Resource Allocation scheme for Multi - hop D2D communication in 5G Networks

Kalivaraprasad., B

Assistant Professor, Dept. of. ECE(BES-II), Koneru Lakshmaiah Education Foundation (KLEF), Vaddeswaram, Guntur(dt)-522302. baditakali@gmail.com

Abstract

Multi-hop D2D communications increases proximity and an efficient capacity scaling technique for 5G networks. In this letter, we proposed a novel resource allocation scheme for multi-hop D2D communication for 5G networks. This scheme comprises in three phases: in first phase, a Q learning mechanism is used for finding the best routing path between multi-hop D2D communication. In second phase, mean-max distance (MMF) resource allocation scheme is proposed for best routing path. In the last phase, we proposed a hybrid optimization technique by using particle swarm optimization (PSO) and grey wolf optimization (GWO) technique (HPSOGWO). The analysis framework and the results guarantee that the proposed schemes improve the system throughput.

Introduction:

Device-to-Device (D2D) communication will play a major role in 5G networks because most of communication will be locally in future[1]. The major advantage over D2D communication is that they have not required dedicated resource block for their communication. They reuse the resource block of CU users for their communication. Thus, D2D has attracted attention from the researchers. However, the transmission distance and transit power of D2D nodes in direct D2D communication are really depressed. The multi-hop D2D communication is an extended version of D2D communication where a minimum of two connections must be formed between D2D pairs (source-to-relay and relay-to-destination)[2]. The base station (BS)/ENodeB (ENB)provides resource block for each link with a minimum interference of cellular users (CU) because these links reuse the resource block of CU users. Fig.1 shows the multi-hop D2D communication where source (D5) communicate with destination (D1) with the help of D3 device.

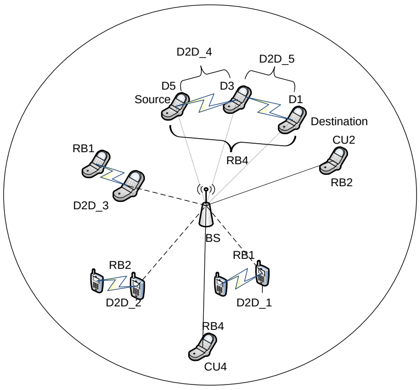

Figure 1 multi-hop D2D communication

In multi-hop D2D communication, there are two major challenges: relay selection and resource allocation. In the relay selection, the BS selects best device (known as relay) among the available devices between source-to-destination. Number of proposals for relay selection has been investigated by the authors in [3-5]. In [3], the authors have discussed formation of multi-hop D2D. The SINR based relay selection scheme is proposed in [4]. In [5], the authors have conceded several factors for choosing the relay selection ahead of SINR.On the other hand, in resource allocation, the BS allocates the resource block to source-to-relay and relay-to-destination. In [6, 7], the authors proposed resource allocation scheme for multi-hop D2D communication. In [6], the authors proposed hybrid resource allocation scheme for multi-hop D2D communication for 5G networks. In [7], the authors have proposed a resource allocation scheme along with route finding. In this letter, we introduced a Q learning model for route determination between source-to destination device for multi-hop D2D based on the overall reward (total distance between the D2D nodes). The primary contributions are as follows.:

a. We provide a path selection scheme based on Qlearning algorithm for multi-hop D2D communication.
b. b. Max Mean Far algorithm-based Resource allocationscheme is proposed for selected route between multi-hop D2D communication.
c. We have proposed a hybrid optimized scheme (HPSOGWO) for enhancing the throughput.

System Model and Problem Formulation:TheBS is placed in the cell center and it is surrounded by the randomly distributed cellular users CU CU_m $where\ m=\{1,2,3...,i,..M\}\ i\subseteq M$ and $D2D$ users D_n $where\ n=\{1,2,3...,j,..N\}\ j\subseteq N$. Let there are K number of resource blocks are available $k=\{1,2,3,...\mathrm{K}\}$The number of possible of nodes in the multi hop path is defined as Ml_{mh} $where\ mh=\{1,2,3,...MH\}$.

