

Greedy Based Radio Resource Allocation Algorithm with SARSA Power Control Scheme in D2D Underlying Communication

Astri Wulandari¹, Nachwan Mufti A.², and V.S.W. Prabowo³

^{1,2,3}*School of Electrical Engineering, Telkom University, Bandung, 40287, Indonesia*

^{1,2,3}{*astriwlnr@student., nachwanma@, vinsensiusvsw@*}*telkomuniversity.ac.id*

Manuscript received November 10, 2020; revised December 10, 2020; accepted December 30, 2020

Abstract

Device-to-Device (D2D) underlying communication system is a solution in reducing the workload of eNodeB and increasing the system data rate. This communication system consists of two users, namely Cellular User Equipment (CUE) and D2D pair, where CUE will share its resources with the D2D pair. This sharing resources also causes interference and should be managed using the resource allocation algorithm. In this work, the resource allocation scheme occurs in a single cell with an uplink communication direction. The resource allocation process uses greedy and joint greedy algorithms. After CUE allocates all of its resources, SARSA algorithm performs the power allocation process. The resource allocation process involves the scheduled CUE and D2D pair. After all the resource and power are allocated, parameter performance of the system is calculated. Based on the work results, joint greedy algorithm with power allocation using SARSA algorithm have performance results 1.375×10^7 bps/Watt in energy efficiency, 43.105 bps/Hz in spectral efficiency, and 0.993 in D2D fairness index

Keywords: Device-to-Device; Resource Allocation; SARSA Power Control; Joint Greedy; Greedy

DOI: 10.25124/jmeecs.v7i1.3472

1 Introduction

According to the increasing number of Cellular User Equipment (CUE) nowadays, research on effective communication is being carried out. The increasing number of CUE will effect capacity demand, so that the traffic load on eNodeB will also increase. Therefore, a technology that can reduce the traffic load borne by eNodeB is needed. Device-to-Device (D2D) Communication technology is one of the communication technologies capable to provide solutions in reducing traffic loads on eNodeB.

D2D communication method allows direct communication between CUEs without going through eNodeB. It happens because of utilizing D2D communication reuse of the frequency spectrum from CUE. There are two modes to reuse the CUE frequency spectrum in D2D communication, namely, underlay and overlay. This work will use underlay mode, where D2D users use the same frequency spectrum as CUE.

However, reusing resources will result in interference between users. This interference will interfere the communication link on the connected device. One of the solution to reduce the interference that occurs is arranging the proper power allocation to the D2D device. The increasing transmission power level in a user will increase the performance parameters too.

In work [1], the Adaptive power allocation method using two-phase optimization algorithms successfully increases the system data rate. In work [2], the heuristic power allocation algorithm manages to improve the system data rate as well. In work [3], The distributed Q-learning method for power control of D2D users has successfully updated the Q value both in a distributed and collaborative manner.

This work proposed joint greedy algorithm for resource allocation and a modified Reinforcement Learning algorithm called State Action Reward State Action (SARSA) for power control scheme in D2D commu-

nication. With the right allocated resource and power transmission, the overall system data rate will increase, traffic load will decrease, and Quality of Service (QoS) stay awake.

2 Simulation Process

There are two combinations of algorithms in this work: joint greedy algorithm with SARSA algorithm as power control scheme and greedy algorithm with SARSA algorithm as power control scheme. First of all, greedy and joint greedy allocation algorithms will enter the generated-user results and the obtained gain channel, SINR, and capacity calculations. After the specific CUE allocates all its resources to specific D2D pair, the SARSA algorithm helps manage the allocation of powers used by each allocated user. Figure 1 illustrates the simulation flowchart for this work.

