

Research Article

Clustering Optimization for Out-of-Band D2D Communications

**A. Paramonov,¹ O. Hussain,¹ K. Samouylov,² A. Koucheryavy,¹
R. Kirichek,¹ and Y. Koucheryavy³**

¹The Bonch-Bruевич State University of Telecommunication, 22 Prospekt Bolshhevikov, St. Petersburg, Russia

²Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya St., Moscow, Russia

³Tampere University of Technology (TUT), Korkeakoulunkatu 10, Tampere, Finland

Correspondence should be addressed to R. Kirichek; kirichek@sut.ru

Received 24 April 2017; Revised 12 August 2017; Accepted 12 September 2017; Published 22 October 2017

Academic Editor: Pai-Yen Chen

Copyright © 2017 A. Paramonov et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Significant increase in multimedia traffic challenges 5G networks in terms of capacity and correspondent QoS parameters. Device-to-device communication paradigm has already become an integral part of 3GPP standards; nevertheless it has not yet been widely deployed due to many different reasons. D2D is expected to leverage implementation of many qualitatively new services and to efficiently accomplish it D2D devices are supposed to form clusters. Due to practical limitations, current D2D implementations are mostly out-of-band and use Wi-Fi Direct. In this paper, we propose a novel model for throughput optimization in out-of-band D2D clusters. We delivered numerical results for different typical cluster member distributions and revealed key functional dependencies. Further, for the first time we compare clustering algorithms for out-of-band D2D and identify effective clustering algorithm that increases network resource utilization rate.

1. Introduction and Rationale

Integration of device-to-device (D2D) communication technology became a mainstream direction for fifth-generation (5G) communication networks. Driven by a huge increase in demand of multimedia traffic transfer, D2D communication allows saving scarce network resources by transferring data directly between devices either in-band or out-of-band, and D2D communications allow significantly reducing traffic between base station (BS) and end-user device [1].

According to the 3rd-Generation Partnership Project (3GPP) [2], D2D is a flexible paradigm of direct communication between devices which is open for use and based on cellular communication technologies (in-band D2D communication) and also WLAN technologies which are IEEE 802.11 standardization (out-of-band D2D communication) [3].

The last approach has recently become attractive due to the ease of implementation compared to in-band D2D, where end-user device shall be equipped with appropriate uplink/downlink functionality. In case of in-band D2D communications, the transmission power should be properly regulated so that the D2D transmitter does not interfere

with the cellular UE communication while maintaining the minimum SINR requirement of the D2D receiver [4]. This significantly complicates feasibility of in-band D2D wide-scale implementation at least for the time being. Out-of-band D2D can be easily implemented with network assistance option; hence cellular operators are able to control out-of-band sessions. For the obvious reasons, IEEE 802.11-based Wi-Fi is taken as the transmission technology for implementation of out-of-band D2D functionality [5]. Operators of communication networks can encourage regular users to use D2D technology in order to improve the overall performance of the communication system in return for rewards proportional to their contributions.

Typically, geographically beside D2D nodes can form a cluster (Figure 1), where traffic circulates between cluster nodes directly, and outside-of-cluster traffic is forwarded to BS via relay node, so-called cluster head. A number of algorithms for cluster head selection are available today, e.g., [6, 7]. The decision on selection of a particular cluster member as a cluster head affects at least network efficiency, energy expenditures, and quality of service (QoS) offered to all members within the cluster. Generally, if all data transfer

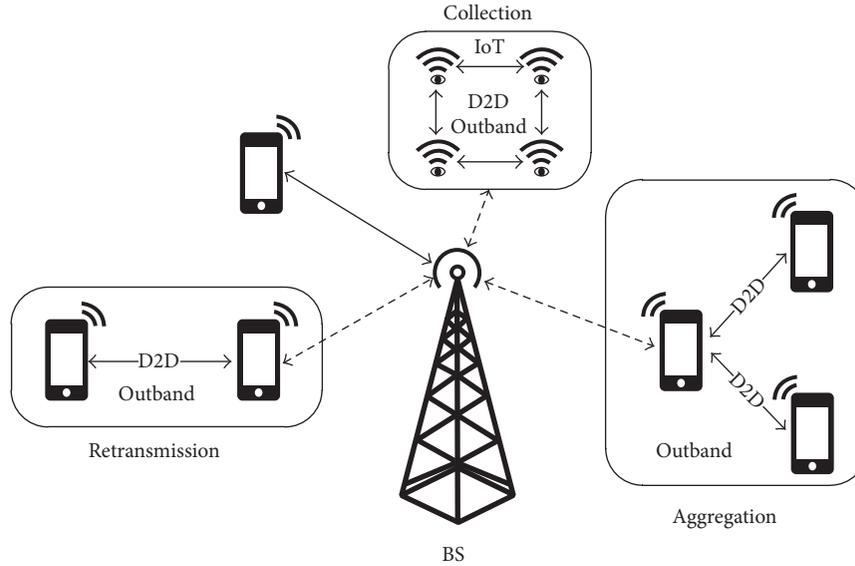


FIGURE 1: Out-of-band clustering of D2D components.

members are the members of the same cluster, the cluster can operate off-line, meaning without connection to network BS.

A larger number of cluster members are expected to lead to larger savings of the network resources. However, the maximum number of members in a cluster is restricted by coverage of selected D2D technology, channel throughput of cluster head and cluster traffic intensity, and cluster members physical location towards cluster head. Existing studies show that D2D clustering in 5G leads to reduction of signaling traffic and provides higher spectral efficiency and better energy performance than conventional cellular systems [7, 8]. Thus, efficient D2D clustering in 5G networks especially with high density of devices is of a paramount importance.

Multiple past works have concentrated on quantitative and qualitative analysis of cluster algorithms for D2D communications. In [1, 3] the authors provide comprehensive analysis of D2D communications. The use of out-of-band D2D communications and D2D clustering is discussed in detail given criteria of cluster head selection based on channel quality between cluster head and BS. In [9], the authors designed clustering algorithm for in-band D2D case, which increases system-level spectral efficiency. Numerical analysis and simulation modeling have shown that this proposal gives 66% gain in terms of throughput compared to traditional solutions, in the case where 20% of users use D2D communication. The authors derived the probability density formula (pdf) for the optimal number of repeater units in the cluster and have come up with the cross-cluster interaction scheme. Also via simulation the authors show that the proposed algorithm provides gains up to 40% in terms of network efficiency of resource use.