For cellular communication: The SINR at the base station of receiving signal form CU_i over k^{th} resource block is given by

$$\delta^{k}_{CU_i 2eNB}=\frac{P_{cu_i 2eNB}(\mathrm{d}_{CU_i 2eNB})^{-\alpha}}{\sum_{j=1}^{n}\rho_{i,j}P_{D_{jTx}}(\mathrm{d}_{CU_i 2eNB})^{-\alpha}+N_o} \quad (1)$$

The Shannon capacity model throughput of cellular communication is calculated as

$$DR_{CU_i}=BW\log_2(1+\delta^{k}_{CU_i 2eNB}) \quad (2)$$

where BW is the bandwidth

For Direct D2D communication: SINR at $D2D_j$ receiver with k^{th} resource block is given by is calculated as

$$\delta^{k}_{D_{jTx,jRx}}=\frac{P_{D_{jTx}}(\mathrm{d}_{D_{jTx}2D_{jRx}})^{-\alpha}}{\sum_{i=1}^{m}\rho_{i,j}P_{CU_i}(\mathrm{d}_{CUi2D_{jTx}})^{-\alpha}+\sum_{l=1,l\neq 1}^{L}\rho_{j.l}P_{D_{lTx}}(\mathrm{d}_{D_{lTx}2D_{jTx}})^{-\alpha}+N_o} \quad (3)$$

Where, $P_{D_{jTx}}, P_{CU_i}, P_{D_{lTx}}$ power of D_{jTx} transmitter and power of cellular user and power of D_{lTx} transmitter. $\mathrm{d}_j, \mathrm{d}_{i2j}, \mathrm{d}_{l2j}$ denotes distance between the D_{jTx} transmitter and receiver, and D_{jRx} transmitter to cellular users and D_{jTx} transmitter to D_{lTx} transmitter respectively. The throughput of D2D user as according to Shannon capacity model is calculated as

$$DR_{D_j}=BW\log_2(1+\delta^{k}_{D_{jTx,jRx}}) \quad (4)$$

For multi-hop D2D communication: The SINR at the multi-hop relay side with k^{th} resource block is given by

$$\delta^{k}_{Ml\_Tx}=\frac{P_{D_{jTx}}\times g_{\mathrm{D}_{jTx},Ml}}{\sigma+\sum_{m=1}^{M}\sum_{n=1}^{N}\sum_{k=1}^{K}\left(I^{D_{jTx}}_{m,k}+I^{D_{jTx}}_{\mathrm{eNB},k}+I^{D_{jTx}}_{\mathrm{n},k}\right)} \quad (5)$$

Similarly, SINR of the relay at the receiver side with k^{th} resource block is given by

$$\delta_{Ml\_Rx}^{k} = \frac{P_{D_{jRx}} \times g_{Ml,D_{jRx}}}{\sigma + \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{k=1}^{K} \left(I_{m,k}^{D_{jRx}} + I_{\mathrm{eNB},k}^{D_{jRx}} + I_{\mathrm{n},k}^{D_{jRx}} \right)} \quad (6)$$

The end-to-end SINR for multi hop D2D is obtained by considering the minimum SINR of the hops, it can be calculated as

$$\delta_{SR} = \min \left\{ \delta_{Ml\_Tx}^{k}, \delta_{Ml\_Rx}^{k} \right\} \quad (7)$$

The throughput of the relay can be expressed as

$$DR_{Ml} = BW \log_2 (1 + \delta_{SR}^{k}) \quad (8)$$

Problem Formulation: The objective function for the maximizing the overall network throughput can be formulated as

$$Tp_{total} = \underset{P_{CU}, P_D}{\arg} \; Max \left(\sum_{m=1}^{M} DR_{CUm} + \sum_{n=1}^{N} DR_{D_n} + \sum_{mh=1}^{MH} DR_{Ml_{mh}} \right) \quad (9)$$