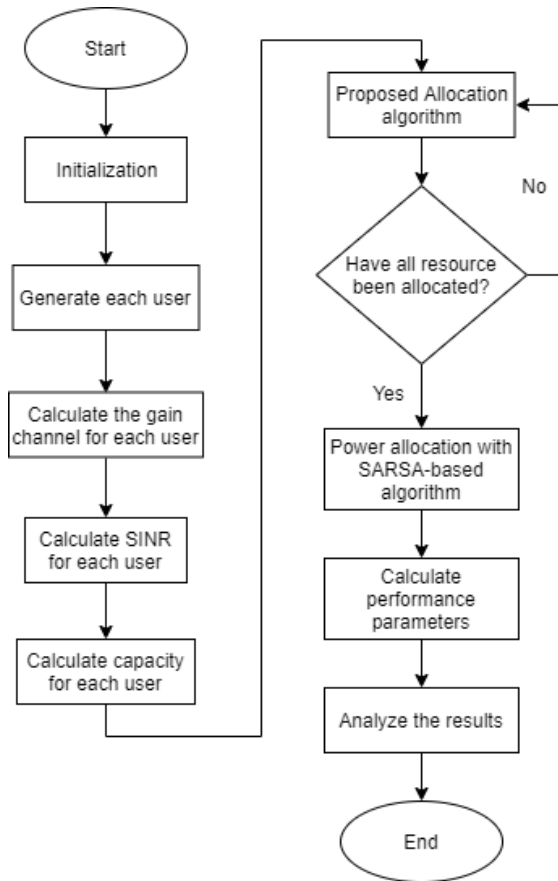


Fig. 1. Simulation Flowchart

2.1 System Model

The system model uses a single cell in which there are Cellular User Equipment (CUE), D2D pairs consisting of a D2D transmitter (D2D TX) and a D2D receiver (D2D RX), and a eNodeB (eNB) device. The transmission direction used is the uplink, where the signal moves from the CUE to eNB. The D2D pair

requires resources from CUE to communicate. Each D2D pair can occupy one resource that is owned by CUE. Because the use of the same resource, interference will occur on D2D RX which comes from CUE. Besides that, interference will occur on the eNB which comes from D2D TX. The system model used be further clarified through Figure 2.

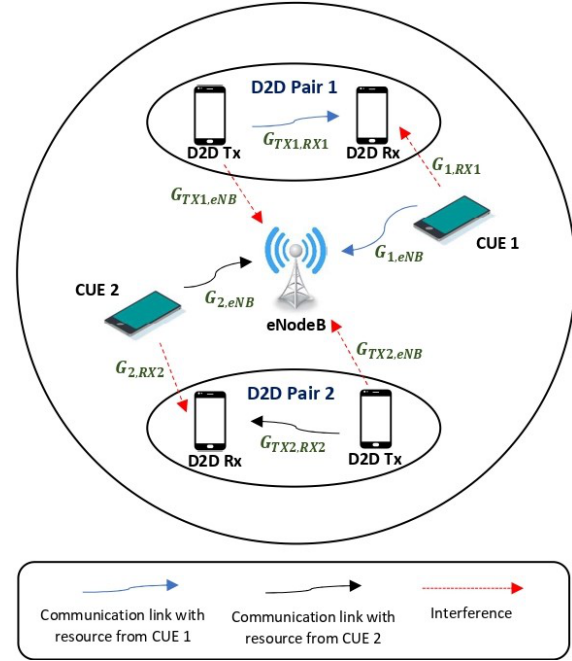


Fig. 2. System Model

The following formula can calculate the gain of a channel [4] :

$$G_{i,j} = PL_{i,j} + X_{\sigma} + \Pi \quad (1)$$

where $PL_{i,j}$ represents the pathloss described in Equation 2 [5], X_{σ} represents the random variable for Log-normal shadowing, and Π represents the random variable for Rayleigh fading.

$$PL_{i,j} = 36.7 \log_{10}(d) + 22.7 + 26 \log_{10}(f_c) \quad (2)$$

where d represents the distance between devices in meter and f_c represents the working frequency in the system in GHz.

SINR is an essential factor in determining channel quality. The following formula can calculate SINR on eNB [6] :

$$\gamma_{i,j} = \frac{P_i \cdot G_{i,eNB}}{N_o + P_j \cdot G_{jTX,eNB}} \quad (3)$$

where i represents CUE, j represents D2D pair, $\gamma_{i,j}$ represents SINR on eNB, $G_{i,eNB}$ represents the channel gain of CUE to eNB, $G_{jTX,eNB}$ represents the channel gain of D2D TX to eNB, N_o represents noise and P_i , P_j represents the CUE transmit power and D2D TX transmit power.

Aside from the eNB, the following formula can calculate SINR on the D2D side [6] :

$$\delta_{i,j} = \frac{P_j \cdot G_{jTX,jRX}}{N_o + P_i \cdot G_{i,jRX}} \quad (4)$$

where $\delta_{i,j}$ represents SINR on D2D, $G_{jTX,jRX}$ represents the channel gain of D2D TX to D2D RX and $G_{i,jRX}$ represents the channel gain of CUE to D2D RX.

