On Social-Aware Content Caching for D2D-Enabled Cellular Networks With Matching Theory

Jun Li, *Senior Member, IEEE*, Miao Liu, Jinhui Lu, Feng Shu, *Member, IEEE*, Yijin Zhang, *Member, IEEE*, Siavash Bayat, *Member, IEEE*, and Dushantha Nalin K. Jayakody, *Member, IEEE*

***Abstract*—In this paper, the problem of content caching in 5G cellular networks relying on social-aware device-to-device communications (DTD) is investigated. Our focus is on how to efficiently select important users (IUs) and how to allocate content files to the storage of these selected IUs to form a distributed caching system. We aim at proposing a novel approach for minimizing the downloading latency and maximizing the social welfare simultaneously. In particular, we first model the problem of maximizing the social welfare as a many-to-one matching game based on the social property of mobile users. We study this game by exploiting users' social properties to generate the utility functions of the two-side players, i.e., content providers (CPs) and IUs. Then we model the problem of minimizing the downloading latency as a many-to-many matching problem. For solving these games, we design a many-to-one IU selection (MOIS) matching algorithm and a many-to-many file allocation (MMFA) matching algorithm, respectively. Simulation and analytical results show that the proposed mechanisms are stable, and are capable of offering a better performance than other benchmarks in terms of social welfare and network downloading latency.**

***Index Terms*—Cellular network, content caching, device-to-device (D2D), download latency, mobile social network, social welfare.**

Manuscript received March 31, 2017; revised June 29, 2017 and August 7, 2017; accepted August 30, 2017. Date of publication September 5, 2017; date of current version February 25, 2019. This work was supported in part by the National Natural Science Foundation of China under Grant 61501238, Grant 61771244, Grant 61702258, Grant 61472190, and Grant 61602245, in part by the Jiangsu Provincial Science Foundation under Project BK20150786, in part by the Specially Appointed Professor Program in Jiangsu Province, in 2015, in part by the Fundamental Research Funds for the Central Universities under Grant 30916011205, in part by the Open Research Fund of National Mobile Communications Research Laboratory, Southeast University under Grant 2017D04 and Grant 2017D09, in part by the China Post-Doctoral Science Foundation under Grant 2016M591852, in part by the Post-Doctoral Research Funding Program of Jiangsu Province under Grant 1601257C, in part by the Natural Science Foundation of Jiangsu Province under Grant BK20150791, and in part by the Open Research Fund of National Mobile Communications Research Laboratory, Southeast University under Grant 2013D02. *(Corresponding author: Jinhui Lu.)*

J. Li and Y. Zhang are with the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China, and also with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210009, China (e-mail: jun.li@njust.edu.cn; yijin.zhang@gmail.com).

M. Liu, J. Lu, and F. Shu are with the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China (e-mail: miao.liu@njust.edu.cn; jinhui.lu@njust.edu.cn; shufeng@njust.edu.cn).

S. Bayat is with the Electronics Research Institute, Sharif University of Technology, Tehran 14588, Iran (e-mail: bayat@sharif.edu).

D. N. K. Jayakody is with the Department of Software Engineering, Institute of Cybernetics, National Research Tomsk Polytechnic University, 634050 Tomsk, Russia (e-mail: nalin.jayakody@ieee.org).

Digital Object Identifier 10.1109/JIOT.2017.2749320

I. Introduction

With the proliferation of smart phones and other derivative intelligent equipments, the network traffic has witnessed a trend of explosive growth. It is expected to increase by 40-fold over the next five years [1], due to mobile video stream and social network traffic. The increasing demands for high-rate transmissions and green power have motivated mobile operators and content providers (CPs) to redesign and seek for more efficient techniques. In this regard, mobile operators and CPs intend to improve their next generation wireless networks by utilizing low-cost, low-power access points, such as small-cell base stations (BSs) and femtocell access points [2]. However, numerous challenges have to be overcome by deploying these access points, such as interference management and efficient resource allocations.

Device-to-device (D2D) communications has emerged as a promising technique to facilitate communicating among mobile devices in the next-generation (5G) wireless communications [3], [4]. Conventional cooperative communication, i.e., relying-based communication technology, is the mainstream network architecture for 4G mobile networks. However, in 5G systems, D2D and small-cell networks are the two potentially dominant architectures. Boccardi *et al.* [5] and Wang *et al.* [6] investigated the direct communication among devices and machine-to-machine communication technology in the evolution of 5G. D2D communication provides high data rates, and reduces end-to-end delay and transmission power. In addition, mobile devices reuse the spectrum via D2D by the control of the BS, which enjoy the benefits of fast access to the radio spectrum [7]–[9]. User equipments (UEs) can obtain data from other mobile devices rather than the cellular BS by employing D2D communications [10], [11]. At the same time, content caching has been adopted as an efficient and low-cost solution to enable local downloading and reduce repetitive transmissions in 5G networks [12]. Thus, combining D2D and content caching can significantly improve network performance. Shanmugam *et al.* [13] selected couple of mobile devices as helper nodes for content delivery in the D2D-enabled wireless network. Although it is a promising technology for 5G wireless communications to meet unprecedented traffic demands, many challenges, such as mutual interference, need to be solved. Also, most research on D2D communications have limited focuses on the physical layer. The social-aware networks among the D2D participators can be investigated to further increase the transmission rate [14]. The social-aware D2D is a promising research topic in 5G networks, in which social characters of users can be integrated in the D2D networks, enabling the joint optimization in both physical domain and social domain. However, in relying systems, social characters cannot be applied.

Social-aware network provides various platforms to users for the purpose of online content sharing with their friends, or searching someone who has common interests in the virtual network. Interestingly, the connection established in the virtual network actually is tightly associated with our offline life. For example, on campus, our connected friends in Facebook, Twitter, YouTube, or Sina Blog, usually have a very close physical distance. In this respect, the influence of social interactions among UEs and mobile devices in wireless network have a very significant impact on resource allocation. Li *et al.* [14] made a detailed summary and analysis on the features of social networks and propose a social-aware D2D communication architecture.

As shown in [14], the social network characteristics consist of ties, community, centrality and bridge. Moreover, eigenvector centrality, closeness centrality and betweenness centrality are commonly used in the identification of social importance. Hu *et al.* [15] presented a novel approach utilizing eigenvector centrality to judge the relationship in social network. Recently, social network has been proposed to combine with the caching mechanism. In [16], the social tie and social similarity are involved in the resource allocation graphs and the placement of popular content is proposed considering the importance of nodes in social layer. Caching in mobile social networks is currently a hot research topic, in which popular contents can be cached in advance utilizing the limited storage of devices. Zhu *et al.* [17] proposed a social-aware caching game to incentivize nodes to cache data for others. However, [17] only one factor of social characteristic is taken into account when calculating the social performance. Ashraf *et al.* [18] proposed a selection algorithm without fully considering the social factor and physical factor. Also, rewarding mechanism of caching is not considered in both [17] and [18]. Moreover, Bai *et al.* [19] combined social-aware system with D2D communication and exploit mobility behaviors and social characteristics of mobile users to achieve the framework. However, how to efficiently select the important users (IUs) and how to match the contents with users in a joint manner remains a challenge. Especially, resource allocation in the wireless network keeps changing from the centralized mechanisms to self-organizing methods.

Recently, matching theory, a Nobel-prize winning framework, has become an effective approach for solving the combinatorial problems that utilizes mathematically tractable solutions, where the matching players are partitioned in two distinct sets. According to [20], the classification of matching problem includes one-to-one matching, many-to-one matching and many-to-many matching. A typical case of one-to-one matching is the stable marriage problem, where each player can be matched with at most one player in the opposite set. Many-to-one problem can be exemplified as the college admission problem. And in the many-to-many matching, the members in both sets can be matched with multiple players in the opposite set. These matching approaches are widely applied to wireless communications, such as resource allocation problem and power optimization problem [21].

While most of the existing works focus on resource allocations [22], [23], e.g., spectrum and power, there are few studies on content allocations, especially by using many-to-many matching game. Hamidouche *et al.* [24] considered the files allocation problem without taking into account the social performance of the caching placement. In [25], the many-to-one matching game is applied to content allocations. However, the authors only consider transmissions in a fixed framework and also ignore the social performance of caching placement. To be specific, Gale and Shapley [26] illustrated the matching problem and prove the stability by using deferred acceptance procedure. The many-to-one matching has been utilized in spectrum allocation, where players have different preferences toward network resources [27]. Bayat *et al.* [28] proposed a novel solution for femto-cell access points allocation and spectrum allocation by using many-to-one matching. Moreover, Xu *et al.* [29] proposed a content distributed method based on matching theory in social network. However, many-to-many matching has a high complexity compared with many-to-one matching, and thus has not been widely used.

In this paper, our focus is on a two-layer matching game. The first matching is to solve the IUs selection problem and the second one is to solve the social-aware content allocation problem in wireless cellular networks. In the first problem, CPs and mobile users are assumed as the players in the matching game, where we optimize the selection of IUs to be the D2D nodes for data offloading. The game is determined as a many-to-one matching game. In the second problem, the matching between IUs and files is modeled as a many-to-many matching game. Our main contributions can be summarized as follows.

1) We first propose a many-to-one IUs selection (MOIS) matching game to solve the problem of IUs selection to obtain the D2D nodes for files transmission based on the content popularity, social connection features, and the wireless physical layer metrics.
2) Then we propose a many-to-many files allocation (MMFA) matching game to solve the content allocation problem. In this game, the two sides of players establish their preferences toward each other considering social network transmission performance.
3) Finally, the stability, the convergence and the optimality of the proposed matching algorithm are investigated and proved. Moreover, simulations are carried out to evaluate the performance of the proposed algorithms.

The rest of this paper is organized as follows. In Section II, we describe the system model. The IUs selection problem and the content allocation problem are formulated in Section III. In Section IV, we propose the many-to-one and many-to-many matching games to optimize the system performance, and design two novel matching algorithms based on the two games. In Section V, the stability, convergence, and the optimality of two algorithms are analyzed. Simulation results are conducted in Section VI and the conclusion is drawn in Section VII.

