

Resource Control in Network Assisted Device-to-Device Communications: Solutions and Challenges

Lei Lei, Yiru Kuang, Xuemin (Sherman) Shen, Chuang Lin, and Zhangdui Zhong

ABSTRACT

Network assisted Device-to-Device (D2D) communications is envisioned as a promising technology to facilitate the discovery of geographically close cellular devices and reduce the communication cost between these devices. In order to fully take advantages of the freedom in resource management introduced by D2D communications, it is essential to design optimal resource control policies. In this article, we first review and classify the state-of-the-art research based on different considered network scenarios. We then introduce our preliminary studies on delay-aware resource control with bursty traffic, in which we propose an optimization framework and formulate a general queuing model for performance evaluation and optimization. Finally, the challenges and open issues are highlighted to serve as guidelines for future research.

INTRODUCTION

Device-to-device (D2D) communications are commonly referred to as a type of technology that enable devices to communicate directly with each other without an infrastructure, for example, access points or base stations. Bluetooth and WiFi-Direct are the two most popular D2D techniques, both working in the unlicensed 2.4 GHz industrial, scientific, and medical (ISM) bands. Cellular networks, on the other hand, do not support direct over-the-air communications between user devices. However, with the emergence of context-aware applications and the accelerating growth of Machine-to-Machine (M2M) applications [1], D2D communications plays a more important role since it facilitates the discovery of geographically close devices and reduces the communication cost between these devices. To seize the emerging market that requires D2D communications, the mobile operators and vendors are exploring the possibilities of introducing D2D communications in cellular networks [2]. Such network assisted D2D communications work in the licensed band of cellular networks with more controllable interference. Moreover, the network infrastructure can assist the user equipments (UEs) in various key functions of

D2D communications, such as new peer discovery, physical layer procedures, and radio resource control, which make it different from traditional D2D technologies, such as WiFi direct.

Resource control in network assisted D2D communications mainly includes three actions: mode selection, power control, and resource allocation. Compared with the resource control problem in traditional cellular networks, more degrees of freedom are introduced by the D2D communications. First, a pair of D2D users can either communicate directly over-the-air, or communicate via the base station (BS). Moreover, the direct over-the-air link or D2D link may reuse radio resources with other cellular or D2D links. Finally, as signaling can be exchanged directly between a pair of D2D users due to their proximity, some of the resource control actions can be performed distributively and offloaded to the D2D users from the BS. The above flexibility provides both opportunities and challenges in designing optimal resource control policies. Related work in existing literature usually assumes that the D2D and cellular users are saturated with infinite backlogs and focuses only on optimizing the PHY layer performance metrics such as sum throughput and power consumption [3–8]. However, the data arrival process at the users is dynamic, and performance metrics such as delay and dropping probability are also important, especially for real-time services. Therefore, both the channel state information (CSI) and the queue state information (QSI) should be taken into account in resource control policies. Although the problem of cross-layer performance evaluation and resource optimization in conventional cellular networks and wireless multihop networks has been addressed in the literature [9, 10], these research issues in network assisted D2D communications are still open.

The objective of this article is to provide a systematic exploration of the existing research and trigger various potentially significant research in this topic. We first identify the fundamental design principles for the three resource control actions and classify the existing work with infinite backlog assumptions. As an initial attempt to deal with the resource control problem with bursty traffic, we then develop a formal network

Lei Lei, Yiru Kuang, and Zhangdui Zhong are with Beijing Jiaotong University.

Xuemin (Sherman) Shen is with University of Waterloo.

Chuang Lin is with Tsinghua University.

model, where three routing modes are defined. The routing modes are different from the transmission modes in D2D communications (i.e. D2D mode and cellular mode) in that it explicitly characterizes the semi-static and dynamic mode selection actions. We propose an integrated optimization framework for the three resource control actions, where a queuing model is formulated to evaluate the performance of different routing modes under a given power and resource allocation policy, so that the optimal mode can be determined based on the evaluation result. Finally, we outline our preliminary solutions and highlight the open issues in performance evaluation and optimization of delay-aware resource control policies under our framework.

RESOURCE CONTROL WITH INFINITE BACKLOG TRAFFIC: SURVEY AND CLASSIFICATION

Consider a cellular wireless network with D2D communications capability, where there are D D2D UE pairs, C_u cellular UEs (CUEs) with uplink communications and C_d CUEs with downlink communications in a single cell. A D2D UE pair consists of a source D2D UE (src. DUE) and a destination D2D UE (dest. DUE) within direct over-the-air communications range with each other, which is formed through the various neighbor/peer/service discovery mechanisms proposed in literature. Figure 1 illustrates a simple example with $D = C_u = C_d = 1$, that is, there are one pair of DUEs (src. DUE 1 and dest. DUE 2), one uplink CUE (CUE 3), and one downlink CUE (CUE 4). Time is slotted and each time slot has an equal length.

A D2D pair can either work in the D2D mode through direct communications, or in the cellular mode where the BS relays the data from the src. DUE to the dest. DUE as in traditional cellular networks. Compared with the cellular mode, the D2D mode can achieve the “hop gain” as only a single hop is needed instead of two hops in the cellular mode. However, if the distance between either the source or destination DUE and the BS is shorter than that between the DUE pair, the D2D mode may have negative “proximity gain,” which can result in the D2D mode performing worse than the cellular mode in spite of the “hop gain.” Therefore, mode selection is needed for a D2D pair, which can be performed either semi-statically at the time-scale of connection establishment/ release, or dynamically per time slot. Dynamic mode selection can capture and utilize the fast fading effects of wireless channels opportunistically, while semi-static mode selection has the advantage of saving computation and communication overhead. When a D2D pair works in the D2D mode, a D2D link exists from the transmitter of the src. DUE to the receiver of the dest. DUE. In each time slot, a resource unit (e.g. a time slot in TDMA system), a subchannel (a group of successive subcarriers) in an OFDMA system, etc.) can be assigned to one or more D2D links and at most one cellular link in order to achieve the “reuse gain.” Since interference exists between the set of links with frequency

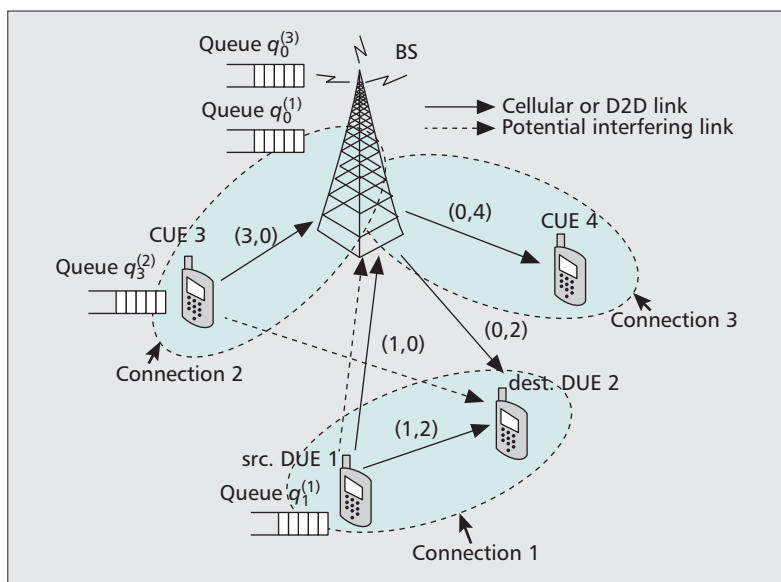


Figure 1. Wireless cellular networks with D2D communications capability.

reuse, whether the “reuse gain” can be achieved depends on the trade-off between the improvement in resource utilization and the degradation in spectral efficiency due to interference. Therefore, resource allocation is needed at each time slot to determine the sets of links for transmissions using uplink and downlink resources, respectively. Moreover, power control is needed to determine the transmit power of each link with frequency reuse. The amount of “reuse gain” should also be taken into account in the mode selection phase.