2.2 Proposed Algorithm

This work will compare performance parameter using joint greedy algorithm with SARSA power control scheme, greedy algorithm with SARSA power control scheme, joint greedy algorithm, and greedy algorithm.

2.2.1 Joint Greedy Algorithm

Joint greedy algorithm is a development of the greedy algorithm. There are two phases in the allocation process in this algorithm. In the first phase, the allocation process sees the largest capacity of eNodeB, which will then become the source of the allocation process in the second phase. In the second phase, The allocation process sees the largest capacity of D2D based on the results of allocation process in the first stage. The D2D pair will receive the largest capacity resource allocation at these two phases. The other D2D pair can not use resources that were already used by the previous D2D pair. This allocation will continue until all D2D pairs get resources from CUE. The flowchart of joint greedy algorithm is described in Figure 3 [7].

2.2.2 Greedy Algorithm

Greedy algorithm is a resource scheduling algorithm in D2D communication, where the resource is only allocated to the user with the best channel status. The order of the scheduling process is determined based on the order in which the user arrives. The user in the first order can choose the best channel to use. The other D2D pair can not use resources that were already used by the previous D2D pair. This allocation will continue until all the D2D pair get resources from CUE. The flowchart of greedy algorithm is described in Figure 4 [7].

2.2.3 SARSA Algorithm

State Action Reward State Action Algorithm or abbreviated as SARSA, is a reinforcement learning algorithm for power control scheme in D2D communication [4]. This algorithm takes the channel gain obtained from the D2D pair based on the channel condition to calculate the allocated power for each D2D user after receiving the resource allocation. The power

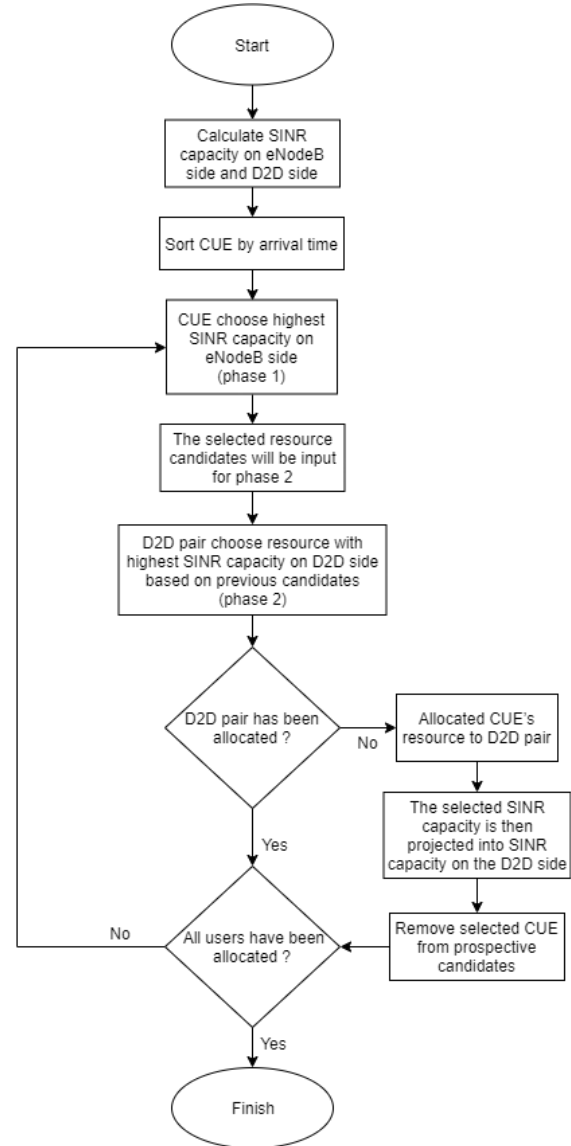


Fig. 3. Flowchart of joint greedy algorithm

allocation process is carried out after ensuring that all resources have been allocated using the resource allocation algorithm to calculate the transmit power of each D2D pair. The following formula can calculate the transmit power for each D2D pair [4].

$$\tau_j = \frac{G_{jRX,jTX}}{\sum_{j=1}^D G_{jRX,jTX}} \cdot P_D \cdot D \quad (5)$$

where τ_j represents the transmit power on D2D TX after resource allocation, P_D represents the transmit power on D2D TX before resource allocation, and D represents the number of D2D users.

