shown to perform better than frequency hopping with both the standard and optimal jammers. This is quite intuitive, since the design goal of the proposed scheme is utility maximization. A more interesting observation is shown in Fig. 8(b) with respect to the achieved sum rate. It can be observed that, with the optimal jammer, the proposed cooperative anti-jamming scheme achieves similar sum rate to the frequency hopping scheme. Recall that the objective of the optimal jammer is to minimize the sum rate of all the affected links. This implies that cooperation between different users is unable to increase the achievable sum rate if the jammer is designed to minimize it. Note that this conclusion is valid in a statistical sense, i.e., the sum rate is the same when averaged over a considerable amount of randomly generated topologies. For each individual topology, cooperative anti-jamming may achieve higher or lower sum rates than frequency hopping. The primary advantage of cooperation is that it allows different users to trade achievable capacity among them. As for the standard jammer, cooperative anti-jamming achieves higher sum rate than frequency hopping.

Combining this with the previous observation on utility, we can conclude that: (i) in the worst case (with the optimal jammer), the cooperative anti-jamming scheme can achieve better fairness, with similar sum rate; (ii) with standard jammers, the cooperative anti-jamming scheme outperforms frequency hopping on both sum rate and fairness. Considering that, in reality, it is impossible for a jammer to know all the information to operate in the optimal way defined in (26), the benefits of the cooperative anti-jamming scheme are significant.

Another interesting observation concerns physical layer cooperation. In Fig. 9, the mean and maximum cooperation probability with both jammers are shown. Clearly, there is no

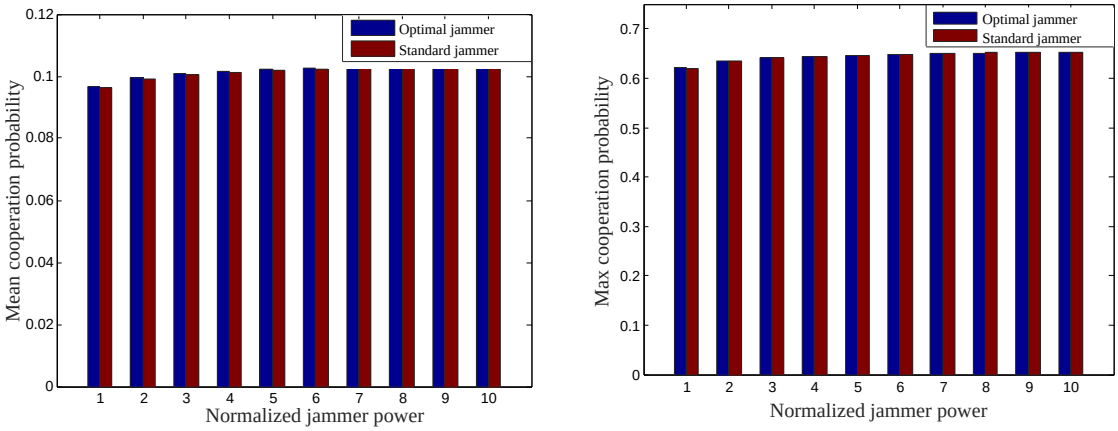
1536-1276 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2016.2569083, IEEE Transactions on Wireless Communications



(a) (b)

Fig. 8: (a) Sum rate vs jamming power; (b) Utility vs jamming power.



(a) (b)

Fig. 9: (a) Mean cooperation probability vs jamming power; (b) Max cooperation probability vs jamming power.

obvious difference for different jamming models or jamming power. This is consistent with the way in which physical layer cooperation works. In fact, the opportunity for physical layer cooperation is determined by the relative location of the nodes involved and the corresponding channel gains. Neither of these two factors has anything to do with jamming models or power. Considering that Fig. 8 shows that the cooperative anti-jamming scheme achieves different results with different jamming models, it can be inferred that MAC layer cooperation plays a major role in the overall scheme.

