Multi-Objective Approach to Improve Load Balance and Blockage in Millimeter Wave Cellular Networks

Masoud Zarifneshat, Proteek Roy, Li Xiao
Department of Computer Science and Engineering Michigan State University
East Lansing, Michigan
Email: {zarifnes, lxiao}@cse.msu.edu, royprote@egr.msu.edu

***Abstract*—One of the main enabling technologies of 5G wireless networks is to use mm-Wave spectrum band. Despite its large and wide frequency bandwidth, the obtained data rate can be diminished due to link blockage in this frequency band. In this paper, we formulate a bi-objective optimization problem to optimize user association in cellular networks with mm-Wave enabled base stations. The two objectives to minimize are maximum base station utility and blockage score (to indicate the chance of a link getting blocked). We simulate three different scalarization methods to turn a bi-objective vector into a scalar. Since the combinatorial bi-objective problem is NP-Hard, we conduct Lagrangian dual analysis on all of the scalarization methods. Solving the dual problem decreases the time complexity of the solver algorithm, but the solution has a distance from the optimal point created by solving the primal. We also solve the primal optimization problem with a single objective optimization tool. Compared to the time complexity of the primal problem of scalarization methods, the time complexities of solutions to the dual problems are lower. The results show that our solution to bi-objective optimization problem has a better outcome in terms of the number of link blockage and the maximum base station utility compared to optimizing each objective alone.**[1]

***Index Terms*—Millimeter Wave, Cellular Networks, Link Blockage, Load Balancing, Bi-Objective Combinatorial Optimization**

I. Introduction

The increasing demand for wireless network capacity in terms of frequency bandwidth has created the need to find wide bandwidths that can handle vast amounts of data required for future data-hungry applications. One technology that is able to provide wireless networks with such bandwidth is to use high-frequency mm-Wave spectrum. The natural difference of mm-Wave spectrum with low-frequency spectrum introduces a new set of challenges that require new approaches to tackle problems like high path loss and sensitivity to blockage [1].

The link breakage due to low diffraction of mm-Wave is an important issue especially in mm-Wave cellular networks with potential blocking objects (e.g., humans). mm-Wave signals experience dramatic attenuation when going through the human body. When a transmitter transmits by beams with widths 3.4° and 30° and beams are fully blocked by a close-by human body, the receiver senses no signal [2]. Since the human body is composed of a significant amount of water, the beam blockage severity is quite high. Zhu et al. [3] show that groups of people can block a communication link. The reflected beam off the walls or other hard surfaces cannot contain the required energy to be used as high rate carrier [1]. Therefore, a method to switch to different base stations is needed.

The existing user association algorithms for reducing blockage and increasing link availability in mm-Wave networks do not consider load balancing base stations. The blockage reduction algorithms aim to minimize the blockage effect which may result in overloading some base stations (BS) due to having more user equipments (UE) associated to those BSs with links that are considered to have a lower chance of getting blocked by blocking objects. Therefore, if the UEs associate themselves to BSs only according to the blockage objective, some BSs will be overloaded while others be underutilized, which may lead to performance degradation [4]. In an unbalanced network, some BSs may have extra capacity to accept new UEs but are less attractive for UEs to connect because they are assigned with high blockage chance based on the blockage reduction algorithm. In this case, since the blockage chance of a link is only an estimation of future blockage incidents based on previous blockages, the actual blockage may not happen for a BS whose blockage chance is estimated to be high and consequently becomes underutilized. In general, the blockage reduction algorithms divert UEs to BSs with less blockage chance, which may result in overloaded or underutilized BSs and degrades network performance.

This problem calls for a more holistic view of the blockage control and load balancing. Improving the performance based on one metric results in deterioration of the other, a framework that accounts for both blockage and load balancing at the same time is needed to improve the overall performance.

In this paper, we consider both blockage reduction and load balance in our user association algorithm design and formulate a bi-objective combinatorial optimization problem to optimize user association in mm-Wave cellular networks according to the two objectives. The bi-objective problem is then transformed into a single objective problem using three scalarization methods. For all of the scalarization methods, both primal and dual problems are solved. The challenge of the problem is formulating the optimization problem as a single objective problem from biobjective problem and capturing the true relation between two objectives. The objectives are of discrepant nature in the way their parameters are collected and influence the performance of the network. The involvement of

[1]This work was partially supported by the U.S. National Science Foundation under grants CNS-1547015 and CNS-1617412.

978-1-7281-2376-9/19/$31.00 ©2019 IEEE

each of the objectives can be adjusted when needed. In this way, the solver can prioritize an objective over the other.

The contributions of this paper include defining an optimization problem and the solution to the optimization problem. The solution to the optimization problem will give an opportunity to the network to adjust its performance policy via user association with respect to the objective functions. The second contribution is a subgradient solution to different single objective problems of the bi-objective problem, which has polynomial time complexity compared to the primal solver that has exponential time complexity.

The contributions of this paper are summarized as follows:

1) We define a biobjective optimization problem to optimize load balance and blockage score in the network.
2) We use scalarization methods to convert vector of objective to a scalar.
3) We develop a solution for each scalarization method within polynomial time complexity (compared to exponential time complexity of current methods).

The rest of this paper is structured as follows. In Section II, the state-of-the-art of the user association algorithms with combinatorial optimization problems is discussed. Section III is dedicated to system model and objective functions used in this paper. In Section IV, we formulate our bi-objective optimization problem. We solve the single objective ptimization problems generated by scalarization methods in Section IV. Simulation results are discussed in Section V. Finally, the paper concludes with Section VI.

II. Related Work

In this section, we discuss the literature of user association problem in wireless networks on joint optimization/improvement domains. Liu et al. [5] have proposed a taxonomy for User Association (UA) methods in wireless networks. According to the methodology used in our paper, we review some of the works related.

The user association can be done to improve or optimize the outage/coverage probability. Dhillon et al. [6] have proposed a framework for computing the coverage probability in multi-tier HetNets. Cheung et al. [7] have introduced a model to compute success probability in each tier considering different spectrum allocation policies.

The combinatorial approach for user association is used in [8]. The authors formulated a joint optimization problem of user association and channel allocation decision between macrocells and small cells in a HetNet. As a follow-up, Ghimire et al. [9] have proposed a framework to analyze the performance of HetNets by optimizing resource allocation, transmission coordination and user association with the objective of maximizing data rate throughput. Joint optimization of base station sleep mode, user association and subcarrier allocation to maximize total power consumption are discussed in [10].

In the area of the mm-Wave spectrum, Xu et al. [11] have discussed the problem of joint optimization of user association and relaying traffic to other clients. The objective is to maximize the total network throughput. In [12] the authors defined an optimization problem to minimize the maximum load across all base stations with the goal of optimizing user association. Sakaguchi et al. [13] have proposed a user association method in mm-Wave networks that considers supported available rate and number of users in each cell. For energy harvesting networks [14] have formulated an optimization problem to maximize network utility while the energy consumption of each base station does not exceed the harvested amount.

All of the discussed literature share the fact that they try to optimize one or more variables based on single objectives. The factor that distinguishes our work with the ones discussed here is that in our work, we want to optimize user association in mm-Wave networks considering two objectives at the same time. Since the input variables are binary, the problem we are considering is a bi-objective combinatorial optimization problem. To the best of our knowledge, the joint consideration of load balancing and blockage avoidance in the same optimization problem is not considered in any other work.

