

Connectivity Mode Management for User Devices in Heterogeneous D2D Networks

S. Sinem Kafiloğlu, Gürkan Gür and Fatih Alagöz

Abstract—Device-to-Device (D2D) paradigm is one of the key enablers for boosting the energy efficiency (EE) of networks. For the actualization of EE in multi-mode heterogeneous D2D networks, resource allocation should be executed accordingly. In that regard, we study resource allocation management in heterogeneous D2D networks where there is a LEO satellite and a cellular base station (BS) with multi-mode user devices (satellite, BS and D2D connectivity modes) dispersed in the terrestrial cell. We formulate the EE optimization of mode selection constrained by a designated total network service capacity threshold. We also implement a sub-optimal pattern search algorithm (PSA) with the aim of enhancing EE while keeping total system service capacity at an acceptable level and compare it to alternative resource management schemes. Based on our simulation results, PSA achieves the most beneficial EE value in contrast to other schemes while still managing to meet the goodput threshold. Additionally, we discuss the multi-mode nature of our heterogeneous network from EE and goodput perspectives.

I. INTRODUCTION

Heterogeneous networks consist of different network types (e.g. satellite, cellular, personal-area networks) that co-exist together for improving network efficiency and user experience. They enable flexible services and connectivity according to dynamic system needs. The customization ability of such networks stems from the capability of the component integration/disintegration and mode selection mechanism. For emerging and future networks like 5G and Beyond 5G, Device-to-Device (D2D) paradigm is becoming more widely utilized for the sake of EE improvement [1], [2]. In this context, we study the optimal connectivity mode management in *heterogeneous D2D* networks in this work. Specifically, we have three service modes in the investigated network: *i) D2D*, *ii) cellular* and *iii) satellite*. The satellite mode can essentially provide 5G services to locations under inadequate cellular network support or with no coverage at all due to cost concerns [3]. Additionally, the satellite has a significant caching-and-broadcasting gain due to its cache capability and coverage [4]. However, the satellite link is more prone to channel impairments such as phase noise, nonlinearity, and tropospheric effects [5], and has capacity saturation problem. On the other hand, D2D mode has a potential to improve service rate and boost network throughput [6]. Besides, it is preferable in terms of EE if properly managed [7]. But devices have limited cache capacity with intermittent cooperation opportunities, and to overcome

this issue we can make use of the cellular counterpart that has greater cache capacity. Cellular service links have longer transmission range than D2D links. They are also more stable due to one non-mobile transmission end-point (BS). However, BS communications suffer more transmission energy consumption in contrast to the D2D approach due to large BS power and longer service durations [8]. When we look at all these contradicting factors, we need to analyze the *mode selection* rigorously for a complete trade-off analysis in terms of system service capacity and EE.

In our previous study [8], we built a Markovian model to monitor service characteristics of a cellular D2D network integrated with a satellite. We essentially focused on the integration of universal source concept and the usage of overlaying in D2D mode. Now, as a contribution, we formulate an optimization problem in these heterogeneous D2D networks. We look for the mode weight assignment that achieves improved EE constrained by an acceptable system capacity. We analyze the complexity of our optimization problem and propose a sub-optimal mode assignment scheme by pattern search algorithm (PSA) as another novel extension. We compare the PSA service capacity and EE results to alternative mode selection approaches and show resulting EE improvement.

II. SYSTEM MODEL

In our network, users have hybrid-mode devices that can operate in *a) primary* and *b) secondary* modes in *satellite* and *terrestrial* links, respectively. These users are called hybrid users (HUs). The satellite is of type LEO and the corresponding link has $N_{f_{sat}}$ frequencies while the terrestrial link has $N_{f_{ter}}$. In the terrestrial segment, our users can get service via D2D communications or by the conventional cellular connectivity. For improving the system performance in terms of EE, in D2D mode overlaying is used with at most D_{max} concurrent operations at dedicated terrestrial frequencies. BS mode operates at terrestrial frequencies except for D2D dedicated ones. Without loss of generality, we take the D2D mode dedicated terrestrial frequency as f_1 . The frequency selection reason is provided in our previous work [8]. The universal source concept is introduced to be able to serve non-present contents from the original servers. The operation modes, transitions and derived EE and goodput functions are elaborated in detail in our previous studies [8], [9]. A summary of our system model with mode details is provided in Fig. 1.

In our optimization formulation, we utilize the expected consumed energy per successfully received bits $EPB_{HU}(r)$ and overall system goodput $G_{HU}(r)$ functions in our previous work [8]. In this section, we recapture the Markov model and

S. Sinem Kafiloğlu and Fatih Alagöz are with Bogazici University, Dept. of Computer Engineering, 34342, Istanbul, TR. e-mail:{sinem.kafiloglu, fatih.alagoz}@boun.edu.tr

Gürkan Gür is with Zurich University of Applied Sciences (ZHAW), Institute of Applied Information Technology, Switzerland. e-mail:gueu@zhaw.ch

This work was supported by the Scientific and Technical Research Council of Turkey (TUBITAK) under Grant 116E245.

