Abstract—In this paper, we formulate a novel throughput aware user association scheme for ultra dense terahertz (THz) networks. In more detail, we introduce a user association problem, which takes into account the THz channel particularities, the directivity of the BSs’ and UEs’ antennas, as well as their position and the UEs’ minimum rate requirements. Moreover, we provide the solution framework, which is based on the grey wolf optimizer (GWO) and returns the optimal user association table. Finally, we present comparative simulation results, which validate the superiority of the proposed framework against the commonly-used particle swarm optimizer (PSO) approach.

I. INTRODUCTION

The increased data rates demand, in the beyond fifth generation (5G) wireless systems, as well as the fact that the spectral efficiency of the microwave links is approaching its fundamental limits have motivated the exploitation of higher frequency bands that offer abundance of communication bandwidth [1]–[4]. In this end, terahertz (THz) band wireless communications have been recognized as an attractive candidate for providing an order of magnitude capacity improvements [5]–[7]. Moreover, THz communications bring novel unique challenges requiring to rethink the conventional user association mechanisms [1]. In more detail, the ultra-wideband extremely directional nature of the communications links in combination with the non-uniform user equipment (UE) spatial distribution may lead to inefficient user association, when the classical minimum-distance criterion is employed. As in the case of millimeter wave communications [8], THz networks are considered to be noise-limited, because the high path-loss attenuates the interference. Hence, user association metrics designed for interference limited homogenous systems are not well suited to THz systems [9]. On the other hand, user association should be designed to meet the dominant requirements of throughput. Additionally, user orientation has a substantial impact on the performance of THz links, due to the fact that directional transmission is required for mitigating the high path-loss. As a consequence, users may not be associated with the geographically closest BS, since a better directional link may exist for a farther away BS.

Scanning the open technical literature, several dynamic UE-BS associations schemes have been presented [10]–[14]. For example, in [10], the authors presented a load-aware cell association method and distributed algorithm for downlink heterogeneous networks aiming in maximizing the total throughput, whereas, in [11], the associations objective was to minimize the total power consumption. Additionally, in [12], a load balancing user association problem is presented and solved for heterogeneous networks deployments, while, in [13], the authors investigated the energy efficient user association problem in HetNets, and formulated a network logarithmic utility maximization problem. Finally, in [14], an optimization-based framework was proposed for energy-efficient global radio resource management in heterogeneous wireless network. However, none of the above presented mechanisms take into account the particularities of the THz channel, as well as the analog beamforming constraints.

Motivated by the above, in this paper, we present a novel user association scheme for ultra dense THz networks, which aims at maximizing the total throughput. In particular, we formulate a UE association problem, which takes into account the THz channel particularities, the directivity of the BSs and UEs antennas, as well as their position and the UEs’ minimum rate requirements. Moreover, we provide the solution framework, which is based on the grey wolf optimizer (GWO) [15] and returns the optimal user association table. GWO is a swarm intelligence meta-heuristic algorithm that models the grey wolf behavior in hunting prey. Finally, we present comparative simulation results, which validate the superiority of the proposed framework against the commonly-used particle swarm optimizer (PSO) approach.

The remainder of this paper is organized as follows. In Section II, the system and channel model is presented, while, Section III is devoted in formulating the optimization problem and providing the solution framework. Finally, respective numerical results and discussions are provided in Section IV, whereas closing remarks are presented in Section V.

II. SYSTEM AND CHANNEL MODEL

As illustrated in Fig. 1, we consider a downlink fully-directional communication scenario in a THz wireless cellular network, which consists of \( N_u \) UEs that can be served by \( N_b \) BSs. By \( \mathcal{U} \), we denote the UEs set, whereas, \( B \) represents the set of the \( N_b \) BSs. Moreover, the transmission power of each BS is assumed to be constant and equals \( P_b \), while the power of the white additive Gaussian noise is \( N_0 \) and the channel gain between the BS \( i \) and UE \( j \) is \( h_{ij} \). Since the association will execute on a large time scale compared to the instantaneous channel fluctuations, without loss of generality, we can assume that the channel gain captures only the impact of the path-loss, while the effect of fast fading can be neglected. By assuming that each UE uses a different fraction of time-frequency resources, the baseband equivalent received signal