Different aspects of the out-of-band D2D communication are presented in [4, 10, 11]. In [4, 10], the authors developed analytical model of the network unloading for different D2D scenarios using stochastic geometry. The authors estimated

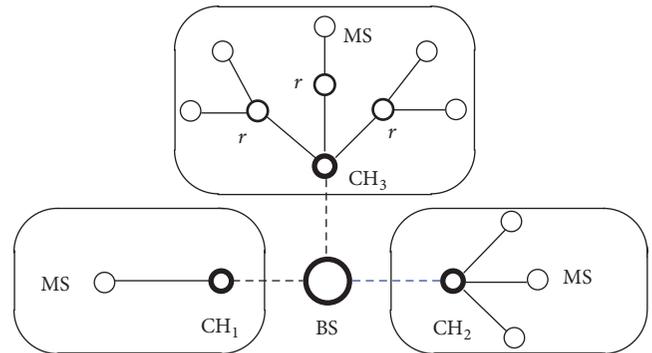


FIGURE 2: Possible cluster structures.

potential opportunities of the out-of-band D2D communication, using both the system level and the mathematical analysis. They show that at 30% of clustering productivity and energy network performance increase up to four and two times, respectively. In [11] the authors studied problems of implementation of network-assisted D2D communication while interacting in social networks. Besides, they use the existing experimental LTE testbed [12] for implementation of D2D system and show its performance evaluation in terms of latency and users satisfaction.

D2D transmission technology selection is still rather limited to Wi-Fi and Bluetooth due to wide implementation of those in consumer devices. Most recent of works, e.g., [13, 14], consider D2D devices forming clusters by using Wi-Fi Direct (see Figure 2). Due to features of radio channel, resources of channel between cluster member and cluster head may drastically vary for different nodes within one cluster. Therefore, while forming cluster, we suggest the way cluster head is selected shall be based primarily on anticipated

QoS parameters, not distance as many studies suggested up to date [15, 16]. The same applies to selection of cluster members. Clustering algorithm can be implemented for different target parameters such as cumulative throughput of cluster as a whole, maximum number of cluster nodes, and quality of service.

Summarizing the results of the analysis of publications, we can conclude that most of them are devoted to analysis and evidence of the effectiveness of using both in-band and out-of-band clustering. The efficiency of spectrum utilization is not the only criterion that should be considered when solving the clustering problem. Therefore, further in this paper we design clustering algorithm, which is characterized by bandwidth of the channels between cluster members and the cluster head.

The remainder of this paper is organized as follows. In Section 2, the reference scenario for D2D is defined and analytical model for cluster is introduced. The proposed model is delivered in Section 3. The required analytical and simulation performance evaluation campaign are reported in Section 4. Section 5 discusses clustering algorithms, whereas the concluding remarks and future research directions appear in Section 6.

2. Reference Scenario Definition

The use of D2D communication allows the increase in system effectiveness of cellular communication; moreover D2D directly influences at system level both efficiency and energy. The users are distributed on the BS coverage area randomly. Generally, network planning takes into account distribution of nodes in the geographical area letting operator provide at least wanted coverage and required throughput and QoS. The possible cluster structures are presented in Figure 2.

We assume that mobile *stations* (MS) can interact directly with base station (BS), through transit node (CH₁) or head node of a cluster (CH₂, CH₃). Generally, one cluster can have a star-like structure with single transit node, head node of cluster, or tree-like structure, including both head node and other transit nodes (r). The choice of the cluster structure can be done without the involvement of network functionality (for mobile stations) and involving the network functionality (for the BS). Shaping of the cluster consists of *the MS group* choice and distribution of their functionality within the cluster (terminal node, transit node, or head node of the cluster). To shape a cluster one needs to define the indicators that characterize the decision for cluster shaping (status indicators) and the criterion validity (quality) of decision and control settings that affect performance status and also method of finding the valid (optimal) solution.

The authors in [3, 4] suggested approach for the analysis and shaping of tree-like structure of clusters. Following the proposal, in this paper we focus attention on clusters with star-like structure with one head node; such structure is very useful for high density wireless environments such as apartment block houses, offices, and stadiums.

Quality of traffic service within a cluster and between cluster head node and BS depends on channels throughput

between cluster participants (b_{ij}) and between head node and BS (b_k) and also on traffic intensity (a_i) produced by users. We suppose that in BS service area there are n of MS that support D2D mode. We indicate the set of MS as $M = \{m_1, m_2, \dots, m_n\}$. Then the task of clustering consists of estimating a quantity of clusters (k) and choosing their structure when the best possible QoS is provided. We assume the efficiency of the solution is higher, if in a service area of BS there are a smaller number of channels (BS-CH) (i.e., quantity of clusters (k)) and a greater number of MS which are using D2D mode. At the same time QoS for the participants of clusters should not be below target value (the rule) b_{ij}^0 . Generally, target values can be different for different users. They depend on characteristics of the produced traffic, that is, the type of services required.

We indicate the set of clusters in BS service area as $C = \{c_1, c_2, \dots, c_k\}$. We assume that all clusters owned by multitude C are formed by elements of multitude M , and not all elements of M should be included in clusters, and the clusters have no general elements (they form disjoint subsets) $C \subseteq M, \forall c \in C$ and $\cup C = \{\}$.

As it was said above, cluster shaping can be performed by various methods, the choice of which depends on desired outcome. Further we consider the possibility of use of certain centroid methods for the task of clustering objects [17, 18]. The solution of the clustering problem represents the solution to the optimization problem where certain metric $d(m, p_m)$ is minimized or maximized. This metric characterizes the "distance" between cluster participant and cluster center $p_m = (1/|c|) \sum_{m \in c} m$.

Distance, throughput, time delay, and so forth can be used as optimization metric. Generally, it is the task of nonconvex optimization which may not have a unique solution. As a rule, to solve this problem dynamic programming is required, which can minimize the parameter $d^2(m, p_m)$ for all clusters.

$$C = \min \sum_{c \in C} \sum_{m \in c} d^2(m, p_m). \quad (1)$$

Well-known clustering algorithms allow us to find the particular (near-optimal) solution. As it was said above, clustering can be chosen in different scenarios, depending on goals and restrictions. Therefore, to solve the considered clustering task the analysis of possible solutions and approach to selecting of criteria and clustering method is required.