There are several works focused on D2D communications, which have studied the above three resource control actions, that is, mode selection, power control, and resource allocation with the infinite backlog traffic model based on different assumptions of network scenario, as summarized in Table 1.

RESOURCE CONTROL ACTIONS

The three resource control actions in D2D communications are either separately studied [3, 4] or jointly optimized [5–8]. For example, [5, 6] focused on power control with mode selection over the target Signal to Interference plus Noise Ratio (SINR), while [7, 8] presented the joint scheme combining mode selection with power control and resource allocation, which proved to be more effective in realizing the optimal system performance.

INTERFERENCE SCENARIOS

Addressing the interferences between BS, CUEs, and DUEs caused by resource sharing plays a key role in performance optimization. Most of the related works focus on intra-cell interference due to the complexity involved in handling interference between multiple cells and also between the cellular and the D2D layers [5]. In addition, the interference scenario that was considered mostly assumes that each cellular link can be allocated the same resources with at most one D2D link in each time slot [4, 5, 8], while in [6, 7], a more general scenario involving resource sharing among CUEs and multiple

Refs	Control Function			Multiple Access Methods	Duplex Scheme	Channel Model			Interference Scenario	Performance Metrics			Distributed Algorithm
	M	P	R			PL	SF	FF		rate	power	SINR	
[3]			✓	OFDMA	TDD/FDD	✓	✓		D,D	O			✓
[4]			✓	OFDMA	TDD/FDD	✓		✓	C,s-D	O			✓
[5]	✓	✓		OFDMA	TDD/FDD	✓	✓	✓	C,s-D	C	O	C	✓
[6]	✓	✓		CDMA	TDD	✓	✓		C,m-D		O	C	
[7]	✓		✓	OFDMA	TDD/FDD	✓	✓		C,m-D	O	C		
[8]	✓	✓	✓	TDMA/OFDMA	TDD/FDD	✓	✓	✓	C,s-D	O	C		✓

Notes:

M: mode selection P: power control R: resource allocation

PL: path loss SF: shadowing FF: fast fading

D,D: interference between D2D pairs

C,s-D: interference between a D2D pair and a cellular UE

C,m-D: interference between multiple D2D pairs and a cellular UE

O: optimization objective C: constraint

Table 1. Comparison of existing work in literature.

D2D pairs is given. In [3] scheduling protocols based on the interference solely between D2D pairs are studied. In addition, with non-orthogonal sharing, the aggressors and victims of the interference are different when sharing the cellular users' uplink and downlink resources. D2D transmitter causes interference to the cellular UE receiver when reusing the downlink resources and to the immobile BS when reusing the uplink resources. Although it is conceptually feasible for the D2D links to use either the uplink or downlink radio resources, it is much easier to use the uplink resources from the protocol design perspective [2].

PERFORMANCE METRICS

D2D communications underlaying a cellular infrastructure has been proposed to increase throughput, spectrum efficiency, and resource utilization. Generally, throughput is one of the most common performance metrics with constraints such as maximum power, SINR, and fairness [3, 4, 7, 8]. Besides, [5, 6] jointly obtain the optimal transmission power and the mode maximizing the power-efficiency with respect to the system capacity. The packet-level performance such as the average delay or dropping probability cannot be addressed under the infinite backlog assumption.

DELAY-AWARE RESOURCE CONTROL WITH BURSTY TRAFFIC: MODEL FORMULATION AND OPTIMIZATION FRAMEWORK

NETWORK MODEL

Delay-aware resource control with bursty traffic has received little attention in network assisted D2D communications due to its design complex-

ity compared with the peer problem with infinite backlog traffic [11]. In this section we develop a general network model for this problem. Based on the above discussion on network scenario, we make the following assumptions. First, we consider an OFDM system. The whole uplink or downlink spectrum is divided into multiple equal-sized subchannels. A subchannel in the uplink (resp. downlink) spectrum shall be referred to as uplink (resp. downlink) subchannel in the rest of the article. Moreover, we only consider intra-cell interference and assume that the D2D links only use the uplink resources.

Nodes, Links, and Connections — Consider a general network model with a set \mathcal{N} of nodes and a set \mathcal{L} of transmission links, as illustrated in Fig. 2. Define $\mathcal{N} = \{0, 1, \dots, N\}$, where node 0 represents the base station (BS) and nodes 1, ..., N represent the UEs. Let \mathcal{N}_D be the set of DUEs of the D D2D pairs, where nodes $2c - 1$ and $2c$ denote the src. and dest. DUEs of pair c ($c \in \{1, \dots, D\}$), respectively. Let \mathcal{N}_{Cu} and \mathcal{N}_{Cd} be the set of C_u uplink CUEs and C_d downlink CUEs, respectively. For example, in Fig. 1, we have $\mathcal{N}_D = \{1, 2\}$, $\mathcal{N}_{Cu} = \{3\}$, and $\mathcal{N}_{Cd} = \{4\}$.

Each transmission link represents a communication channel for direct transmission from a given node i to another node j , and is labeled by (i, j) (where $i, j \in \mathcal{N}$). Note that link (i, j) is distinct from link (j, i) . The link set \mathcal{L} is composed of three non-overlapping subsets, where \mathcal{L}_D is the set of D2D links, \mathcal{L}_{Cu} is the set of cellular uplinks, and \mathcal{L}_{Cd} is the set of cellular downlinks. The three types of links are represented by solid lines, dashed lines, and dotted lines in Fig. 2, respectively. For example, in Fig. 1 we have $\mathcal{L}_D = \{(1, 2)\}$, $\mathcal{L}_{Cu} = \{(3, 0), (1, 0)\}$, and $\mathcal{L}_{Cd} = \{(0, 4), (0, 2)\}$.

All data that enters the network are associated with a particular connection $c \in \mathcal{C}$, which

D2D communications underlying a cellular infrastructure has been proposed to increase throughput, spectrum efficiency and resource utilization. Generally, throughput is one of the most common performance metrics with constraints such as maximum power, SINR and fairness.