–

–
at the BS $i$ from the UE $j$ can be obtained as
\begin{equation}
    r_{ij} = \psi_{ij} h_{ij} s_j + n_i,
\end{equation}
where $h_{ij}$ and $s_j$ respectively stand for the channel coefficient and the transmitted signal, while $n_i$ represents the zero mean AWGN process. Finally, $\psi_{ij}$ is a binary variable, which describes the association of the BS $i$ with the UE $j$ and can be expressed as
\begin{equation}
    \psi_{ij} = \begin{cases} 
    1, & \text{if the UE } j \text{ is associated with the BS } i \\
    0, & \text{otherwise.}
    \end{cases}
\end{equation}

The channel model utilized in this paper comprises of the free space path and molecular absorption gains, i.e., the total path gain for the channel between the BS $i$ and UE $j$, which is established in the frequency $f_{ij}$, can be evaluated as
\begin{equation}
    L(f_{ij}, d_{ij}) = L_f(f_{ij}, d_{ij}) L_a(f_{ij}, d_{ij}),
\end{equation}
where $L_f(f_{ij}, d_{ij})$ and $L_a(f_{ij}, d_{ij})$ respectively represent the free space and molecular absorption path gains, while $d_{ij}$ is the distance between the BS $i$ and UE $j$.

The free space path gain can be evaluate, by radially expanding wavefront adjusted with the BS and UE antenna gains, as
\begin{equation}
    L_f(f_{ij}, d_{ij}) = \frac{c^2}{(4\pi d_{ij} f_{ij})^2} G_b(\Theta_i) G_u(\Theta_j),
\end{equation}
where $G_b(\Theta_i)$ and $G_u(\Theta_j)$ represent the BS $i$ and UE $j$ antenna gain, respectively, and $c$ is the speed of light. Note that the BS and UE antenna gains are usually dependent on the angles of incident $\Theta_i$ and reception $\Theta_j$. By assuming that the TX and RX are aligned and analog pencil beamforming is employed, we can obtain that when the BS and UE antennas are aligned, $G_b(\Theta_i) = G_b$ and $G_u(\Theta_j) = G_u$.

The molecular absorption loss in THz frequency regions can be modeled according to the Beer-Lambert’s law as [16]
\begin{equation}
    \tau(f_{ij}, d_{ij}) = \frac{P_u(f_{ij}, d_{ij})}{P_b(f_{ij})} = e^{-\kappa_a(f_{ij}) d_{ij}},
\end{equation}
where $\tau(f_{ij}, d_{ij})$ stand for the transmittance, $P_b(f_{ij})$ represents the transmitted powers, and $\kappa_a(f_{ij})$ denotes the absorption coefficient describing the relative absorbing area of the molecules in the medium per unit volume. The main cause of absorption loss in millimeter and THz frequencies is the water vapor that causes discrete, but deterministic loss to the signals in the frequency domain. Other molecules, such as oxygen, also cause some level of loss to the signals, but the water vapor dominates the overall molecular absorption loss above 200 GHz frequencies.

In order to evaluate the absorption loss, we utilize a simplified model for the molecular absorption loss in 275–400 GHz band, which was initially presented in [17]. According to this model, the absorption coefficient can be evaluated as
\begin{equation}
    \kappa_a(f_{ij}) = y_1(f_{ij}, \mu) + y_2(f_{ij}, \mu) + g(f_{ij}),
\end{equation}
whereas, based on (5) and (6), the molecular absorption gain can be estimated as
\begin{equation}
    L_a(f_{ij}, d_{ij}, \mu) = \exp(-d_{ij}(y_1(f_{ij}, \mu) + y_2(f_{ij}, \mu) + g(f_{ij}))),
\end{equation}
where $\mu$ denotes the volume of the mixing ratio of water vapor.