III. System Model and Objective Functions

In this section, we discuss the specifics of the system that we consider for our proposed method. Then, we introduce the two objectives, load balance and blockage score, on which we want to optimize the user association variable. We also briefly introduce three scalarization methods used to transform 2-dimensional objective vector to a scalar.

A. System Model

In our system, we have N base stations (BS), M user equipment (UE) and P blocking objects (BO). The set of BSs is defined as $\mathcal{N} = \{1, 2, \dots, N\}$ and the set of the UEs is defined as $\mathcal{M} = \{1, 2, \dots, M\}$.

M_i is defined as the set of UEs that are served by BS_i. Analogously, N_j is defined as the set of all BSs that are in the UE_js communication range. The communication range of UE and BS is assumed to be equal to R meters.

The antennas on BSs are assumed to be able to rotate all 360° degrees or in the interval $[-\pi, +\pi]$ in radian.

B. Objective Functions

We want to optimize user association based on two objectives. The first one is *Load Balance (t)* objective which is defined as the maximum total utility of one base station across BS $i \in \mathcal{N}$. The second objective is *Blockage Score (B)* that is defined as total blockage score of all links in the network. In the following two subsections, we discuss these two objective functions in more details.

1) Load Balance: For the first objective, we use the load balancing optimization problem definition and solution provided by Athanasiou et al [12]. The authors have defined the channel utilization between BS_i and UE_j as $\beta_{ij} = \frac{Q_j}{R_{ij}}$ where R_{ij} is the rate of the link between BS_i and UE_j and Q_j is the demanded data rate of UE_j. The user association variable that is going to be optimized is defined as,

$$x_{ij} = \begin{cases} 1, & \text{if } UE_j \text{ associated to } BS_i \\ 0, & \text{Otherwise} \end{cases} \tag{1}$$

The combinatorial optimization problem for optimizing user association variable x is formulated as,

$$\begin{aligned} &\text{minimize} \quad t && && \text{(2a)} \\ &\text{subject to} \quad \sum_{j \in \mathcal{M}_i} \beta_{ij} x_{ij} \leq t, && \forall i \in \mathcal{N} && \text{(2b)} \\ &\qquad \sum_{i \in \mathcal{N}_j} x_{ij} = 1, && \forall j \in \mathcal{M} && \text{(2c)} \\ &\qquad x_{ij} \in \{0, 1\}, && \forall j \in \mathcal{M}, i \in \mathcal{N}_j && \text{(2d)} \end{aligned}$$

2) Blockage Score: For the second objective, we use the idea introduced in [15]. In order to capture the chance of a link between BS and UE to be blocked, the authors of [15] introduced the concept of Blockage Score. The blockage score is used to have an estimate for likelihood of a link being blocked in the future based on previous blockage incidents. The blockage score is a function of distance from previous blockage incidences. When an new UE comes into the area, its blockage score with respect to all BSs in its range is computed.

In their paper, the blockage score shown as γ_{ij}, is the chance of link between UE_j and BS_i being blocked in future. The authors of [15] defined objective function for blockage score of the whole network is defined as follows,

$$\begin{aligned} &\text{minimize} \quad B = \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{M}_i} \gamma_{ij} . x_{ij} \\ &\text{subject to} \quad \sum_{i \in \mathcal{N}_j} x_{ij} = 1, && \forall j \in \mathcal{M} \\ &\qquad x_{ij} \in \{0, 1\}, && \forall j \in \mathcal{M}, i \in \mathcal{N}_j \end{aligned} \tag{3}$$

C. Scalarization Methods

To solve the bi-objective optimization problem, one needs to convert it to single objective function. Scalarization methods are used to do the conversion. We have used three different scalarization methods including Achievement Scalarization Function (ASF) [16], Normal Constraint (NC) [17] and Weighted Sum (WS) [18]. Due to space limitation, we do not elaborate the mentioned methods, and the interested reader may read the cited papers.

To generate the weight vectors for the scalarization methods, we have used Das and Dennis method [19]. This method generates equally spaced weight vectors in any number of dimensions.

IV. Bi-Objective Optimization Problem

In this section, we first formulate the bi-objective optimization problem. The single objective optimization problem of bi-objective optimization problem (4) is formulated by the scalarization methods introduced in Section III-C. Then, the transformed single objective problem is solved by appropriate methods. We borrow the general approach used in [12] to solve the single objective problem generated by the three scalarization methods via dual analysis.

To optimize user association based on both load balancing and blockage score objectives, we formulate the optimization problem as,

$$\begin{aligned} &\text{minimize} \quad f(t, B) && && \text{(4a)} \\ &\text{subject to} \quad \sum_{j \in \mathcal{M}_i} \beta_{ij} . x_{ij} \leq t, && \forall i \in \mathcal{N} && \text{(4b)} \\ &\qquad \sum_{i \in \mathcal{N}_j} x_{ij} = 1, && \forall j \in \mathcal{M} && \text{(4c)} \\ &\qquad x_{ij} \in \{0, 1\}, && \forall j \in \mathcal{M}, i \in \mathcal{N}_j && \text{(4d)} \end{aligned}$$

In Equation (4), the objective B is total blockage score in the network. The problem formulated in (4) is a Bi-Objective Combinatorial Optimization (BOCO) problem. User association problem is combinatorial. The algorithms for solving these problems have exponential time complexity [12]. We transform the problem (4) into single objective problem by using three scalarization methods introduced in Section III-C. Then we solve the resulting primal single objective problems. We also provide the Lagrangian dual solutions to the single objective problems which have lower time complexity than the exponential primal solution.

A. Formulation and Solution of ASF Scalarization Method

In this section, we formulate the single objective optimization problem of the bi-objective problem defined in (4) generated by ASF scalarization method. Then, we use dual analysis to solve the single objective optimization problem for ASF scalarization method.

1) ASF Generated Single Objective Problem: ASF method needs an ideal minimum point $\mathbf{z}$. The ideal point in our defined problem (4) is $\mathbf{z} = (\mathbf{0}, \mathbf{0})$ for objective vector (t, B). Since both objectives of maximum BS utilization and total blockage score cannot get negative values, thus assuming the ideal point as zero for both objectives is logical. The ideal point may or may not be reachable. We formulate the ASF generated single objective problem as,

$$\begin{aligned} &\text{minimize} \quad \max\left(\frac{t}{w_0}, \frac{B}{w_1}\right) \\ &\text{subject to} \quad \sum_{j \in \mathcal{M}_i} \beta_{ij} . x_{ij} \leq t, && \forall i \in \mathcal{N} \\ &\qquad \sum_{i \in \mathcal{N}_j} x_{ij} = 1, && \forall j \in \mathcal{M} \\ &\qquad x_{ij} \in \{0, 1\}, && \forall j \in \mathcal{M}, i \in \mathcal{N}_j \end{aligned} \tag{5}$$

In problem (5), B is total blockage score, $\mathbf{w} = (w_0, w_1)$ is the weight vector. Since there are two objectives the size of weight vector is 2. The problem in (5) can be reformulated to be linear as authors in [12] did to formulate min-max problem. The problem can be reformulated as,

$$\begin{aligned} &\text{minimize} \quad S && \text{(6a)} \\ &\text{subject to} \quad \frac{t}{w_0} \leq S && \text{(6b)} \end{aligned}$$

$$\frac{B}{w_1} \leq S \tag{6c}$$

$$\sum_{j\in\mathcal{M}_i} \beta_{ij}.x_{ij} \leq t \qquad \forall i \in \mathcal{N} \tag{6d}$$

$$\sum_{i\in\mathcal{N}_j} x_{ij} = 1, \qquad \forall j \in \mathcal{M} \tag{6e}$$

$$x_{ij} \in \{0,1\}, \qquad \forall j \in \mathcal{M}, i \in \mathcal{N}_j \tag{6f}$$

In problem (6), is a linear combinatorial optimization problem. The objective is the maximum of two terms of original ASF formulation. The primal problem in (6) is solved by a linear program solver to get to its optimum solution.