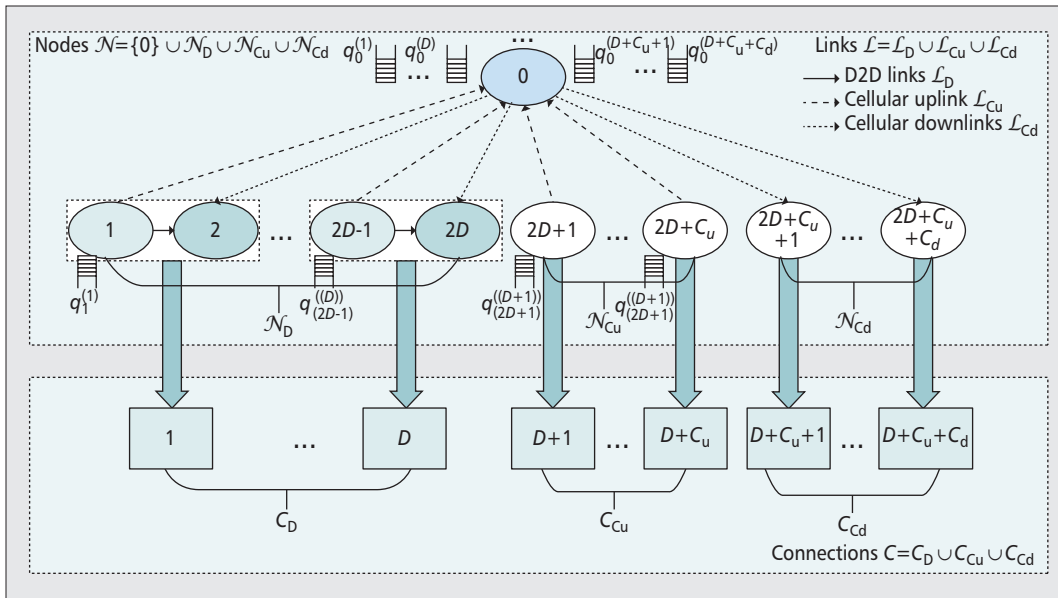


Figure 2. General network model.

defines the source and destination of the data. Let C_D represent the set of D D2D connections, C_{Cu} represent the set of C_u cellular uplink connections, and C_{Cd} represent the set of C_d cellular downlink connections, as illustrated in Fig. 2. For example, in Fig. 1 there are $C = 3$ connections, with $C_D = \{1\}$, $C_{Cu} = \{2\}$, and $C_{Cd} = \{3\}$.

Define the link constraint set for a connection as the set of all links that the connection data is allowed to use. Obviously, the link constraint set for any cellular connections consists of only one link, since there is only a single-hop route between the CUE and the BS for these connections. For example, in Fig. 1 the link constraint sets of connections 2 and 3 are $\{(3, 0)\}$ and $\{(0, 4)\}$, respectively. However, the link constraint set for any D2D connection depends on which of the following Routing Modes (RMs) is chosen, that is, whether the single hop route of D2D link or the two-hop route of cellular links or both can be used:

- **D2D RM** means the connection c data is always transmitted via the D2D link, and there is only one D2D link in the link constraint set.
- **Cellular RM** means the connection c data is always transmitted via cellular links, and there is one cellular uplink and one cellular downlink in the link constraint set.
- **Hybrid RM** means the connection c data can be transmitted either via the D2D or cellular links, and the decision is made dynamically at each time slot. There are one D2D link, one cellular uplink, and one cellular downlink in the link constraint set.

For example, in Fig. 1 the link constraint set of connection 1 is $\{(1, 2)\}$ in D2D RM, $\{(1, 0), (0, 2)\}$ in Cellular RM, and $\{(1, 2), (1, 0), (0, 2)\}$ in Hybrid RM.

When the RM and link constraint set are determined for every D2D connection, the set of eligible links in the network is a subset of all the transmission links, which is equal to the union of the link constraint sets for all the connections.

This is because the corresponding D2D link (resp. cellular links) does not exist for a D2D connection in Cellular RM (resp. D2D RM). In the rest of the article we will use \mathcal{L} to denote the set of eligible links instead of all the transmission links by default.

Queuing Dynamics — We assume that the amount of new connection c data¹ that exogenously arrives at its source node during time slot t is i.i.d. over time slots following general distribution. The data is transmitted hop by hop along the route(s) of the connection to its destination node. Each node i along the route(s) of connection c maintains a queue $q_i^{(c)}$ for storing its data except for the destination node, since the data is assumed to exit the network once it reaches the destination, as shown in Fig. 2. We assume each queue has a finite capacity. Note that the number of queues maintained at the BS for the D2D connections depends on their selected RMs, since no queue needs to be maintained for a D2D connection in D2D RM. For example, in Fig. 1 there are four queues $\{q_1^{(1)}, q_0^{(1)}, q_3^{(2)}, q_0^{(3)}\}$ if Cellular or Hybrid RM is chosen for D2D connection 1. On the other hand, if D2D RM is chosen for D2D connection 1, there are only three queues $\{q_1^{(1)}, q_3^{(2)}, q_0^{(3)}\}$.

We define a Resource Reuse Group (RRG) as the subset of links that can be scheduled for transmission simultaneously on any subchannel in a time slot. Therefore, an RRG for an uplink subchannel may contain at most one cellular uplink and one or more D2D links. On the other hand, an RRG for a downlink subchannel can contain one and only one cellular downlink. For example, in Fig. 1 there are four RRGs in the uplink, that is, $\{(1, 2)\}$, $\{(3, 0)\}$, $\{(1, 2), (3, 0)\}$, and $\{(1, 0)\}$. Moreover, there are two RRGs in the downlink, that is, $\{(0, 2)\}$, $\{(0, 4)\}$. In each time slot, a subchannel can be allocated to at most one RRG for uplink or downlink transmissions. The subchannel allocation for an RRG at time slot t can be represented by a vector of 0-1

¹ The data can take units of bits or packets. The latter is appropriate when all the packets have fixed length.

The instantaneous data rate of a scheduled link on its allocated subchannel depends on which RRG is scheduled on this subchannel, since different amount of interference will arise when different RRGs containing the considered link are scheduled.