Note that the volume of the mixing ratio of water vapor is not equal to its relative humidity and it can be evaluated as
\begin{equation}
    \mu = \frac{\phi}{\Phi_p(T, p)},
\end{equation}
where $\phi$ and $\Phi_p$ respectively stand for the relative humidity and the pressure, whereas $\Phi_p(T, p)$ is the saturated water vapor partial pressure in temperature, $T$, measured in °C, and, according to Buck equation [18], can be calculated as
\begin{equation}
    \Phi_p(T, p) = w_1 (w_2 + w_3 p) \exp \left( \frac{w_4 T}{w_5 + T} \right),
\end{equation}
where $p$ is given in Pa and $w_1 = 6.1121$, $w_2 = 1.0007$, $w_3 = 3.46 \times 10^{-8}$, $w_4 = 17.502$ and $w_5 = 240.97$. Moreover, $y_1 (f_{ij}, \mu) = \frac{C_1(\mu)}{B(\mu) + A(\mu)}$, $y_2 (f_{ij}, \mu) = \frac{C_2(\mu)}{B(\mu) + A(\mu)}$ and $g (f_{ij}) = p_1 f_{ij}^3 + p_2 f_{ij}^2 + p_3 f_{ij} + p_4$, where $c_1 = 10.835$ cm$^{-1}$, $c_2 = 12.664$ cm$^{-1}$, $p_1 = 5.54 \times 10^{-37}$ Hz$^{-2}$, $p_2 = - 3.94 \times 10^{-25}$ Hz$^{-2}$, $p_3 = 9.06 \times 10^{-14}$ Hz$^{-2}$, $p_4 = -6.36 \times 10^{-3}$, and $A(\mu) = 0.2205\mu(0.1303\mu + 0.0294)$, $B(\mu) = (0.4093\mu + 0.0092)^2$, $C_1(\mu) = 2.014\mu(0.1702\mu + 0.0303)$, $C_2(\mu) = (0.537\mu + 0.0956)^2$. This model was shown to have high accuracy for up to 1 km links in standard atmospheric conditions, i.e., temperature of 296 °K and pressure of 101325 Pa [17]. Moreover, since the impact of the temperature and pressure can be modeled through the water vapor mixing ratio, $\mu$, (7) can describe the THz link molecular absorption pathloss beyond the standard atmospheric conditions. Finally, note that the parameters $c_1$, $c_2$, $p_1$, $p_2$, $p_3$, and $p_4$ can be considered relatively independent of the atmospheric conditions [17].

Based on (3), (4), and (7), the total path gain can be rewritten as
\begin{equation}
    L(f_{ij}, d_{ij}, \mu) = \frac{c^2}{(4\pi d_{ij} f_{ij})^2} G_b G_u \times \exp(-d_{ij}(y_1(f_{ij}, \mu) + y_2(f_{ij}, \mu) + g(f_{ij}))).
\end{equation}
From (9), it is evident that the total path gain depends not only on the operation frequency, \( f_{ij} \), and the distance between the BS \( i \) and UE \( j \), but also on the BS and UE antenna gains, as well as the atmospheric conditions.

### III. Problem Formulation & Solution

According to (1) and (9), the achievable data rate of the link between the BS \( i \) and UE \( j \) can be obtained as

\[
    r_{ij} = \log_2 (1 + \gamma_{ij}),
\]

where

\[
    \gamma_{ij} = \frac{P G_i^b C_j^b L(f_{ij}, d_{ij}, \mu)}{N_0}.
\]

For a fixed network topology, which is assumed to be known to every BS \( i \) and UE \( j \), the optimal cell formation should satisfy the following optimization problem

\[
    \max_{\Psi, C} \sum_{j \in \mathcal{U}, i \in \mathcal{B}} r_{ij} c_{ij}
\]

s.t.\n
\[
    \begin{align*}
    C_1 & : \sum_{j \in \mathcal{U}} c_{ij} \leq 1, \ \forall i \in \mathcal{B}, \\
    C_2 & : \sum_{i \in \mathcal{B}} \psi_{ij} = 1, \ \forall j \in \mathcal{U}, \\
    C_3 & : 0 \leq \psi_{ij} \leq \psi_{ij}, \ \psi_{ij} \in \{0,1\}, \ \forall i \in \mathcal{B}, \\
    C_4 & : \psi_{ij} = 0 \text{ if } R_{ij} \leq R_{j,\min}, \\
    C_5 & : 0 \leq \phi_{ij}^b \leq 2\pi, \ \forall i \in \mathcal{B}, \\
    C_6 & : 0 \leq \phi_{ij}^\theta \leq 2\pi, \ \forall j \in \mathcal{U},
\end{align*}
\]