D2D allocated transmit power used in the SARSA algorithm for power control scheme. The power control scheme is carried out by looking at the reward value (R), which is the difference between D2D transmit power before resource allocation and D2D trans-

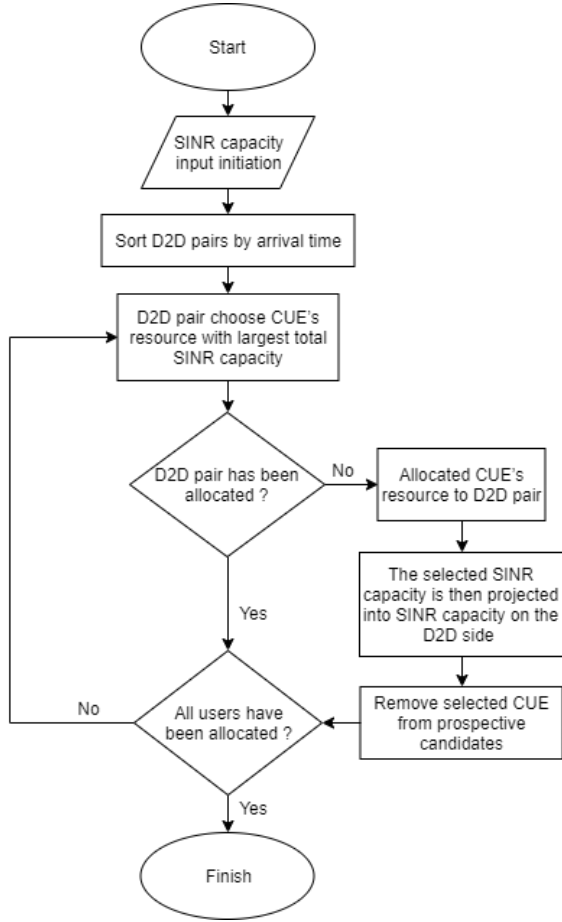


Fig. 4. Flowchart of greedy algorithm

mit power after resource allocation. D2D transmit power will increase if the reward value is positive but decrease if the reward value is negative. The reward will be recalculated with the transmit power of D2D to achieve a new balanced D2D transmit power (P_N). Table 1 describes in more detail the power control scheme using SARSA algorithm [9].

2.3 Simulation Parameter

The proposed algorithm is analyzed and simulated by Python. Simulation with greedy and joint greedy resource allocation without power control scheme will be compared to greedy and joint greedy resource allocation with power control scheme. Table 2 shows the simulation parameters used.

Performance parameters used in this simulation are energy efficiency, spectral efficiency, and D2D fairness index.

Energy efficiency is the number of bits that can be sent from the transmitter every second in the total power allocated to the system. The following formula can calculate energy efficiency [8].

$$\alpha_{i,j} = \frac{\sum_{i=1}^I \sum_{j=1}^J \mu_{i,j}}{P_{i,j}} \quad (6)$$

Table 1: SARSA Algorithm

Power Control Scheme

- 1: Initialize all parameters in Table 2
- 2: **loop**
- 3: Calculate the channel gain based on pathloss (Eq.1)
- 4: Calculate the transmitted power on D2D pair (Eq.5)
- 5: $R = \tau_j - P_D$
- 6: Calculate the new transmitted power for D2D pair (P_N)
- 7: $P_N = R + P_D$
- 8: **end loop**

where I represents the number of CUE, J represents the number of D2D pair, $\alpha_{i,j}$ represents energy efficiency, $\mu_{i,j}$ represents sumrate at which the CUE i distributes resources among D2D pairs j and $P_{i,j}$ represents total transmit power.

Spectral efficiency explains how much bits per second (bps) that can be sent in each hertz. The following formula can calculate spectral efficiency [7].

$$\beta_{i,j} = \frac{\sum_{i=1}^I \sum_{j=1}^J \mu_{i,j}}{rb.B} \quad (7)$$

where $\beta_{i,j}$ represents spectral efficiency, rb represents the number of resource blocks in the system, and B represents resource bandwidth.