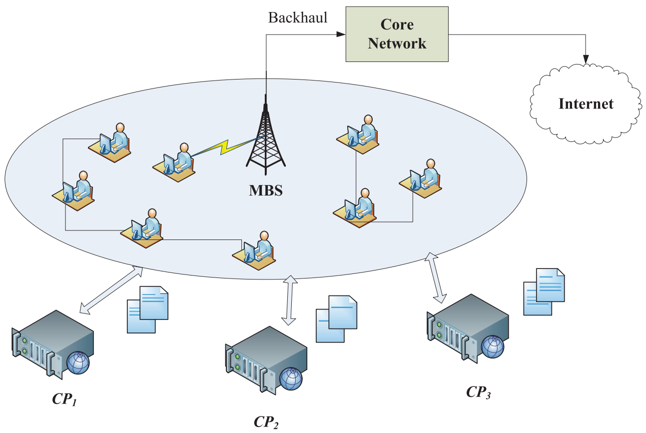

Fig. 1. Social-aware caching system relaying on D2D-enabled cellular networks, which consists of CPs and mobile users. A social virtual relationship exists between a pair users. Moreover, one BS owned by a operator can guarantee the normal communications apart from D2D.

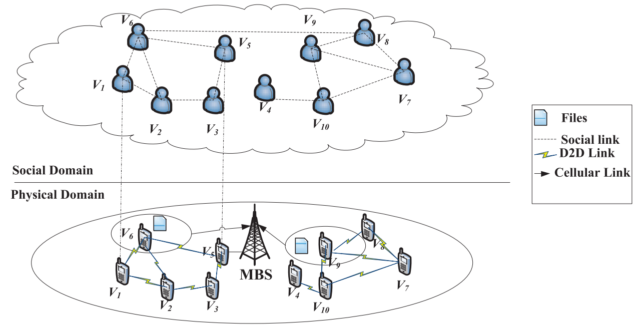

Fig. 2. Illustration of network system for combining social domain and physical domain under the cellular network.

II. System Model

A. System Description

As shown in Fig. 1, the considered caching system consists of C CPs, and a number of UEs that are randomly distributed in the network. A macro-cell BS (MBS) owned by a certain operator is deployed with a fixed transmit power. We denote the set of M UEs by $\mathcal{U} = \{U_1, U_2, U_3, \ldots, U_M\}$, where U_m, $m \in \mathcal{M} = \{1, 2, 3, \ldots, M\}$, represents the mth UE, and denote by $\mathcal{CP} = \{\mathrm{CP}_1, \mathrm{CP}_2, \ldots, \mathrm{CP}_C\}$ the set of the CPs. In addition, the CPs provide the file library $\mathcal{F} = \{f_1, f_2, \ldots, f_L\}$, where $f_l \in \mathcal{F}$ denotes the lth file and L is the total number of the files. In addition, the IUs are selected from the UEs as the caching nodes for D2D communications, which are presented by $\mathcal{I} = \{I_1, I_2, I_3, \ldots, I_N\}$, where N is the total number of IUs with I_n denoting the nth IU. Note that an IU can also request data from other IUs if this IU does not cache the data it wants.

The transmission channels are modeled as Rayleigh fading. A dedicated frequency band of bandwidth W is allocated to the channels for contents distribution. Additionally, the spectrums used by D2D is different from the spectrum used by MBS. Hence, there is no interference between IUs and the MBS in the cellular network. We set P_n as the transmission power of nth IU and σ^2 the noise power at each UE. The path loss between nth IU and the mth UE U_m is modeled as $k_{n,m}^{-\alpha}$, where α is the path-loss exponent and $k_{n,m}$ is the physical distance between I_n and U_m.

We assume that the connections between UEs are detected according to the physical distance and the frequency of social network communications. This means that the pairs of linked UEs need to have a deep social trust and adjacent geographical position for D2D communications. For the purpose of data offloading, we consider that D2D communications are more preferable for file downloading than the MBS transmissions. That is, the UEs prefer to offload the demanded data from trustworthy friends rather than from the MBS.

In order to encourage the UEs to offload demanded files via D2D rather than directly accessing the MBS, some incentives are needed to motivate the UEs. For example, the cost of downloading the files from D2D approach is much cheaper than via the MBS. Additionally, we assume that the MBS is far away from the mobile users. So the transmission rate supported by the MBS is never higher than the lowest download rate supported by the IUs underlaying our wireless network framework.

Moreover, each IU has a limited cache size and limited transmission range, while the MBS has a large enough storage capacity to accommodate the entire library. In order to avoid traffic congestion at the backbone network due to repetitive data downloading requests, CPs are willing to hire some IUs by providing monetary payment to the selected IUs for content transmissions. The IUs can cache some popular files utilizing their limited storage space and UEs can obtain required contents from the IUs via D2D transmissions or the MBS via cellular connection. The limited storage capacity of IUs is denoted by $\mathcal{Q} = \{q_1, q_2, q_3, \ldots, q_N\}$, where q_n denotes the number of files that each IU can store. In the case that the demanded file is cached at an IU which is connected with the requesting UE, the UE can obtain the demanded file from this IU via D2D. Otherwise, the UE obtains his desired file from the MBS, leading to a higher transmission delay.

The social community is divided according to the general positions of users during a certain period. Moreover, the UEs are the customers of different CPs and tend to choose the services from their preferred CPs. In addition, for simplicity, we assume that there is no communications between a pair of users crossing two communities. In order to improve personal satisfactory, CPs are willing to hire some IUs for data caching and relaying for its customers. In this case, once an IU is hired by a CP, it will give up its previous preference and remain neutral for D2D transmissions. Users within a community may have similar interests and demand for similar contents. Thus, data mining from their daily behavior is important for us to find beneficial information. The first thing that we need to do is detecting their social relationship in the community, such as Weibo and YouTube. Then we focus on the combination of social link and real link in physical layer.

A detailed description of combination of social network and D2D communication is shown in Fig. 2. Such a system can be divided into two domains: the physical domain and social domain. In the social domain, the virtual connection between social users can reflect their offline behaviors, which can be speculated from the platform, such as Weibo, Twitter, Facebook, and so on. Thus, we can infer the close degree of user's relationship by observing their behaviors in social network platforms. On the other hand, in the physical domain, users can access to the cellular network or establish D2D connections. Moreover, the performance of D2D communication is mainly determined by transmission distance. Obviously, each connection in social domain can be projected onto a real link in the physical domain.

B. Content Popularity

We now model the distribution of the content popularity. We suppose that the probability of requesting the file f_l, $\forall l$, follows the Zipf distribution over a period of statistical time, which is defined as:

$$p_l = \frac{1/l^\gamma}{\sum_{i=1}^{L} 1/i^\gamma} \qquad \forall l = 1, \ldots, L \tag{1}$$

where γ characterizes the steepness of the distribution, reflecting different content popularity. Generally, users have different appetites to various files. Thus, the content popularity matrix for all users is given by $\mathbf{P} \in R^{M\times L}$ where each entry $P_{m,l}$ represents the probability that the mth user requests the content f_l and the relation between $P_{m,l}$ and p_l is illustrated in [30] in detail.

Due to the limited storage capacity of the selected IUs and increasing demand during traffic peak periods, we assume that each IU can cache multiple files within their storage. In turn, one file can be cached by multiple IUs. For simplicity, all of these files have the same size. The purpose of the IUs to cache popular files is for minimizing the average download delay from cellular network via D2D communications. Generally speaking, social users are selfish. Thus, to motivate the IUs to cooperate with the transmissions, the CPs are willing to provide monetary compensations.

C. Social Characters and Physical Characters

We use $\mathcal{Z} = (\mathcal{U}, \mathcal{E})$ to denote a social relationship graph, in which $\mathcal{U} = \mathcal{I} \cup \mathcal{G}$ and $\mathcal{E}$ is the set of edges connecting I_n and U_m. Adjacent nodes (I_n, U_m) are connected by means of a bidirectional edge $e(I_n, U_m) \in \mathcal{E}$. For simplicity, the existence of $e(I_n, U_m)$ represents the connection between I_n and U_m, while the nonexistence of $e(I_n, U_m)$ represents the disconnection of them.

In general, the MBS is assumed to be far away from the UEs. This will not only encourage the UEs to download data from the D2D transmissions first, but also effectively offload the data traffic from the MBS to the IUs. In addition, the UEs can only communication with IUs that they are connected with, considering the trust cost risks. In the case that the UEs request data from the IUs that have no link with them, we suggest that the access delay is infinite for simplicity.

According to the wireless network model, the transmission rate of direct cellular network connection is denoted by a fixed value R_0, while the D2D connection can be expressed by

$$R_{n,m} = W \log_2\left(1 + \frac{P_n h_{n,m}^2 k_{n,m}^{-\alpha}}{\sum_{I_{n'}\in\mathcal{I}, n\neq n'} P_{n'} h_{n',m}^2 k_{n',m}^{-\alpha} + \sigma^2}\right) \tag{2}$$

where W is the bandwidth of transmission channel, and $h_{n,m}$ is the channel coefficient between I_n and U_m. As we all known, desirable transmission rate is a key factor to improve the satisfaction of social users, since whether files can be successfully delivered or not depends on the transmission rate and the stability of connections.

However, the quality of a connection between a pair of IU and UE is not only determined by the transmission rate, but also by the connect duration at a temporary moment. In specific, the connect duration is defined as the time frame, beginning with the initial contact of the two sides in a D2D pair, while ending with their moving out of each other's transmission range. In this paper, we adopt contact duration as one of the metrics for IU selections, denoted by T_c, which can be approximated by the exponential distribution [15].

Accordingly, we assume that the contact duration follows exponential distribution with rate λ, and contact processes between connected pairs are independent. Thus, we write the distribution of the connection duration as

$$f(T_c) = \begin{cases} \lambda e^{-\lambda T_c}, & \text{if } T_c \geq 0 \\ 0, & \text{if } T_c < 0. \end{cases} \tag{3}$$

Furthermore, we assume that the success of files delivery depends on the following constraints.