3. The Proposed Model Description

Quality of traffic service is instantiated by probability and time indexes as the probability of availability, discard probability, and data delivery time. These indexes depend on the traffic parameters and bandwidth of the network connection. Thus, for given traffic characteristics, the amount of bandwidth best describes the results of decision that is made in terms of quality of communication services. Under the throughput we assume the achievable data rate. The throughput is not a complete metric for the quality of service description. The quality of service depends on the traffic characteristics. To analyze such quality characteristics as

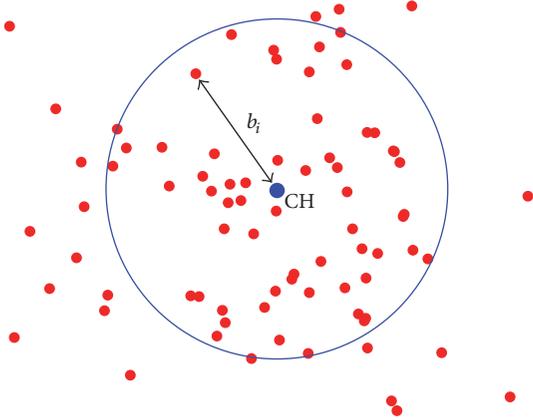


FIGURE 3: Cluster model.

packet loss ratio and delay it is necessary to build a complete model based on queuing theory. At this stage we concentrate on the initial model where known parameters such as number of users and distribution of users in the service area are available. We suppose that the initial clustering solution shall be done taking into account only throughput parameter in assumption that all users generate equal traffic flows. In fact, analytical results given below can be used for different types of traffic flows. As a target metric we consider throughput between network elements b_{ij} . In our analysis we consider the case when head node is already defined. We assume the communication area of head node represents as a circle with R radius, centered at the location of the head CH node, as it is shown on Figure 3.

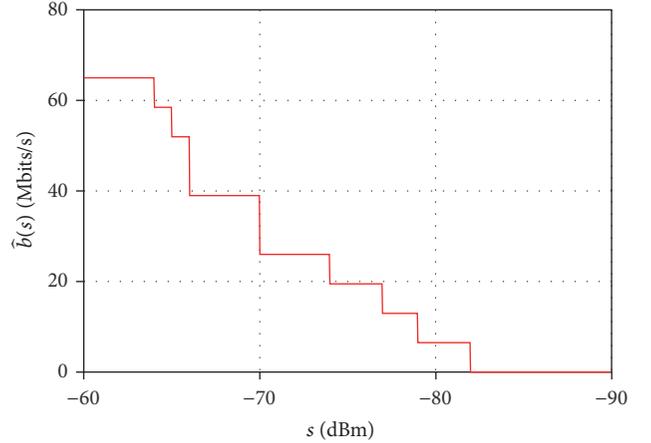
Considering the IEEE 802.11 family standards as communication technology between D2D nodes, one shall define the type of dependence b_{ij} . According to [12], data transmission rate between two cluster members is defined by making the choice of modulation and coding scheme (MCS) according to receiving conditions (i.e., radio channel quality). These conditions are evaluated by signal power on receiver input $s = p_{rx}$ or the signal-to-noise ratio (SNR) or signal/noise + noise (SINR), where, in addition to the useful signal, the noise power is total power of all received signals, including noise. The noise power is the power created by nodes from neighboring clusters. According to [12] the model dependence of data transmission rate on signal power at the receiver input s is a jump function which increases with power growth. Figure 4 shows an example of dependence between data transfer rate and magnitude s for the IEEE 802.11n standard using the 20 MHz channel width.

Similar dependence can also be constructed for values of SNR and SINR. Signal power, SNR, and SINR are estimated by the equation given below.

$$s = p_{rx} = 10 \lg \left(\frac{P_{rx}}{1 \text{ mW}} \right) \text{ dBm},$$

$$\text{SNR} = 10 \lg \frac{P_{rx}}{P_N} = p_{rx} - 10 \lg \left(\frac{P_N}{1 \text{ mW}} \right) \text{ dB}, \quad (2)$$

$$\text{SINR} = 10 \lg \frac{P_{rx}}{P_N + P_I} = p_{rx} - 10 \lg \left(\frac{P_N + P_I}{1 \text{ mW}} \right) \text{ dB},$$

FIGURE 4: The data transfer rate dependence from the magnitude s .

where p_{rx} is the receiver input power (dBm); P_{rx} is the receiver input power (W); P_N is the input noise intensity (W); P_I is the receiver input interference power (W). The receiver input signal power can be described by the RCPI power indicator of receiving channel. According to [19] the value of this indicator is measured with an accuracy of ± 5 dB (95% of confidential interval) taking into account the band noise, corresponding to the channel strip.

Along with the specified metrics, signal power on an input of the receiver can be described by the indicator of power of the accepted RSSI signal. For this value the exact compliance with a power of the accepted signal is not defined. According to the IEEE 802.11 standard it can vary from the minimum to the maximum value.

Each of the parameters specified above affects throughput of the channel and can be selected as a metric in the task of a clustering. The choice of specific parameter depends on possibility of its assessment and statements of the problem. In this article we generally restrict ourselves by the analysis of throughput in a cluster without considering a specific method of clusters shaping (this is a topic of further research). In the presented analysis as metrics we selected the input signal power on the receiver s which can be estimated by RCPI parameter.

The value of throughput which can be defined by this model (see Figure 4) depends on receiving conditions and, in practice, has the considerable dispersion. Taking this fact into account for analyzing the bandwidth we can assume that the jump model approximation by the continuous function does not propagate a significant error in the results but significantly simplifies the task.

Considering that most of D2D users are concentrated indoors, we describe the signal attenuation using the model recommended in [20] for indoor application (ITU-R 1238), thus considering out-of-band D2D:

$$L(d) = 20 \lg(f) + N \lg(d) + P_f(n) - 28 \text{ dB}, \quad (3)$$

where d is the distance in meters; f is the frequency in MHz; N is the back-off power; n is the number of obstacles; $P_f(n)$ is the loss of power parameter while passing over the

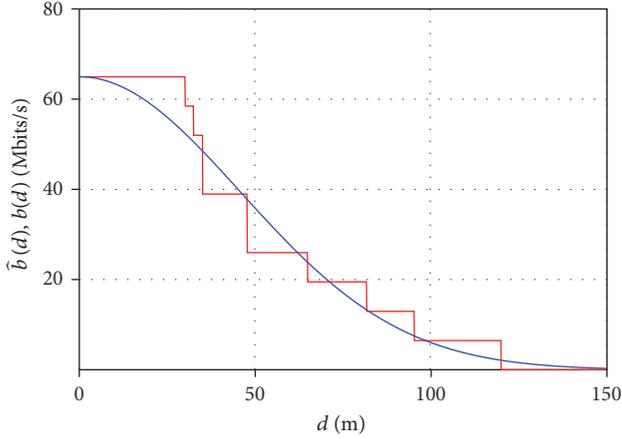


FIGURE 5: Dependence of data transfer rate on the distance to the device.

obstacle (dB). Considering the attenuation model we describe dependence of throughput on the distance of jump function (see Figure 5):

$$\widehat{b}(d) = \widehat{b}(L(d)). \quad (4)$$

The jump function was approximated by normal distribution [21], which reflects the trend of the throughput at varying distance.

$$b(d) = \begin{cases} 0 & d < 0 \\ b_{\max} e^{-d^2/2c^2} & 0 \leq d \leq R \text{ Mbps} \\ 0 & d > R, \end{cases} \quad (5)$$

where d is the distance in meters; D is the constant; b_{\max} is the maximum possible data transfer rate (Mbps); c is the half-width of the curve in meters.

$$R = \arg\{\widehat{b}(d) = 0\} \text{ (m)}. \quad (6)$$

Generally, mobile stations are distributed across the service area in a random way; therefore, value of distance d and value of signal attenuation $L(d)$ in between are also a random value. In this case we do not consider signal depression, which also affects the character of a random value of attenuation and consequently throughput. In this analysis we consider only a factor of a relative positioning of users $b(d)$.