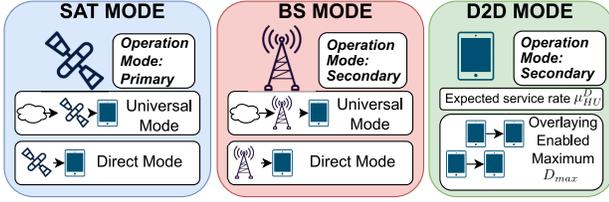


Fig. 1. System model and different connectivity modes.

the mode selection while the essence of the mode weight vector \mathbf{r} in our system is revealed. In addition, we present the definitions of $EPB_{HU}(\mathbf{r})$ and $G_{HU}(\mathbf{r})$ functions.

The channel state utilized in our Markov model is defined as $(i_{HU}^{sat}, i_{HU}^{sat(u)}, i_{PU}^{ter(f_1)}, i_{HU}^{BS}, i_{HU}^{BS(u)}, i_{PU}^{ter(f_1)}, i_{HU}^{D(f_1)}, i_{HU}^{sat}, i_{HU}^{sat(u)}, i_{PU}^{ter(f_1)}, i_{HU}^{BS}, i_{HU}^{BS(u)}, i_{PU}^{ter(f_1)}, i_{HU}^{D(f_1)})$. i_{HU}^{sat} and $i_{HU}^{sat(u)}$ are the number of active HU services from the satellite cache and from the universal source through the satellite, respectively. $i_{PU}^{ter(f_1)}$ gives the number of active primary users (PUs) in the BS occupied terrestrial link while i_{HU}^{BS} and $i_{HU}^{BS(u)}$ are the number of active HU services from the BS cache and from the universal source across the BS, respectively. The indicator for PU existence that is blocking D2D is $i_{PU}^{ter(f_1)}$ and $i_{HU}^{D(f_1)}$ defines the number of actively operating HUs in D2D mode simultaneously. At the mode selection phase, we first check the content availability at system units (e.g. satellite, a device in the vicinity, etc.). We observe the channel state and available frequencies to select the accessed frequencies according to the mode weight vector $\mathbf{r} = [r_{sat}, r_{BS}, r_{dev}]$. This vector contains element r_{sat} showing the weight factor per each idle satellite mode frequency, r_{BS} the weight factor per each idle BS mode frequency and r_{dev} the weight factor for D2D dedicated frequency as long as the maximum overlaying limit D_{max} is not reached. e.g. assume a requested content is available in the BS cache and in some device in the reception range with two idle BS frequencies and D2D frequency is not reached to D_{max} . Then with probability $r_{dev}/(r_{dev} + 2 \cdot r_{BS})$, D2D mode is selected. In that regard, the \mathbf{r} vector determines all HU arrival transitions. With a PU/HU starting a content reception (arrival), we increment the corresponding number of users in the selected mode. For instance, when a device starts reception in D2D mode, we increment $i_{HU}^{D(f_1)}$ by one as long as it does not reach the maximum allowable concurrent D2D services D_{max} . After a complete content retrieval, the user departs the channel and the corresponding indicator of that channel is decreased by one. With the total set of arrival and departure transitions (further transition rate details provided in our previous work [9]), we calculate the channel steady state probabilities π_x 's. The state-probability weighted transition rates yield the effective arrival rates to different service modes (e.g., $\lambda_{eff}^{D2D}(HU) := \sum_{x \in S} \Gamma_{HU}^{D(f_1)}(x) \pi_x$ is the effective arrival rate to D2D mode with $\Gamma_{HU}^{D(f_1)}(x)$ being the expected arrival rate to D2D mode at channel state x .) The dropping probability of HUs, p_{drop}^{BS} in BS mode and p_{drop}^{D2D} in D2D mode, and the probability of an HU having a local hit p_{local} are calculated by utilizing the channel model as well. We refer the readers to our previous work [9] for detailed derivations in that regard.