1) Once a connection between an IU and a UE is determined, the two pair's relationship is trustworthy and harmless to guarantee the files transmission.
2) The connection duration between requesters and linked IUs must be longer than the required transmission time that is defined as $Y/R_{n,m}$, where Y is the file size.

As a result, the success probability for delivery a file f_l from I_n to a file-requester U_m during a contact duration is given by

$$F_{T_c}(Y/R_{n,m}) = P_{n,m}^l(T_c \geq Y/R_{n,m}) \tag{4}$$

where $P_{n,m}^l(T_c \geq Y/R_{n,m})$ denotes the probability of successful obtain desired content f_l from helper I_n by U_m. We further have

$$\begin{aligned} P_{n,m}^l\big(T_c \geq Y/R_{n,m}\big) &= \Pr\big(T_c \cdot R_{n,m} \geq Y\big) \\ &= \int_{Y/R_{n,m}}^{\infty} \lambda e^{-\lambda T_c}\, \mathrm{d}T_c \\ &= e^{-\lambda Y/R_{n,m}}. \end{aligned} \tag{5}$$

In this paper, we assume that if the probability of successful delivery the content $F_{T_c}(Y/R_{n,m})$ is lower than a threshold Γ, the connect is unstable and lead to a downloading failure.

Furthermore, we consider the connectivity of a pair in $\mathcal{U}$, which can be represented by an adjacency $M \times M$ symmetric matrix $\boldsymbol{A}$. In specific, as an element in $\boldsymbol{A}$, the connectivity between an IU I_n and UE U_m, denoted by $a_{n,m}$, represents the joint social and physical connection, that is

$$a_{n,m} = \begin{cases} 1, & e(I_n, U_m) \text{ exists and } T_c \geq Y/R_{n,m} \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

III. Problem Formulation

In this section, we formulate two problems. In the first problem, we need to solve the IU-selection issue. Generally, the

IUs are selfish and greedy to increase their profits. In turn, CPs hire IUs to improve their business impact and social satisfaction. Assume that once an IU is hired by one CP, it will not accept other CPs' invitations. The purpose of selecting IUs is maximizing their social utility function and to find the game equilibrium for price allocation. In the second problem, CPs make decisions on optimal file placement on the IUs for minimizing the overall downloading latency.

A. First Problem

To encourage the IUs to participate in the cooperation, CPs would like to offer some rewards for inspiration. We define the utility function of CPs as

$$S_i^{\mathrm{CP}} = S_i^{\mathrm{gain}} - S_i^{\mathrm{rent}}, \quad \text{for} \quad i = 1, \ldots, C \tag{7}$$

where $S_i^{\mathrm{gain}} = \sum_{m=1}^{Q_i} S_{i,m}^{\mathrm{gain}}$ determines the total gains via renting the IUs to the ith CP. Additionally, $S_{i,m}^{\mathrm{gain}}$ is the satisfaction of selecting U_m which is related to social characters and Q_i represents the maximizing rental number of IUs for the ith CP. Furthermore, $S_i^{\mathrm{rent}} = \sum_{m=1}^{Q_i} S_{i,m}^{\mathrm{rent}}$ is the total rental costs and $S_{i,m}^{\mathrm{rent}}$ is the cost by renting the mth UE for ith CP. The concrete values of $S_{i,m}^{\mathrm{rent}}$ will be discussed later. Similarly, the utility function of the mth UE is defined as

$$S_m^{\mathrm{UE}} = S_{i,m}^{\mathrm{gain}} - S^{\mathrm{cost}} \tag{8}$$

where $S_{i,m}^{\mathrm{gain}}$ is the received price from CP_i. S^{cost} denotes the UE's individual maintenance costs and is a constant. In the first problem, the purpose of selecting IUs for CPs is to find a equilibrium such that both selection of CPs and IUs is stable and make the benefit maximize. The optimization problem can thus be written as

$$\begin{aligned} \max_{\boldsymbol{X}} \quad & \sum_{i=1}^{C}\sum_{m=1}^{M} x_{i,m} S^{\mathrm{social}} \\ \text{s.t.} \quad & ① \; x_{i,m} \in \{0,1\} \\ & ② \sum_{m=1}^{M} x_{i,m} \le Q_i, \quad i = 1, \ldots, C \\ & ③ \sum_{i=1}^{C} x_{i,m} \le 1, \quad m = 1, \ldots, M \end{aligned} \tag{9}$$

where S^{social} denotes the social welfare, which is the sum of the utility of CPs and the selected UEs, and the calculation will be introduced in the next section. $\boldsymbol{X}$ is the selecting matrix and the element $x_{i,m}$ denotes that the mth UE is selected as IU by ith CP. Condition ① guarantees that $x_{i,m}$ is a binary variable and condition ② states that a maximum number of UEs can be selected by a CP. Moreover, condition ③ denotes a maximum number of CPs that a UE can match. Observing from above problem, we find that the optimization problem (9) is an NP-hard combinatorial problem. Moreover, the optimization problem is a binary linear programming problem, which is a special case of integer problem, and in this case, the variable can only be zero or one. It means that there exists a solution of this optimization problem in which the value of $x_{i,m}$ will be one or zero. Hence, for solving this problem, we need to optimize the matrix $\boldsymbol{X}$. In addition, there exists many feasible solution for the selection problem, but among all the results, the solution $\boldsymbol{X}$ is called optimistic solution if

$$\sum_{i=1}^{C}\sum_{m=1}^{M} x_{i,m} S^{\mathrm{social}} > \sum_{i=1}^{C}\sum_{m=1}^{M} \bar{x}_{i,m} S^{\mathrm{social}} \tag{10}$$

where $\bar{x}_{i,m}$ does not belong to the optimal solution $\boldsymbol{X}$. The above analysis results in that there exists only one optimal solution and guarantees the existence of $\boldsymbol{X}$.

Notice that the optimization problem for selecting IUs is for profits maximization. However, in real life, the performance of UEs subjects to many restrictions such as storage capacity, battery capacity and hardware quality. For the sake of simplicity, we assume that the storage of UEs is unified and the battery capacity is adequate for transmission tasks. Meanwhile, the hardware quality is homogeneous and CPs consider the factor of transmission characteristics.

After all the CPs select their desired IUs, we need to formulate the problem of how to resolve the optimum content allocation between the IUs and files.

B. Second Problem

In the second problem, the CPs make decision on content allocation to IUs so that UEs can obtain demanded files from the IUs rather than from the MBS for the purpose of lower offloading delay. To formulate the allocation problem between the files and IUs, we set up a file-distribution matrix $\boldsymbol{D}^{L\times N}$. The entry $d_{l,n} \in \{0, 1\}$ in $\boldsymbol{D}$ indicates whether f_l is cached by the nth IU or not. Thus, we have

$$d_{l,n} = \begin{cases} 1, & \text{if } f_l \text{ is cached by } n\text{th IU} \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

In the second problem, the strategy of each IU is to select caching files for sake of minimizing the average transmission delay of its connected UEs. In a community, UEs may have different interests on various contents during a certain time. Based on the recommendation of social software, the content with a higher click rate and downloads is more frequently accessed by the UEs. Recall the contact duration we have defined previously. According to the consideration, we denote the offloading delay of different request methods in the network.

First, we define $\mathcal{H}_m$ as the set of IUs connected with the mth UE. The delay of downloading the file f_l by the UE U_m can be calculated as

$$T_{m,l} = \begin{cases} \frac{Y}{\max_{g\in\mathcal{H}_m}\{R_{g,m}\}}, & \exists d_{l,g}\neq 0,\ a_{g,m} = 1 \quad \forall g \in \mathcal{H}_m \\ \infty, & \exists d_{l,g}\neq 0,\ a_{g,m} = 0 \quad \forall g \in \mathcal{H}_m \\ \frac{Y}{R_0}, & \text{otherwise.} \end{cases} \tag{12}$$

There are three cases in the above equation. The first case is that if the demanded content is cached at I_n and the social and physical constraints between I_n and U_m are satisfied, the delay for downloading the file f_l is based on the minimum D2D delay. In addition, the value of D2D delay is related to the D2D transmission rate. The second case is that if the demanded file is cached at I_n but the social and physical constraints between

I_n and U_m are not satisfied, the UE access I_n with an infinite delay because the untrustful connection. In addition, if the outage is beyond the scope of the threshold, the requester will not access the IU even it has satisfactory transmission delay. In this case, the constraints avoid the unsafe connection. The third case is that demanded file is not cached at the neighboring IUs, hence the UE U_m has to obtain the content via the MBS.

Based on the transmission delay and the popularity distribution of the files, the delay for user U_m to download a file from $\mathcal{F}$ can be written as

$$T_m = P_{m,l} \cdot T_{m,l}. \tag{13}$$

Thus, the content allocation strategy can be obtained by solving the following optimization problem:

$$\begin{aligned} \min_{\boldsymbol{D}} \quad & \sum_{l=1}^{L}\sum_{m=1}^{M} d_{l,n}T_m \\ \text{s.t.} \quad & ① \sum_{l=1}^{L} d_{l,n} \le q_n, \quad I_n \in \mathcal{I} \\ & ② \sum_{n=1}^{N} d_{l,n} \le N_l, \quad N_l \le N \\ & ③\, F_{T_c}(Y/R_{n,m}) > \Gamma \end{aligned} \tag{14}$$

where N_l is the number of IUs that files can be cached. Specifically, condition ① represents that a maximum number of files can be stored for an IU and condition ② guarantees the maximum number of IUs that a file can be cached in. Also, condition ③ denotes a connection limitation of D2D communication. Similarly, the optimization problem (14) is an NP-hard combinatorial problem as the problem (9). The common character of the two problem is that they can be modeled as matching problems. However, the first problem is a many-to-one matching problem, and the second problem is a many-to-many matching problem.