Subjected to the following constraints

$\delta_{CU2eNB} \le \delta_{Threshold}$ **(C1)**

$\delta_{D2eNB} \le \delta_{Threshold}$ **(C2)**

$0 < P_D \le P_{\mathrm{D,max}}$ **(C3)**

$0 < P_{CU} \le P_{CU,\max}$ **(C4)**

The (C1) and (C2) constraint details that the interference occurred in the cell should be always less than the threshold interference.The (C3, C4) constraints defines about the power of cellular user and D2D should be less his maximum power.

***Proposed Scheme*:**The proposed model has been divided into three phases. In first phase, Q learning algorithm-based path selection in multi-hop D2D communication is proposed. In second phase, a mean max far distance-based resource allocation scheme is proposedfor path selected form the first stage. Finally, in the third phase we proposed a hybrid optimization technique for maximizing the throughput by considering power as a particle.

Phase1: path selection for multi hop D2D communication: Q-learning uses the action selection policies for the agents. The optimal policy at the end of the learning is what actions to take in order to maximize the reward. The present state as S and the agent takes an action

A. in response to this action generates a reward signal R and agent moves to the state S_1 .The Q-value function given as:

$$Q(S,A) \leftarrow Q(S,A) + \alpha(R + \gamma max_{A1}(Q(S_1, A_1)) - Q(S,A)) \quad (10)$$

The distance between two nodes points $D2D_{Tx}(x1, y1)$ and $D2D_{Rx}(x2, y2)$ is taken here as the reward function

$$R = \max\left(\sum_{j=1}^{J} \sqrt{\left(x_j 1 - x_j 2\right)^2 + \left(y_j 1 - y_j 2\right)^2} \right) \quad (11)$$

Phase2:Resource allocation: we proposed a max mean far (MMF) algorithm-based resource allocation for selected routing path by Q learning-based scheme. To avoid the interference, the BScalculates the mean distance from selected routed path with source and destination D2D to identified cellular users as follows: the BS calculate the distance from source D2D node to CU user, distance from all intermediate node (relay) to CU user and destination D2D node to cellular users. Once, the BS calculate all the distance form cellular user, it takes means average value and stored into list. Similarly, the BS calculate the distance from all identified CU users and stored into list in descending order. The BS selects the top most element and find the correspond cellular users., The BS allocate the same resource block to selected CU user's resource block to newly setup multi-hop D2D communication

Phase3: Through Maximization: In resource allocation scheme, we have considered that each CU and D2D users uses fixed power for their communication (24 dB for cellular communication and 22 dB for D2D communication). Due to sharing of resource block between D2D and CU users, they create interference to each which degrade the overall system throughput. Therefore, in this scheme, we have considered the variable power of D2D and CU users. To select the power individual users, we have used hybrid optimization scheme. The hybrid optimization scheme is combination of Particle Swarm Optimization (PSO) and Gray Wolf Optimization (GWO) [8]. The PSO technique provides maximum exploration and GWO techniques increasing the exploitation.

According to the PSO algorithm, the position of particles z_i $(i = 1, 2, ..., \mathrm{i})$ is set to the power allocated to CUE and D2D users. In this paper, the fitness function is seeking maximum throughput.

$$Z_i^k = \left[P_{D2D_{1_i}}^k, P_{D2D_{2_i}}^k, P_{D2D_{3_i}}^k, ..., P_{D2D_{m_i}}^k, P_{CU_{1_j}}^k, P_{CU_{2_j}}^k, P_{CU_{3_j}}^k, ...P_{CU_{n_j}}^k \right] \quad (12)$$