To find the total system service rate, we calculate the service rate of different operation modes as elements in (1) with the definitions provided in Table I. By multiplying the effective arrival to a service mode with its non-dropping rate and the

mean multimedia content size $\overline{s(c)}$, we get the service rate of each mode. For instance, $Th_{HU}^D := \lambda_{eff}^{D2D}(HU) \cdot (1 - p_{drop}^{D2D}) \cdot \overline{s(c)}$ gives the total service rate of D2D mode operations. Our HUs operate in primary mode over the satellite link, thereby no dropping occurs due to its prioritized nature. Also for the local services, once a service starts, it does not cease the operation.

$$G_{HU}(\mathbf{r}) := G_{HU}^{local}(\mathbf{r}) + Th_{HU}^{sat}(\mathbf{r}) + Th_{HU}^{sat(u)}(\mathbf{r}) + Th_{HU}^{BS}(\mathbf{r}) + Th_{HU}^{BS(u)}(\mathbf{r}) + Th_{HU}^D(\mathbf{r}) \quad (1)$$

For the total system power consumption, we calculate the power consumption of different modes as subfunctions in (2). The satellite is solar-powered, so it is not included in (2). The power consumption in a mode (apart from local hits) is the addition of power consumption for dropped and non-dropped services. The former power consumption is the multiplication of the effective arrival rate to that mode, the drop rate and the corresponding mean energy consumption (by taking the half of a complete service in that mode). P_{dev}^{tx} is the HU transmission power and μ_{HU}^D is the D2D mode HU service rate. e.g. $P_{D2D}^d := \lambda_{eff}^{D2D}(HU) \cdot p_{drop}^{D2D} \cdot \frac{P_{dev}^{tx}}{2 \cdot \mu_{HU}^D}$ is the D2D mode power consumption for dropped services. For the non-dropped, we consider complete service durations for the energy consumption calculation and take the non-dropped arrivals. $P_{D2D}^{n-d} := \lambda_{eff}^{D2D}(HU) \cdot (1 - p_{drop}^{D2D}) \cdot \frac{P_{dev}^{tx}}{\mu_{HU}^D}$ is the D2D mode power consumption for non-dropped services. Their sum $P_{D2D} := P_{D2D}^d + P_{D2D}^{n-d}$ gives the power consumption of D2D mode. Similar to the service capacity calculations, no local hit drops after a local service starts its operation. With all the components, the total energy consumption is

$$P_{all}(\mathbf{r}) := P_{loc}(\mathbf{r}) + P_{BS}(\mathbf{r}) + P_{BS(u)}(\mathbf{r}) + P_{D2D}(\mathbf{r}) \quad (2)$$

Finally, we get the expected energy consumed per successfully transmitted bit $EPB_{HU}(\mathbf{r}) := P_{all}(\mathbf{r})/G_{HU}(\mathbf{r})$.

TABLE I
DEFINITION OF FUNCTIONS.

Function	Definition
$G_{HU}(\mathbf{r})$	The total system service rate (goodput)
$G_{HU}^{local}(\mathbf{r})$	The total local hit rate
$Th_{HU}^{sat}(\mathbf{r})$	The total service rate directly from the satellite cache
$Th_{HU}^{sat(u)}(\mathbf{r})$	The total service rate from the universal source across the satellite
$Th_{HU}^{BS}(\mathbf{r})$	The total service rate directly from the BS cache
$Th_{HU}^{BS(u)}(\mathbf{r})$	The total service rate from the universal source across the BS
$Th_{HU}^D(\mathbf{r})$	The total service rate of D2D mode operations
$P_{all}(\mathbf{r})$	The total power consumption in the system
$P_{loc}(\mathbf{r})$	The power consumption of local hits
$P_{BS}(\mathbf{r})$	The BS power consumption for transmissions directly from its cache
$P_{BS(u)}(\mathbf{r})$	The BS power consumption for transmissions from the universal source across the BS
$P_{D2D}(\mathbf{r})$	The power consumption of D2D mode operations
$EPB_{HU}(\mathbf{r})$	The expected energy consumed per successfully received bits

III. CONNECTIVITY MODE ASSIGNMENT

In our connectivity mode managing resource allocation (RA), we utilize system unit caches (the cache of requester device, satellite, BS and the caches of other HU devices in predetermined vicinity) and channel states. As the main control, we have a mode weight vector $\mathbf{r} := [r_{sat}, r_{BS}, r_{dev}]$ consisting of r_{sat} , r_{BS} and r_{dev} elements designating the weight factor for each idle satellite mode operating frequencies

(r_{sat}), for each idle BS mode operating frequencies (r_{BS}) and for D2D mode operating frequency as long as its maximum overlaying limit D_{max} is not reached (r_{dev}). The assigned values of \mathbf{r} are used to configure the selection rate of different network modes which in turn determines the channel steady state probabilities, effective arrival rates, dropping probabilities and local hit probability. Accordingly, $EPB_{HU}(\mathbf{r})$ and $G_{HU}(\mathbf{r})$ are determined as explained in Section II regarding the RA configuration. In this paper, we find the sub-optimal assignment of \mathbf{r} such that the energy consumed per successfully transmitted bit EPB_{HU} decreases (the decrease in EPB_{HU} contributes to EE) while overall goodput G_{HU} is kept above a threshold. Our optimization problem is as follows:

$$\begin{aligned} \min_{\mathbf{r}} \quad & EPB_{HU}(\mathbf{r}) \\ \text{s.t.} \quad & G_{HU}^{th} \leq G_{HU}(\mathbf{r}) \\ & r_{sat} + r_{BS} + r_{dev} = 1 \\ & 0 \leq r_x \leq 1, \quad r_x \in \mathbb{R}, \quad x \in \{sat, BS, dev\} \end{aligned}$$