It should be noted that the optimization problems (9) and (14) have some limitations concerned about matching condition. In fact, CPs must provide wireless service even if UEs cannot access IUs during congestion phase or the performance of D2D is low in some ares so that cellular connection is a good choice. Moreover, in problem (14), there exists a many-to-many matching game between files and IUs. In order to efficiently utilize the storage, the IUs can cache multiple required files and a file can be cached in multiple IUs. In addition, a file can be cached only by N_l IUs to guarantee that more files are cached for transmission capability.

From the above analysis of the proposed two problems, we find that they are coincident with a unknown binary matrix and both of them can be regarded as NP hard problems. Thus in the next section, we propose two matching algorithms to solve the two optimization problems, respectively.

IV. Matching Algorithms

This paper involves many-to-one and many-to-many matching games. The first problem belongs to many-to-one problem and the second problem belongs to many-to-many problem. Thus, in this section, we will first introduce the definitions of many-to-one matching as well as many-to-many matching. Then, we propose two matching algorithms for solving these two problems.

A. Matching Related Definitions

The many-to-one matching is introduced from *college admission*, in which the students apply for their college based on their preference and in turn, the college decides wether to accept or refuse the application from students. Formally, let us assume that there are two finite and disjoint sets: $\mathcal{U} = \{U_m\}_{m=1}^{|\mathcal{U}|}$ and $\mathcal{CP} = \{\mathrm{CP}_i\}_{i=1}^{|\mathcal{C}|}$. Here $\succ_{\mathcal{U}} = \{\succ_{U_m}\}$ and $\succ_{\mathcal{CP}} = \{\succ_{\mathrm{CP}_i}\}$ denote, respectively, the set of preference relations of two players. Let $V_m(i)$ denote the gain of the user U_m obtained from CP_i, while $V_i(m)$ denote the gain of CP_i obtained from U_m. Given these utilities, we can get the following instructions.

Definition 1: A many-to-one matching function Φ is defined as a function from the set $\mathcal{U} \cup \mathcal{CP}$ based on the preference list.

1) $\Phi(U_m) \subseteq (\mathcal{CP} = \{\mathrm{CP}_i\}_{i=1}^{|\mathcal{C}|})$ and $|\Phi(U_m)| \le Q_i$.
2) $\Phi(\mathrm{CP}_i) \in (\mathcal{U} = \{U_m\}_{m=1}^{|\mathcal{U}|})$ and $|\Phi(\mathrm{CP}_i)| \in \{0, 1\}$.
3) $\Phi(U_m) = \mathrm{CP}_i \Leftrightarrow \Phi(\mathrm{CP}_i) = U_m$.

The Φ is determined on account of the utility, respectively. For example, if a UE prefers a content provider CP_i, it can be express as

$$V_m(i) > V_m\big(i'\big) \Leftrightarrow \mathrm{CP}_i \succ_{U_m} \mathrm{CP}_{i'}. \tag{15}$$

Also, the CP_i's preference over users can be described by

$$V_i(m) > V_i\big(m'\big) \Leftrightarrow U_m \succ_{\mathrm{CP}_i} U_{m'}. \tag{16}$$

Furthermore, the many-to-many matching is introduced for creating partnerships in D2D communications generally. In the game, each player within the two disjoint sets can select more than one opponents in the other set for improving their utility by matching. Given the set of players $\mathcal{I} = \{I_n\}_{n=1}^{|\mathcal{I}|}$ and the set of players $\mathcal{F} = \{f_l\}_{l=1}^{|\mathcal{F}|}$, which are finite and disjoint.

Definition 2: A many-to-many matching function Ψ is defined as a function from the set $\mathcal{I} \cup \mathcal{F}$ based on their preference lists.

1) $\Psi(f_l) \subseteq (\mathcal{I} = \{I_n\}_{n=1}^{|\mathcal{I}|})$ and $|\Psi(f_l)| \le N_l$.
2) $\Psi(I_n) \in (\mathcal{F} = \{f_l\}_{l=1}^{|\mathcal{F}|})$ and $|\Psi(I_n)| \le q_n$.
3) $\Psi(f_l) = I_n \Leftrightarrow \Psi(I_n) = f_l$.

The expression of preference is similar to the define of the many-to-one matching, so we omit the details of description. Observing from above two definitions, we denote matching function $\Phi : (\mathcal{U}, \mathcal{CP}, Q_i)$ and $\Psi : (\mathcal{I}, \mathcal{F}, N_l, q_n)$ as many-to-one matching and many-to-many matching, respectively. Moreover, note that the establishment of a matching relationship is successful if and only if $\Phi(U_m) = \mathrm{CP}_i$ as well as $\Phi(\mathrm{CP}_i) = U_m$, which guarantees that the matching game is a bilateral agreement.

B. MOIS Matching Algorithm

In this section, we propose an MOIS matching algorithm to solve the optimal problem (9). The first problem consists of two types of game players, including UEs and CPs. To design a algorithm for solving the first problem, where UEs and CPs maximize their utility function strategically, we first focus on the preference lists of two players. For inspiring UEs to cooperate, the CPs will provide some compensation to hire storage spaces of the UEs for files allocation and D2D transmission. In the first problem, a CP can hire multiple UEs as IUs for serving its users, while an UE can only be rented by one CP to avoid interference and privacy leaks. In this case, once an UE is rented by a CP, the UE must give up its initial preference over the CPs and keep impartial to provide services to the employer.

Since all the CPs and UEs would like to maximize their social welfare utility function, we now focus on how to formulate the utility function. The utility function of a CP toward UEs is related to the UE's social characteristics and it should be increasing with UE's social importance and decreasing with the CP's costs. Let us consider the UEs' social importance as follows.

Taking both social interaction and interest similarity into account to choose the IUs, we define the following importance measurement between U_j and U_k as:

$$w_{j,k} = \mu b_{j,k} + \nu s_{j,k} + \upsilon c_{j,k} \tag{17}$$

where $b_{j,k}$, $s_{j,k}$, and $c_{j,k}$ denote, respectively, the social trust insex, interest similarity, and relative battery capacity between U_j and U_k. Additionally, μ, ν, and υ are adjustable parameters with constraint $\mu + \nu + \upsilon = 1$. Betweenness centrality is one commonly used way to measure the nodes social trust. The edge betweenness centrality between node U_j and U_k can be calculated as

$$\overline{b}_{j,k} = \sum_{j,k \in \mathcal{V}} \frac{d_{jk}(e)}{d_{jk}} \tag{18}$$

according to [31]. In this equation above, d_{jk} is the number of shortest distance paths of connecting from node U_j to U_k, and $d_{jk}(e)$ is the number of geodesic paths including edge e. In order to facilitate the calculation, a normalized element is as follows:

$$b_{j,k} = \frac{\overline{b}_{j,k}}{(M-1)^2}. \tag{19}$$

1) Similarity Matrix: In [32], for a pair of nodes, (U_j, U_k), their similarity matrix is defined as

$$\overline{s}_{j,k} = \begin{cases} \sum_{z \in M(j) \cap M(k)} \frac{1}{k(z)}, & \text{if } j \text{ is connected with } k \\ 0, & \text{otherwise} \end{cases} \tag{20}$$

where $M(j)$ is the set of neighbors of U_j, and $z \in M(j) \cap M(k)$ denotes the set of the common neighbors between node U_j and U_k. $k(z)$ is the number of nodes connected with z. Similarly, in order to facilitate the calculation, the simple additive weighting method is considered. Also, the normalized entries are

$$s_{j,k} = \frac{\overline{s}_{j,k}}{\max \overline{s}_j} \tag{21}$$

where $\overline{s}_j$ denotes the jth row of $\overline{s}_{j,k}$. In (17), based on the above definition, we can formulate the preference function of CPs over UEs.

According to the analysis from above, we can formulate the utility function of CPs and UEs. To design the utility function, $\mathcal{N}_m = \{N_{m1}, N_{m2}, \ldots, N_{mh}\}$ denotes the index set of nodes connected with UE U_m. The utility function of CPs can be rewritten as

$$S_i^{\mathrm{CP}} = \zeta \sum_{j=1}^{|\mathcal{N}_m|} F_{T_c} w_{m,j} - \sum_{m=1}^{Q_i} S_{i,m}^{\mathrm{rent}}, \text{ for } i = 1, \ldots, C \tag{22}$$

where $S_{i,m}^{\mathrm{rent}}$ is the rent price of CP_i to in exchange for the service of U_m previously involved in Section II and ζ is a fixed coefficient effected by the number of customers of CPs. The utility function of CP is affected by the importance of UEs and the price of rent. Besides, the importance of UEs can be calculated by considering the interest similarity and social trust. Moreover, by sorting the utility function in a descending order, we can obtain the preference list, denoted by $\mathcal{PR}(\mathrm{CP})$.

Recall that the utility function of UEs U_m over CPs can be calculated as $S_m^{\mathrm{UE}} = S_{i,m}^{\mathrm{gain}} - S^{\mathrm{cost}}$, where the exact amount of rent price $S_{i,m}^{\mathrm{gain}}$ is the received price from CP and S^{cost} is the maintenance costs of U_m. Observing from the equation, we can find that the utility of UEs is almost based on the price of rent by CPs. Besides, the price will be determined by sorting the utility function in a descent order. We can thus obtain the preference list of UE, denoted by $\mathcal{PR}(\mathrm{UE})$. The social welfare is the sum of utilities of all the CPs and the selected UEs in the network and is given by

$$\begin{aligned} S^{\mathrm{social}} &= \sum_{i=1}^{C} \sum_{m=1}^{Q_i} x_{i,m} \left(S_i^{\mathrm{CP}} + S_m^{\mathrm{UE}} \right) \\ &= \zeta \sum_{i=1}^{C} \sum_{m=1}^{M} x_{i,m} \left[\sum_{j=1}^{|\mathcal{N}_m|} F_{T_c} w_{m,j} - S^{\mathrm{cost}} \right]. \end{aligned} \tag{23}$$

The MOIS matching algorithm is displayed in Table I. We further briefly describe the matching algorithm during a specific iteration. The strategy of both CPs and UEs is to maximize the social welfare in the network. Each CP makes a price allocation number price $S_{i,m}^{\mathrm{rent}}$ to the UEs based on their importance and the number of consumers of each CP. We denote Ω_m as the number of CPs that request U_m. The social characteristics can be comprehensively evaluated by $b_{j,k}$, $s_{j,k}$, and the $c_{j,k}$. The initial price is determined before the matching game based on the social characteristics of UEs and the number of customer of CPs. The utility of CP is based on the social characteristics and the maintenance cost. In turn, the profit of IUs is the revenue from CPs. Both of two players have their preference to each other. The optimal optimization target is maximizing the social welfare in the entire network.