D2D Fairness is equal distribution of sumrates for each D2D pair obtained. The following formula can calculate the D2D fairness index [10].

$$FI = \frac{\left(\sum_{i=1}^I \sum_{j=1}^J \mu_{i,j} \right)^2}{n \cdot \sum_{i=1}^I \sum_{j=1}^J (\mu_{i,j})^2} \quad (8)$$

where n represents the number of D2D user in the system.

3 Results and Discussion

3.1 Performance Result of Energy Efficiency

Comparison performance results of each algorithm in energy efficiency shown in Figure 5. Based on Figure 5, the addition of the D2D user into the system with power control scheme using SARSA algorithm increases energy efficiency in the system. The increase in energy efficiency is due to an increase in data rate due to user's addition so that SINR increases.

Table 3 shows the value of energy efficiency for each algorithm and it shows that the higher energy

Table 2: Simulation Parameter

Parameter	Value
TTI	900
Number of CUE	50 users
Number of D2D pairs	10,20,30,40,50 users
Number of resources	50
CUE transmit power	0.5 Watt
D2D transmit power	0.1 Watt
Cell radius	500 m
D2D radius	40 m
Frequency	1.8 GHz
Resource bandwidth	180 KHz
Pathloss model	UMi pathloss
Channel model	Rayleigh model with $\mu = 1$ and $\sigma = 1$
Shadowing model	Lognormal model with $\mu = 0$ and $\sigma = 1$

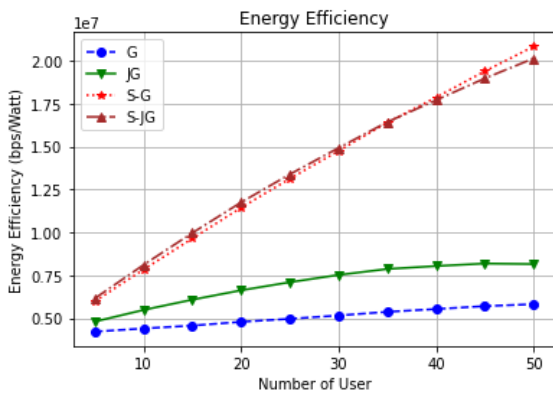


Fig. 5. Energy Efficiency Performance Result

efficiency takes place in joint greedy algorithm with power control scheme using SARSA algorithm.

Based on Table 3, the highest energy efficiency value occurs in the joint greedy algorithm with power control scheme using SARSA algorithm with 1.375×10^7 bps/Watt. This value is almost the same as greedy algorithm with power control scheme using SARSA algorithms with increase only 0.003×10^7 , increase 0.677×10^7 than joint greedy algorithm, and increase 0.871×10^7 than greedy algorithm. It happens because joint greedy algorithm reviews resourced allocation from the largest capacity of CUE and D2D side. Coupled with the allocation of power using the SARSA algorithm also helps increase power consumption effectiveness for each user in the system.

Table 3: Energy Efficiency for each algorithm

Algorithm	Energy Efficiency (bps/Watt)
SARSA-Joint Greedy (S-JG)	1.375×10^7
SARSA-Greedy (S-G)	1.372×10^7
Joint Greedy (JG)	0.698×10^7
Greedy (G)	0.504×10^7

3.2 Performance Result of Spectral Efficiency

Comparison performance results of each algorithm in spectral efficiency shown in Figure 6. Based on Figure 6, the addition of the D2D user to the system with the power control scheme process using SARSA algorithm also positively impacts the spectral efficiency value. The value of spectral efficiency depends on the value perceived by the user.

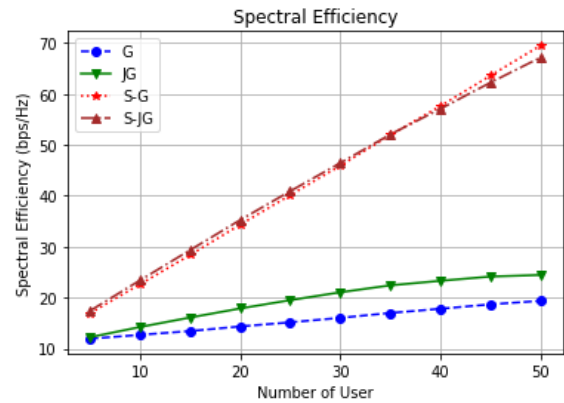


Fig. 6. Spectral Efficiency Performance Result

Table 4 shows the value of spectral efficiency for each algorithm and it shows that the higher spectral efficiency takes place in joint greedy algorithm with power control scheme using SARSA algorithm.