The space complexity for this problem is calculated by counting the number of all channel states provided in Section II. The upper bound of active satellite mode HU services are N_{fsat}^2 . At non-D2D mode operating terrestrial frequencies, the upper bound of active operations for both HU (BS mode) and PU services is $(N_{fter} - \alpha)^3$ with α being the number of dedicated terrestrial frequencies for D2D mode. Finally, $(D_{max})^\alpha$ gives the maximum number of concurrently operating D2D services. Hence, $O((D_{max})^\alpha (N_{fter} - \alpha)^3 N_{fsat}^2)$ is the space complexity of channel states. We take $\alpha = 1$, so this complexity is $O(D_{max}(N_{fter} - 1)^3 N_{fsat}^2)$. Solving the system of balance equations is $O(n^3)$ with total n channel states.

To determine the optimization approach based on our problem definition, we inspect the linearity and convexity of our objective function $EPB_{HU}(r)$. A function is linear if it satisfies *i) superposition principle* and *ii) homogeneity*. When we look at our objective, $EPB_{HU}(0 \cdot r) = 1.96$ nJbp $\neq 0 \cdot EPB_{HU}(r)$ for any r . Property-*ii* is not satisfied and our objective function is non-linear. Next, we delve into the convexity examination. A function $f: \mathbb{R}^N \rightarrow \mathbb{R}$ is convex if for $\forall x, y \in \mathbb{R}^N$ and $\forall c \in [0, 1]$, $f(c \cdot x + (1 - c) \cdot y) \leq c \cdot f(x) + (1 - c) \cdot f(y)$ is satisfied. We investigate $r_1 = [0.3, 0.2, 0.5]$, $r_2 = [0.3, 0.3, 0.4]$ and $c = 0.2$ scenarios. In our system, we consider weight vector scenarios with the largest D2D mode weight r_{dev} (0.5 and 0.4) among all weights since our preliminary studies [8], [9] have presented that D2D communication is a low energy consuming operation with considerably good channel conditions and hence service capacity. Since the satellite is solar-powered, we consider scenarios with the satellite mode weight r_{sat} (0.3) as large as or larger than the BS mode weight r_{BS} (0.2 and 0.3). In these realistic scenarios, our objective does not satisfy the convexity condition with $EPB_{HU}(r_1) = 181.1$ nJpb and $EPB_{HU}(r_2) = 186.6$ nJpb. $EPB_{HU}(0.2 \cdot r_1 + (1 - 0.2) \cdot r_2) = 185.57$ nJpb $\not\leq 0.2 \cdot EPB_{HU}(r_1) + (1 - 0.2) \cdot EPB_{HU}(r_2) = 185.50$ nJpb. By showing at least one instance not holding convexity property, our objective function is proven to be non-convex. Hence, our optimization problem is of non-convex non-linear programming (NLP) type. It is NP-hard and hence

to solve it in feasible time, we employ a heuristic for finding a sub-optimal solution, namely *Pattern Search Algorithm (PSA)*.