In this matching game, there are three cases. The first case is that some UEs receive request from multiple CPs. In this case, the UEs will choose the most preferred CP. The second case is that an UE only receives the request from one CP. In this case, the UE will be matched with the CP. The third case is that UEs do not receive any request from a CP. In this case, they do not change their price and wait for next round. At the end of the matching game, all the CPs have their matching IUs.

TABLE I
PROPOSED MOIS MATCHING ALGORITHM

Algorithm 1 :

Input: Q_i;
Output: Matching pairs;
Steps:
1: Each CP calculates comprehensive characteristic of UEs according to (17) and then get the importance about all UEs. CPs propose initial rent price $S_{i,m}^{\text{rent}}$;
2: S_i^{CP} is calculated based on the equation (22) and CPs determine their preference lists $\mathcal{PR}(CP)$;
3: CPs request the most preferred unmatched UE in their preference lists;
4: S_m^{UE} are calculated based on (8) and UEs formulate the preference lists, Ω_m represents the number of CPs that request U_m;
5: **if** $\Omega_m = 1$
6: The requested unmatched UE will be matched with the requester;
7: **end if**
8: **if** $\Omega_m = 0$
9: The UE will keep unmatched;
10: **end if**
11: **if** $\Omega_m > 1$
12: The unmatched UE will select the most preferred CP based on the preference lists. The CPs that are rejected by UE will choose the next preferred unmatched UE in their preference lists under the limitation of Q_i and return to **Step 4**;
13: **end if**
14: End the algorithm when all CPs choose Q_i UEs as IUs;
15: Obtain the optimal matching pairs;

TABLE II
PROPOSED MMFA MATCHING ALGORITHM

Algorithm 2 :

Input: $\mathcal{H}_n$, q_n, N_l;
Output: Matching pairs;
Steps:
1: **for** $l = 1$ to L **do**
2: $p_l = \frac{1/l^{\alpha}}{\sum_{i=1}^{L} 1/i^{\alpha}}$;
3: $P_{m,l}$;
4: **end for**
5: Each IU calculates the $V_n(l)$ based on the equation (24);
6: $V_n(l)$ are sorted in descending way;
7: Every IU creates the request matrix and selects the most preferred file according to $V_n(l)$;
8: The requested file calculates utility function and makes decision to choose the most preferred N_l IU based on the preference lists related to equation (25);
9: **if** quota$<N_l$ **do**
10: The most popular file is cached by the most preferred IU, and the selected IU is removed from file's preference lists, quota=quota+1 and return **Step 8**.
11: **end if**
12: **if** quota$==N_l$ **do**
13: IUs are rejected and select the second preferred file and repeat the **Step 8**;
14: **end if**
15: The game end when all IUs are matched with q_n files;
16: Obtain the optimal matching pairs.

C. MMFA Matching Algorithm

In this section, we propose an MMFA algorithm to solve the problem (14). The many-to-many matching game is further comprised of two types of game players including IUs and files regulated by the CPs, which means that based on the above definition, in the system model, limited by the storage capacity of mobile users, an IU can save multiple files or file sets. Similarly, one file can be stored at multiple IUs for D2D transmissions.

The strategy of both IUs and files controlled by the CPs is to maximize their respective profits in matching algorithm based on the preference over opposite sets. Moreover, a CP makes its file allocation decision based on its local information without relying on a central coordinator.

The preference relationship between files and IUs can be determined as the above definition. We denote $V_n(l)$ and $V_l(n)$ as the gains of I_n obtained from f_l and the gains of f_l obtained from I_n. Based on this consideration, the utility function over the file f_l is defined as

$$V_n(l) = \frac{1}{|\mathcal{H}_n|} \sum_{h \in \mathcal{H}_n} P_{h,l} \tag{24}$$

where $\mathcal{H}_n$ is defined same as $\mathcal{N}_m$ and denotes the number of UE connected with I_n. $P_{h,l}$ represents preference degree of the connected node h to the file f_l and the above equation illustrates that the IUs' preference over files is ranked based on the average degree of content popularity of its connected UEs. Furthermore, the favorite file can be obtained by sorting the utility function in a descending order. Similarly, the utility for the file $f_l \in \mathcal{F}$ to be matched with the nth IU can be written as

$$V_l(n) = \frac{1}{|\mathcal{H}_n|} \sum_{h \in \mathcal{H}_n} P_{h,l} T_{h,l}. \tag{25}$$

The utility function over D2D users is affected by the average transmission delay and social network structure. Besides, by sorting the utility function in a ascending order, we can obtain the preference list.

The many-to-many matching problem proposed in this paper is not a traditional matching game, since the preference lists of files and IUs depend not solely on the information available locally but on the character of social-layer architecture. Our proposed matching problem exhibits externality such as peer effects, which means that the users and files may change their preferences during the game, due to the constantly updated social relationship among users. But once the network structure is determined during a time, the preference lists will not change. Therefore, we develop a new algorithm to find a stable solution of this problem in this paper.

The many-to-many matching algorithm considers one CP, and is the solution to the problem in (14). The algorithm is displayed in Table II. The MMFA algorithm concentrates on obtaining the minimum download delay of UEs. In this algorithm, after the CP determines the selected IUs, they will focus on solving the problem of file allocation. The preference of files is related to the transmission delay of the IUs, while the preference of the IUs is related to the popularity of files. In the following, we describe the process of the algorithm briefly. The preferences are calculated by IUs and files, respectively. Then, IUs make proposals to the most preferred files, and in turn, the CP's files decide to accept or reject these proposals based on their preference lists. For a particular user, if it requests for the top of file f_l within the set $\mathcal{F}$, the file f_l updates its utility and accepts the request if the action do not yield a degradation of its utility.

V. Analysis of Proposed Algorithm

A. Stability of Many-to-One Matching

Let $\mathbb{A}(\mathcal{CP}, \mathcal{U})$ denotes the set of ultimate matching pairs, and $\eta(\mathrm{CP}_i, U_m)$ denotes the subset of $\mathbb{A}(\mathcal{CP}, \mathcal{U})$, where (CP_i, U_m) are matched. Thus, the concept of blocking pair and stability is introduced as follows. A matching η is blocked by a pair if: 1) it is not blocked by only CP or UE and 2) CP_i and U_m can both achieve a higher utility if they are matched comparing with their current matching. Condition 1) states that blocking pair must improve the utility of both matching pair. Condition 2) guarantees that the utility of blocking pair must higher than the current matching.

Definition 3: A matching $\eta^*(\mathrm{CP}_i, U_{m'}) \nsubseteq \mathbb{A}(\mathcal{CP}, \mathcal{U})$, but comparing with matching pair $\eta(\mathrm{CP}_i, U_m)$, it has a higher utility value if they match together. There exists relationship that $\eta^* \succ_{\mathrm{CP}_i} \eta$, that it to say, the current matching does not maximize the utility and both $(\mathrm{CP}_i, U_{m'})$ can achieve a higher utility. This implies that $V_i(m) > V_i(m')$ and we define this matching pair $\eta^*(\mathrm{CP}_i, U_{m'})$ as the blocking pair. If and only if there is no blocking pair, the proposed matching algorithm is stable.

In the many-to-one matching, every CP can match multiple UE. Thus, we need to consider the concept of group stability, which consists of multiple stable pairs. The matching game is stable if all the matching pairs are stable.

We prove the stability of the algorithm proposed in Table I. Here, we merely discuss the situation in a stable community, which means all the nodes may not readily add or remove any connections established between them during a period. This condition guarantees that the peer effects cannot make any change in community. Observing from our algorithm, the preference is strictly monotone and subjects to (9). In this case, the blocking pairs cannot exist because all the players select their matching pairs based on the preference. Moreover, the number of selection for CPs is finite and our matching pair selecting method always adheres to the utility maximum principle. Accordingly, under the situation of a stable community, our proposed algorithm is stable.

B. Stability of Many-to-Many Matching

In the many-to-many matching game, we take the concept of pairwise stability into account to analyze the stability [21]. Moreover, we denote the $S(f, \Psi(I_n))$ as the I_n's selection sets and f is the set of files matched with I_n. Similarly, $Z(I, \Psi(f_l))$ is the f_l's selection sets. The pairwise stability is defined as follows [33].

Definition 4: A matching outcome $\mathbb{A}'$ is pairwise stable if the players of the game are individually rational and there does not exist a matching of a pair (I_n', f_l') with $I_n' \nsubseteq I$ and $f_l' \nsubseteq f$ such that $\{S(f, \Psi(I_n)) \cup (f_l')\} \succ_{I_n} S(f, \Psi(I_n))$ and $\{Z(I_n, \Psi(f_l)) \cup (I_n')\} \succ_{f_l} Z(I, u)$.

In the second problem, the files and IUs are always interested in the utility they can achieve from the opposite set. For example, once the preference list of an I_n is determined rationally, it will not change its selection and always first cache the most popular files. Moreover, the preference list is monotonous and once if the IUs find the current selection does not get higher utility, it will change the selection observing from our proposed MMFA matching algorithm.

C. Convergence of the Two Matching Games

Regarding the convergence of the proposed two matching algorithm, we have the following results.

Theorem 1: The MOIS and MMFA games converge to the same matching obtained from optimization problem in (9) and (14), respectively.