$$V_{ic}^{k+1} = \omega^K + V_{ic}^K + c_1 r_1 (\mathrm{pb}_1^K - z_i^K) + c_2 r_2 (\mathrm{gb}^K - z_i^K) \quad (13) \quad V_{iD2D}^{k+1} = \omega^K + V_{ic}^K + c_1 r_1 (\mathrm{pb}_1^K - z_i^K) + c_2 r_2 (\mathrm{gb}^K - z_i^K) \quad (14)$$

$$V_i^k = V_{ic}^{k+1}, V_{iD2D}^{k+1} \qquad (15)$$

According to GWO algorithm the encircling behavior is given by the following mathematical equations

$$\mathrm{P}_{DC} = \left[P_{D2D_{1_i}}^k, P_{D2D_{2_i}}^k, P_{D2D_{3_i}}^k, ..., P_{D2D_{m_i}}^k, P_{CU_{1_j}}^k, P_{CU_{2_j}}^k, P_{CU_{3_j}}^k, ...P_{CU_{n_j}}^k \right] = Z_i^k \quad (16) \quad d = \left[c.\mathrm{P}_{DC} - \mathrm{P}_{DC}(\mathrm{t}) \right] \quad (17) \quad \mathrm{P}_{DC}(\mathrm{t}+1) = \left[\mathrm{P}_{DC} - a.d \right]$$

$$(18)$$

$$\begin{aligned} V_{ic}^{k+1} = \omega^K + V_{ic}^K + c_1 r_1 (P_{DC_1}^K - P_{DCi}^K) + c_2 r_2 (P_{DC_2}^K - P_{DCi}^K) + \\ c_2 r_2 (P_{DC_2}^K - P_{DCi}^K) + c_3 r_3 (P_{DC3}^K - P_{DCi}^K) \end{aligned} \quad (19)$$

Velocity and updated HPSOGWO equation are proposed as follows:

$$\begin{aligned} V_{iD2D}^{k+1} = \omega^K + V_{ic}^K + c_1 r_1 (P_{DC_1}^K - P_{DCi}^K) + c_2 r_2 (P_{DC_2}^K - P_{DCi}^K) + \\ c_2 r_2 (P_{DC_2}^K - P_{DCi}^K) + c_3 r_3 (P_{DC3}^K - P_{DCi}^K) \end{aligned} \quad (21) \quad P_{DCi}^{K+1} = P_{DCi}^K + \left(V_{ic}^{k+1}, V_{iD2D}^{k+1} \right) \qquad (22)$$

***Simulation Results*:**Simulation result in Matlab is performed for the performance of proposed scheme and compared with another optimization scheme. ThePSO parameters and GWO parameters used in the performance evaluation are briefed in Table.1

Table.1 Simulation Parameters

Description	Value
Learning factor α	0.9
Discount factor γ	0.9
Inertial w	[0.4, 0.9].
Maximum power of CU and D2D users	24 dB, 22 dB
Maximum number of multi-hops D2D users	22

Form the Figure.2 the experimental results of the Q-leaning model for path selection in comparison with the Dijkstra model algorithm. It addresses the problem with reward and stochastic transmissions. The proposed Q leaning based scheme required less time compared to Dikata algorithm.

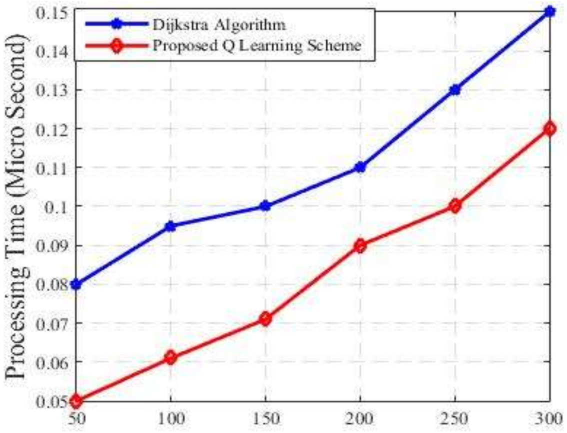

Figure 2 Processing time for large number of requests

Fig. 3 illustrates the comparison of throughputsfor PSO and GWO with the proposed HPSOGWO scheme. From the figure, it is observed that theproposedscheme performs better than other optimization scheme with low computation (less iteration).