Algorithm 1 PSA Polling Mechanism

```

1 % Construct a sequence  $\{r_0, r_1, \dots, r_n\}$  with decreasing  $EPB(r_x)$  values
2  $r_0 \leftarrow r_{in}$ ; % Set initial point as the first sequence point  $r_0$ 
3  $r_{prev} \leftarrow r_0$ ;
4  $epb_{prev} \leftarrow EPB(r_{prev})$ ;
5  $r_{cur} \leftarrow r_0$ ; % Pin it as current sequence point
6  $s_{mesh} \leftarrow s_{in}$ ; % Set the mesh size  $s_{mesh}$  to the initial
7  $s_{prev} \leftarrow 0$ ;
8 while (1) do
9    $found = 0$ ; % Denotes if a new sequence point is found
10  % Basis vectors  $b_0 = [1 \ 0 \ 0]$ ,  $b_1 = [0 \ 1 \ 0]$ ,  $b_2 = [0 \ 0 \ 1]$ ,  $b_3 =$ 
11  %  $[-1 \ 0 \ 0]$ ,  $b_4 = [0 \ -1 \ 0]$ ,  $b_5 = [0 \ 0 \ -1]$ .
12  for ( $i = 0:1:5$ ) do
13     $p_i \leftarrow s_{mesh} \cdot b_i$ ; % Calculate pattern vector  $p_i$ 
14     $m_i^{cur} \leftarrow p_i + r_{cur}$ ; % Then get current mesh point  $m_i^{cur}$ 
15     $epb_i^{cur} \leftarrow EPB(m_i^{cur})$ ; % Compute its objective function value
16
17    % If current mesh point has decreasing EPB value
18    if ( $epb_i^{cur} < epb_{prev}$ ) then
19       $found \leftarrow 1$ ; % New sequence point is found
20       $\delta_{obj} \leftarrow epb_{prev} - epb_i^{cur}$ ;
21       $\delta_{const} \leftarrow \|G(r_{prev}) - G(m_i^{cur})\|$ ;
22       $\delta_{mesh} \leftarrow \|s_{prev} - s_{mesh}\|$ ;
23       $r_{cur} \leftarrow m_i^{cur}$ ; % Set  $m_i^{cur}$  as the new current sequence point
24       $[r_0, \dots, r_{prev}, r_{cur}] \leftarrow \text{Insert}([r_0, \dots, r_{prev}], r_{cur})$ ;
25       $r_{prev} \leftarrow r_{cur}$ ; % Mark the current sequence point as previous
26       $epb_{prev} \leftarrow epb_i^{cur}$ ; % Mark its objective function value
27       $s_{prev} \leftarrow s_{mesh}$ ;
28       $s_{mesh} \leftarrow f_c \cdot s_{mesh}$ ; % Increase mesh size by expansion factor
29      break;
30
31  % If no new sequence point is found
32  if ( $found == 0$ ) then
33     $s_{prev} \leftarrow s_{mesh}$ ;
34     $s_{mesh} \leftarrow f_c \cdot s_{mesh}$ ; % Decrease mesh size by contraction factor
35
36  else
37    % Stop PSA if change in objective, constraint or mesh size is less
38    % than corresponding tolerance
39    if ( $(\delta_{obj} < T_{obj}) \vee (\delta_{const} < T_{const}) \vee (\delta_{mesh} < T_{mesh})$ ) then
40      break;
41
42 return  $r_{cur}$ ; % Return the point with lowest EPB value

```

PSA is a derivative-free (black-box) search method [10]. We inspect the constraint *overall goodput* G_{HU} . For $r = [0.3, 0.2, 0.5]$ and $c = 0.1$, $G_{HU}(c \cdot r) = 47.89$ Mbps $\neq c \cdot G_{HU}(r) = 0.1 \cdot 47.89$ Mbps indicating the violation of homogeneity property of linearity for the G_{HU} . Due to the non-linear constraint, we employ Augmented Lagrangian PSA for our non-linearly constraint problem [11]. PSA uses polling mechanism provided in Algorithm 1 to create a sequence of vector \mathbf{r} for converging to a lower objective EPB_{HU} . In polling, at each iteration the PSA computes EPB_{HU} values for each mesh point of the current sequence point r_{cur} (lines 11-34 in Algorithm 1). In a successful poll (lines 15-27), current mesh point m_i^{cur} improving the objective EPB (line 16) is set as the new sequence point r_{cur} (line 21) and added to the sequence (line 22), current sequence point r_{cur} and its objective value are marked as previous for the next iteration (lines 23-24) and the mesh size is increased (line 26). In an unsuccessful poll (lines 28-31), no improvement is observed so no addition occurs and the mesh size is decreased (line 31). PSA algorithm stops when changes in objective, constraint or mesh size less than specified values is observed (lines 32-34). The tolerance values and other PSA parameters are provided in Table II.

PSA does not guarantee global solution but starting runs from several initial points, solution can be improved. In that

TABLE II
PARAMETERS FOR RESOURCE ALLOCATION SCHEMES.

Par.	Explanation	Value
G_{HU}^{th}	The minimum overall system goodput required from the network	48 Mbps
T_{const}	Tolerance on constraint G_{HU}	$1e^{-8}$
T_{obj}	Tolerance on objective function EPB_{HU}	$1e^{-9}$
T_{mesh}	Tolerance on mesh size	$1e^{-7}$
s_{in}	Initial mesh size	0.02
f_c	Contraction factor	0.5
f_e	Expansion factor	2

regard, for the initial points, we use all possible on/off settings for different modes depicted in Fig. 2. The corners of the triangle in Fig. 2 (the first three cases in the legend) show the initial points with only one mode ON. The lines drawn between corners (4^{th} to 6^{th} cases in the legend) give the set of initial points with one mode OFF and two modes ON. e.g., the yellow line drawn between “only D2D ON” and “only BS ON” cases depicts the scenario of “only satellite OFF”. N_α many random points on this line (5^{th} case in the legend) are selected to start PSA at satellite inactive but other modes active scenarios. Finally, we draw N_α many random initial points depicted as dots in Fig. 2 (7^{th} case in the legend) from the surface bounded by the three lines. These points surveil all modes ON scenarios. As a special initial point scenario, the point with equal weights for each mode in the 8^{th} case is used. We start PSA polling from all of these initial points to improve the performance. In the next section, we present and discuss performance results.