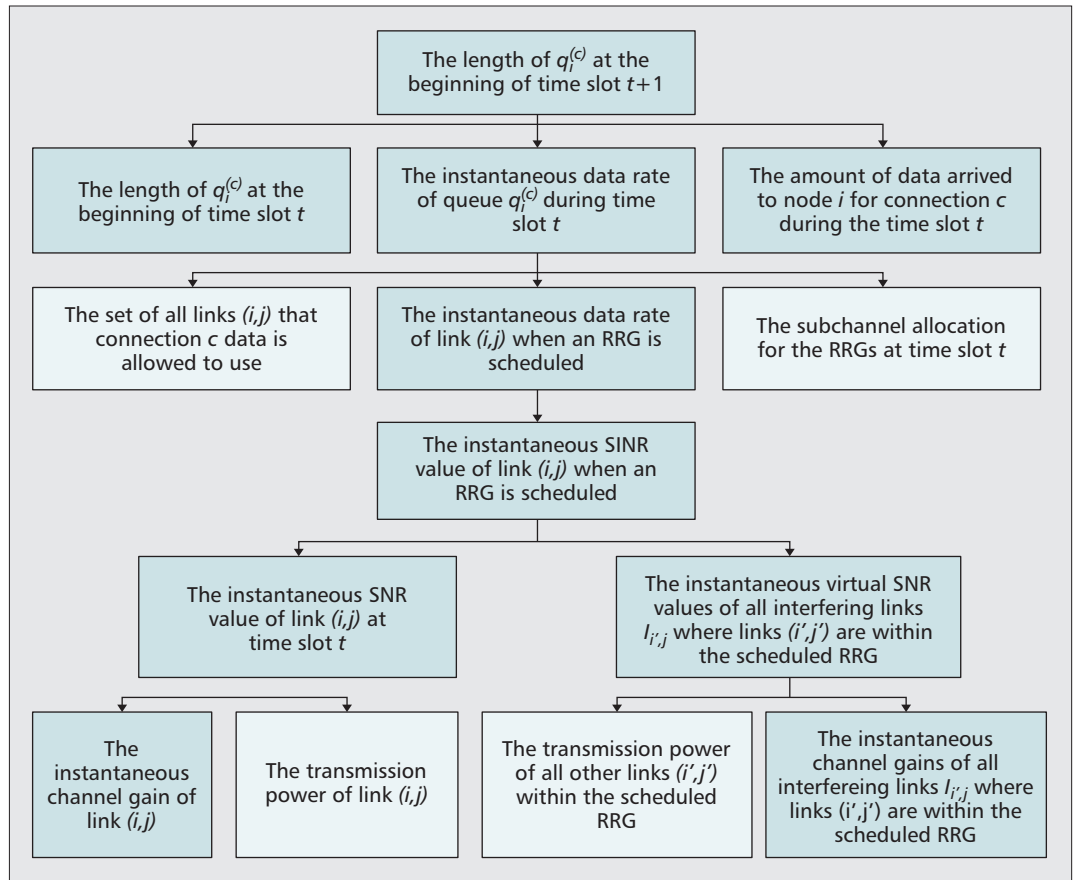


Figure 3. The dependency relationships between the variables in the network model (an arrow means that the variable in the source block depends on the variable in the destination block).

variables, where an element equals 1 if the corresponding subchannel is allocated to the considered RRG, and 0 otherwise. We assume that an RRG is scheduled for transmission only when all its links have non-empty queues.

A queue $q_i^{(c)}$ is scheduled in time slot t when at least one link (i, j) within the link constraint set of connection c is scheduled on any subchannel. Note that more than one link with source node i exist within the link constraint set of connection c when node i is the src. DUE and the D2D connection c is in Hybrid RM. For example, in Fig. 1 $q_1^{(1)}$ is scheduled if either D2D link $(1, 2)$ or cellular uplink $(1, 0)$ is scheduled on any subchannel when D2D connection 1 is in Hybrid RM. Only when a queue is scheduled and non-empty shall it move the data out of the queue for transmission. The instantaneous data rate or transmission capability² of a scheduled queue $q_i^{(c)}$ during time slot t is equal to the sum of the instantaneous data rate of all its scheduled links (i, j) on their allocated subchannels at time slot t . The instantaneous data rate of a scheduled link on its allocated subchannel depends on which RRG is scheduled on this subchannel, since different amounts of interference will arise when different RRGs containing the considered link are scheduled. For example, in Fig. 1 the instantaneous data rate of D2D link $(1, 2)$ is larger when the scheduled RRG is $\{(1, 2)\}$ instead of $\{(1, 2), (3, 0)\}$, since interference from link $(3, 0)$ exists for link $(1, 2)$ in the latter case.

If the queue length of $q_i^{(c)}$ at the beginning of time slot t is less than its transmission capability during time slot t , padding bits shall be transmitted along with the data. Arriving data are placed in the queue throughout the time slot t and can only be transmitted during the next time slot $t + 1$. If the queue length reached the buffer capacity, the subsequent arriving data will be dropped. The amount of data arrived at node i for connection c during time slot t is equal to the amount of new connection c data if node i is the source node of connection c , or else it depends on the data departure process of the corresponding uplink transmission. According to the above assumption, the queuing process evolves recursively where the queue length of $q_i^{(c)}$ at the beginning of time slot $t + 1$ is determined by its queue length at the beginning of time slot t , along with its transmission capability and the amount of data arrival during time slot t [9].

Instantaneous Data Rate — From the above discussion, we have to determine the instantaneous data rate of a link within a scheduled RRG on the allocated subchannel in order to derive the instantaneous data rate of a scheduled queue.

For any D2D link or cellular uplink (i, j) , we define its potential interfering link $I_{i',j}$ as the communication channel from the transmitter of any link (i', j') that belongs to the same RRG with link (i, j) to the receiver of node j . For

² The instantaneous data rate can take units of bits/slot or packets/slot. The latter is appropriate when all the packets have fixed length and the achievable data rates are constrained to integral multiples of the packet size.

example, in Fig. 1 we have I_{10} (resp. I_{32}) as the potential interfering link of link (3, 0) (resp. link (1, 2)), since cellular uplink (3, 0) and D2D link (1, 2) can reuse the same radio resources. An interfering link is ‘potential’ since it only exists when the corresponding RRG is scheduled for transmission. Since there are two categories of links, that is, transmission links and potential interfering links, all links mentioned are referred to the transmission links by default in the rest of the article.

Assume that the instantaneous channel gain comprising the path loss, shadowing, and fast fading effects of the wireless channel from the transmitter of node $i \in \mathcal{N}$ to the receiver of node $j \in \mathcal{N}$ on any subchannel remains constant within a time slot. Similar to the Signal to Noise Ratio (SNR) of a transmission link, we define the virtual SNR of a potential interfering link $I_{i'j}$ on a subchannel as the ratio between the received interference power at node j from the transmitter of node i' and the noise power [9]. Obviously, the SNR value of a transmission link or the virtual SNR value of a potential interfering link on any subchannel is dependent on the transmission power from the transmitter of the source node on the considered subchannel, along with the instantaneous channel gain of the transmission link or potential interfering link. The SINR value of a link within a scheduled RRG on the allocated subchannel may or may not equal its SNR value, depending on whether there are any other links within the same scheduled RRG and thus causing interference to each other. Specifically, the SINR value of a link on any subchannel is determined by its own SNR value on the considered subchannel, and also the virtual SNR values of the interfering links from the transmitters of all the other links within the scheduled RRG to the receiver of the considered link on that subchannel [11].

The corresponding instantaneous data rate of a link within the scheduled RRG on the allocated subchannel is a function of its corresponding SINR value. We assume that Adaptive Modulation and Coding (AMC) is used, where the SINR values are divided into multiple non-overlapping consecutive regions. If the SINR value of a link falls within a certain region, its corresponding data rate is a fixed value according to the selected modulation and coding scheme in this state.

Figure 3 summarizes the dependency relationships between the variables in our network model as introduced in this section, where an arrow means that the variable in the source block depends on the variable in the destination block. It can be observed that the queue length as an important indicator of network performance depends on the subchannel and power allocation variables, and also the link constraint set that is determined by the selected RM for D2D connections. These variables are different under different resource control functions.