Proof: We first analysis the convergence of the MOIS matching game. In the matching game, there exists contention among CPs when the UE receives requests from more than one CP. At the start of each contention, the price of UEs have initial value and would increase by the step in the process of contention. Finally, we can obtain the results of selection based on the maximum social welfare

$$\hat{X} = \arg\max_{X} \zeta \sum_{i=1}^{C} \sum_{m=1}^{M} x_{i,m} \left[\sum_{j=1}^{|\mathcal{N}_m|} F_{T_c} w_{m,j} - S^{\mathrm{cost}} \right]. \tag{26}$$

Observing from the above equation, S^{cost} is a constant. In the MOIS algorithm, the price of UEs would gradually increase by the step when the contention exists among CPs and the UE would be allocated to the CP that is the last one remaining in the request queue. That means if the price exceeds the CPs' expected value, they have to give up competition. Thus we conclude the contention must come to an end within a finite steps. Moreover, the MOIS algorithm finish when all the CPs are assigned IUs. Therefore, we can conclude the matching algorithm process within finite iterations.

The MMFA algorithm is a special case of MOIS because it is a matching game without monetary transfer between the two sets players. The matching game terminates when every IUs have been allocated their wanted files or have been rejected by every files to which they are willing and permitted to apply. ■

D. Optimality of the Two Matching Games

In the following, we first introduce the definition of weak Pareto based on [29].

Definition 5: For a multiobjective function, if a change allocation can improve the utility and other players can approve the change, it is a Pareto improvement. Moreover, if in the process of matching game, there exists no Pareto improvement, the allocation results a weak Pareto optimal.

Theorem 2: The MOIS and MMFA matching are weak Pareto optimal for IUs selection and content allocation problem in D2D transmission.

Proof: Observing two proposed matching algorithm, the pointer of one set players in its preference list moves when the next choosing can obtain a better utility and whatever the matching results of two matching game, the players cannot achieve a higher utility by moving back. Thus, for the two matching game, if there exists a higher utility than the current selection, they would reject the current selection. One case is that $U_{m'} \nsubseteq \Phi(U_m)$ for CP_i, but there has a improvement when $U_{m'}$ match with CP_i, thus, CP_i and $U_{m'}$ prefer to be matched and form a blocking pair. According to the stability we proved before, the case is contradictory. Moreover, the two matching converge to a stable matching based on the above. Thus, when the two matching come to an end, there exists no Pareto improvement. We can conclude that the MOIS and MMFA matching is weak Pareto optimal. ■

E. Implementation Issue

In the first proposed matching algorithm, each CP's selection of potential IUs from the UEs can be implemented via negotiations between this CP and the UEs. At the beginning, UEs broadcast their social information and physical information to the CPs. Upon receiving the UEs' information, the CPs calculate UEs' comprehensive characteristic according to (17) and then obtain the importance about all the UEs. Afterwards, CPs propose initial rent price, related to UEs' importance and the percentage of their customer in the total UEs, and then calculate the utility function based on (22). The CPs rank the UEs based on the calculation results and request their most preferred UE. In turn, when UEs receive the requests from CPs, there are two cases. One case is that the unmatched UE just has one CP requester, then it is matched with this CP. The other case is that when an UE has multiple requesters, the UE calculates the utility function according to S_m^{UE} and ranks the CPs requesters to select the most preferred CP. In this case, if CPs is rejected by the UE, they will request the second preferred UE in their preference lists. The game will terminate when all CPs obtain Q_i UEs as their IUs.

In the second proposed algorithm, after IUs are selected via the first proposed algorithm, the matching pairs between CPs and IUs are determined. We consider how to allocate the files to the selected IUs belonging to a CP. Each IU analyses files' popularity and calculates the utility function according to (24) to obtain the ranking of files. Then the IUs request the most preferred files via the files' ranking. Accordingly, there exists two cases. One case is that if the file receives a request, it will be matched with the requester within the limit N_l. The other case is that if the file receives requests from multiple IUs, it will calculate the utility function based on (25) and select the preferred IUs under the limit N_l. Moreover, the rejected IUs will request the next preferred files in their preference lists under the limit q_n. The algorithm will terminate when all IUs finish matching files.

VI. Simulation Results

In this section, we study a wireless network consisting of two CPs $\mathcal{CP} = \{\mathrm{CP}_1, \mathrm{CP}_2\}$. This cellular network is owned by one operator. Our simulations are based on the tool of MATLAB. The detailed configurations of the simulation environment are given as follows. The UEs are posited within an area of the 300×300 square meters with their locations being uniform distributed. Furthermore, the contact duration between two UEs obeys a exponential distribution with a index $\lambda = 4$. The the size of a file is set to 10^5 bits. Each edge in the social online relationship graph is generated by a given probability, i.e., an edge exists with a probability 0.6, and does not exist with a probability 0.4.

TABLE III
Parameter Descriptions

Description	Value
Transmission power of user equipment P_n	2W
Coefficient of channel fading α	4
Noise coefficient σ^2	10^{-10}
Number of user equipments M	20-60
Size of a file Y	10^5 bit
Bandwidth W	10^5HZ
Zipf parameter γ	0.9
Number of files L	10
Fixed coefficient with unit revenue ζ	1
Index of contact duration λ	4

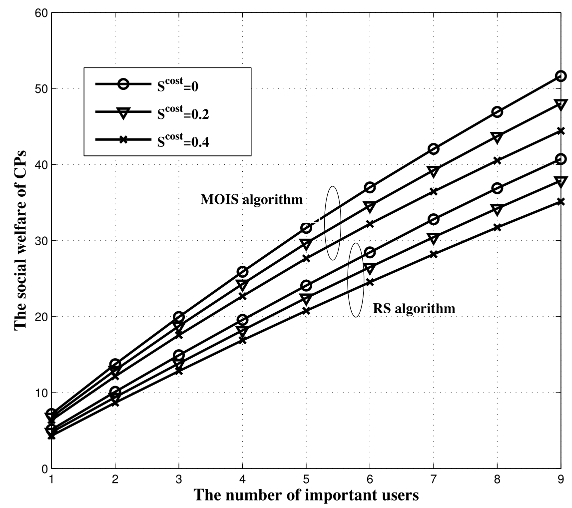

Fig. 3. Number of IUs versus the social welfare of CPs, with $M = 20$ and $\zeta = 1$.

In our system, there are two CPs who rent the IUs as D2D nodes. Besides, the tunable variables μ, υ, and ν are set to $1/3$. The simulation parameters are elaborated in Table III in detail.

In this simulation, we use random selection (RS) and random allocation (RA) as the benchmark to compare with the two proposed mechanisms, i.e., the proposed MOIS matching algorithm and the MMFA matching algorithm. In the RS algorithm, we assign IUs randomly to CPs while not exceeding the quota. In the RA algorithm, we allocate files randomly to the IUs for caching while satisfying the capacity constraints. Moreover, we compare our proposed MMFA matching algorithm with the many-to-one (MOFA) matching algorithm in which each content is cached only once in order to efficiently utilize the storage while the caching nodes can storage multiple contents [30], then we research the difference between the two mechanisms. In addition, the exhaustive searching algorithm is taken into account for comparison with the proposed algorithms. In the exhaustive searching algorithm, the allocation between files and IUs is via exhaustive searching to solving the optimization problem while the exhaustive searching algorithm requires high computational complexity and can be viewed as a centralized solution [28].

In our proposed MOIS algorithm, there are two CPs to rent IUs and each IU can be chosen by one CP. The number of IUs that one CP can choose increase within 1–9. Moreover,

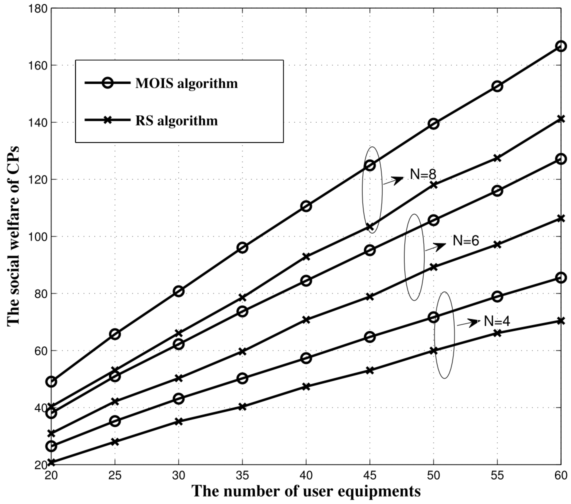

Fig. 4. Number of UEs versus the social welfare of CPs, with different number of IU and $\zeta = 1$.

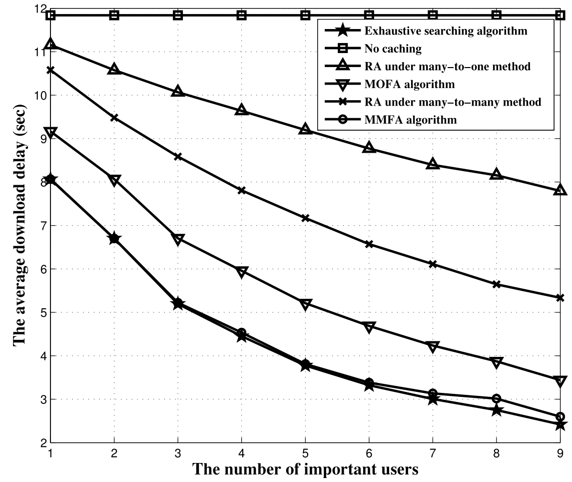

Fig. 5. Number of user IUs versus the average download delay, with the total number of UEs 20.

in the proposed MMFA algorithm, we assume $q_n = 2$ which each IU can cache two files and $N_l = 2$.

Fig. 3 shows the social welfare under RS and MOIS mechanism by increasing the number of IUs in the social network. In Fig. 3, we fix the number of total UEs to be 20 and vary the number of IUs in the system. We can observe the proposed MOIS algorithm achieve a significant proportion of the social welfare compared with the RS algorithm. It is obviously that as the increase of the number of IUs, the social welfare presents a rising trend. Moreover, we also observe that with the increasing of the repair cost of IUs, the value of social welfare will decrease. For example, when $S^{\text{cost}} = 0.4$ and $N = 6$, the MOIS achieves 131% of the social welfare obtained by RS algorithm.