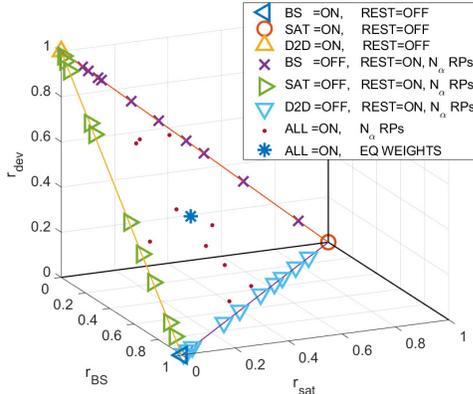


Fig. 2. Set of initial points $[r_{sat}, r_{BS}, r_{dev}]$ (RP:random point, EQ: equal).

IV. PERFORMANCE EVALUATION

Building on our model [8], we look at the suboptimal assignment of \mathbf{r} by PSA. We assume $N_{fsat} = 2$ and $N_{fter} = 3$. The simulation parameter list is provided in Table III. The suboptimal assignment of \mathbf{r} by PSA is $r_{PSA} = [0.09 + \epsilon_1, \epsilon_2 - \epsilon_1, 0.91 - \epsilon_2]$ where $0 < \epsilon_i$ for $i = 1, 2$, $\epsilon_i \rightarrow 0$ such that $\epsilon_2 > \epsilon_1$. The objective of PSA is $EPB_{HU} = 0.166 \mu\text{Jpb}$ and the constraint $G_{HU} = 48.43\text{Mbps}$. Note that the convergence of a mode weight to zero does not necessarily mean corresponding mode selection is zero. For r_{PSA} , even the BS mode weight converges to zero the BS mode selection rate is 15.7%. All mode selection rates of assignment configurations among non-blocked HU services are given in Table IV.

We explore three mode availability options for the assignment of PSA: (i) BS mode is an option and there are also other

TABLE III
SIMULATION PARAMETERS.

Par.	Explanation	Value
N_{fsat}	The total number of satellite frequencies	2
N_{fter}	The total number of terrestrial frequencies	3
λ_{HU}	The arrival rate of hybrid users for content request	$2.4 \frac{user}{sec}$
P_{sat}^{ch}	Per channel transmission power of the satellite	48 W
P_{BS}^{ch}	Per channel transmission power of the BS	6 W
P_{dev}^{tx}	The transmission power of a hybrid user device	80mW
d_{sat}	The distance from LEO satellite to earth	300km
d_{BS}	Mean distance of an HU to the BS	150 m
d_{D2D}	Mean distance between receiver and sender HU devices	30 m
D_{max}	The maximum number of concurrent D2D operations allowed by network	5
N_α	The number of random initial points for only one mode OFF scenarios and for all modes ON scenarios	10

option(s) available (satellite mode, D2D mode) (ii) BS mode is the only available option (iii) BS mode is not an option. In case (i), as the BS mode weight r_{BS} is very small compared to other weights, BS mode will be selected with a very small probability. By decreasing the selection of BS among all system units, EPB_{HU} is decreased since the satellite is solar-powered and D2D mode operations consume less energy than BS mode operations (The D2D mode operations consume lower power P_{dev}^{tx} compared to BS mode transmission power level P_{BS}^{ch} . Additionally, they have shorter durations compared to BS mode operations). In case (ii), for serving the requester HUs, BS mode (directly or indirectly) is used. We utilize BS cache and corresponding frequencies, which is a contributing factor for keeping G_{HU} above the designated threshold.

Case (iii) can occur either due to the requested content unavailability in the BS cache and/or the terrestrial frequencies that can be used for BS operation are not idle. Then this scenario forks into three availability options: (a) only satellite mode (b) only D2D mode (c) both satellite and D2D mode. In (c), r_{sat} and r_{dev} are determinants of content traffic in our network. With the sub-optimal solution, r_{sat} attains lower value than r_{dev} . Although HU energy consumption in D2D mode is higher than that in the satellite mode (satellite is solar powered but HU devices consume P_{dev}^{tx}/μ_{HU}^D energy per successful service and on average $P_{dev}^{tx}/(2 \cdot \mu_{HU}^D)$ per unsuccessful service), the HU service durations over the satellite link are long and thereof, the satellite link saturates. So PSA returns larger r_{dev} ($0.91 - \epsilon_2$) compared to r_{sat} ($0.09 + \epsilon_1$).