where \( c_{ij} \) is the fraction of resources that the BS \( i \) employs to serve the UE \( j \). Furthermore, \( \phi_{ij}^b \) and \( \phi_{ij}^\theta \) respectively stand for the boresight angles of the BS \( i \) and UE \( j \), whereas \( R_{j,\min} \) is the minimum required rate for the UE \( j \). Finally, \( \Psi \) and \( C \) are matrices that collects all the user association variables, \( \psi_{ij} \), and fraction of resources, \( c_{ij} \), used by the BS \( i \) to serve the UE \( j \), respectively.

In (12), the constraint \( C_2 \) guarantees the association of the UE \( j \) to exactly one BS, whereas the constraints \( C_3 \) and \( C_4 \) ensures that the minimum acceptable QoS for every UE. Additionally, \( C_3 \) ensures that hat every BS \( i \) will provide a positive resource share only to its associated UEs. Note that the solution of (12) provides a long-term association policy along with proper orientation and operating beamwidths for fully-directional wireless THz communications. This solution guarantees the optimal UE-BS association as long as the inputs of the optimization problem, namely network topology and UE demands, are unchanged. If a UE requires more resources or loses its connection, the optimization problem in (12) has to be re-executed.

In order to find the optimal solution for the association problem, described in (12), we present Algorithm 1, which is based on the swarm intelligence meta-heuristic algorithm called GWO [15] and returns the optimal user association matrix. According to the GWO, the solutions are grouped into four groups, namely \( \alpha, \beta, \delta, \) and \( \omega \), based on their optimality. In more detail, \( x^\alpha \) vector represents the best solution, while \( x^\beta \) and \( x^\delta \) are respectively the second and third optimal solution vectors. Finally, \( \omega \) is the set of the rest feasible solutions. Moreover, in GWO we define the coefficient vectors \( A \) and \( K \). The definition of the vector components can be found in the expressions below. Furthermore, we define the vector \( \alpha \) whose components are linearly decreased from 2 to 0 with the number of iterations. A brief description of the GWO algorithm for the user association problem is given in Algorithm 1. In Algorithm 1, \( N_P \) denotes the population size, and \( G_{\text{max}} \) is the maximum number of generations. We notice that these are the only inputs for a given problem, since all the other parameters are randomly generated. Thus, GWO does not require additional control parameter setting like other algorithms. The population of \( N_P \) vectors is initialized randomly from a uniform distribution. After the objective function value of each vector is calculated the algorithm sorts the vectors descending according to objective function value. The best solution is selected as the \( x^\alpha \) vector, while the second best is the \( x^\beta \), and the third best is the \( x^\delta \) vector respectively.

Then the algorithm main loop starts.

In the \( G + 1 \) generation, the position of the \( k \)-th vector in the \( n \)-th dimension is evaluated as

\[
    x_{nk}(G + 1) = x_{nk}^\alpha(G) + \frac{x_{nk}^\beta(G) - x_{nk}^\delta(G)}{3},
\]

where

\[
    x_{nk}^\beta(G) = x_{nk}^\alpha(G) - A_{nk}^j(G)D_{nk}^j(G), \ \text{for} \ \{i, j\} \in \{\alpha, 1\}, (\beta, 2), (\delta, 3) \}
\]

and

\[
    D_{nk}^j(G) = |K_{nk}(G)x_{nk}^j(G) - x_{nk}(G)|,
\]

with \( A_{nk}^j(G) = \alpha \left(2f_{nk}^j(G) - 1\right) \) and \( K_{nk}(G) = 2f_{nk}^j(G) \). Note that \( f_{nk}^j(G) \) and \( f_{nk}^\beta(G) \) are uniformly distributed randomly selected numbers in the range of \([0, 1]\). The algorithm updates the \( \alpha \), \( A \) and \( K \) vectors according.
to the above relations and evaluates the objective function value of the new vectors. It must be noted that the \( x_{nk} \) values obtained are also checked using a boundary handling mechanism to ensure they lie within the allowable limits. Then again the new \( x^0, x^1, x^2 \) vectors are selected if better than the previous values. The stopping criterion is the maximum number of generations.