Since throughput is a function of a random value, the distribution function b can be determined according to [22] as

$$F(b) = \iint_{D_b} f(x, y) dx dy, \quad (7)$$

where D_b is the range of b values and $f(x, y)$ is the function of the user distribution within the circle with radius R .

The probability density of b can be determined as

$$f(b) = \frac{dF(b)}{db}. \quad (8)$$

The mathematical expectation of b is

$$M(b) = \int_0^{b_0} b \cdot f(b) db. \quad (9)$$

Further we consider two kinds of functions of the users distribution on the service area: (i) uniform distribution and (ii) normal probability distribution.

3.1. Uniform Distribution. The uniform distribution is described as a set inside of a circle, the square is S and the radius is R ($S = \pi R^2$), and the interval is $0 \leq r \leq R$. The radius R of the circle is defined as $R = \arg\{\widehat{b}(d) = 0\}$ (m).

Probability density function $f(r)$ inside the circle is constant:

$$f(r) = \frac{1}{S} = \frac{1}{\pi R^2}. \quad (10)$$

If the function expressing dependence of throughput b from distances to the base station has the form (5), that is,

$$b(d) = b_{\max} e^{-d^2/2c^2}, \quad (11)$$

where R is the radius of the service area, which is defined from the model of attenuation, we can express from (5) $d = c\sqrt{-2\ln(b/b_{\max})}$ (m).

Throughput distribution function b inside the circle R , according to (8), can be expressed as

$$\begin{aligned} F(b) &= \int_0^{2\pi} \int_{c\sqrt{-2\ln(b/b_{\max})}}^R \frac{1}{S} r dr d\theta \\ &= \frac{1}{2\pi R^2} r^2 \Big|_{c\sqrt{-2\ln(b/b_{\max})}}^R \cdot 2\pi \\ &= \frac{1}{R^2} \left(R^2 + 2c^2 \ln\left(\frac{b}{b_{\max}}\right) \right) \\ &= 1 + \frac{2c^2}{R^2} \ln\left(\frac{b}{b_{\max}}\right). \end{aligned} \quad (12)$$

The probability density function according to (9) is

$$f(b) = \frac{dF_b(r)}{dr} = \frac{d}{dr} \left(1 + \frac{2c^2}{R^2} \ln\left(\frac{b}{b_{\max}}\right) \right) = \frac{2c^2}{R^2 b}. \quad (13)$$

Throughput distribution and probability density functions are shown on Figure 6.

For example, for the IEEE 802.11g standard the mathematical expectation of the throughput $M(b)$ in the service area (circle R) according to (9) can be determined as

$$M(b) = \int_{b_{\min}}^{b_{\max}} b \frac{2c^2}{R^2 b} db = 2 \frac{c^2}{R^2} (b_{\max} - b_{\min}) \text{ Mbps}. \quad (14)$$

By choosing approximation throughput function in the distance and using uniform users distribution in the service area, mathematical expectation of throughput, for IEEE 802.11n we obtain 19.51 Mbps.

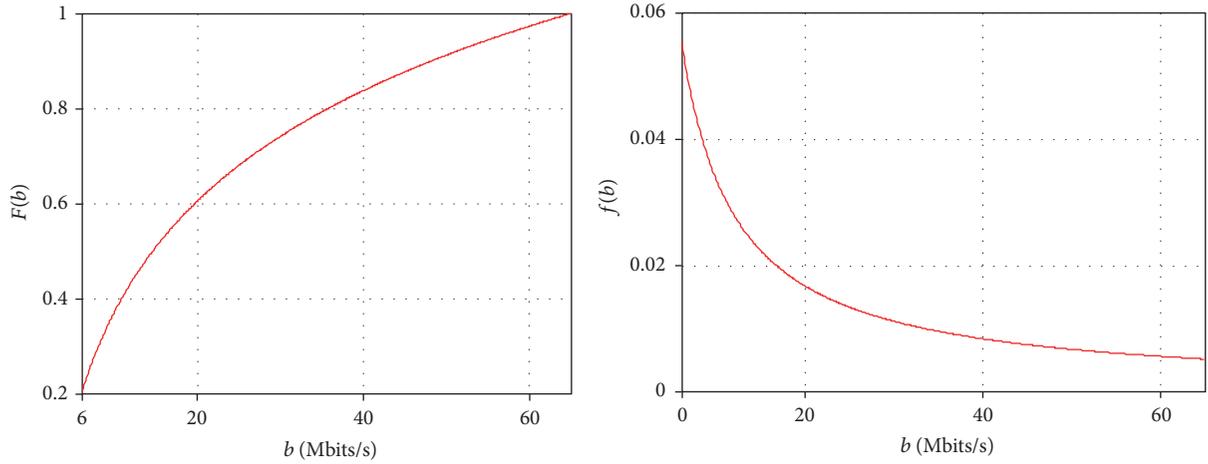


FIGURE 6: Throughput probability distribution and the probability density functions.

3.2. *Normal Distribution.* It is supposed that the traffic intensity in each point on the surface is a random value and is given by random, independent x and y coordinates. Then, the probability density function will be determined by the joint distribution function of random x and y . For a normal distribution with a center of dispersion in the center of the circle and circular dispersion (equality of the variance of x and y), the density distribution is equal to

$$f(x, y) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}, \quad (15)$$

$$r = \sqrt{x^2 + y^2},$$

where σ is the root-mean-square deviation.

For throughput analysis, we assume that the probability of penetration inside the circle R (which is the service area) is equal to 1. The normal law of probability distribution is infinite according to the x and y values. In those conditions, the probability distribution law cannot be normal. However, with sufficient accuracy it can be described by normal truncated distribution [21] as

$$f(x, y) = K(\sigma, R) \cdot \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}, \quad (16)$$

where

$$K(\sigma, R) = \frac{1}{\iint_{S_R} (1/2\pi\sigma^2) e^{-(x^2+y^2)/2\sigma^2} dx dy}, \quad (17)$$

where S_R is the area that is restricted by the circle with radius R .

The b is throughput distribution function inside the circle R .