2) Solution to ASF Generated Single Objective Problem via Dual Analysis: In this section, we elaborate a solution to ASF generated single optimization problem formulated in Equation (6). The problem in (6) is a mixed integer linear program (MILP) and its complexity is proved to be NP-hard [12]. In order to have a distributed low complexity method of solving bi-objective problem defined in Equation (4), we discuss Lagrangian dual analysis of single objective optimiztion problem generated by ASF scalarization method.

First, we eliminate variable t from Equation (6) and replacing B with its equivalent. In order to do that, the constraint (6b) can be removed by using it in constraint (6d) and replacing variable B with its value defined in [15], making the resulting optimization problem as,

$$\text{minimize} \quad S \tag{7a}$$

$$\text{subject to} \quad \frac{\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}_i} \gamma_{ij}.x_{ij}}{w_1} \leq S \tag{7b}$$

$$\frac{\sum_{j\in\mathcal{M}_i} \beta_{ij}.x_{ij}}{w_0} \leq S \qquad \forall i \in \mathcal{N} \tag{7c}$$

$$\sum_{i\in\mathcal{N}_j} x_{ij} = 1, \qquad \forall j \in \mathcal{M} \tag{7d}$$

$$x_{ij} \in \{0,1\}, \qquad \forall j \in \mathcal{M}, i \in \mathcal{N}_j \tag{7e}$$

We solve Equation (7) as the optimization problem for ASF scalarization method.

In order to obtain Lagrangian of an optimization problem, the objective function and the constraints of the primal problem go together in another optimization problem. Then, the Lagrangian dual is derived by computing infimum of Lagrangian with respect to all variables other than Lagrangian multipliers. The Lagrangian of problem in Equation (7) using only objective function (7a) and first two constraints (7b) and (7c) is formulated as,

$$\begin{aligned} L(S,\mathbf{x},\boldsymbol{\lambda}) = S.\Big(1-\lambda_{N+1}-\sum_{i\in\mathcal{N}}\lambda_i\Big)+ \\ \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}_i} x_{ij}.\big(\frac{\gamma_{ij}}{w_1}.\lambda_{N+1}+\frac{\beta_{ij}}{w_0}.\lambda_i\big) \end{aligned} \tag{8}$$

In Equation (8), the Lagrangian multipliers are gathered in vector $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_{N+1})$. There are N Lagrangian multipliers, one for each constraint in (7c) and another one for constraint (7b). We also swapped the index of nested sum operators according to equivalence $\{(i,j) \mid i \in \mathcal{N}, j \in \mathcal{M}_i\} \equiv \{(n,m) \mid m \in \mathcal{M}, n \in \mathcal{N}_m\}$ used in [12]. We also borrowed from [12] the definition of a vector space that input $\mathbf{x}$ satisfies constraints (7d) and (7e). The vector space $\mathcal{X}$ is Cartesian product of $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_M$. Each $\mathcal{X}_j$ is defined as,

$$\mathcal{X}_j = \{\boldsymbol{x}_j = (x_{ij})_{i\in\mathcal{N}_j} \mid \sum_{i\in\mathcal{N}_j} x_{ij} = 1, x_{ij} \in \{0,1\}, i \in \mathcal{N}_j\} \tag{9}$$

According to Equation (9), each $\mathcal{X}_j$ is j^{th} column in matrix $\mathbf{x}$ that has a single one and all other entries are zeros. The single one can be in any N entries of the column j. Now, we are ready to define the Lagrangian dual problem. It is defined as,

$$g(\boldsymbol{\lambda}) = \inf_{\substack{S\in\mathbb{R}\\ \mathbf{x}\in\mathcal{X}}} L(S,\mathbf{x},\boldsymbol{\lambda}) \tag{10a}$$

$$= \begin{cases} \inf_{\mathbf{x}\in\mathcal{X}} \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}_i} x_{ij}.\big(\frac{\gamma_{ij}}{w_1}.\lambda_{N+1}+\frac{\beta_{ij}}{w_0}.\lambda_i\big), \\ \sum_{i=1}^{N+1}\lambda_i = 1 \\ -\infty, \qquad \text{otherwise} \end{cases} \tag{10b}$$

$$= \begin{cases} \sum_{j\in\mathcal{M}} \inf_{\mathbf{x}_j\in\mathcal{X}_j} \sum_{i\in\mathcal{N}_j} x_{ij}.\big(\frac{\gamma_{ij}}{w_1}.\lambda_{N+1}+\frac{\beta_{ij}}{w_0}.\lambda_i\big), \\ \sum_{i=1}^{N+1}\lambda_i = 1 \\ -\infty, \qquad \text{otherwise} \end{cases} \tag{10c}$$

$$= \begin{cases} \sum_{j\in\mathcal{M}} g_j(\boldsymbol{\lambda}), & \sum_{i=1}^{N+1}\lambda_i = 1 \\ -\infty, & \text{otherwise} \end{cases} \tag{10d}$$

In Equation (10), the condition for case one of Equations (10b) and (10c) go to the next line since the whole line could not be fitted in the column width. The infimum of Lagrangian in Equation (8) is formulated with respect to two variables S and $\mathbf{x}$. The infimum of the term containing S with respect to S in Equation (8) is $-\infty$. Therefore, it is unbounded. In order to have a bounded S term, we need to make the coefficient of S, zero in Equation (8). That is made possible by having coefficient of S equal to zero as condition in dual problem. The vector space we defined in (9) helps us to deduce Equation (10c). In Equation (10c), the infimum is computed over all $\mathbf{x}_j$ members of the set $\mathcal{X}_j$. The Lagrangian dual problem is formulated as,

$$\text{maximize} \quad g(\boldsymbol{\lambda}) = \sum_{j\in\mathcal{M}} x^*_{ij}.\big(\frac{\gamma_{ij}}{w_1}.\lambda_{N+1}+\frac{\beta_{ij}}{w_0}.\lambda_i\big) \tag{11a}$$

$$\text{subject to} \quad \sum_{i=1}^{N+1}\lambda_i = 1 \tag{11b}$$

$$\lambda_i \geq 0 \tag{11c}$$

In Equation (11a), x^*_{ij} is computed using the following formula,

$$x^*_{ij} = \begin{cases} 1, & i = \arg\min_{k \in \mathcal{N}_j} \big(\frac{\gamma_{kj}}{w_1}.\lambda_{N+1} + \frac{\beta_{kj}}{w_0}.\lambda_k\big) \\ 0, & \text{otherwise} \end{cases} \tag{12}$$

The Equation (12), is resulted from the fact that the infimum in $g_j(\boldsymbol{\lambda})$ is over the set $\mathcal{X}_j$. Therefore, the entries of x^*_{ij} are all zeros except the i^{th} entry that has minimum $\frac{\gamma_{ij}}{w_1}.\lambda_{N+1} + \frac{\beta_{ij}}{w_0}.\lambda_i$ for UE_j.