TABLE IV
ENERGY EFFICIENCY RESULTS AND MODE SELECTION RATES.

Configuration	EPB _{HU}	Sat mode	BS mode	D2D mode
i) PSA	0.166 μJpb	5.4%	15.7%	78.9%
ii) EQ_{ALL}	0.189 μJpb	6.6%	18.4%	75.1%
iii) ACT_{ALL}	0.192 μJpb	6.6%	18.7%	74.7%
iv) CAP [12]	0.352 μJpb	24.6%	75.4%	-
v) CAP_{EX}	0.177 μJpb	5.5%	16.9%	77.6%
vi) $LQ(T_d=0.1)$ [13]	0.176 μJpb	-	17.0%	83.0%
vii) $LQ_{EX}(T_d=0.1)$	0.171 μJpb	4.23%	16.23%	79.54%

A. EE and different policies

We investigate EE of different assignment configurations listed in Table IV for a comprehensive comparison to baseline configurations (EQ_{ALL} , ACT_{ALL}) and existing work in the literature (CAP , LQ). In baseline configurations, EQ_{ALL} is the assignment with equal weights $[r_{sat}, r_{BS}, r_{dev}] = [1/3, 1/3, 1/3]$ while ACT_{ALL} is the average of 10 random assignments with all units on $[r_{sat}, r_{BS}, r_{dev}] = [\gamma_{sat}^i, \gamma_{BS}^i, 1 - \gamma_{sat}^i - \gamma_{BS}^i]$ where $1 > \gamma_{sat}^i > 0$, $1 > \gamma_{BS}^i > 0$, $i \in \{1, 2, \dots, 10\}$. For mode selection mechanisms in the literature, in [12] one of the utilized techniques chooses the flow through

the satellite/cellular links by the ratio of the corresponding capacity over the total capacity. We denote this scheme as *CAP*. We integrate the D2D mechanism into the *CAP* scheme and update the mode selection scheme accordingly leading to the adapted version *CAP_{EX}*. In [13], with the aim of selecting the better quality link, D2D mode is selected if $T_d \cdot r_d^{-n_{D2D}} > r_{BS}^{-n_{BS}}$ and BS mode is selected otherwise. r_d and r_{BS} denote the distance between the transmitter and receiver in the D2D and BS modes, respectively. n_{D2D} and n_{BS} are the corresponding path loss exponents. T_d is the BS to D2D mode offloading factor. Due to the link quality based nature we call this scheme *LQ*. We extend this scheme by utilizing the satellite mode as long as other modes are not available (due to content unavailability/ busy channel conditions) since satellite links are not as stable as terrestrial counterparts. This extended version is called *LQ_{EX}*. We have looked at the achieved energy efficiency and goodput results of each *ACT_{ALL}* scenario. Among them the scenario with the best EE has goodput 48.43Mbps. We select a slightly lower goodput threshold to enforce the system for improved EE.

In terms of EE, the sub-optimal assignment *PSA* achieves improvement with 13.8% lower *EPB_{HU}* (0.166 μ Jpb) over the average *EPB_{HU}* value (0.192 μ Jpb) of *ACT_{ALL}*. Similarly, *PSA* has *EPB_{HU}* value EE-wise better than *EQ_{ALL}*: 0.166 μ Jpb vs. 0.189 μ Jpb, a 12.4% lower figure. *CAP* scheme does not utilize D2D mode and hence it reaches to very poor EE with 0.352 μ Jpb compared to *PSA*. For *CAP_{EX}*, observed *EPB_{HU}* is 0.177 μ Jpb and *PSA* achieves 6.5% improvement over *CAP_{EX}*. *PSA* also attains 6.0% lower *EPB_{HU}* figure (0.166 μ Jpb) than that of *LQ* with D2D offload factor $T_d = 1$ (0.176 μ Jpb), thereby outperforming *LQ* in terms of EE. When we extend *LQ* with the satellite component, the EE difference decreases but still *PSA* has slightly lower *EPB_{HU}* value. Thus, we can conclude that our sub-optimal assignment scheme reaches the best EE result compared to all of these aforementioned schemes.

B. Discussion

According to sub-optimal mode management scheme *PSA*, the highest mode weight is assigned to D2D paradigm. Even though the satellite is solar-powered, due to the fast saturating nature of the link, D2D mode selection is shown to be more beneficial in terms of EE. Furthermore, cellular services require high energy. Consequently, BS mode selection occurs when it is the only available mode option to complete a content request. Otherwise, this mode is mostly ignored for improving system-wise EE of our hybrid D2D network. D2D communication improves EE due to its low power consumption and high speed transmission enabling larger channel rates. Thus, sub-optimal assignment verifies the importance of D2D systems by the large D2D mode weight assignment.