From (5) it can be seen that $d = c\sqrt{-2\ln(b/b_{\max})}$, and from (5) according to (16) and (17)

$$F(b) = K(\sigma, R) \int_0^{2\pi} \int_{c\sqrt{-2\ln(b/b_{\max})}}^R \frac{1}{2\pi\sigma^2} e^{-r^2/2\sigma^2} r dr d\theta$$

$$= K(\sigma, R) \frac{1}{2\pi} e^{-r^2/2\sigma^2} 2\pi \Big|_{c\sqrt{-2\ln(b/b_{\max})}}^R \quad (18)$$

$$= K(\sigma, R) \left(\left(\frac{b}{b_{\max}} \right)^{c^2/\sigma^2} - e^{-R^2/2\sigma^2} \right).$$

The throughput probability density according to (8) is

$$f(b) = K(\sigma, R) \frac{c^2}{\sigma^2 b_{\max}} \left(\frac{b}{b_{\max}} \right)^{c^2/\sigma^2 - 1}. \quad (19)$$

Throughput distribution function and probability density results are presented on Figure 7.

The mathematical expectation of the throughput value according to (19) is

$$M(b) = \int_0^{b_0} b \cdot f(b) db$$

$$= K(\sigma, R) \frac{c^2}{\sigma^2 + c^2} \left(b_{\max} - b_{\min} \left(\frac{b_{\min}}{b_{\max}} \right)^{c^2/\sigma^2} \right). \quad (20)$$

The values of $M(b)$ are 55.72; 47.28; and 25.15 Mbps when the root-mean-square deviation values are 20, 30, and 80 m, respectively. Thus, the average throughput using the normal law of user traffic distribution inside the service area depends on the variance (scattering), when some bigger throughput values happen while there are lower values of variance.

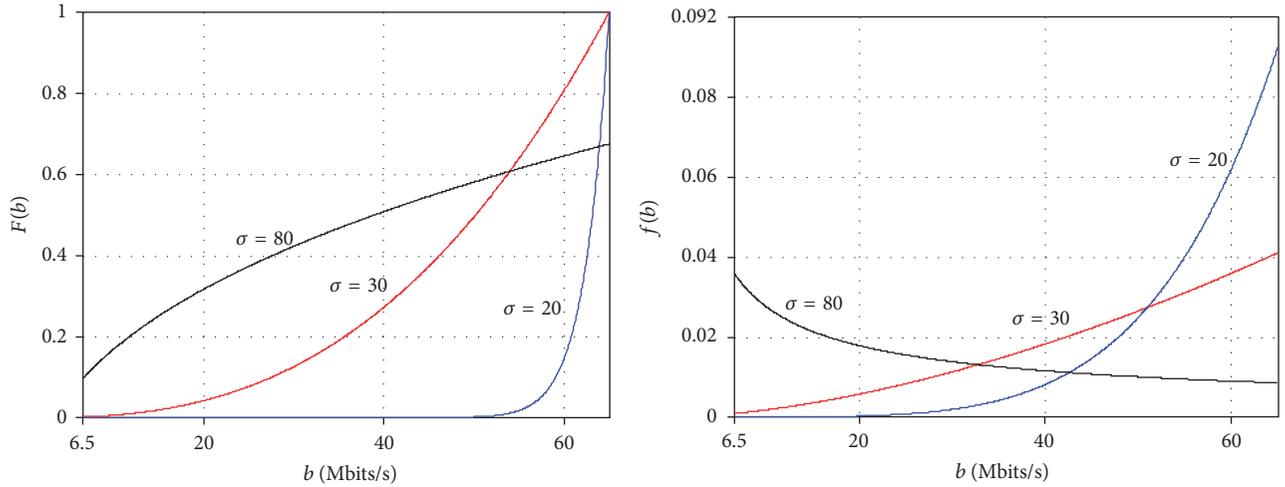


FIGURE 7: $F(b)$: throughput distribution function and throughput probability density $f(b)$.

Average throughput in the case of normal distribution always exceeds analogical value of uniform distribution.

From the analysis of the examples we can assume that the uniform distribution is the worst one according to the average amount of throughput criteria in comparison with other types of distributions, where the coordinates of the access point are coincident with the mathematical expectation of traffic distribution. The mathematical expectation of traffic distribution coordinate is the most appropriate base station installation point from the position of maximum throughput support (potential capacity), at least for unimodal distribution laws.

Using (12) or (18) depending on type of distribution of users we can estimate probability that the throughput not less than the given value:

$$P(b \geq B_{\min}) = 1 - F(b), \quad (21)$$

where B_{\min} is the minimum throughput value permitted for a cluster member and $F(b)$ is the probability distribution given by (12) or (18).

In the condition of B_{\min} restriction this probability (21) presents part of users included in clusters. It allows estimating a number of clustering users, choosing parameters of the network from (12) or (18). If in the network service area a number of randomly distributed clusters formed, using (21) we can estimate the number of users served with throughput no lower than B_{\min} .

4. Performance Evaluation

To verify the models above, a simulation campaign was performed. The following data represent the results of throughput simulation between user terminals and terminals within the communication area with uniform distribution. During simulation modeling there are permissible variations about the level of received signal that are verified by short

fading effect and represent random value with a Nakagami distribution [22].

$$f(x, \alpha, \beta) = \frac{2}{\Gamma(\alpha)} \left(\frac{\alpha}{\beta}\right)^\alpha x^{2\alpha-1} e^{-(\alpha/\beta)x^2}, \quad (22)$$

where β is the shape parameter; α is the scale parameter; α and β are associated with the throughput by the formula given below:

$$\bar{b} = \frac{\Gamma(\alpha + 1/2)}{\Gamma(\alpha)} \sqrt{\frac{\beta}{\alpha}}. \quad (23)$$

Figure 8 represents analytical and simulation results for the case of uniformly distributed users: dependence between throughput and distance between cluster head and cluster member (a); probability density function (b), obtained via simulation (red) and analysis according to the formula (13) (blue). The average throughput for this case is 17.2 Mbps. Figure 9, in turn, represents simulation results for the case of normally distributed users: dependence between throughput and distance between cluster head and cluster member (a) and probability density function (b), obtained via simulation (red) and analysis according to formula (19) (blue). The average throughput for this case is 47.4 Mbps. Therefore, one can conclude that in case of normal distribution the average throughput between cluster head and cluster member is larger than 2.5 times that for the case with uniform distribution. Also, as it can be seen from the simulation and analytical distributions, the probability density function of throughput is sufficiently close to the simulation modeling results. This allows us to make a conclusion on credibility of obtained models for uniform and normal distributions of user devices. Hence, both analytical and simulation models allow defining the throughput in clusters with different user distribution in the service area of the head node.