OPTIMIZATION FRAMEWORK

Based on the above network model, the resource control actions, including mode selection, resource allocation, and power control, can be optimized. As an initial attempt to address this problem, we introduce an optimization frame-

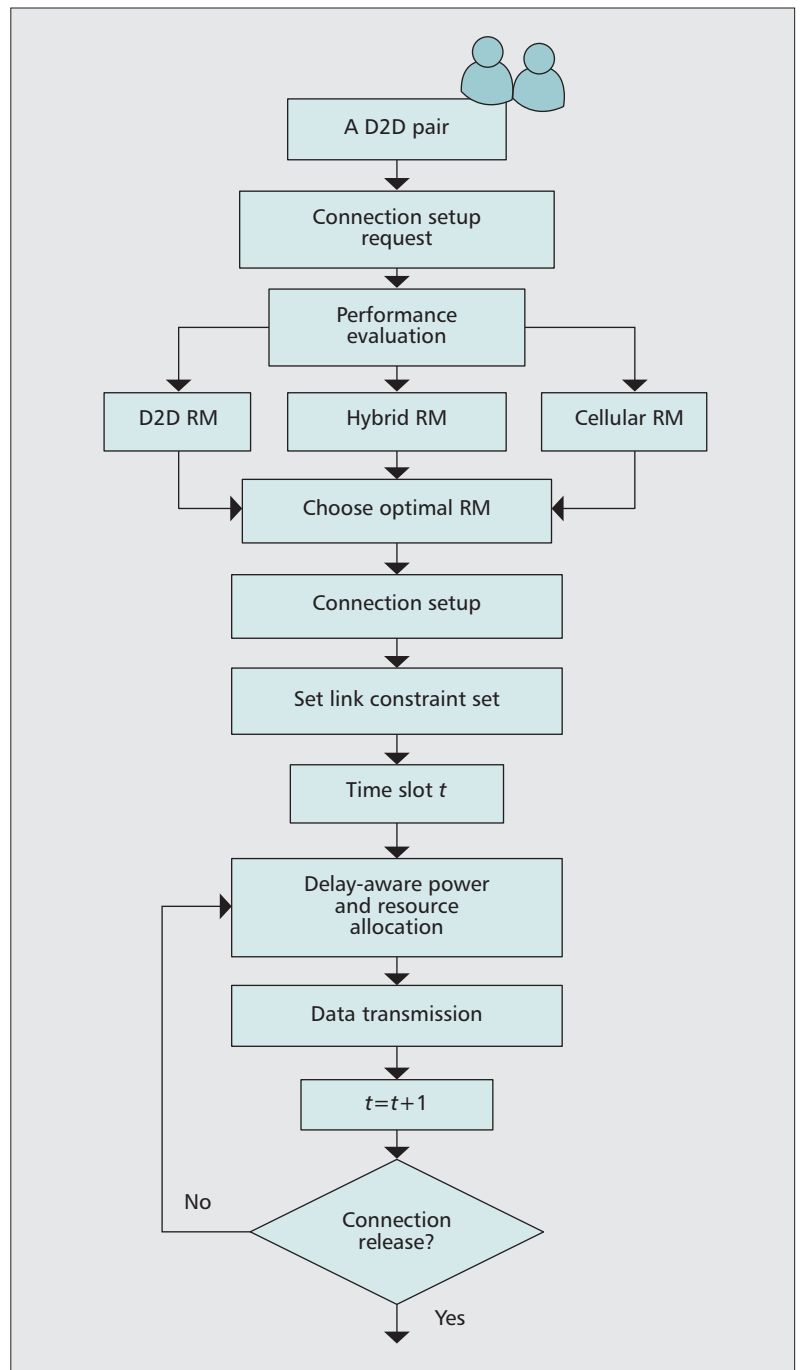


Figure 4. Optimization framework.

work as illustrated in Fig. 4. When a D2D pair wants to set up a connection between them, the BS first evaluates its performance such as average delay and dropping probability under D2D RM, Cellular RM, and Hybrid RM, respectively, assuming that the data arrival pattern of all the connections are known. Although the Hybrid RM generally achieves better performance than the D2D RM and Cellular RM due to its ability to exploit the channel variation opportunity, it involves larger computation complexity and signaling overhead, since more than one route exists for the D2D connection. Therefore, the selected RM should consider the trade-off between performance and complexity. Next, the

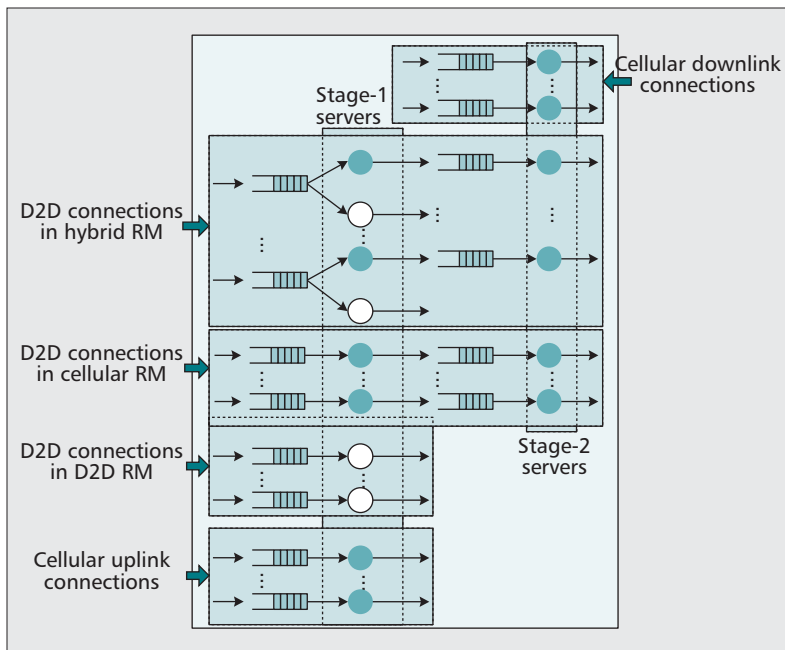


Figure 5. Queuing model for the general network model.

D2D connection is set up with the proper link constraint corresponding to the chosen RM. After that, data transmission is performed slot-by-slot until the connection is released. In each time slot, the optimal delay-aware resource and power allocation is performed for the D2D connection under the link constraint. Note that if D2D RM or Cellular RM is chosen, the D2D mode or Cellular mode is chosen semi-statically for the D2D pair. Otherwise, if the Hybrid RM is chosen, mode selection is performed dynamically and implicitly in each time slot by the resource allocation function, since mode selection reduces to whether to choose the D2D link or cellular uplink in each time slot.

QUEUING MODEL

In order to realize the optimal resource control in the above framework, a queuing model (QM) needs to be formulated for performance evaluation and optimization, which is illustrated in Fig. 5. We use a solid and a white circle to illustrate a server corresponding to a cellular link and a D2D link, respectively.