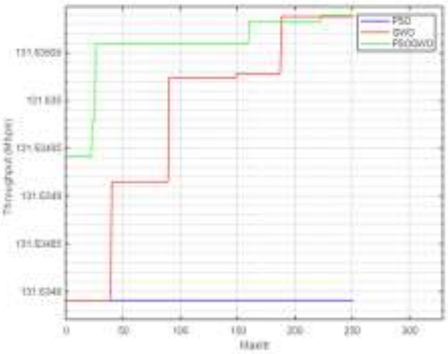

Figure 3 Throughput Vs Maximum iteration=300

Figure.4describes the total throughput of MMF resource allocation and the combined performance of the MMF and HPSOGWO program, which is far more substantial than that of the MMF resource management scheme alone.

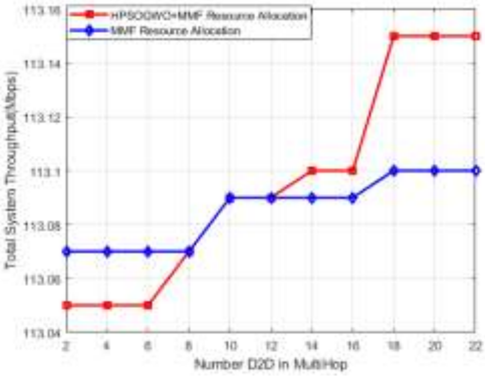

Figure 4. System Throughput comparison

Conclusion: In this letter, an efficient route selection for multi-hop D2D communication is proposed with a Q learning algorithm. We also proposed a novel max-mean-far resource allocation scheme for resource allocation. We adapted a HPSOGWO scheme for maximization of throughput. The simulation results conclude outperformance the results in comparison with the previously proposed techniques.

Acknowledgement

This research work is supported by SERB under the Early Carrier Research Award (ECRA) scheme project grant number is ECR/2017/002314.

Panduranga Ravi Teja, Surbhi Rai, Pavan Kumar Mishra
Department of IT, N.I.T Raipur, India Pin-492010
tejapanduranga@gmail.com, surbhi.rai1998@gmail.compavanmishra.it@nitrr.ac.in

References:

[1] M. N. Tehrani, et al., "Device-to-device communication in 5G cellular networks: challenges, solutions, and future directions," IEEE Communications Magazine, vol. 52, pp. 86-92, 2014.

[2] F. S. Shaikh and R. Wismüller, "Routing in multi-hop cellular device-to-device (D2D) networks: A survey," IEEE Communications Surveys & Tutorials, vol. 20, pp. 2622-2657, 2018.

[3] Y. Li, et al., "A dynamic graph optimization framework for multihop device-to-device communication underlaying cellular networks," IEEE Wireless Communications, vol. 21, pp. 52-61, 2014.

[4] G. Feng, et al., "Energy efficiency resource allocation for D2D communication network based on relay selection," Wireless Networks, pp. 1-11, 2015.

[5] P. K. Mishra, et al., "Relay Selection Scheme for Dynamic Network Scenario in Multi-hop D2D Communication," in 2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS), 2019, pp. 538-542.

[6] P. K. Mishra, et al., "Hybrid Resource Allocation Scheme in Multi-hop Device-to-Device Communication for 5G Networks," Wireless Personal Communications, vol. 103, pp. 2553-2573, 2018.

[7] S. Alwan, et al., "Joint routing and wireless resource allocation in multihop LTE-D2D communications," in 2018 IEEE 43rd Conference on Local Computer Networks (LCN), 2018, pp. 167-174.

[8] N. Singh and S. Singh, "Hybrid algorithm of particle swarm optimization and grey wolf optimizer for improving convergence performance," Journal of Applied Mathematics, vol. 2017, 2017.