The problem in (11), is a concave function. It also is non-smooth which leads up to use of subgradient method because it is a non-differentiable function. Since the dual problem has constraint (11b), the projected version of subgradient method is used. In this version, in order to make sure that the variable values are chosen from the constraint polygon, an Euclidean projection of current point on the constraint polygon is generated as the next solution to the optimization problem. There are $N+1$ variables of λ_i that each will have a subgradient. Therefore, subgradient of $-g$ (since subgradient is for minimization problems) is a vector with $N+1$ elements. The subgradient of $-g$ in a feasible $\boldsymbol{\lambda}$ is

$$u_i = \begin{cases} -\sum_{j \in \mathcal{M}_i} \frac{\beta_{ij}}{w_0}.x^*_{ij}, & i \in \mathcal{N} \\ -\sum_{j \in \mathcal{M}} \sum_{i \in \mathcal{N}_j} \frac{\gamma_{ij}}{w_1}.x^*_{ij}, & i = N+1 \end{cases} \tag{13}$$

Then the next value of vector $\boldsymbol{\lambda}$ is computed by,

$$\boldsymbol{\lambda}^{(k+1)} = P(\boldsymbol{\lambda}^{(k)} - \alpha_k.\boldsymbol{u}^{(k)}) \tag{14}$$

In Equation (14), $P(\boldsymbol{\lambda}^{(k)} - \alpha_k.\boldsymbol{u}^{(k)})$ is Euclidean projection of point $\boldsymbol{\lambda}^{(k)} - \alpha_k.\boldsymbol{u}^{(k)}$ in $N+1$ dimensional space into simplex defined by $\{\boldsymbol{\lambda} \mid \sum_{i=1}^{N+1} \lambda_i = 1\}$. The parameter α_k is a step size variable in k^{th} iteration.

The subgradient algorithm works iteratively. In each iteration (k), it considers vector $\boldsymbol{\lambda}^{(k-1)}$ and computes a new vector $\boldsymbol{\lambda}^{(k)}$ for current iteration. Vector $\boldsymbol{\lambda}^{(k)}$ is then used to compute x^*_{ij} via Equation (25). Then the value of x^*_{ij} is used to compute the subgradient vector. This cycle continues for a constant number of times defined by iteration number of subgradient algorithm.

The subgradient method does not find the optimal solution but can find a near optimal solution within reasonable time complexity. In order to compute the time complexity of subgradient algorithm, we analyze all the steps mentioned earlier. The computation of x^* matrix is of order $O(M.N)$. Computing subgradient vector also costs $O(M.N)$. There are different algorithms to compute the Euclidean projection on simplex. The one we used has a time complexity of order $O(N \log(N))$. If we call iteration number of subgradient method K, the total time complexity of subgradient method is $O(K.M.N)$. The worst order is $O(\max(K, M, N)^3)$ which is of polynomial time complexity compared to exponential time complexity for methods that solve primal problem. This difference is very important for systems with limited processing resources and real time response expectations.

B. Formulation and Solution of NC Scalarization Method

In this section, we formulate and solve the linear single objective optimization problem that is generated by NC scalarization method.

Normal Constraint (NC) method requires two Utopian points to do the scalarization. These two points are the optimum values for all objectives when solved separately. The NC is the same as ASF method, but the difference is in how they draw the reference vectors. In NC, unlike ASF, the reference vectors are drawn in parallel. The reformulated version for NC generated single objective problem is shown as,

$$\text{minimize} \quad \max\Big(NF_1 - (w_0 - 0.5), NF_2 - (w_1 - 0.5)\Big) \tag{15a}$$

$$\text{subject to} \quad \sum_{j \in \mathcal{M}_i} \beta_{ij}.x_{ij} \le t \qquad \forall i \in \mathcal{N} \tag{15b}$$

$$\sum_{i \in \mathcal{N}_j} x_{ij} = 1, \qquad \forall j \in \mathcal{M} \tag{15c}$$

$$x_{ij} \in \{0,1\}, \qquad \forall j \in \mathcal{M}, i \in \mathcal{N}_j \tag{15d}$$

$$\text{In which} \quad NF_1 = \frac{t - f_1(X_l)}{f_1(X_r) - f_1(X_l)} = \frac{t - f_1(X_l)}{C_1} \tag{15e}$$

$$NF_2 = \frac{B - f_2(X_r)}{f_2(X_l) - f_2(X_r)} = \frac{B - f_2(X_r)}{C_2} \tag{15f}$$

In problem (15), NF_1 and NF_2 are normalized objective values for first and second objectives. X_l is the optimal solution to the optimization problem that minimizes maximum load t in the network. Similarly, X_r is the optimal solution to optimization problem that minimizes total blockage score B. $f_1(x)$ is the objective value for maximum load of solution x and $f_2(x)$ is the objective value for total blockage score B. The values in max function are normalized objective values. Like ASF formulation (6), we want to reformulate problem (15) to make it a linear optimization problem. The result is formulated as,

$$\text{minimize} \quad S \tag{16a}$$

$$\text{subject to} \quad NF_1 - (w_0 - 0.5) \le S \tag{16b}$$

$$NF_2 - (w_1 - 0.5) \le S \tag{16c}$$

$$\sum_{j \in \mathcal{M}_i} \beta_{ij}.x_{ij} \le t, \qquad \forall i \in \mathcal{N} \tag{16d}$$

$$\sum_{i \in \mathcal{N}_j} x_{ij} = 1, \qquad \forall j \in \mathcal{M} \tag{16e}$$

$$x_{ij} \in \{0,1\}, \qquad \forall j \in \mathcal{M}, i \in \mathcal{N}_j \tag{16f}$$

The problem in (16) is a linear optimization problem that can be solved by LP solvers.

1) Solution to NC Generated Single Objective Problem via Dual Analysis: The solution to NC generated signal objective problem uses the same methodology as the solution to ASF scalarizartion method.

In order to get to the linear optimization problem, the t variable needs to be eliminated and NF_1 and NF_2 parameters are replaced with extreme constants as in Equations (15e) and (15f). Parameter B is also replaced by it equivalent according to [15]. The variable elimination is done through using constraint (16b) in constraint (16d). Therefore, variable t is eliminated and the reformulated problem is shown as,

$$\text{minimize} \quad S \tag{17a}$$

$$\text{subject to} \quad \frac{\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}_i}\gamma_{ij}.x_{ij}}{C_2} - \frac{f_2(X_r)}{C_2} - w_1 + 0.5 \leq S \tag{17b}$$

$$\frac{\sum_{j\in\mathcal{M}_i}\beta_{ij}.x_{ij}}{C_1} - \frac{f_1(X_l)}{C_1} - w_0 + 0.5 \leq S, \ \forall i \in \mathcal{N} \tag{17c}$$

$$\sum_{i\in\mathcal{N}_j} x_{ij} = 1, \qquad \forall j \in \mathcal{M} \tag{17d}$$

$$x_{ij} \in \{0,1\}, \qquad \forall j \in \mathcal{M}, i \in \mathcal{N}_j \tag{17e}$$