As the set of connections can be divided into five non-overlapping subsets, that is, cellular uplink connections, cellular downlink connections, D2D connections in D2D RM, in Cellular RM, and in Hybrid RM, the queues and servers in the QM can also be divided accordingly. For any cellular uplink or downlink connection, or D2D connection in D2D RM, since there is only one single-hop route, its QM has a single queue with an exogenous data arrival process, and a single server. For any D2D connection in Cellular RM, since there is one two-hop route, the system can be formulated as a two-stage tandem queuing model. Specifically, there is a queue at the src. DUE having a stage-1 server corresponding to cellular uplink and a queue at the BS having a stage-2 server corresponding to cellular down-

link. The new data arrives at the first queue, and joins the second queue immediately after it receives service from the stage-1 server, and leaves the system upon completion of service at the stage-2 server. For any D2D connection in Hybrid RM, since the data can be either transmitted via the one-hop route or two-hop route, the system can be formulated as a two-stage tandem queuing model as well. The only difference from the tandem queuing model of the Cellular RM is that there are two stage-1 servers for the first queue at the src. DUE, corresponding to D2D link and cellular uplink, respectively. The data in the first queue served by the server corresponding to cellular uplink will join the second queue immediately after it receives service from the stage-1 server, and leaves the system upon completion of service at the stage-2 server. On the other hand, the data in the first queue served by the server corresponding to D2D link will leave the system directly upon completion.

DELAY-AWARE RESOURCE CONTROL WITH BURSTY TRAFFIC: PRELIMINARY SOLUTIONS AND OPEN PROBLEMS

SYSTEM STATE REPRESENTATION

Preliminary Solutions — The global system state of the above QM at time slot t can be characterized by the aggregation of the CSI and QSI, that is, $\mathbf{S}_t = (\mathbf{H}_t, \mathbf{Q}_t)$. Let \mathcal{S} be the full system state space. The QSI \mathbf{Q}_t is defined as a vector consisting of the lengths of all the queues at the beginning of time slot t . The CSI is defined as $\mathbf{H}_t = \{\mathbf{H}_{ij,t} | (i, j) \in \mathcal{L}\}$, where $\mathbf{H}_{ij,t}$ denotes the CSI of link (i, j) in time slot t . We refer to the aggregate CSI for all the cellular uplinks and D2D links as uplink CSI, and the aggregate CSI for all the cellular downlinks as the downlink CSI. We will discuss the definition of $\mathbf{H}_{ij,t}$ for uplink CSI and downlink CSI, respectively.

It is much easier to represent the downlink CSI than the uplink CSI, since there is no interference between the cellular downlinks. Assume that every scheduled link on a downlink subchannel always transmits at constant power. Since every subchannel of a cellular downlink is flat fading, the downlink CSI per link (i, j) can be represented by a vector consisting of the local channel states of the link on every subchannel, where the local channel state on any subchannel depends on which region the received SNR falls in as defined by the AMC scheme.

Open Issues — The representation of the uplink CSI for the general network model is an open problem. Generally speaking, the definition of CSI should meet the following two requirements

- 1 The resource controller can make decisions based on the observation of the CSI and QSI.
- 2 Given the current CSI and QSI along with the control decision (e.g. the power and subchannel allocation decision) at any time slot t , the instantaneous data rate of every queue in time slot t can be determined.

When every scheduled link transmits at constant power, we can define $\mathbf{H}_{ij,t}$ as a vector, where each element represents the local channel state of link (i, j) on a subchannel when an RRG consisting of link (i, j) is scheduled. The local channel state depends on which region the SINR value of link (i, j) falls in, given the scheduled RRG and the allocated subchannel. Since the subchannel allocation decision can be made based on the observation of CSI and QSI, and a fixed transmission rate is determined by a local channel state according to the AMC scheme, both requirements 1 and 2 are satisfied. However, since the channel state space size for any link depends on the number of RRGs that it belongs to, this representation method is applicable only when the number of RRGs containing link (i, j) is small for every link.

When uplink power allocation is performed, the above defined channel state and transmission rate no longer have a one-to-one mapping, so that requirement 2 cannot be satisfied. Moreover, since a channel state corresponds to a SINR region instead of an exact SINR, optimal power allocation may not be possible due to the information loss, so that requirement 1 may not be satisfied. Therefore, the uplink CSI representation when power allocation is considered is of interest.

PERFORMANCE EVALUATION WITH A GIVEN CONTROL POLICY

Preliminary Solutions

Control Policy — In our formulated QM for the general network model, the RM for every D2D pair is assumed to be determined and the corresponding link constraint is set. Therefore, resource control reduces to power and subchannel allocation. In each time slot, the resource controller observes the system state \mathbf{S}_t and chooses an action from the set of allowable actions in the action space \mathcal{A} . An action \mathbf{a} is composed of a power allocation action \mathbf{p} and a subchannel allocation action \mathbf{x} . The power allocation action is a vector comprising of the allocated transmission power for all subchannels and all links, and the subchannel allocation action is a vector comprising the subchannel allocation for all RRGs and all subchannels.

A control policy prescribes a procedure for action selection in each state at all decision epochs t . We consider stationary Markovian control policies. A control policy can be either *deterministic* or *randomized*. A deterministic control policy given by Ω^{MD} is a mapping $\mathcal{S} \rightarrow \mathcal{A}$ from the state space to the action space, which is given by $\Omega^{\text{MD}}(\mathbf{S}) = \mathbf{a} \in \mathcal{A}, \forall \mathbf{S} \in \mathcal{S}$. A randomized control policy given by Ω^{MR} is a mapping $\mathcal{S} \rightarrow \mathcal{P}(\mathcal{A})$ from the state space to the set of probability distributions on the action space, which is given by $\Omega^{\text{MR}}(\mathbf{S}) = \{\phi_{\mathbf{S}}(\mathbf{a}) | \mathbf{a} \in \mathcal{A}\}, \forall \mathbf{S} \in \mathcal{S}$.

A deterministic control policy may be regarded as a special case of a randomized control policy in which the probability distribution on the set of actions is degenerate. Many existing channel-aware and/or queue-aware scheduling algorithms for wireless networks usually select the user with the largest utility, which is a function of the CSI and/or QSI. Examples are

MaxWeight, Exponential (Exp) rule, and Log rule scheduling algorithms. Since more than one user with the largest utility may exist, one of these eligible users can be randomly selected, and this category of control policies becomes randomized. We will discuss the performance evaluation method of randomized control policy in the following section, treating deterministic control policy as a special case.

Steady-State Probability and Performance Metrics Derivation

— The system behavior of the above QM can be represented by the discrete-time Markov chain (DTMC) $\{\mathbf{S}_t\}_{t=0,1,\dots}$. Given a system state \mathbf{S}_t and an action \mathbf{a} at time slot t , the state transition probability of the DTMC is given by $\text{Pr}\{\mathbf{S}_{t+1} | \mathbf{S}_t, \mathbf{a}\}$.

Given a deterministic control policy Ω^{MD} , since the action \mathbf{a}_t under every system state \mathbf{S}_t is determined, we can directly derive $\text{Pr}\{\mathbf{S}_{t+1} | \mathbf{S}_t, \Omega^{\text{MD}}(\mathbf{S}_t)\}$. Given a randomized control policy Ω^{MR} , we can derive $\text{Pr}\{\mathbf{S}_{t+1} | \mathbf{S}_t, \Omega^{\text{MR}}(\mathbf{S}_t)\}$ as the expectation of the state transition probability given an action w.r.t. the probability measure $\phi_{\mathbf{S}_t}(\mathbf{a})$ of every action under the given policy.