Next, we present Algorithm 2, which returns the objective function value, for a feasible solution.

**Algorithm 2** Calculate objective function

1: Input a possible solution vector \( y \)
2: for \( i = 1 \) to \( N_u \) do
3: Calculate rate \( R_{ij} \) for the \( i \)-th user connecting to the \( j \)-th BS
4: if \( R_{ij} \geq R_{j,\text{min}} \) then
5: if the \( j \)-th BS can support the desired rate then
6: calculate \( c_{ij} \)
7: end if
8: end if
9: end for
10: Calculate objective function value according to (12)
11: Return objective function value

---

**IV. RESULTS & DISCUSSION**

This section is devoted in presenting simulation results that evaluate the performance of the solution framework. We consider a scenario in which 120 UEs are randomly deployed and served by 6 BSs. Each UE requires a minimum data rate that is randomly selected within \([1, 10 \text{ Gb/s}]\) with bandwidth equals 1 GHz. Moreover, we assume that the UE and the BS are placed in a uniform random way within a circle of radius equals 50 m, while \( 10 \log_{10} \left( \frac{PG_iG_j}{N_o} \right) = 120 \text{ dB} \). Additionally, a temperature, relative humidity and pressure of 25°C, 50% and 101325 are respectively assumed. The population size is to 200 and the number of maximum generations is set to 150 for both algorithms. Finally note that in order to compare the effectiveness of the proposed method, we compare it with the PSO approach.

We executed the GWO and the PSO algorithms for 20 independent runs. In each case we generate a random topology of \( N_u \) users and \( N_b \) BS. It must be pointed out that both algorithms are applied to the same topology in order to better compare results.

Fig. 2 depicts the average convergence rate as a function of the number of iterations, for both the GWO-based and PSO algorithms. From this figure, it is observed that GWO-based algorithm was able to converge slightly faster than PSO and cover less at a higher objective function values. Additionally, the convergence rate for the first about 30 generations of GWO-based algorithm is quite faster than the PSO.

The cumulative distribution function (CDF) of the date rates obtained by both algorithms is depicted in Fig. 3. It is evident that the PSO obtained values are better for smaller data rates and probabilities less than 0.2. On the other hand, as the average user rate increases, the GWO-based algorithm outperforms the PSO. As a consequence, the GWO-based algorithm achieves faster and higher objective values compared to PSO. This indicates the effectiveness of the proposed solution framework. Moreover, Fig. 4 shows the utility function value obtained by both algorithm in each algorithm run. We notice the GWO outperforms PSO in 17 out of the 20 cases. The percentage of users served by the network is depicted in Fig. 5. Again GWO clearly outperforms PSO in most of the cases. Next we evaluate the algorithms performance for increasing user number and we set the circle
radius to 10 m and obtain results from 10 to 260 users with step 10 using both algorithms. Fig. 6 depicts the results for this case. We notice that both algorithms obtain results that are very close. PSO obtains in general better results for small user number. GWO manages to obtain good results for a larger number, which indicates that it is better suited for higher dimensional problems. We also notice that as the number of users increases the utility function value tends close to 10 Gbps while the percentage of users served tends to about 43%. The algorithms can both serve all the users in every run, when the user number is below 60.

V. CONCLUSIONS

In this paper, we presented a user association problem for ultra dense THz networks, and we provided an appropriate solution framework, which is based on the grey wolf optimizer (GWO) and returns the optimal user association table. Additionally, comparative simulation results, which validate the superiority of the proposed framework against the commonly-used particle swarm optimizer (PSO) approach, were given. The results revealed that, although, both algorithms can be used as powerful optimizers for the user association problem, the GWO-based solution outperforms the PSO. Another benefit of GWO is that the only control parameter that is required is the population size and number of iterations.

ACKNOWLEDGMENT

This work has received funding from the European Commission’s Horizon 2020 research and innovation programme under grant agreement No. 761794.

REFERENCES