VIII. Conclusions

We proposed and designed a cooperative anti-jamming scheme by introducing the notion of cooperative diversity into anti-jamming. There are two levels of cooperations. At the medium access control layer, a cooperative channel access scheme is proposed where the channel access probabilities of different users are optimally regulated so that users degraded severely by jammers have an increased share of air time; at the physical layer, users able to enhance the link capacity of another user through cooperative transmission cooperate as relays with a certain probability. We designed a pricing-based distributed algorithm to jointly optimize these two levels of cooperations. We proved that the algorithm always converges, even if the centralized optimization problem cannot be proven to be convex. Compared to non-cooperative algorithms, our algorithm achieves considerable gains. By comparing it with a newly-designed branch-and-bound based centralized algorithm, we also showed that the proposed distributed algorithm achieves close-to-global optimality with a moderate number of users. The gain is shown to increase with increasing network traffic and with jamming power. Our results also demonstrate significant cooperative gains when a network is experiencing very low throughput. For the performance against different jamming strategies, we show that, when the jammer is sophisticated enough to be able to bound the total achievable rate of all sessions, the proposed scheme achieves better fairness than frequency hopping while maintaining the same performance in terms of sum rate; if the jammer is less sophisticated, which is the case in practice, the proposed scheme outperforms frequency hopping in both metrics.

Appendix

A. Verification of The Assumption on Cooperation Probability

Here, we verify the assumption that cooperation opportunities are sparse. We consider a network with moderate traffic,

1536-1276 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2016.2569083, IEEE Transactions on Wireless Communications

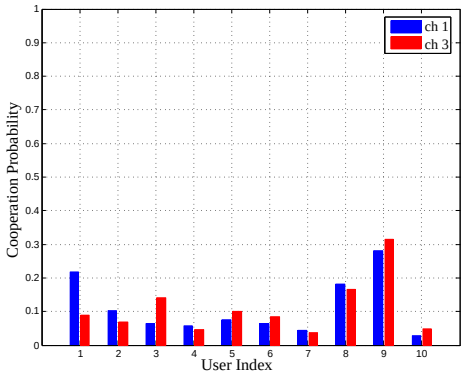


Fig. 10: Probability of positive cooperative gains.

i.e., with 10 sessions and 4 channels. All other settings are the same as in Section VII. The experiment is repeated with multiple randomly generated topologies. Figure 10 shows the average probability that a user can be assisted by a cooperative relay with positive gain, i.e., the probability that a cooperation opportunity exists. For readability, only selected channels are shown. We observe that the assumption that opportunities for cooperation are sparse is verified. In fact, for most users on most channels, the average probability that another user can provide cooperative gain is fairly low (below 0.1). Moreover, since the probability of cooperative transmission is the product of the probability shown in the figure and the probability that the cooperator does not sense any channel, the resulting cooperation probability is even lower.

B. Reverse Waterfilling

The optimization problem (26) can be viewed as a reverse water-filling problem. Rewrite the objective function as

$$f(\boldsymbol{p}_j) = \sum_n \sum_f w_n^f \log\left(1 + \frac{S_n^f}{p_j^f H_n^f + D_n^f}\right) \tag{44}$$

where D_n^f represents the sum of noise and non-jammer interference at node n and H_n^f is the channel coefficients from jammer to node n, we have the KKT conditions:

$$p_j^f \geq 0, \forall f \in \mathcal{F}, \tag{45}$$

$$\sum_{f\in\mathcal{F}} p_j^f = p_j^{\max}, \tag{46}$$

$$\lambda_f \geq 0, \forall f \in \mathcal{F}, \tag{47}$$

$$\lambda_f p_j^f = 0, \forall f \in \mathcal{F}, \tag{48}$$

$$-g_f(p_j^f) - \lambda_f + \nu = 0, \forall f \in \mathcal{F}, \tag{49}$$

where $g_f(p_j^f) = \sum_{n\in\mathcal{N}} w_n^f \frac{S_n^f H_n^f}{(D_n^f + p_j^f H_n^f)(S_n^f + D_n^f + p_j^f H_n^f)}$ is a decreasing function of p_j^f.

From (49), we have $\lambda_f = \nu_f - g_f(p_j^f)$. Substituting λ_n in (47) and (48), we have

$$p_j^f \geq 0, \forall f \in \mathcal{F}, \tag{50}$$

$$\sum_{f\in\mathcal{F}} p_j^f = p_j^{\max}, \tag{51}$$

$$\nu - g_f(p_j^f) \geq 0, \forall f \in \mathcal{F}, \tag{52}$$

$$p_j^f \left(\nu - g_f(p_j^f)\right) = 0, \forall f \in \mathcal{F}. \tag{53}$$

If $\nu > g_f(0)$, because of the decreasing property of $g_f(p_j^f)$, we have $\nu - g_f(p_j^f) > 0, \forall p_j^f \geq 0$. According to (53), this implies $p_j^f = 0$.