We compute the Lagrangian of optimization problem in (17), like we did it for ASF in Section IV-A2. We use objective function (17a) along with constraints (17b) and (17c). Therefore, the Lagrangian is formulated as,

$$\begin{aligned} L(S,\mathbf{x},\boldsymbol{\lambda}) = S.\Big(1 - \lambda_{N+1} - \sum_{i\in\mathcal{N}}\lambda_i\Big) + \\ \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}_i} x_{ij}.\big(\frac{\gamma_{ij}}{C_2}.\lambda_{N+1} + \frac{\beta_{ij}}{C_1}.\lambda_i\big) - \\ \Big(\frac{f_2(X_r)}{C_2} + w_1 - 0.5\Big).\lambda_{N+1} - \\ \Big(\frac{f_1(X_l)}{C_1} + w_0 - 0.5\Big).\sum_{i\in\mathcal{N}}\lambda_i \end{aligned} \tag{18}$$

From Lagrangian in (18), we can deduce Lagrangian dual optimization problem for (17). The Lagrangian dual is formulated as,

$$g(\boldsymbol{\lambda}) = \inf_{\substack{S\in\mathbb{R}\\ \mathbf{x}\in\mathcal{X}}} L(S,\mathbf{x},\boldsymbol{\lambda}) \tag{19a}$$

$$= \begin{cases} \inf_{\mathbf{x}\in\mathcal{X}} \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}_i} x_{ij}.\big(\frac{\gamma_{ij}}{C_2}.\lambda_{N+1} + \frac{\beta_{ij}}{C_1}.\lambda_i\big) \\ -\Big(\frac{f_2(X_r)}{C_2} + w_1 - 0.5\Big).\lambda_{N+1} \\ -\Big(\frac{f_1(X_l)}{C_1} + w_0 - 0.5\Big).\sum_{i\in\mathcal{N}}\lambda_i, & \sum_{i=1}^{N+1}\lambda_i = 1 \\ -\infty, & \text{otherwise} \end{cases} \tag{19b}$$

$$= \begin{cases} \sum_{j\in\mathcal{M}} \inf_{\mathbf{x}_j\in\mathcal{X}_j} \sum_{i\in\mathcal{N}_j} x_{ij}.\big(\frac{\gamma_{ij}}{C_2}.\lambda_{N+1} + \frac{\beta_{ij}}{C_1}.\lambda_i\big) \\ -\Big(\frac{f_2(X_r)}{C_2} + w_1 - 0.5\Big).\lambda_{N+1} \\ -\Big(\frac{f_1(X_l)}{C_1} + w_0 - 0.5\Big).\sum_{i\in\mathcal{N}}\lambda_i, & \sum_{i=1}^{N+1}\lambda_i = 1 \\ -\infty, & \text{otherwise} \end{cases} \tag{19c}$$

$$= \begin{cases} \sum_{j\in\mathcal{M}} g_j(\boldsymbol{\lambda}), & \sum_{i=1}^{N+1}\lambda_i = 1 \\ -\infty, & \text{otherwise} \end{cases} \tag{19d}$$

In (19b) and (19c), the whole term before each colon is the term for the first case and the term after colon is the condition. For spacing reasons the term for the first case in those two equations are extended to three lines. Similar to the ASF solution, we get different formulations for Lagrangian dual.

The Lagrangian dual optimization problem is,

$$\begin{aligned} \text{maximize} \quad g(\boldsymbol{\lambda}) = \sum_{j\in\mathcal{M}} x^*_{ij}.\big(\frac{\gamma_{ij}}{C_2}.\lambda_{N+1} + \frac{\beta_{ij}}{C_1}.\lambda_i\big) \\ -\Big(\frac{f_2(X_r)}{C_2} + w_1 - 0.5\Big).\lambda_{N+1} \\ -\Big(\frac{f_1(X_l)}{C_1} + w_0 - 0.5\Big).\sum_{i\in\mathcal{N}}\lambda_i \\ \text{subject to} \quad \sum_{i=1}^{N+1}\lambda_i = 1 \\ \lambda_i \geq 0 \end{aligned} \tag{20}$$

In Equation (20), x^*_{ij} is computed as,

$$x^*_{ij} = \begin{cases} 1, & i = \arg\min_{k\in\mathcal{N}_j} \big(\frac{\gamma_{kj}}{C_2}.\lambda_{N+1} + \frac{\beta_{kj}}{C_1}.\lambda_k\big) \\ 0, & \text{otherwise} \end{cases} \tag{21}$$

The subgradient of optimization problem in (20) is formulated as,

$$u_i = \begin{cases} -\sum_{j\in\mathcal{M}_i}\frac{\beta_{ij}}{C_1}.x^*_{ij} + \frac{f_1(X_l)}{C_1} + w_0 - 0.5, & i \in \mathcal{N} \\ -\sum_{j\in\mathcal{M}}\sum_{i\in\mathcal{N}_j}\frac{\gamma_{ij}}{C_2}.x^*_{ij} + \frac{f_2(X_r)}{C_2} + w_1 - 0.5, & i = N+1 \end{cases} \tag{22}$$

By using the subgradient in Equation (22) and subgradient method it is possible to solve the optimization problem in (20). The general approach is the same as described in Section IV-A2.

C. Formulation and Solution of WS Scalarization Method

Weighted sum method multiplies objective values by a weight vector and then sums over the results. For our specific problem, we formulated weighted sum scalarized optimization problem as,

$$\text{minimize} \quad w_0.t + w_1.B \tag{23a}$$

$$\text{subject to} \quad \sum_{j\in\mathcal{M}_i}\beta_{ij}.x_{ij} \leq t, \qquad \forall i \in \mathcal{N} \tag{23b}$$

$$\sum_{i\in\mathcal{N}_j} x_{ij} = 1, \qquad \forall j \in \mathcal{M} \tag{23c}$$

$$x_{ij} \in \{0,1\}, \qquad \forall j \in \mathcal{M}, i \in \mathcal{N}_j \tag{23d}$$

The objective function in weighted sum method is a linear function, so there is no need for reformulation of problem (23). Due to space limitation, we skip the analysis that leads to the Lagrangian dual problem. The Lagrangian dual problem is formulated as,

$$\text{maximize} \quad g(\boldsymbol{\lambda}) = \sum_{j\in\mathcal{M}} x^*_{ij}.\big(w_1.\gamma_{ij} + \lambda_i.\beta_{ij}\big) \tag{24a}$$

$$\text{subject to} \quad \sum_{i\in\mathcal{N}} \lambda_i = w_0 \tag{24b}$$

$$\lambda_i \geq 0, i \in \mathcal{N} \tag{24c}$$

In Equation (24a), x^*_{ij} is computed using the following formula,

$$x^*_{ij} = \begin{cases} 1, & i = \arg\min_{k\in\mathcal{N}_j} \big(w_1.\gamma_{kj} + \lambda_k.\beta_{kj}\big) \\ 0, & \text{otherwise} \end{cases} \tag{25}$$

For subgradient method we compute the subgradient of optimization problem in (24). The subgradient of $-g$ in a feasible $\boldsymbol{\lambda}$ is $\mathbf{u} = (u_j)_{i\in\mathcal{N}}$ and $u_i = -\sum_{j\in\mathcal{M}_i} x^*_{ij}.\beta_{ij}$. Then the next value of vector $\boldsymbol{\lambda}$ is computed by recursive subgradient method formulation in (14).

In Equation (14), $P(\boldsymbol{\lambda}^{(k)} - \alpha_k.\boldsymbol{u}^{(k)})$ is Euclidean projection of point $\boldsymbol{\lambda}^{(k)} - \alpha_k.\boldsymbol{u}^{(k)}$ in N dimensional space into simplex defined by $\{\boldsymbol{\lambda} \mid \sum_{i\in\mathcal{N}} \lambda_i = w_0\}$.

