Inbound Handover Interference-Based Margin for Load Balancing in Heterogeneous Networks

Mohanad Alhabo, Li Zhang, and Obinna Oguejiofor
School of Electronic and Electrical Engineering, University of Leeds, UK.
E-mail: {elmdj, l.x.zhang, eloso}@leeds.ac.uk

***Abstract*—Inbound handover (HO) or hand-in is accomplished when the user equipment (UE) performs HO from macrocell (MC) to a small cell (SC) [1] [2]. When the UE connects to a SC with a time of stay (ToS) less than a predefined time threshold, this will result in frequent unnecessary HOs and also increase service interruption which in turn will degrade the end user quality of service (QoS). In this paper, we propose an inbound HO method for the purpose of throughput enhancement and load balancing in SC heterogeneous networks (HetNets). The impact of interference from both MC and SC tiers is considered so that the UE is offloaded from congested MC and forced to perform the HO to the SC tier that supplies a sufficient data rate by selecting a proper SC target, which has the highest signal to interference plus noise ratio (SINR), from a reduced neighbour cell list (NCL). The proposed method uses a modified A3 HO triggering condition taking into account the interference and cell load. Results show that our proposed method can perform inbound HO while keeping the throughput to the maximum level. Moreover, the proposed method has significantly minimized the unnecessary inbound HOs and radio link failures compared to the competitive methods. With different network load factors, the proposed method can significantly give a good performance which yields higher throughput for the user and the network as well.**

I. Introduction

The incorporating of dense SCs to the HetNet environment has a tremendous effect on the network performance in terms of maximizing the coverage and capacity, in addition to reducing the burden of traffic on the MC by offloading the UEs from the MC tier to the SC tier. However, the SC densification introduces new problems into the network, such as interference and mobility issues [3]. In traditional homogeneous networks, the UE performs HO to the neighbour cell if its power is offset greater than that of the serving cell for a period of time known as a time to trigger (TTT). Moreover, these parameters, i.e. the offset and TTT, are identical to all cells in the network [4]. Therefore, the traditional homogeneous HO methods are not sufficient for SC heterogeneous networks. Since the MC transmits at much higher power compared to the SCs, the users will always prefer to be connected to MC rather than a SC. Hence, the proper utilization of SCs is not achieved. This will lead to a severe congestion in the MC tier and eventually ends up with a lower network throughput.

The mobility of the UE in ultra-dense SCs HetNets is a big challenge because the UE has a vast number of target cells for HO [5]. Therefore, reducing the number of possible targets for HO is an appropriate strategy to minimize the signalling overhead in the network. Different methods that deal with this issue have been conducted in the literature. In [6], the authors proposed a method to automatically adjust the HM for outbound HO from SC to MC. This method adjusts the HM according to the UE speed so that for fast moving UE the HM is decreased (avoiding late HO), and for low speed UE the HM is increased (avoiding early HO). The method has helped in avoiding late and early HOs in addition to the reduction in the radio link failures for different UE speeds. However, no mechanism for adjusting the traffic load between the SC and MC tiers is considered which may lead to a severe congestion in the MC tier, hence high call dropping rate is expected. The authors in [7] proposed a method to minimize the unnecessary HOs by reducing the number of scanned SCs. The building of the SC list is based on the downlink received power and ToS criteria which avoid the SCs with a short time of stay. The UE performs HO to the SC with the strongest downlink received power from the list. However, the interference scenario and cell load are not taken into account in this work which may lead to throughput unaware HO strategy and radio link failures. In [8], the authors proposed a method to reduce the number of target SC NCL and minimize the probability of unnecessary HO in HetNet. The NCL is constructed using the distance between the UE and the SC in addition to the UE angle of movement. The average human walking speed is used to prevent high-speed UEs from performing HO to SCs. Results show a good performance in terms of SC NCL reduction, unnecessary HO minimization and network throughput enhancement. The authors in [9] proposed a method to minimize the probability of unnecessary HO and HO failure. An estimated ToS criteria is used to exclude SC from the target HO NCL. The HO is performed to the SC which provides the higher SINR and has enough capacity to deliver services to the UE. Results revealed that this method has minimized the unnecessary HO and HO failure.