Define the transition probability matrix \mathbf{P}^{Ω} and the steady-state probability matrix π^{Ω} . Each element of the transition probability matrix \mathbf{P}^{Ω} can be derived depending on whether the control policy is deterministic or randomized. Then the stationary distribution of the ergodic process $\{\mathbf{S}_t\}_{t=0,1,\dots}$ can be uniquely determined from the balance equations. Given π^{Ω} , the end-to-end performance measures such as the mean throughput, the average delay, and the dropping probability for all the connections can be derived.

The computation of the steady-state distribution of $\{\mathbf{S}_t\}_{t=0,1,\dots}$ from the balance equations faces the challenge of exponentially enlarged state space, which makes it unacceptable for a large number of connections. In order to solve the above problem, the model decomposition and iteration method can be used to simplify the analysis [11].

Open Issues — The model decomposition and iteration method can deal with the exponential growth of system state space with the number of connections. However, it cannot be applied to tackle the exponential growth of CSI state space with the number of subchannels when a frequency-selective fading channel is considered. This is because a single queue is maintained for any connection $c \in \mathcal{C}$ on any node $i \in \mathcal{N}$, instead of one queue per subchannel. This means that the submodel for each queue cannot be further decomposed. Therefore, the aggregation of the state space to reduce its size with respect to the number of subchannels needs in-depth investigation.

CONTROL POLICY OPTIMIZATION

Preliminary Solutions — We have considered a power and subchannel allocation control policy to be fixed, and evaluated the system performance under every RM. The optimal RM is chosen based on the evaluation result. In order to further optimize system performance, the power

The MDP approach is a more systematic approach in dealing with delay-optimal resource control in general delay regime. The optimization of resource and power allocation policy under a RM can be solved by formulating a dynamic optimization problem using the MDP model.

and subchannel allocation control policy should also be optimized. In [9] three approaches to deal with the delay-aware resource control problem in wireless networks is given: the equivalent rate constraint (ERC) approach, the stochastic Lyapunov stability drift (SLS) approach, and the Markov Decision Process (MDP) approach. While the ERC approach allows potentially simple solutions, the resulting control policies are only functions of the CSI, and such policies are good only for the large delay regime where the probability of empty queues is small. The SLS approach can be directly extended to the D2D scenario. Specifically, for each link (i, j) , define its weight at each time slot as the maximum value of 0 and its differential backlog. Given the link weight, we can select a RRG with maximum sum over all its links of the product of link weight and link transmission rate. Compared with the ERC approach, the derived policies in SLS approach is adaptive to both the CSI and QSI and is throughput-optimal (in a stability sense). However, stability is only a weak form of delay performance, and derived policies may not have good delay performance especially in the small delay regime.

The MDP approach is a more systematic approach in dealing with delay-optimal resource control in a general delay regime. The optimization of resource and power allocation policy under a RM can be solved by formulating a dynamic optimization problem using the MDP model. To apply the MDP model, we need to define four elements: state space \mathcal{S} , action space \mathcal{A} , state transition probability $\Pr\{\mathcal{S}_{(t+1)} | \mathcal{S}_t, \mathbf{a}\}$, and reward function $g(\mathcal{S}, \mathbf{a})$. The former three elements have been defined above. Assuming that a reward function is given and only deterministic control policy is considered, we can define the infinite horizon average reward optimization problem as $\min_{\Omega} \mathbb{E}^{\pi(\Omega)}[g(\mathcal{S}, \Omega(\mathcal{S}))]$ under any unichain policy, where $\mathbb{E}^{\pi(\Omega)}[x]$ denote the expectation operation taken w.r.t. the unique steady-state distribution induced by the given policy Ω . The optimization problem can be solved by the Bellman Equation [9, 10]. However, the brute force value iteration method for the solution of the Bellman Equation is too complicated to compute due to the curse of dimensionality, that is, the exponential growth of the cardinality of the system state space and the large dimension of the control action space involved. Moreover, for the centralized solution, BS requires the global knowledge of CSI and QSI from all UEs in order to derive the optimal control action. However, these CSI and QSI observations are often measured locally at each UE, and huge signaling overhead is required to deliver all these local CSI and QSI to the BS. Therefore, a distributed implementation of the control policy with proper help from the BS is preferable.

Since the above-mentioned problems are general for delay-aware resource control in wireless networks, solutions have been proposed [9, 10]. First, a linear structure is employed that approximates the global value function of the Bellman equation by the sum of per-node value functions. Then a distributive online stochastic learning algorithm using stochastic approximation is

applied, where the control actions are computed locally by a few simple computations based on the local CSI and QSI measurements and the per-node value functions. Under some mild conditions, it is almost certain that the proposed solution will converge to the optimal solution.

Open Issues

Existing literature has used the distributed online stochastic learning algorithm to solve the delay-aware resource control problems in uplink OFDMA [9] and multi-user multi-relay scenarios [10]. However, our considered D2D scenario is more complex than the scenarios in the above two studies, since we consider intra-cell interference, multi-hop structure, and the OFDMA system simultaneously. Apart from the scenario differences, the above studies have made several assumptions that simplify the modeling and solution of the optimization problems. For example, in [10] it is assumed that the time slot duration is substantially smaller than the average packet interarrival time as well as the average packet service time. Therefore, the queue dynamic of each link becomes a controlled birth-death process, which simplifies the state transition probability computation of the queuing system. In [10] although two-hop transmission is also involved as in our MDP model, the control policy of the first phase of transmission is optimized independently of the second phase, while control policy for the second phase is optimized given the departure process of the first phase. This sequential optimization method is much simpler compared to the joint optimization of the two phases. However, it may not be globally optimal in the general case without the assumption of negligible blocking probability. For example, when the queue length of the second hop of a connection is large, it is better to restrain its transmission in the first hop even if its channel condition is good. Due to the above scenario and assumption differences, the problem of solving our MDP model for D2D communications is non-trivial and still open.

CONCLUDING REMARKS

In this article we have surveyed the existing resource control schemes for network assisted D2D communications, where the resource control actions, interference scenarios, and performance metrics studied in existing literature are summarized. Since most of the existing work focuses on the infinite backlog traffic model due to its simplicity, we have made an initial attempt at establishing an optimization framework for the delay-aware resource control policies with more realistic bursty traffic. The queuing model for performance evaluation and optimization are formulated within the framework. Furthermore, we outline our preliminary solutions and highlight the open issues.