D. Semi-Distributed Algorithm for Dual Analysis

In this section, we introduce a semi-distributed algorithm that is used to do the computations of the solution to the dual problem discussed in previous sections. We need to compute the subgradient of the dual problem and the step in subgradient method to solve the dual problem for three single objective problems generated by all three scalarization methods in Equation (14). The evaluation of this equation requires the evaluation of subgradients of three scalarization methods. It also needs to compute x^*_{ij} for the scalarization methods.

The algorithm is a semi-distributed one as there are some parts of it need to be run by a centralized entity like CRAN [20]. The rest of the operations can be done in a distributed manner between BSs and UEs. The semi-distributed algorithm for user association (SDA) is elaborated in Algorithm 1.

In Algorithm 1, there are two different initializations, central and distributed. In central initialization, CRAN is responsible to distribute the general parameters to BSs. These parameters include weight vectors for all scalarization methods and λ_{N+1} for ASF and NC scalarization methods and C_1 and C_2 for NC scalarization method. The general parameters can only be computed by an entity that has a holistic view of the network. That is why CRAN is responsible for it. According to this algorithm, a change in number of BS or UE should not incur a considerable load on distributed process. For central process, the CRAN can handle the change properly as it is designed to do so.

Algorithm 1 Semi-distributed Algorithm for User Association (SDA)

1: Initialize: Central {CRAN distributes the parameters to BSs}
2: Initialize: Distributed {Each BS_i distributes $\boldsymbol{\lambda}^{(k)}$ at time k and γ_{ij} to UEs $j \in \mathcal{M}_i$}
3: UE_j computes its associated BS_i by Equations (12), (21) and (25)
4: UE_j only signals its associated BS_i and BS_i compute subgradient according to subgradient equations for each scalarization method
5: $\boldsymbol{\lambda}^{(k+1)}$ computed by communication between BSs for the next step
6: if stopping criterion not met go to 3

V. Simulation Results

In this section, we discuss the simulation results of different scalarization methods discussed in the previous section based on two different objective functions defined. We also used a primal problem solver called Gurobi [21]. Version 7.0.1 of Gurobi solver is used in our simulations.

A. Parameters and Metrics

We use OMNet++ [22] network simulator to simulate the environment. The simulation area is a rectangle of 100×500 square meters. There are $M = 100$ user equipments, $N = 50$ base stations and $P = 130$ blocking objects.

The user association methods studied in this simulation section are Achievement Scalarizing Function (ASF), Normal Constraint (NC) and Weighted Sum (WS) methods. We also solved the optimization problem for each of load balance and blockage score objective functions separately. Two different solvers used for solving different methods. Subgradient algorithm is used to solve the dual optimization problem. This algorithm is used to solve ASF, NC and WS generated single objective problems and single objective optimization problem for load balance objective function. We also use primal problem solver to solve all single objective problems at hand. It includes single objective scalarized optimization problems and the load balance and blockage score optimization problems separately. Gurobi is the name of the solver, and it is shown on figures.

There are several metrics that we use to measure and compare the performance of our bi-objective optimization approach to single objective approaches. The first metric is *Average Blockage per Blocking Object*. Each time a blockage occurs in the network that is recorded. The blockage count is cumulative over time. Then the total number of blockages in the network is divided by the number of blocking objects. The second metric is *Average Handover per User Equipment*. A counter increments when a handover occurs. The total is divided by M. This metric is cumulative over time as well. We also compute rate and SINR for all active links in the network. We use the formulation in [23]. In this formulation, the SINR for a link in the mm-Wave network is a function of antenna

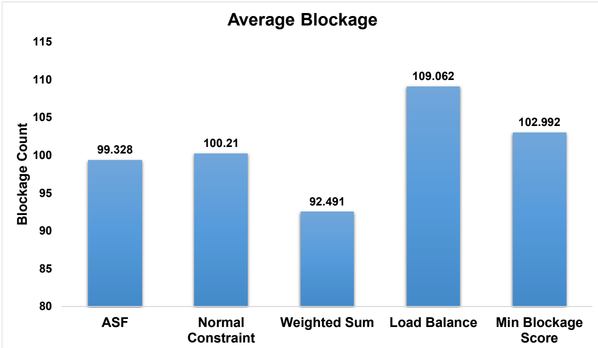

Fig. 1. Average Blockage Count

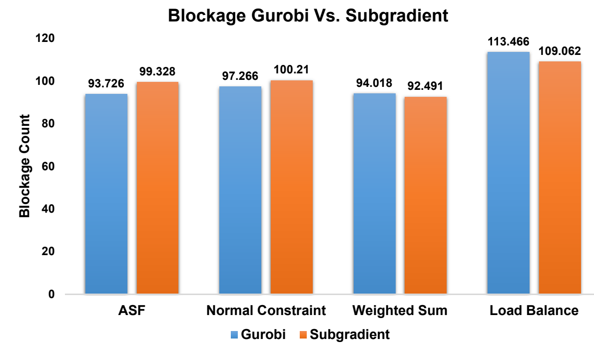

Fig. 2. Average Blockage Count Comparison Gurobi vs. Subgradient

orientation, path loss model, independent Nakagami fading model for small-scale fading and thermal noise of propagation environment. We also use the objective values of load balance and total blockage score for all user association methods we discussed.

The values for these parameters are absolute quantities. All of them are the results of simulations with the equal duration. Therefore, the results show the differences of the performance of the algorithms under the same conditions. Thus, the values of the parameters are used to fairly comparing the performance of the algorithms.

There are two metrics named *False Positive* and *False Negative*. They are defined based on optimal behavior of handover. The simulator knows the optimal BS for a UE to connect. false positive and false negative ratios are defined based on whether a UE is connected to an optimal BS or not. If based on the optimal behavior the UE should be connected to a BS, but it does not, it counts as one false negative. On the other hand, if a UE should stay in the current BS, but it is connected to another BS, it counts as one false positive. The negativity is detected based on handover event. The optimal behavior is defined based on the blockage. since the blokage objective is heuristic, it is possible the after handover to a BS, the UE experiences a blockage. Therefore, this handover is not optimal. In general the optimal behavior is when doing or not doing a handover does not cause blockage and does not degrade the load balance objective of the network.

B. Results and Discussion

In the following, we depict figures for different metrics we discussed.

All of the scalarization methods have 40 subproblems. It means that we generated 40 different weight vectors and then solved the single optimization problem generated by each of scalarization methods using each weight vector. Unless specifically mentioned, the comparison between load balance and minimum blockage score and scalarization methods is made with subgradient solution of scalarization method with 40 subproblems. Unless stated otherwise, the solver for all methods is subgradient. Whenever there is a comparison involving primal solver, the number of the subproblem is 40

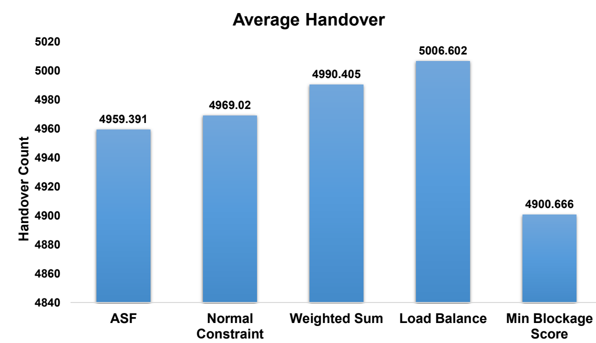

Fig. 3. Average Handover Count

as well. For load balance method, we used the subgradient method used in [12]. As depicted in Fig. 1, the three scalarization methods have lower average blockage score than each load balance and minimum blockage score methods. Among the scalarization methods, the weighted sum has the best performance while ASF and normal constraint are close to each other. The normal constraint has the worst performance than the other two scalarization methods. Since blockage score is a heuristic, minimizing that does not guarantee the best performance.