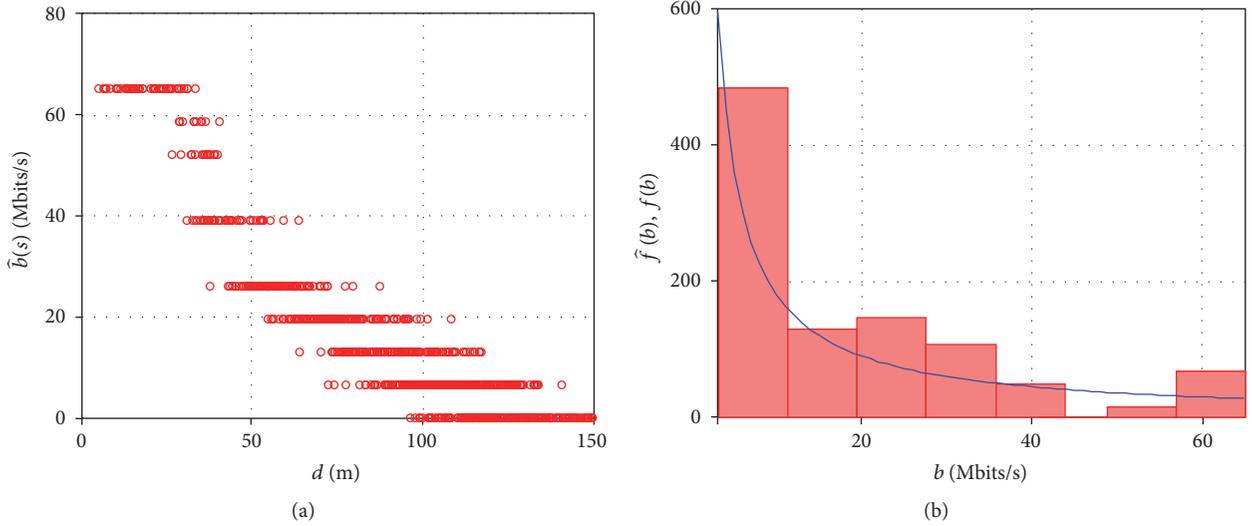


FIGURE 8: Dependence of throughput on the distance and probability density in case of uniform distribution of user devices.

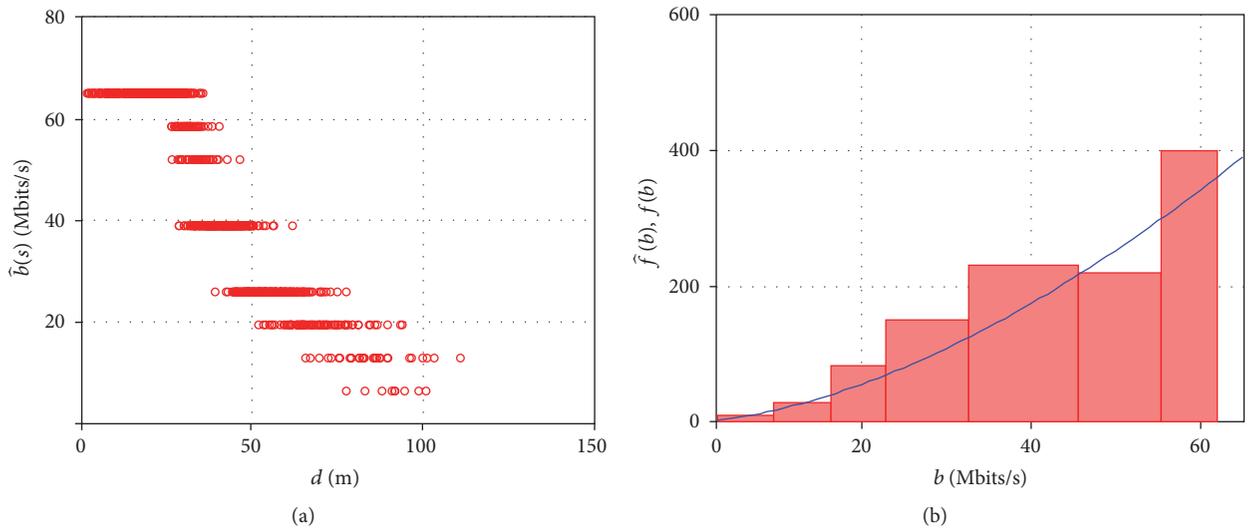


FIGURE 9: Dependence of throughput on the distance and probability density in case of normal distribution of user devices.

5. Clustering Method Selection

Figure 10 shows an example of clustering of 5,000 objects using two different algorithms [23, 24] (k -means algorithm on the left and FOREL algorithm on the right), obtained as a result of our simulation modeling. This is a theoretical example that does not reflect the real sizes of clusters in the network. However, this method quite clearly demonstrates the difference of the result, by choosing different methods of clustering.

As it can be seen from Figure 10, with an equal number of clusters (25 clusters in total) the shape of clusters is marked visually differently in the first and second cases.

To compare these methods we performed analysis of cluster members distribution processed relative to the cluster

centers. Distribution was obtained by simulation modeling. For center-of-mass we used the following expression:

$$s = gm_j - Cm_i, \quad i = 1 \cdots |C|, \quad j = 1 \cdots n_i, \quad m_j \in c_i, \quad (24)$$

where Cm_i is the center-of-mass coordinate of cluster i ; gm_j is the coordinate of m_j element of c_i cluster; c_i is cluster i ; m_j is element j of cluster i ; n_i is the number of elements within the cluster i ; $|C|$ is the number of clusters.

Figure 11 shows the empirical probability density of cluster members coordinates relative to cluster center (histogram) and its approximation by the normal distribution (smooth curve) for k -means and FOREL cases, respectively. The diagrams show that in both the first and second cases the distribution of elements within clusters (relative to the centers

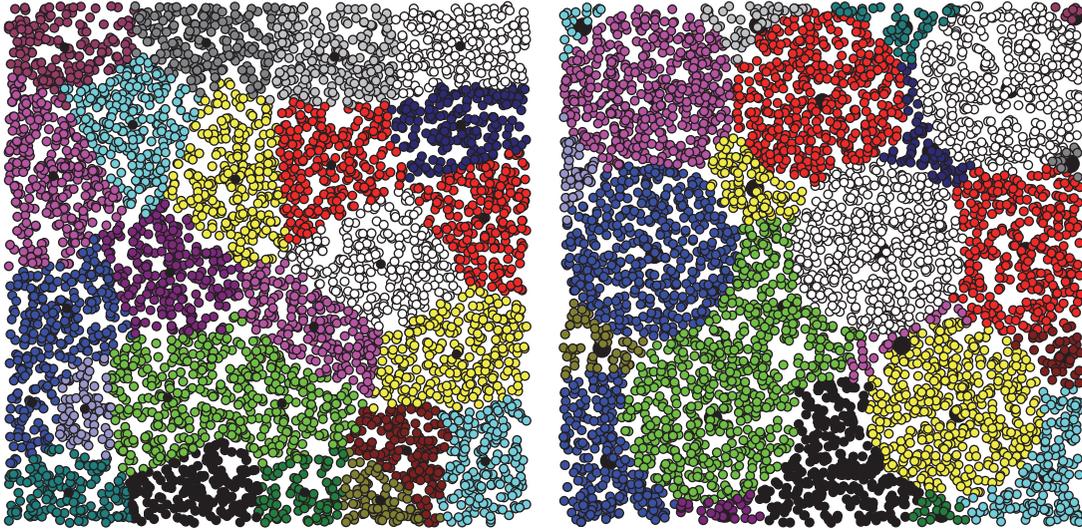


FIGURE 10: Clustering of 5,000 of elements using two different algorithms.