ACKNOWLEDGMENT

This work was supported by the National Basic Research Program of China (973 program: 2010CB328105), the National Natural Science Foundation of China (No. 61272168, No. U1334202), the State Key Laboratory of Rail

REFERENCES

- [1] K. Zheng *et al.*, "Radio Resource Allocation in LTE-Advanced Cellular Networks with M2M Communications," *IEEE Commun. Mag.*, vol. 50, no. 7, July 2012, pp. 184–92.
- [2] L. Lei *et al.*, "Operator Controlled Device-to-Device Communications in LTE-Advanced Networks," *IEEE Wireless Commun. Mag.*, vol. 19, no. 3, June 2012, pp. 96–104.
- [3] X. Wu *et al.*, "FlashLinQ: A Synchronous Distributed Scheduler for Peer-to-Peer Ad Hoc Networks," *48th Annual Allerton Conf.*, Illinois, USA, Sept. 2010.
- [4] R. Zhang, L. Son, and Z. Han, "Distributed Resource Allocation for Device-to-Device Communications Underlying Cellular Networks," *IEEE ICC 2013- Ad-hoc and Sensor Networking Symp.*, 2013, pp. 482–6.
- [5] G. Fodor and N. Reider, "A Distributed Power Control Scheme for Cellular Network Assisted D2D Communications," *Global Telecommun. Conf. (GLOBECOM 2011)*, IEEE, Houston, Texas, USA, 5–9 Dec. 2011.
- [6] M. Jung, K. Hwang, and S. Choi, "Joint Mode Selection and Power Allocation Scheme for Power-Efficient Device-to-Device (D2D) Communication," *IEEE 75th Vehic. Tech. Conf.*, Yokohama, Japan, 6–9 May 2012.
- [7] C.-P. Chien, Y.-C. Chen, and H.-Y. Hsieh, "Exploiting Spatial Reuse Gain Through Joint Mode Selection and Resource Allocation for Underlay Device-to-Device Communications," *2012 15th Int'l. Symp. Wireless Personal Multimedia Commun. (WPMC)*, Sept. 2012, pp. 80–84.
- [8] C.-H. Yu *et al.*, "Resource Sharing Optimization for Device-to-Device Communication Underlying Cellular Networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, Aug. 2011, pp. 2752–63.
- [9] Y. Cui *et al.*, "A Survey on Delay-Aware Resource Control for Wireless Systems—Large Deviation Theory, Stochastic Lyapunov Drift, and Distributed Stochastic Learning," *IEEE Trans. Info. Theory*, vol. 58, no. 3, Mar. 2012, pp. 1677–701.
- [10] M. Moghaddari, E. Hossain, and L. B. Le, "Delay-Optimal Distributed Scheduling in Multi-User Multi-Relay Cellular Wireless Networks," *IEEE Trans. Wireless Commun.*, vol. 61, no. 4, Apr. 2013, pp. 1349–60.
- [11] L. Lei *et al.*, "Performance Analysis of Device-to-Device Communications with Dynamic Interference using Stochastic Petri Nets," *IEEE Trans. Wireless Commun.*, vol. 12, no. 12, Dec. 2013, pp. 6121–41.

BIOGRAPHIES

LEI LEI received a B.S. degree in 2001 and a Ph.D. degree in 2006, respectively, from Beijing University of Posts & Telecommunications, China, both in telecommunications engineering. From July 2006 to March 2008, she was a postdoctoral fellow at Computer Science Department, Tsinghua University, Beijing, China. She worked for the Wireless Communications Department, China Mobile Research Institute from April 2008 to August 2011. She has been an Associate Professor with the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, since Sept. 2011. Her current research interests include performance evaluation, quality-of-service and radio resource management in wireless communication networks.

YIRU KUANG received her B.E. degree in Beijing Jiaotong University (BJTU), China, in 2013. She is currently studying for a master's degree in Beijing Jiaotong University, BJTU. Her research interests lie in wireless communications, especially radio resource optimization and device-to-device communications.

XUEMIN (SHERMAN) SHEN [M'97, SM'02, F'09] received the B.Sc.(1982) degree from Dalian Maritime University (China)

and the M.Sc. (1987) and Ph.D. degrees (1990) from Rutgers University, New Jersey (USA), all in electrical engineering. He is a Professor and University Research Chair, Department of Electrical and Computer Engineering, University of Waterloo, Canada. He was the Associate Chair for Graduate Studies from 2004 to 2008. His research focuses on resource management in interconnected wireless/wired networks, wireless network security, wireless body area networks, vehicular ad hoc and sensor networks. He is a co-author/editor of six books, and has published more than 600 papers and book chapters in wireless communications and networks, control and filtering. He served as the Technical Program Committee Chair for IEEE VTC'10 Fall, the Symposia Chair for IEEE ICC'10, the Tutorial Chair for IEEE VTC'11 Spring and IEEE ICC'08, the Technical Program Committee Chair for IEEE Globecom'07, the General Co-Chair for Chinacom'07 and QShine'06, the Chair for IEEE Communications Society Technical Committee on Wireless Communications, and P2P Communications and Networking. He also serves/served as the Editor-in-Chief for *IEEE Network*, *Peer-to-Peer Networking and Application*, and *IET Communications*; a Founding Area Editor for *IEEE Transactions on Wireless Communications*; an Associate Editor for *IEEE Transactions on Vehicular Technology*, *Computer Networks*, and *ACM/Wireless Networks*, etc.; and the Guest Editor for *IEEE JSAC*, *IEEE Wireless Communications*, *IEEE Communications Magazine*, and *ACM Mobile Networks and Applications*, etc. He received the Excellent Graduate Supervision Award in 2006, and the Outstanding Performance Award in 2004, 2007 and 2010 from the University of Waterloo, the Premier's Research Excellence Award (PREA) in 2003 from the Province of Ontario, Canada, and the Distinguished Performance Award in 2002 and 2007 from the Faculty of Engineering, University of Waterloo. Dr. Shen is a registered Professional Engineer of Ontario, Canada, an Engineering Institute of Canada Fellow, and a Distinguished Lecturer of IEEE Vehicular Technology Society and Communications Society.

CHUANG LIN is a professor of the Department of Computer Science and Technology, Tsinghua University, Beijing, China. He is a Honorary Visiting Professor, University of Bradford, UK. He received the Ph.D. degree in Computer Science from the Tsinghua University in 1994. His current research interests include computer networks, performance evaluation, network security analysis, and Petri net theory and its applications. He has published more than 500 papers in research journals and IEEE conference proceedings in these areas and has published five books. Professor Lin is a senior member of the IEEE. He serves as the Technical Program Vice Chair, the 10th IEEE Workshop on Future Trends of Distributed Computing Systems (FTDCS 2004); the General Chair, ACM SIGCOMM Asia workshop 2005 and the 2010 IEEE International Workshop on Quality of Service (IWQoS 2010); the Associate Editor, *IEEE Transactions on Vehicular Technology*; and the Area Editor, *Journal of Computer Networks*.

ZHANGDUI ZHONG is a professor and advisor of Ph.D. candidates in Beijing Jiaotong University. He is now a director of School of Computer and Information Technology and a Chief Scientist of State Key Laboratory of Rail Traffic Control and Safety in Beijing Jiaotong University. He is also a director of the Innovative Research Team of Ministry of Education, and a Chief Scientist of Ministry of Railways in China. He is an executive council member of Radio Association of China, and a deputy director of Radio Association of Beijing. His interests are wireless communications for railways, control theory and techniques for railways, and GSM-R system. His research has been widely used in the railway engineering, such as Qinghai-Xizang railway, Datong-Qinhuangdao Heavy Haul railway, and many high-speed railway lines of China. He has authored/co-authored seven books, five invention patents, and over 200 scientific research papers in his research area. He received MaOYiSheng Scientific Award of China, Zhan-TianYou Railway Honorary Award of China, and Top 10 Science/Technology Achievements Award of Chinese Universities.

Since most of the existing work focuses on the infinite backlog traffic model due to its simplicity, we have made initial attempt at establishing an optimization framework for the delay-aware resource control policies with more realistic bursty traffic.