In Fig. 2, the comparison is made between all scalarization methods and load balance method solved by subgradient and gurobi solvers. As can be seen in this figure, subgradient solver has a slightly better performance than gurobi solver for weighted sum and load balance methods.

The average handover count is showed in Fig. 3. As shown in this figure, the handover count of scalarization methods are nearly the same but more than minimum blockage score method. Load balance has the worst performance in this metric. Minimum blockage score on the other hand has the best performance.

False positive ratio is showed in Fig. 4. The false positive count for three scalarization methods are less than the separate objectives. Among separate objectives, load balance has better ratio. We see almost the same pattern for false negative ratio in Fig. 5. Within scalarization methods, weighted sum has the best performance of all.

The CDF of SINR is depicted in Fig. 6. It is evident that

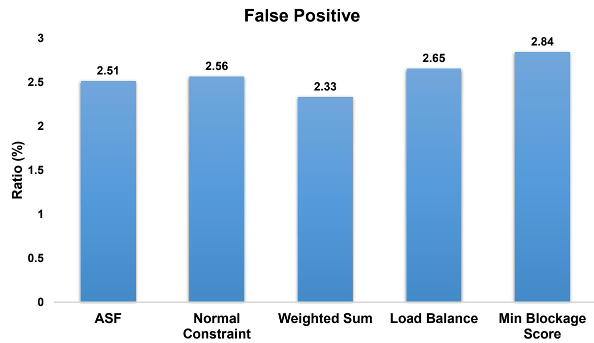

Fig. 4. False Positive Ratio

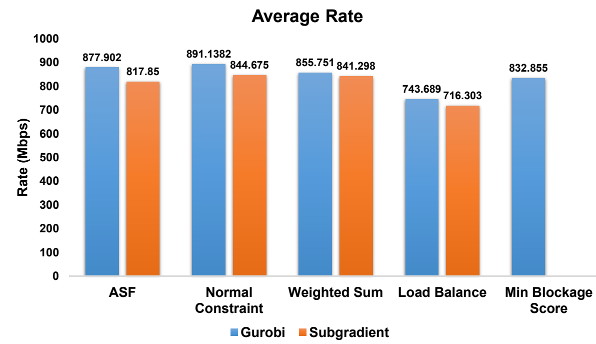

Fig. 7. Average Rate for Different User Association Algorithms

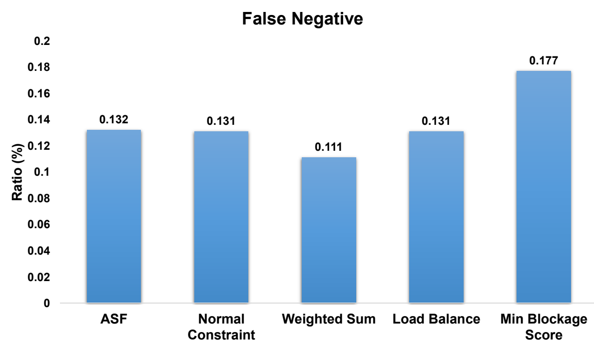

Fig. 5. False Negative Ratio

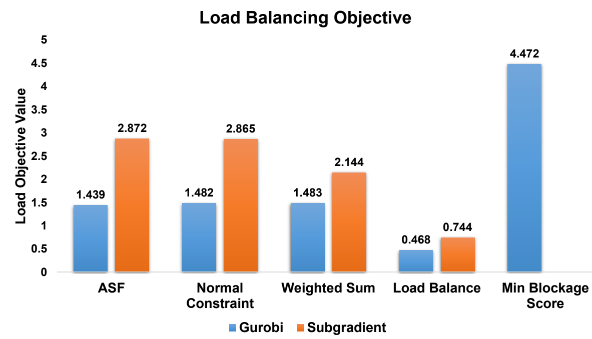

Fig. 8. Average Maximum Load for Different User Association Algorithms

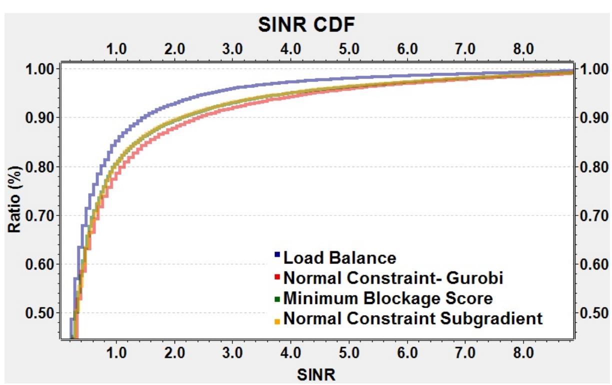

Fig. 6. Cumulative Distribution Function for SINR

normal constraint solved with gurobi solver has links with high SINR with higher frequency than two objective functions separately. Minimum blockage score and normal constraint solved with subgradient solver are the same in terms of SINR. Load balance has the worst performance.

The column diagram of the average rate for different user association algorithms is presented in Fig. 7. In this figure, we include the result for both solvers. However, for minimum blockage score method, the solver is just a simple, primal solver and we did not need to use either gurobi or subgradient solvers. The normal constraint is the best algorithm according to rate metric. The scalarization methods are performing better than blockage score and load balance methods except for ASF subgradient. Load balance has the worst metric value.

As shown in Fig. 8, the load balance algorithm has the best performance for minimizing the maximum load in the network. We have the same structure as Fig. 7 here. Three scalarization methods have close performance while all out-performing blockage score algorithm. For subgradient versions of scalarization methods, the weighted sum has the best performance compared to other two methods.

Fig. 9 shows how the number of subproblems in bi-objective solver will affect resulting load objective value. There are solutions with 10, 20, 30 and 40 subproblems for each scalarization and single objective methods that are solved by either subgradient or gurobi solvers. As shown in the figure, the general trend of load objective reduces when the number of subproblems (weight vectors) increases. It means that when there are more weight vectors, it is more probable that the solver finds better solutions. We can also see that the gurobi solver results are better than subgradient. For selecting the best solution from s solutions (s is the subproblem size), we choose a solution that its objective values are closest to the ideal point, $(0, 0)$ point.

The comparison between SINR values for two different solutions to normal constraint scalarization method is presented in Fig. 10. The figure shows that the primal solutions have better SINRs for all sub-problem sizes. However, the trend within each curve is almost increasing with increase in subproblem size.