If $\nu \leq g_f(0)$, because (52) and the decreasing property of $g_f(p_j^f)$, it must be satisfied that $p_j^f \geq g_f^{-1}(\nu)$. Considering (53), the equality must hold, i.e., $p_j^f = g_f^{-1}(\nu)$.

To sum up, we have

$$p_j^f = g_f^{-1}\left(\min\left\{\nu, g_f(0)\right\}\right). \tag{54}$$

(54) suggests that an algorithm similar to water-filling algorithm can be proposed to solve the problem.

REFERENCES

[1] L. Zhang, Z. Guan, and T. Melodia, "Cooperative Anti-jamming for Infrastructure-less Wireless Networks with Stochastic Relaying," in *Proc. of IEEE Conference on Computer Communications (INFOCOM)*, Toronto, Canada, April 2014, pp. 549–557.

[2] W. Xu, W. Trappe, Y. Zhang, and T. Wood, "The Feasibility of Launching and Detecting Jamming Attacks in Wireless Networks," in *Proc. of ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, NY, USA, May 2005.

[3] W. Xu, K. Ma, W. Trappe, and Y. Zhang, "Jamming Sensor Networks: Attack and Defense Dtrategies," *IEEE Network*, vol. 20, no. 3, pp. 41–47, May 2006.

[4] A. Mpitziopoulos, D. Gavalas, C. Konstantopoulos, and G. Pantziou, "A Survey on Jamming Attacks and Countermeasures in WSNs," *IEEE Communications Surveys Tutorials*, vol. 11, no. 4, pp. 42–56, Fourth Quarter 2009.

[5] L. Ding, T. Melodia, S. Batalama, J. Matyjas, and M. Medley, "Cross-layer Routing and Dynamic Spectrum Allocation in Cognitive Radio Ad Hoc Networks," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 4, pp. 1969–1979, May 2010.

[6] K. Pelechrinis, C. Koufogiannakis, and S. V. Krishnamurthy, "Gaming the Jammer: Is Frequency Hopping Effective?" in *Proc. of the International Conference on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, Seoul, Korea, Jun. 2009, pp. 187–196. [Online]. Available: http://dl.acm.org/citation.cfm?id=1715782.1715816

[7] A. Sampath, H. Dai, H. Zheng, and B. Zhao, "Multi-channel Jamming Attacks using Cognitive Radios," in *Proc. of International Conference on Computer Communications and Networks (ICCCN)*, Honolulu, HI, Aug. 2007, pp. 352–357.

[8] M. Strasser, S. Capkun, C. Popper, and M. Cagalj, "Jamming-resistant Key Establishment using Uncoordinated Frequency Hopping," in *Proc. of IEEE Symposium on Security and Privacy*, Oakland, CA, USA, May 2008.

[9] Q. Ling and T. Li, "Message-driven Frequency Hopping: Design and Analysis," *IEEE Transactions on Wireless Communications*, vol. 8, no. 4, pp. 1773–1782, April 2009.

[10] C. Li, H. Dai, L. Xiao, and P. Ning, "Communication Efficiency of Anti-jamming Broadcast in Large-scale Multi-channel Wireless Networks," *IEEE Transactions on Signal Processing*, vol. 60, no. 10, pp. 5281–5292, Oct. 2012.

[11] Q. Wang, P. Xu, K. Ren, and X. Li, "Delay-bounded Adaptive UFH-based Anti-jamming Wireless Communication," in *Proc. of IEEE International Conference on Computer Communications (INFOCOM)*, Shanghai, China, April 2011.

[12] K. Xu, Q. Wang, and K. Ren, "Joint UFH and Power Control for Effective Wireless Anti-jamming Communication," in *Proc. of IEEE International Conference on Computer Communications (INFOCOM)*, Orlando, FL, USA, March 2012.

[13] L. Zhang, H. Wang, and T. Li, "Anti-jamming Message-driven Frequency Hopping Part I: System Design," *IEEE Transactions on Wireless Communications*, vol. 12, no. 1, pp. 70–79, Jan. 2013.

[14] L. Zhang and T. Li, "Anti-jamming Message-driven Frequency Hopping Part II: Capacity Analysis under Disguised Jamming," *IEEE Transactions on Wireless Communications*, vol. 12, no. 1, pp. 80–88, Jan. 2013.