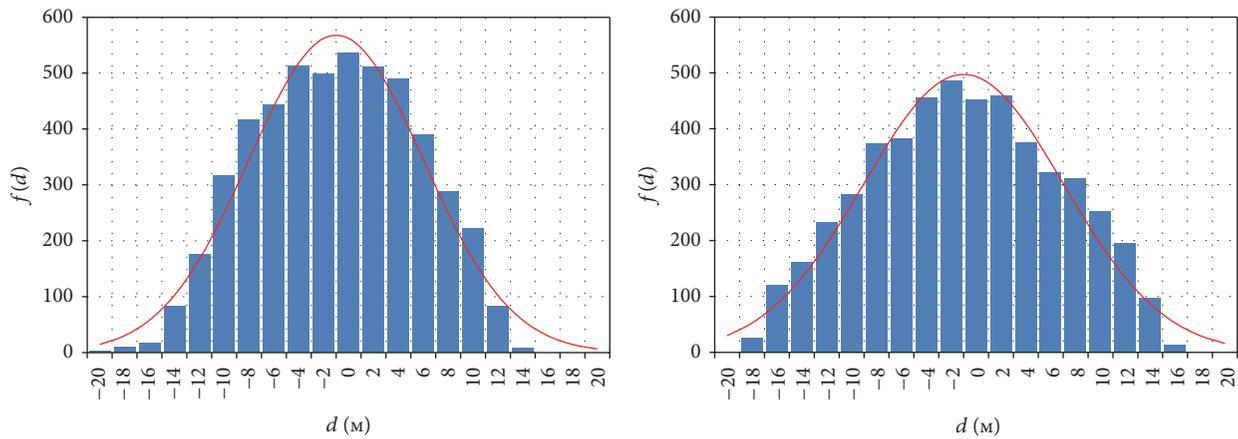


FIGURE 11: The node distribution within the clusters by using different methods of clustering (25 clusters in total).

of clusters) is close enough to the normal. It is worth noting that for the considered clustering algorithms dispersion of elements in clusters is remarkably different, which can be definitely seen from Figure 11.

Figure 12 shows an example for small number of clusters, 2 and 5 clusters, respectively. The figure shows that different algorithms form clusters in different ways. k -means split two clusters of a similar size, FOREL in turn has combined most of the elements in one cluster of a given size, and for the remaining elements 4 clusters were formed.

As it is shown on Figure 13, with a small number of clusters, the distribution of the elements inside them is closer to uniform than to normal, and more specifically it is closer to the original distribution of elements in a clustering area. It should be noted that the shape of the boundaries of the clusters when using k -means is close to the Voronoi diagram [17], which is constructed relative to the cluster centers. Therefore, in case of large number of clusters the distribution

of cluster members can be approximated by normal law, in case of small number of clusters by uniform distribution.

6. Conclusions and Future Work

In this paper, we propose and evaluate a novel model for throughput estimation in out-of-band D2D clustering. Compared to existing studies where a distance between cluster head and cluster members was used, we suggest using a throughput. We obtained the numerical results for different types of typical cluster members distributions, uniform and normal. We delivered closed-form analytical expressions for probability distribution of throughput in the cluster for a given distribution and density of users. Through analytical and simulation studies, we show that the average throughput between cluster head and cluster member for the normal distribution is 2.5 times larger than for the case with uniform distribution. The obtained results show that known clustering

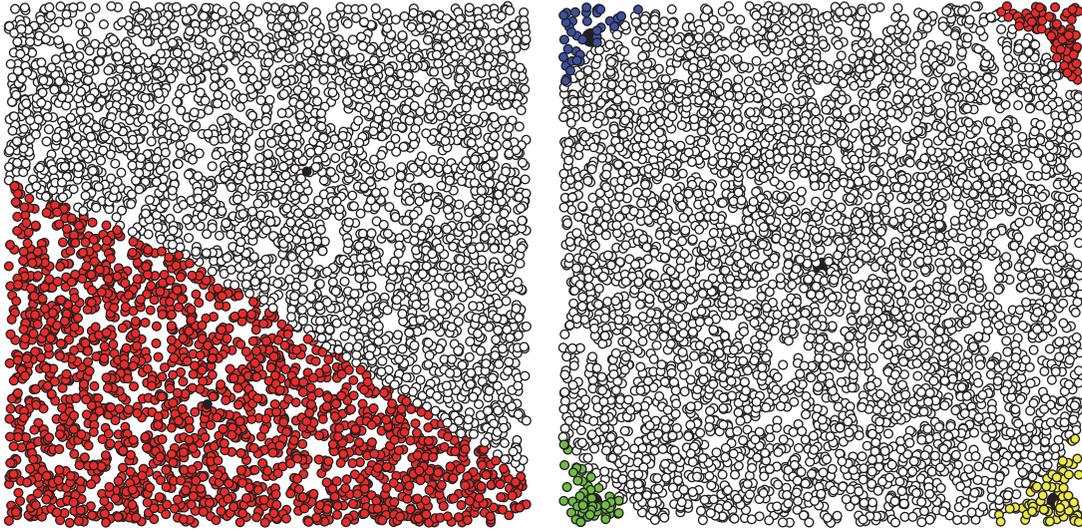


FIGURE 12: The node distribution within clusters by using different clustering methods, 2 and 5 clusters.

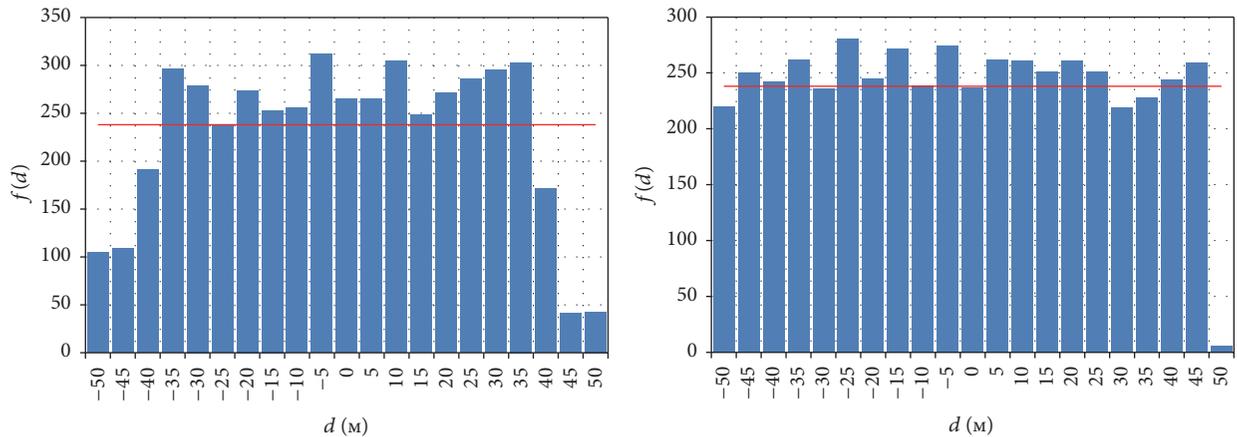


FIGURE 13: The node distribution within clusters by using different clustering methods, 2 and 5 clusters.

algorithms can be used to choose cluster head which provides near-optimal solution for throughput of channels between cluster head and cluster members.