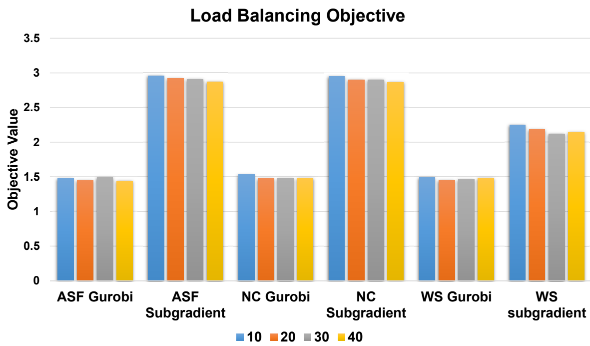

Fig. 9. Average Maximum Load for Different Subproblem Sizes

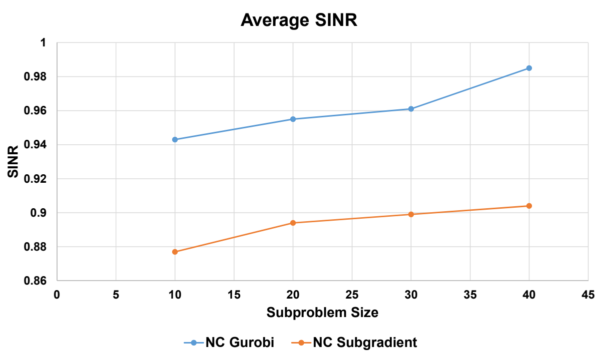

Fig. 10. SINR Comparison for Normal Constraint Method

VI. Conclusion

In this paper, we study the user association problem in mm-Wave cellular networks. We considered two objective functions, one for minimizing blockage in the network and the other for minimizing maximum load across all base stations in the network. The results show that the bi-objective approach achieves better performance on reducing both load and blockage score, on increasing SINR and keeping base stations in the networks less loaded. For weighted sum scalarization method, we showed that solution to the dual problem via subgradient method has better time complexity than solving the primal problem, but with the cost of losing some SINR value. A possible Future direction of this research can be an investigation of other factors affecting link blockage chance and the importance of their role in this area.

References

[1] Y. Niu, Y. Li, D. Jin, L. Su, and A. V. Vasilakos, "A survey of millimeter wave communications (mmwave) for 5g: opportunities and challenges," *Wireless Networks*, vol. 21, no. 8, pp. 2657–2676, 2015. [Online]. Available: http://dx.doi.org/10.1007/s11276-015-0942-z

[2] S. Sur, V. Venkateswaran, X. Zhang, and P. Ramanathan, "60 ghz indoor networking through flexible beams: A link-level profiling," in *Proceedings of the 2015 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems*, ser. SIGMETRICS '15. New York, NY, USA: ACM, 2015, pp. 71–84. [Online]. Available: http://doi.acm.org/10.1145/2745844.2745858

[3] Y. Zhu, Z. Zhang, Z. Marzi, C. Nelson, U. Madhow, B. Y. Zhao, and H. Zheng, "Demystifying 60ghz outdoor picocells," in *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking*, ser. MobiCom '14. New York, NY, USA: ACM, 2014, pp. 5–16. [Online]. Available: http://doi.acm.org/10.1145/2639108.2639121

[4] K. Sui, S. Sun, Y. Azzabi, X. Zhang, Y. Zhao, J. Wang, Z. Li, and D. Pei, "Understanding the impact of ap density on wifi performance through real-world deployment," in *2016 IEEE International Symposium on Local and Metropolitan Area Networks (LANMAN)*, June 2016, pp. 1–6.

[5] D. Liu, L. Wang, Y. Chen, M. Elkashlan, K. K. Wong, R. Schober, and L. Hanzo, "User association in 5g networks: A survey and an outlook," *IEEE Communications Surveys Tutorials*, vol. 18, no. 2, pp. 1018–1044, Second quarter 2016.

[6] H. S. Dhillon, R. K. Ganti, F. Baccelli, and J. G. Andrews, "Modeling and analysis of k-tier downlink heterogeneous cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, pp. 550–560, April 2012.

[7] W. C. Cheung, T. Q. S. Quek, and M. Kountouris, "Throughput optimization, spectrum allocation, and access control in two-tier femtocell networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, pp. 561–574, April 2012.

[8] D. Fooladivanda, A. A. Daoud, and C. Rosenberg, "Joint channel allocation and user association for heterogeneous wireless cellular networks," in *2011 IEEE 22nd International Symposium on Personal, Indoor and Mobile Radio Communications*, Sept 2011, pp. 384–390.

[9] J. Ghimire and C. Rosenberg, "Resource allocation, transmission coordination and user association in heterogeneous networks: A flow-based unified approach," *IEEE Transactions on Wireless Communications*, vol. 12, no. 3, pp. 1340–1351, March 2013.

[10] L. Su, C. Yang, Z. Xu, and A. F. Molisch, "Energy-efficient downlink transmission with base station closing in small cell networks," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2013, pp. 4784–4788.

[11] Y. Xu, G. Athanasiou, C. Fischione, and L. Tassiulas, "Distributed association control and relaying in millimeter wave wireless networks," in *2016 IEEE International Conference on Communications (ICC)*, May 2016, pp. 1–6.

[12] G. Athanasiou, P. C. Weeraddana, C. Fischione, and L. Tassiulas, "Optimizing client association for load balancing and fairness in millimeter-wave wireless networks," *IEEE/ACM Transactions on Networking*, vol. 23, no. 3, pp. 836–850, June 2015.

[13] K. Sakaguchi, G. K. Tran, H. Shimodaira, S. Nanba, T. Sakurai, K. Takinami, I. Siaud, E. C. Strinati, A. Capone, I. Karls, R. Arefi, and T. Haustein, "Millimeter-wave evolution for 5g cellular networks," *IEICE TRANSACTIONS on Communications*, vol. E98-B, no. 3, pp. 388–402, March 2015.

[14] B. Xu, Y. Chen, M. Elkashlan, T. Zhang, and K. K. Wong, "User association in massive mimo and mmwave enabled hetnets powered by renewable energy," in *2016 IEEE Wireless Communications and Networking Conference*, April 2016, pp. 1–6.

[15] M. Zarifneshat, C. J. Liu, and L. Xiao, "A protocol for link blockage mitigation in mm-wave networks," in *2017 IEEE 14th International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*, Oct 2017, pp. 215–223.

[16] A. Pascoletti and P. Serafini, "Scalarizing vector optimization problems," *Journal of Optimization Theory and Applications*, vol. 42, no. 4, pp. 499–524, 1984.

[17] A. Messac, A. Ismail-Yahaya, and C. Mattson, "The normalized normal constraint method for generating the pareto frontier," *Structural and Multidisciplinary Optimization*, vol. 25, no. 2, pp. 86–98, 2003.

[18] M. Ehrgott and X. Gandibleux, *Multiobjective Combinatorial Optimization — Theory, Methodology, and Applications*. Boston, MA: Springer US, 2002, pp. 369–444.

[19] I. Das and J. E. Dennis, "Normal-boundary intersection: A new method for generating the pareto surface in nonlinear multicriteria optimization problems," *SIAM J. on Optimization*, vol. 8, no. 3, 1998.

[20] C. L. I, J. Huang, R. Duan, C. Cui, J. . Jiang, and L. Li, "Recent progress on c-ran centralization and cloudification," *IEEE Access*, vol. 2, pp. 1030–1039, 2014.

[21] "Gurobi optimization - the state-of-the-art mathematical programming solver," http://www.gurobi.com/, accessed: 21-5-2018.

[22] "Omnet++ discrete event simulator - home," https://www.omnetpp.org/, accessed: 21-5-2018.

[23] T. Bai and R. W. Heath, "Coverage and rate analysis for millimeter-wave cellular networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 2, pp. 1100–1114, Feb 2015.