[15] L. Ding, K. Gao, T. Melodia, S. Batalama, D. Pados, and J. Matyjas, "All-spectrum Cognitive Networking through Jointly Optimal Distributed Channelization and Routing," *IEEE Transactions on Wireless Communications*, vol. 12, no. 11, pp. 5394–5405, Nov. 2013.

1536-1276 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2016.2569083, IEEE Transactions on Wireless Communications

[16] Y. Liu, P. Ning, H. Dai, and A. Liu, "Randomized Differential DSSS: Jamming-resistant Wireless Broadcast Communication," in *Proc. of International Conference on Computer Communications (INFOCOM)*, March San Diego, CA, USA, Mar. 2010, pp. 1–9.

[17] R. Gohary, Y. Huang, Z.-Q. Luo, and J.-S. Pang, "A Generalized Iterative Water-filling Algorithm for Distributed Power Control in The Presence of a Jammer," in *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Taipei Taiwan, April 2009.

[18] Y. Sagduyu, R. Berry, and A. Ephremides, "MAC Games for Distributed Wireless Network Security with Incomplete Information of Selfish and Malicious User Types," in *Proc. of International Conference on Game Theory for Networks (GameNets)*, Istanbul, Turkey, May 2009.

[19] Y. E. Sagduyu, R. A. Berry, and A. Ephremides, "Wireless Jamming Attacks under Dynamic Traffic Uncertainty," in *Proc. of International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)*, Avignon, France, June 2010.

[20] B. Wang, Y. Wu, K. J. R. Liu, and T. Clancy, "An Anti-jamming Stochastic Game for Cognitive Radio Networks," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 29, no. 4, pp. 877–889, April 2011.

[21] Y. Wu, B. Wang, K. Liu, and T. Clancy, "Anti-jamming Games in Multi-channel Cognitive Radio Networks," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 30, no. 1, pp. 4–15, Jan. 2012.

[22] Y. Zhu and Y. Jian, "A Game-theoretic Approach to Anti-jamming in Sensor Networks," in *IEEE International Conference on Parallel and Distributed Systems (ICPADS)*, Shanghai, China, Dec. 2010, pp. 617–624.

[23] Q. Yan, H. Zeng, T. Jiang, M. Li, W. Lou, and Y. Hou, "MIMO-based Jamming Resilient Communication in Wireless Networks," in *Proc. of International Conference on Computer Communication (INFOCOM)*, Toronto, Canada, Apr. 2014, pp. 2697–2706.

[24] S. Zhang, K. Huang, and X. Li, "Adaptive Transmit and Receive Beamforming Based on Subspace Projection for Anti-jamming," in *Proc. of IEEE Military Communications Conference (MILCOM)*, Baltimore, MD, USA, Oct. 2014, pp. 388–393.

[25] G. Noubir, "On Connectivity in Ad Hoc Networks under Jamming Using Directional Antennas and Mobility," in *Proc. of International Conference on Wired/Wireless Internet Communications*, Frankfurt, Germany, Feb. 2004.

[26] H. Mustafa, X. Zhang, Z. Liu, W. Xu, and A. Perrig, "Jamming-Resilient Multipath Routing," *IEEE Transactions on Dependable and Secure Computing*, vol. 9, no. 6, pp. 852–864, Nov 2012.

[27] P. Tague, S. Nabar, J. A. Ritcey, and R. Poovendran, "Jamming-Aware Traffic Allocation for Multiple-Path Routing Using Portfolio Selection," *IEEE/ACM Transactions on Networking*, vol. 19, no. 1, pp. 184–194, Feb 2011.

[28] X. He, H. Dai, and P. Ning, "Dynamic Adaptive Anti-Jamming via Controlled Mobility," in *Proc. of IEEE Conference on Communications and Network Security (CNS)*, Washington, DC, USA, Oct. 2013, pp. 1–9.

[29] J. Huang and A. L. Swindlehurst, "Cooperative Jamming for Secure Communications in MIMO Relay Networks," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4871–4884, Oct 2011.

[30] K.-H. Park, T. Wang, and M.-S. Alouini, "On the Jamming Power Allocation for Secure Amplify-and-Forward Relaying via Cooperative Jamming," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 1741–1750, Sep. 2013.