Fig. 4 illustrates the network social welfare versus the total number of the UEs which are randomly distributed in the network. In Fig. 4, we set the number of IUs to be 4, 6, and 8 and vary the total number of UEs to observe the social welfare. It is important to note that our proposed MOIS algorithm achieves a better performance than the RS algorithm. For example, when the $N = 8$ and the number of UEs is 45, the MOIS algorithm achieves 119% of the social welfare of the RS algorithm. Similar to the results in Fig. 3, the social welfare has a increasing trend with the increase of the total number of UEs. Further more, with the system network scaling, the proposed MOIS algorithm can be a good choice for improving the system social welfare compared with the RS algorithm. Additionally, we note that the more the number of IUs, the higher benefits the social welfare can get.

Figs. 3 and 4 clearly illustrate the advantage of the proposed MOIS algorithm to other benchmarks. It can be observed that our proposed algorithm can achieve a better performance in different networks. Moreover, comparing with the benchmarks, the social welfare in our algorithm has a great improvement.

In Fig. 5, we show the average download delay in the network resulting from the proposed MMFA algorithm and MOFA algorithm, for comparison, we investigated the RA algorithm under many-to-many matching and many-to-one matching scheme. Similar to Fig. 3, the total number of UE is set to be 20. Clearly, from this figure, we can observe that no matter what kind method of matching, the algorithms based on matching can achieve lower download delay than the RA algorithm and no caching method obtain the worst performance because that the UEs only request needed files from cellular network, which lead to a higher download delay. Moreover, It is obviously to see that our proposed MMFA algorithm can obtain the best delay performance in the network compared with MOFA algorithm and RA algorithm. For example, when the number of IUs is 6, the delay performance is 31% lower than the MOFA algorithm and 38% lower than the RA algorithm based many-to-many matching. We can see from Fig. 5 that the complexity of MMFA matching game is higher than the RA algorithm and MOFA algorithm, while it expresses a better performance than the other two benchmark algorithms. Furthermore, it can be also noticed that as the number of IUs increase, the average download delay that UEs request files decrease.

Also, it can be observed in Fig. 5 that our proposed algorithm performs very close to the exhaustive searching algorithm especially when the number of IUs is small and when the number of IUs is large, the performance is also significantly better than other benchmarks. It is reasonable that the proposed algorithm performs slightly worse than exhaustive searching algorithm because the computation complexity of MMFA algorithm is less than the exhaustive searching algorithm. The computation complexity of MMFA algorithm is $\mathcal{O}(N(N-N_l)(L-1))$, where N is the total number of IUs that every round request the files and $(N-N_l)$ is the number of IUs that are rejected each round, and each IU will request for maximum (L-1) files in the worst case. The computation of the exhaustive searching algorithm increases exponentially over the network size. As a result, with the number of IUs scaling, the proposed MMFA algorithm can be a good choice for reducing computation complexity.

Fig. 6 shows the number of UEs versus the average download delay, we can see that the increase in the number of UEs

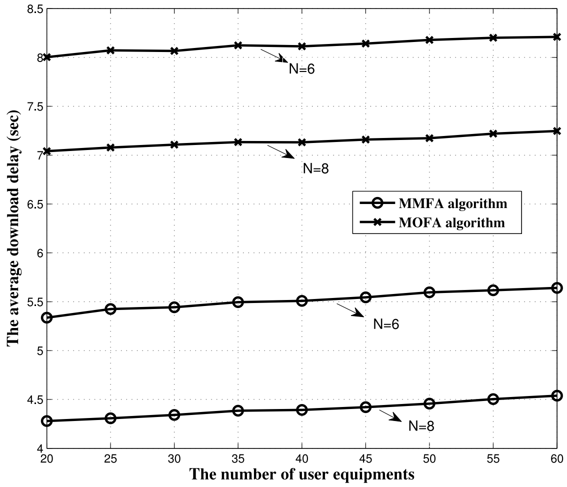

Fig. 6. Number of UEs versus the average download delay, with the number of IUs $N = 6$ and 8.

result in an increase in the download delay with the expansion of network scale while the number of IUs increase under both algorithm, and as the number of IUs increase, the transmission delay decrease. Fig. 6 shows that with the reduced portion of IUs in the total network, the download demands increase and the average download delay shows a rising trend. Moreover, the proposed MMFA algorithm can obtain a better performance.

Fig. 7 depicts the average download delay versus the number of IUs. We compare the proposed selection method with the RS method in two matching system (MMFA algorithm and MOFA algorithm). This figure clearly indicates the significant changes in the average download delay depending on the number of IUs in the network. Note that no matter what kind method of matching system (MMFA algorithm and MOFA algorithm), the proposed IUs selection method achieves a better performance compared with RS method. For example, when the number of IUs is 6, the result delay using proposed selection algorithm in MMFA system is 20% lower than RS method in MMFA system. It also can be obvious that MMFA algorithm can obtain significant improvement than MOFA algorithm because of the storage capacity. In addition, the proposed MOIS algorithm combining MMFA algorithm in Fig. 7 shows a better performance and is only worse than the exhaustive searching algorithm.

Fig. 8 shows the average download delay versus Zipf parameter γ. With the increase of γ, the popularity of top ranked files increase and the average download delay shows an increase trend. We can see from the figure that our proposed algorithm always shows a better performance. To be specific, when γ is small, say, $\gamma = 0.1$, the difference of the popularity among various files are not obvious. Correspondingly, the advantage of our MMFA algorithm compared with the RA algorithm is trivial. However, with the growth of γ, our MMFA algorithm shows a rapidly increasing superiority to the benchmark algorithm. We can also observe in Fig. 8 that the average download delay of proposed algorithms is just higher than the exhaustive searching algorithm while the disparity is small with the increase of zipf parameter distribution.

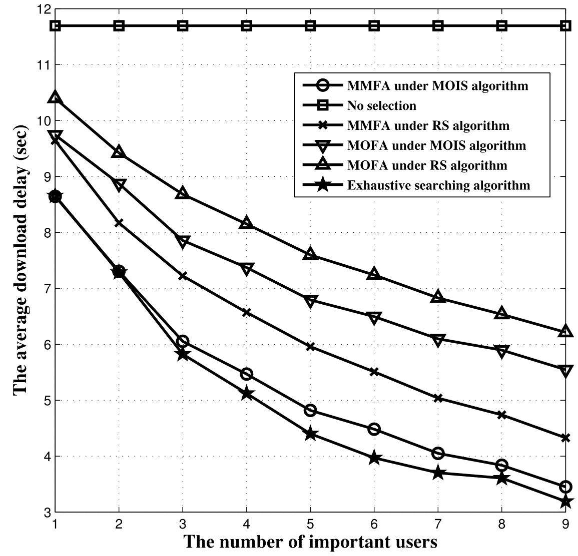

Fig. 7. Number of IUs versus the average download delay, with the total number of UE is 20.

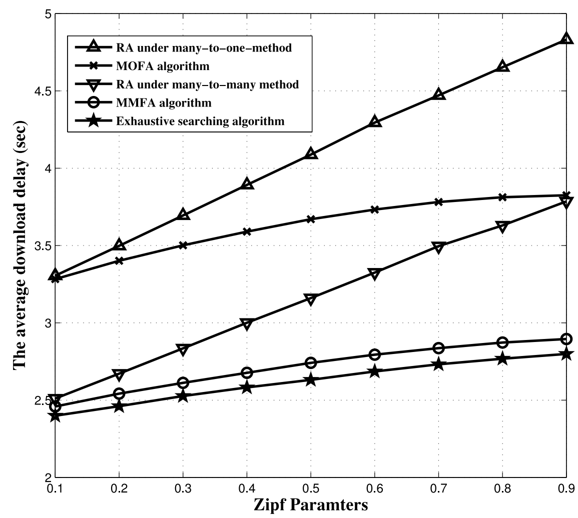

Fig. 8. Average download delay versus Zipf parameter γ.

VII. Conclusion

In this paper, we design a novel IUs selection algorithm to select IUs and design a distributed caching optimization algorithm to solve the cache allocation problem in D2D-enabled underlaid cellular networks with social awareness. We formulate an MOIS matching game combining the social relationship with physical constraint in order to maximizing the social welfare in the network and design an MMFA matching algorithm to minimizing the average transmission delay. Also, we prove the stability, the convergence and the optimality of the two proposed algorithms. The simulation results are provided to demonstrate the validity of the two algorithm that considering the social importance can greatly reduce the transmission delay and the proposed IUs selection algorithm can obtain a significant improvement in social welfare. In our future work, we will study the joint optimization problem in the system, where the selection problem and files allocation problem are jointly considered. Also, the scalable video technology will be considered for satisfying different user community with various request demands.

REFERENCES

[1] *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2013-2018*, Cisco, San Jose, CA, USA, 2012. [Online]. Available: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.pdf

[2] H. Claussen, L. T. W. Ho, and L. G. Samuel, "An overview of the femtocell concept," *Bell Labs Tech. J.*, vol. 13, no. 1, pp. 221–245, Mar. 2008.

[3] K. Doppler, M. Rinne, C. Wijting, C. B. Ribeiro, and K. Hugl, "Device-to-device communication as an underlay to LTE-advanced networks," *IEEE Commun. Mag.*, vol. 47, no. 12, pp. 42–49, Dec. 2009.

[4] S. Hakola, T. Chen, J. Lehtomaki, and T. Koskela, "Device-to-device (D2D) communication in cellular network—Performance analysis of optimum and practical communication mode selection," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Sydney, NSW, Australia, 2010, pp. 1–6.

[5] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[6] C.-X. Wang *et al.*, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 122–130, Feb. 2014.

[7] H. Zhang, Y. Liao, and L. Song, "D2D-u: Device-to-device communications in unlicensed bands for 5G system," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3507–3519, Jun. 2017.