Satellite and cellular communications are also being used if D2D mode is not available due to (1) the requested content being unavailable in HU devices in vicinity (2) channel is not available. With their larger cache capacities compared to end-user devices, they are more successful in keeping contents for services. With their support, overall goodput is kept still at an acceptable level. This phenomenon illustrates the trade-

off between goodput and EE. A key bottleneck is the rapid saturation problem of the satellite link in that regard.

V. CONCLUSION

In this paper, we formulate the optimization problem of mode assignment for system EE constrained by system-wide goodput in hybrid D2D networks. We show our optimization problem is non-convex, non-linear and hence we exercise the sub-optimal resource allocation by pattern search algorithm (PSA). We compare PSA to alternative resource management schemes. The PSA achieves the highest D2D mode selection rate with the lowest satellite and BS mode selection rates among all-modes active resource management profiles except for *LQ_{EX}* and also accomplish the best EE among them while meeting the goodput constraint. For *LQ_{EX}*, despite the fact that its D2D selection rate (79.54%) is greater than that of *PSA* (78.9%), its BS selection rate is also higher and thus *PSA* achieves improved EE over *LQ_{EX}* (0.166 μ Jpb vs. 0.171 μ Jpb - a lower *EPB_{HU}* value).

REFERENCES

- [1] M. Najimi, "Energy-efficient resource allocation in D2D communications for energy harvesting-enabled noma-based cellular networks," *International Journal of Communication Systems*, vol. 33, no. 2, p. e4184, 2020.
- [2] F. H. Panahi, F. H. Panahi, and T. Ohtsuki, "Energy efficiency analysis in cache-enabled D2D-aided heterogeneous cellular networks," *IEEE Access*, vol. 8, pp. 19540–19554, 2020.
- [3] K. Liolis, A. Geurtz, R. Sperber, D. Schulz, S. Watts, G. Poziopoulou, B. Evans, N. Wang, O. Vidal, B. Tiomela Jou, M. Fitch, S. Diaz Sendra, P. Sayyad Khodashenas, and N. Chuberre, "Use cases and scenarios of 5G integrated satellite-terrestrial networks for enhanced mobile broadband: The SaT5G approach," *International Journal of Satellite Communications and Networking*, vol. 37, no. 2, pp. 91–112, 2019.
- [4] S. Kafiloglu, G. Gur, and F. Alagoz, "Modeling and analysis of content delivery over satellite integrated cognitive radio networks," in *2016 14th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, May 2016, pp. 1–8.
- [5] K. Plimon, J. Ebert, K. Plimon, W. Gappmair, M. Angelone, and A. Ginesi, "Interference-dependent performance of multi-user detection in high throughput satellite systems," in *2018 11th International Symposium on Communication Systems, Networks Digital Signal Processing (CSNDSP)*, July 2018, pp. 1–6.
- [6] G. Zhao, S. Chen, L. Qi, L. Zhao, and L. Hanzo, "Mobile-traffic-aware offloading for energy- and spectral-efficient large-scale D2D-enabled cellular networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 6, pp. 3251–3264, June 2019.
- [7] L. Xu, C. Jiang, Y. Shen, T. Q. S. Quek, Z. Han, and Y. Ren, "Energy efficient D2D communications: A perspective of mechanism design," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7272–7285, 2016.
- [8] S. S. Kafiloglu, G. Gur, and F. Alagoz, "A markovian model for satellite integrated cognitive and D2D HetNets," *Computer Networks*, vol. 169, p. 107083, 2020.
- [9] S. S. Kafiloglu, G. Gur, and F. Alagoz, "Analysis of content-oriented heterogeneous networks with D2D and cognitive communications," *CoRR*, vol. abs/1808.01021, 2018.
- [10] L. M. Rios and N. V. Sahinidis, "Derivative-free optimization: a review of algorithms and comparison of software implementations," *Journal of Global Optimization*, vol. 56, no. 3, pp. 1247–1293, Jul 2013.
- [11] R. Lewis and V. Torczon, "Pattern search algorithms for bound constrained minimization," *SIAM Journal on Optimization*, vol. 9, no. 4, pp. 1082–1099, 1999.
- [12] E. Bouttier, R. Dhaou, F. Arnal, C. Baudoin, E. Dubois, and A. Beylot, "Analysis of content size based routing schemes in hybrid satellite / terrestrial networks," in *2016 IEEE Global Communications Conference (GLOBECOM)*, 2016, pp. 1–6.
- [13] H. ElSawy, E. Hossain, and M. Alouini, "Analytical modeling of mode selection and power control for underlay D2D communication in cellular networks," *IEEE Trans. Commun.*, vol. 62, no. 11, pp. 4147–4161, 2014.