[31] H. Kulhandjian, T. Melodia, and D. Koutsonikolas, "Scuring Underwater Acoustic Communications through Analog Network Coding," in *Proc. of IEEE International Conference on Sensing, Communication, and Networking (SECON)*, Singapore, Jun. 2014.

[32] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behavior," *IEEE Trans. on Info. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[33] Z. Guan, T. Melodia, D. Yuan, and D. A. Pados, "Distributed Spectrum Management and Relay Selection in Interference-limited Cooperative Wireless Networks," in *Proc. of ACM Intl. Conf. on Mobile Computing and Networking (MobiCom)*, Las Vegas, Nevada, USA, Sep. 2011. [Online]. Available: http://doi.acm.org/10.1145/2030613.2030639

[34] K. Butler and M. Stephens, "The Distribution of a Sum of Binomial Random Variables," Stanford University, Dept. of Statistics, Tech. Rep. 467, 28 Apr. 1993.

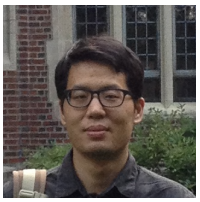
[35] G. Scutari, F. Facchinei, P. Song, D. Palomar, and J.-S. Pang, "Decomposition by Partial Linearization: Parallel Optimization of Multi-Agent Systems," *IEEE Trans. on Signal Processing*, vol. 62, no. 3, pp. 641–656, Feb. 2014.

[36] D. Bertsekas, *Nonlinear Programming*. Belmont, MA, USA: Athena Scientific, 2th Ed., 1999.

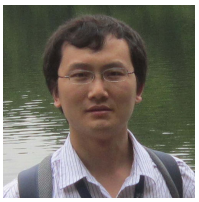
[37] H. D. Sherali and W. P. Adams, *A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems*. Boston: MA: Kluwer Academic, 1999.

[38] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[39] "Raspberry Pi 2 Model B," https://www.raspberrypi.org/products/raspberry-pi-2-model-b/.



Liyang Zhang is a Ph.D. student in the Department of Electrical and Computer Engineering at Northeastern University, Boston, MA, USA. He is currently working in the Wireless Networks and Embedded Systems Laboratory under the guidance of Professor Tommaso Melodia. He received his B.S. and M.S. in Electrical Engineering from Tsinghua University (China) and SUNY Buffalo in 2008 and 2014, respectively. His current research interests are in wireless security.



Zhangyu Guan (M'11) received the Ph.D. degree in communication and information systems from Shandong University, China, in 2010.

He is currently a Postdoctoral Research Associate with the Department of Electrical and Computer Engineering, Northeastern University, Boston, MA, USA. He was a visiting Ph.D. student with the Department of Electrical Engineering, The State University of New York (SUNY) at Buffalo, Buffalo, NY, USA, from 2009 to 2010. He was a Lecturer with Shandong University from 2011 to 2014. He was a Postdoctoral Research Associate with the Department of Electrical Engineering, SUNY Buffalo, from 2012 to 2015. His current research interests are in cognitive radio and software-defined networking, wireless multimedia sensor networks, and underwater networks.

Dr. Guan has served as a TPC member for IEEE INFOCOM 2016-2017, IEEE GLOBECOM 2015-2016, IEEE ICNC 2012-2017, and served as a reviewer for the IEEE/ACM TRANSACTIONS ON NETWORKING, the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, among others.

Tommaso Melodia (M'07) received the Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2007.

He is an Associate Professor with the Department of Electrical and Computer Engineering, Northeastern University, Boston, MA, USA. His research has been supported by the National Science Foundation, Air Force Research Laboratory, and the Office of Naval Research, among others. His current research interests are in modeling, optimization, and experimental evaluation of networked communication systems, with applications to ultrasonic intra-body networks, cognitive and cooperative networks, multimedia sensor networks, and underwater networks.

Prof. Melodia was a recipient of the National Science Foundation CAREER Award and coauthored a paper that was recognized as the ISI Fast Breaking Paper in the field of Computer Science for February 2009 and of Best Paper Awards of ACM WUWNet 2013 and 2015. He was the Technical Program Committee Vice Chair for IEEE Globecom 2013 and the Technical Program Committee Vice Chair for Information Systems for IEEE INFOCOM 2013. He serves on the editorial boards of the IEEE TRANSACTIONS ON MOBILE COMPUTING, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON MULTIMEDIA, and Computer Networks.

1536-1276 (c) 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.