Table 4: Spectral efficiency for each algorithm

Algorithm	Spectral Efficiency (bps/Hz)
SARSA-Joint Greedy (S-JG)	43.105
SARSA-Greedy (S-G)	43.066
Joint Greedy (JG)	19.524
Greedy (G)	15.639

Based on Table 4, the highest spectral efficiency value also occurs in the joint greedy algorithm with power control scheme using SARSA algorithm with

43.105 bps/Hz. This value is almost the same as greedy algorithm with power control scheme using SARSA algorithms with increase only 0.039, increase 23.582 than joint greedy algorithm, and increase 27.466 than greedy algorithm. It happens because joint greedy algorithm with power control scheme using SARSA algorithm provides a higher average system data rate than the other algorithms.

3.3 Performance Result of D2D Fairness Index

Comparison performance results of each algorithm in D2D fairness index shown in Figure 7. Based on Figure 7, the addition of the D2D user to the system with the power allocation process using the SARSA algorithm also positively impacts the fairness value. The value of fairness depends on the user's value and the number of users in the system. The more users in the system, the lower the fairness index.

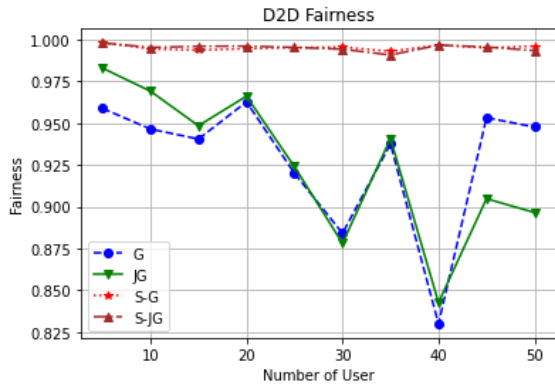


Fig. 7. D2D Fairness Index Performance Result

Table 5 shows the value of D2D fairness index for each algorithm and it shows that the higher D2D fairness index takes place in greedy algorithm with power control scheme using SARSA algorithm, while joint greedy algorithm with power control scheme using SARSA algorithm takes the second place.

Table 5: D2D fairness index for each algorithm

Algorithm	D2D Fairness Index
SARSA-Joint Greedy (S-JG)	0.993
SARSA-Greedy (S-G)	0.996
Joint Greedy (JG)	0.896
Greedy (G)	0.928

Based on Table 5, D2D fairness index of joint greedy algorithm with power control scheme using the SARSA algorithm is 0.993. This value is almost the same as greedy algorithm with power control scheme

using SARSA algorithms with decrease only 0.003, increase 0.097 than joint greedy algorithm, and increase 0.065 than greedy algorithm. This happens because resources allocation in greedy algorithm with power control scheme using SARSA algorithm is allocated based on combination between the largest capacity of eNodeB and the D2D pair. Meanwhile, resource allocation in joint greedy algorithm with power control scheme using SARSA algorithm is allocated based on the largest capacity of eNodeB, than the largest capacity of D2D.

3.4 Overall Performance Analysis

Table 6 shows the performance comparison between joint greedy algorithm with power allocation using SARSA algorithm with three other allocation schemes.

Table 6: Comparison of performance parameters

Algorithm	Energy Efficiency (bps/Watt)	Spectral Efficiency (bps/Hz)	D2D Fairness Index
SARSA-Joint Greedy	1.375×10^7	43.105	0.993
SARSA-Greedy	$-(0.003 \times 10^7)$	$-(0.039)$	$+(0.003)$
Joint Greedy	$-(0.677 \times 10^7)$	$-(23.582)$	$-(0.097)$
Greedy	$-(0.871 \times 10^7)$	$-(27.466)$	$-(0.065)$

Based on Table 6, resource allocation scheme using joint greedy algorithm with power allocation using SARSA algorithm has excellent performance, both in energy efficiency, spectral efficiency, and D2D fairness. This happens because the SARSA algorithm uses D2D power effectively according to channel conditions and D2D transmit power after resource allocation, resulting in a new D2D transmit power that is more fair and effective to increase the value of SINR capacity and performance in the system. However, power allocation using the SARSA algorithm is only carried out on the D2D side, so using SARSA algorithm on the eNB side does not have a significant impact on system performance for the eNB side.