Further, by using well-known clustering algorithms k -means and FOREL we obtained distribution of cluster members for large and small number of clusters. We show that, for the case with large number of cluster members, compared to FOREL k -means algorithm gives remarkably smaller dispersion of elements in clusters. Thus k -means constructs cluster of similar sizes, compared to FOREL that constructs clusters of very different sizes. Therefore, for out-of-band D2D case k -means provide better clustering considering target even resource distribution and increase of network resource utilization rate.

Our future work will concentrate on further enhancements of the delivered model by introducing more QoS metrics such as packet loss ratio, delay, and energy as optimization parameters. Also, a system-level performance evaluation would be needed to understand ultimate implications of the

suggested model. In addition, we plan to consider 3D cases for out-of-band D2D scenarios, as well as cluster member location dynamics, which will allow addressing advanced drone-based scenarios of D2D communications.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

- [1] A. Asadi and V. Mancuso, "Network-assisted outband D2D-clustering in 5G cellular networks: theory and practice," *IEEE Transactions on Mobile Computing*, vol. 16, no. 8, pp. 2246–2259, 2017.
- [2] "3rd generation partnership project; technical specification group services and system aspects; policy and charging control architecture (Release 13)," Tech. Rep. 23., 2015.

- [3] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 4, pp. 1801–1819, 2014.
- [4] S. Andreev, A. Pyattaev, K. Johnsson, O. Galinina, and Y. Koucheryavy, "Cellular traffic offloading onto network-assisted device-to-device connections," *IEEE Communications Magazine*, vol. 52, no. 4, pp. 20–31, 2014.
- [5] A. Gupta and R. K. Jha, "A survey of 5G network: architecture and emerging technologies," *IEEE Access*, vol. 3, pp. 1206–1232, 2015.
- [6] B. Zhou, H. Hu, S.-Q. Huang, and H.-H. Chen, "Intracluster device-to-device relay algorithm with optimal resource utilization," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 5, pp. 2315–2326, 2013.
- [7] M. Ashraf, W.-Y. Yeo, M. Woo, and K.-G. Lee, "Smart energy efficient device-to-multidevice cooperative clustering for multicasting content," *International Journal of Distributed Sensor Networks*, vol. 2016, Article ID 3727918, 9 pages, 2016.
- [8] L. Wang, G. Araniti, C. Cao, W. Wang, and Y. Liu, "Device-to-device users clustering based on physical and social characteristics," *International Journal of Distributed Sensor Networks*, vol. 2015, Article ID 165608, 14 pages, 2015.
- [9] J. Seppälä, T. Koskela, T. Chen, and S. Hakola, "Network controlled Device-to-Device (D2D) and cluster multicast concept for LTE and LTE-A networks," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '11)*, pp. 986–991, Cancún, Mexico, March 2011.
- [10] S. Andreev, O. Galinina, A. Pyattaev, K. Johnsson, and Y. Koucheryavy, "Analyzing assisted offloading of cellular user sessions onto D2D links in unlicensed bands," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 1, pp. 67–80, 2015.
- [11] S. Andreev, J. Hosek, T. Olsson et al., "A unifying perspective on proximity-based cellular-assisted mobile social networking," *IEEE Communications Magazine*, vol. 54, no. 4, pp. 108–116, 2016.
- [12] A. Pyattaev, J. Hosek, K. Johnsson et al., "3GPP LTE-assisted Wi-Fi-direct: Trial implementation of live D2D technology," *ETRI Journal*, vol. 37, no. 5, pp. 877–887, 2015.
- [13] H. Shimodaira, J. Kim, and A. S. Sadri, "Enhanced next generation millimeter-wave multicarrier system with generalized frequency division multiplexing," *International Journal of Antennas and Propagation*, vol. 2016, Article ID 9269567, 11 pages, 2016.
- [14] J. Kim, J.-J. Lee, and W. Lee, "Strategic control of 60 GHz millimeter-wave high-speed wireless links for distributed virtual reality platforms," *Mobile Information Systems*, vol. 2017, Article ID 5040347, 10 pages, 2017.
- [15] K. S. Hassan and E. M. Maher, "Device-to-device communication distance analysis in interference limited cellular networks," in *Proceedings of the Tenth International Symposium on Wireless Communication Systems (ISWCS '13)*, 2013.
- [16] M. Afshang, H. S. Dhillon, P. Han, and J. Chong, "Coverage and area spectral efficiency of clustered device-to-device networks," in *Proceedings of the 2015 IEEE Global Communications Conference (GLOBECOM '15)*, pp. 1–6, San Diego, Calif, USA, August 2015.
- [17] F. Aurenhammer, R. Klein, and D.-T. Lee, *Voronoi Diagrams and Delaunay Triangulations*, World Scientific Publishing Co., Hackensack, NJ, USA, 2013.
- [18] A. Muthanna, P. Masek, J. Hosek et al., "Analytical evaluation of D2D connectivity potential in 5G wireless systems," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Preface*, vol. 9870, pp. 395–403, 2016.
- [19] IEEE Std 802.11™, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," 2012.
- [20] ITU-R P.1238-8, "Propagation data and prediction methods for the planning of indoor radio communication systems and radio local area networks in the frequency range 300 MHz to 100 GHz".
- [21] Wolfram MathWorld, "Gaussian Function," <http://mathworld.wolfram.com/GaussianFunction.html>.
- [22] D. Laurenson, "Nakagami Distribution," *Indoor Radio Channel Propagation Modelling by Ray Tracing Techniques*, 1994.
- [23] L. Kaufman and P. J. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*, John Wiley & Sons, New York, NY, USA, 2005.
- [24] J.-O. Kim and C. W. Mueller, *Factor Analysis: Statistical Methods and Practical Issues*, Sage Publications, Newbury Park, Calif, USA, 1978.



Hindawi

Submit your manuscripts at
<https://www.hindawi.com>