[8] P. K. Mishra, S. Pandey, and S. K. Biswash, "Efficient resource management by exploiting D2D communication for 5G networks," *IEEE Access*, vol. 4, pp. 9910–9922, 2016.

[9] D. Verenzuela and G. Miao, "Scalable D2D communications for frequency reuse >> 1 in 5G," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3435–3447, Jun. 2017.

[10] S. Huang, B. Liang, and J. Li, "Distributed interference and delay aware design for D2D communication in cellular networks," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Washington, DC, USA, 2016, pp. 1–7.

[11] L. Wang, H. Tang, H. Wu, and G. L. Stüber, "Resource allocation for D2D communications underlay in Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 1159–1170, Feb. 2017.

[12] J. Li *et al.*, "Distributed caching for data dissemination in the downlink of heterogeneous networks," *IEEE Trans. Commun.*, vol. 63, no. 10, pp. 3553–3568, Oct. 2015.

[13] K. Shanmugam, N. Golrezaei, A. G. Dimakis, A. F. Molisch, and G. Caire, "FemtoCaching: Wireless content delivery through distributed caching helpers," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 8402–8413, Dec. 2013.

[14] Y. Li, T. Wu, P. Hui, D. Jin, and S. Chen, "Social-aware D2D communications: Qualitative insights and quantitative analysis," *IEEE Commun. Mag.*, vol. 52, no. 6, pp. 150–158, Jun. 2014.

[15] J. Hu, L.-L. Yang, H. V. Poor, and L. Hanzo, "Bridging the social and wireless networking divide: Information dissemination in integrated cellular and opportunistic networks," *IEEE Access*, vol. 3, pp. 1809–1848, 2015.

[16] L. Wang, H. Wu, W. Wang, and K.-C. Chen, "Socially enabled wireless networks: Resource allocation via bipartite graph matching," *IEEE Commun. Mag.*, vol. 53, no. 10, pp. 128–135, Oct. 2015.

[17] K. Zhu, W. Zhi, L. Zhang, X. Chen, and X. Fu, "Social-aware incentivized caching for D2D communications," *IEEE Access*, vol. 4, pp. 7585–7593, 2016.

[18] M. I. Ashraf, M. Bennis, W. Saad, and M. Katz, "Exploring social networks for optimized user association in wireless small cell networks with device-to-device communications," in *Proc. IEEE Wireless Commun. Netw. Conf. Workshops (WCNCW)*, Istanbul, Turkey, Apr. 2014, pp. 224–229.

[19] B. Bai, L. Wang, Z. Han, W. Chen, and T. Svensson, "Caching based socially-aware D2D communications in wireless content delivery networks: A hypergraph framework," *IEEE Wireless Commun.*, vol. 23, no. 4, pp. 74–81, Aug. 2016.

[20] Y. Gu, W. Saad, M. Bennis, M. Debbah, and Z. Han, "Matching theory for future wireless networks: Fundamentals and applications," *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 52–59, May 2015.

[21] S. Bayat, Y. Li, L. Song, and Z. Han, "Matching theory: Applications in wireless communications," *IEEE Signal Process. Mag.*, vol. 33, no. 6, pp. 103–122, Nov. 2016.

[22] D. Wu, Q. Wu, Y. Xu, J. Jing, and Z. Qin, "QoE-based distributed multichannel allocation in 5G heterogeneous cellular networks: A matching-coalitional game solution," *IEEE Access*, vol. 5, pp. 61–71, 2017.

[23] T. LeAnh *et al.*, "Matching theory for distributed user association and resource allocation in cognitive femtocell networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 9, pp. 8413–8428, Sep. 2017.

[24] K. Hamidouche, W. Saad, and M. Debbah, "Many-to-many matching games for proactive social-caching in wireless small cell networks," in *Proc. 12th Int. Symp. Model. Optim. Mobile Ad Hoc Wireless Netw. (WiOpt)*, May 2014, pp. 569–574.

[25] Y. Gu, Y. Zhang, M. Pan, and Z. Han, "Student admission matching based content-cache allocation," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, New Orleans, LA, USA, Mar. 2015, pp. 2179–2184.

[26] D. Gale and L. S. Shapley, "College admissions and the stability of marriage," *Amer. Math. Monthly*, vol. 69, no. 1, pp. 9–15, Jan. 1962.

[27] O. Semiari, W. Saad, S. Valentin, M. Bennis, and H. V. Poor, "Context-aware small cell networks: How social metrics improve wireless resource allocation," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 5927–5940, Nov. 2015.

[28] S. Bayat, R. H. Y. Louie, Z. Han, B. Vucetic, and Y. Li, "Distributed user association and femtocell allocation in heterogeneous wireless networks," *IEEE Trans. Commun.*, vol. 62, no. 8, pp. 3027–3043, Aug. 2014.

[29] C. Xu, C. Gao, Z. Zhou, Z. Chang, and Y. Jia, "Social network-based content delivery in device-to-device underlay cellular networks using matching theory," *IEEE Access*, vol. 5, pp. 924–937, 2017.

[30] Z. Chang, Y. Gu, Z. Han, X. Chen, and T. Ristaniemi, "Context-aware data caching for 5G heterogeneous small cells networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2016, pp. 1–6.

[31] U. Brandes, "On variants of shortest-path betweenness centrality and their generic computation," *Soc. Netw.*, vol. 30, no. 2, pp. 136–145, 2008.

[32] C. Ma, Z. Lin, L. Marini, J. Li, and B. Vucetic, "Learning automaton based distributed caching for mobile social networks," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Doha, Qatar, 2016, pp. 1–6.

[33] A. E. Roth, "Stability and polarization of interests in job matching," *Econometrica*, vol. 52, no. 1, pp. 47–57, 1984.

Jun Li (M'09–SM'16) received the Ph.D. degree in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 2009.

In 2009, he was a Research Scientist with the Department of Research and Innovation, Alcatel Lucent Shanghai Bell, Shanghai, for six months. From 2009 to 2012, he was a Post-Doctoral Fellow with the School of Electrical Engineering and Telecommunications, University of New South Wales, Sydney, NSW, Australia. From 2012 to 2015, he was a Research Fellow with the School of Electrical Engineering, University of Sydney, Sydney. Since 2015, he has been a Professor with the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing, China. His current research interests include network information theory, channel coding theory, wireless network coding, and cooperative communications.

Miao Liu received the bachelor's degree from the School of Engineering, Nanjing Agricultural University, Nanjing, China, in 2015. She is currently pursuing the postgraduate degree at the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing.

Her current research interests include social network, matching theory, and D2D communications.

Jinhui Lu received the M.S. degree in communication and information system from the Nanjing University of Science and Technology, Nanjing, China, in 1985.

He is currently a Professor with the School of Electronic and Optical Engineering, Nanjing University of Science and Technology. His current research interests including millimeter-wave monitoring systems, echo simulation of radar, and passive location technology.

Mr. Lu is a Senior Member of the Chinese Institute of Electronics and the Director of the China Education Society of Electronics.

Feng Shu (M'09) was born in 1973. He received the B.S. degree from Fuyang Teaching College, Fuyang, China, in 1994, the M.S. degree from Xidian University, Xi'an, China, in 1997, and the Ph.D. degree from Southeast University, Nanjing, China, in 2002.

From 2009 to 2010, he was a Visiting Post-Doctoral Research Fellow with the University of Texas at Dallas, Richardson, TX, USA. In 2005, he joined the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing, China, where he is currently a Professor and has supervised Ph.D. and graduate students. He is also with Fujian Agriculture and Forestry University, Fuzhou, China. He has authored or co-authored about 200 papers, of which over 100 are in archival journals, including over 25 papers in IEEE journals and over 50 SCI-indexed papers. He holds four Chinese patents. His current research interests include wireless networks, wireless location, and array signal processing.

Dr. Shu was a recipient of the Mingjian Scholar Chair Professor in Fujian Province. He is currently an Associate Editor for IEEE ACCESS.

Yijin Zhang (M'14) received the B.S. degree in information engineering from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 2004, the M.S. degree in information engineering from Southeast University, Nanjing, in 2007, and the Ph.D. degree in information engineering from the Chinese University of Hong Kong, Hong Kong, in 2010.

Since 2011, he has been an Associate Professor with the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing. His current research interests include sequence design and resource allocation in communication networks.

Siavash Bayat (S'10–M'13) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the University of Sydney, Sydney, NSW, Australia, in 2002, 2005, and 2013, respectively.

He was then a Post-Doctoral Research Fellow with the University of Sydney, prior to joining the Sharif University of Technology, Tehran, Iran, in 2014, where he is currently an Assistant Professor. From 2007 to 2009, he was affiliated with the Sharif University of Technology, where he was a faculty member with responsibility for research in wireless communication networks system design. His current research interests include wireless resource management, wireless communications, signal processing, heterogeneous networks, cognitive radio, game theory, and physical layer security.

Dushantha Nalin K. Jayakody (M'14) received the B.Eng. (First Class Hons.) degree from the Dawood University of Engineering and Technology, Karachi, Pakistan, in 2009, the M.Sc. degree in electronics and communications engineering from the Department of Electrical and Electronics Engineering, Eastern Mediterranean University, Famagusta, Northern Cyprus, in 2010, and the Ph.D. degree in electronics and communications engineering from University College Dublin, Dublin, Ireland, in 2014.

From 2014 to 2016, he was a Post-Doctoral Research Fellow with the Coding and Information Transmission Group, University of Tartu, Tartu, Estonia, and the University of Bergen, Bergen, Norway. Since 2016, he has been a Professor with the Department of Control System Optimization, Institute of Cybernetics, National Research Tomsk Polytechnic University, Tomsk, Russia.

Dr. Jayakody was a recipient of the Merit Position Holder of the University in Pakistan (under the SAARC Scholarship), the University Full Graduate Scholarship and the First Merit Position Holder of the Department of Electronics and Electrical Engineering from Eastern Mediterranean University. He has served as the Session Chair or a Technical Program Committee member for various international conferences such as IEEE PIMRC 2013/2014, IEEE WCNC 2014/2016, and IEEE VTC 2015.