4 Conclusion

In this work, the resource allocation scheme uses two algorithms, namely greedy and joint greedy, with SARSA algorithm as the power control scheme. The simulation uses the number of variations of the D2D pairs in the system. Based on the simulation results, by allocating resources using the joint greedy algorithm and allocating power using the SARSA algorithm have parameter results, 1.375×10^7 bps/Watt in energy efficiency, 43.105 bps/Hz in spectral efficiency, and 0.996 in D2D fairness index. It happens because joint greedy algorithm allocates the resource based on two phases and SARSA algorithm as the power control scheme allocates power to the users based on channel con-

dition and D2D transmit power which makes better performance parameters of the system

References

- [1] Esmat, Haitham H., Mahmoud M. Elmesalawy, and Ibrahim I. Ibrahim. "Adaptive resource sharing algorithm for device-to-device communications underlying cellular networks." *IEEE Communications Letters* 20.3 (2016): 530-533.
- [2] Hajiaghajani, Foad, and Mehdi Rasti. "An adaptive resource allocation scheme for device-to-device communication underlying cellular networks." 2015 *IEEE/CIC International Conference on Communications in China (ICCC)*. IEEE, 2015.
- [3] Luo, Yong, et al. "Dynamic resource allocations based on Q-learning for D2D communication in cellular networks." 2014 11th international computer conference on wavelet active media technology and information processing (ICCWAMTIP). IEEE, 2014.
- [4] Khan, Muhidul Islam, et al. "Cooperative reinforcement learning for adaptive power allocation in device-to-device communication." 2018 *IEEE 4th World Forum on Internet of Things (WF-IoT)*. IEEE, 2018.
- [5] Sousa, Sofia, Fernando J. Velez, and Jon M. Peha. "Impact of propagation model on capacity in small-cell networks." 2017 *International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS)*. IEEE, 2017.
- [6] Song, Xin, et al. "Joint uplink and downlink resource allocation for D2D communications system." *Future Internet* 11.1 (2019): 12.
- [7] Prabowo, V. S. W., et al. "Joint-Greedy Allocation Algorithm on D2D Communication Underlying Networks." 2019 *IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob)*. IEEE, 2019.
- [8] Xu, Hao, et al. "Energy-efficient resource allocation in D2D underlaid cellular uplinks." *IEEE Communications Letters* 21.3 (2016): 560-563.
- [9] Sutton, Richard S., and Andrew G. Barto. *Reinforcement learning: An introduction*. MIT press. (2018).
- [10] Jain, Rajendra K., Dah-Ming W. Chiu, and William R. Hawe. "A quantitative measure of fairness and discrimination." Eastern Research Laboratory, Digital Equipment Corporation, Hudson, MA (1984).



Astri Wulandari was born in Jambi, Indonesia, in 1999. Astri is pursuing a Bachelor's Degree in Telecommunication Engineering at the School of Electrical Engineering, Telkom University. Astri is a research assistant in Mobile Communication Laboratory and a practicum assistant in Electrical Circuit Laboratory. Her research interests include wireless communications and 5G developments.



Nachwan Mufti Adrian-syah received the B.Eng degree from STT Telkom Bandung in 1998, received the M.Eng degree from Institut Teknologi Bandung in 2005, and received the D.Eng degree from Universitas Indonesia in 2016, respectively. He is currently a Senior Lecturer with the School of Electrical Engineering, Telkom University, Indonesia, at Transmission of Telecommunication Research Division. His research interest include electromagnetic theory and applications, optical, and wireless communications.



Vinsensius Sigit Widhi Prabowo was born in Jakarta (Indonesia), in 1993. He received the B.S. degree in Telecommunication engineering from Telkom University, Bandung, Indonesia, in 2015, M.S. degree in Electrical engineering from the Telkom University, Bandung, Indonesia, in 2017. Since 2016, he is Tenured Lecturer with the School of Electrical Engineering, Telkom University. He is the author of *Pengalokasian Sumber Daya Radio pada Sistem Komunikasi Pita Lebar*, and more than 20 indexed articles. His research interests include wireless communication system, radio resources management, and Telecommunication Transmission.