In this paper, we proposed an inbound HO method for the purpose of throughput enhancement and load balancing in SC HetNets. Interference-based HO is considered so as to maximize the throughput. This work considers the reduction of SC NCL by incorporating the interference level, using the SINR, and ToS as HO criteria so as to increase the efficient utilization of SCs and in turn increase the end user QoS by offloading the users from MC base station to SCs. A modified A3 HO triggering condition is proposed by considering the traffic load in the serving MC base station and an equivalent SINR received from a SC within the reduced NCL, which gives a good data rate compared to the serving MC. Results prove that our proposed method yields high throughput for the end user when compared to other works in the literature. The overall network throughput is also improved under different network conditions such as load factor, and different levels of noise.

This paper is organized as follows. System model is given in section II. Section III presents the proposed method process. In section IV the performance of the proposed method and the results are analysed. Finally, section V concludes the paper.

II. System Model

System model in this paper is based on two-tier HetNet which consists of SCs overlaid under the coverage area of MC base station, as depicted in Fig.1. A hexagonal MC is deployed with three sectors. SCs are deployed randomly according to uniform distribution. The MC and SCs are deployed on the same frequency. The UE mobility can be expressed using two parameters: UE velocity, V_{ue}, and UE direction, θ_k. These two parameters can be defined as Gaussian distribution [10] and are updated accordingly by using the following equations

$$V_{ue} = \mathcal{N}(v_m, v_{std}), \tag{1}$$

$$\theta_k = \mathcal{N}(\theta_m, 2\pi - \theta_m \tan(\frac{\sqrt{V_{ue}}}{2})\Delta t), \tag{2}$$

where v_m represents the mean velocity of the user, v_{std} denotes the standard deviation of the user velocity, θ_m is the previous direction of the user, Δt is the period between two updates of the mobility

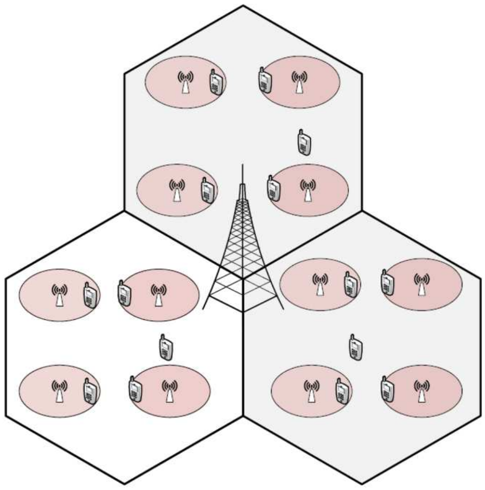

Figure 1: Two-tier HetNet system model

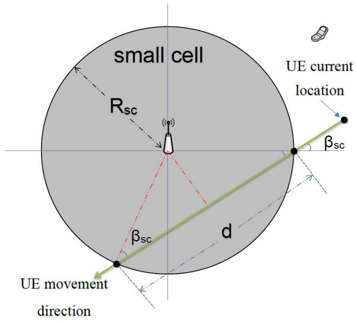

Figure 2: UE ToS measurement

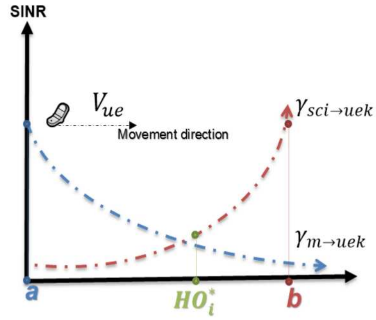

Figure 3: Handover point

model, and $\mathcal{N}(x, y)$ is a Gaussian distribution with mean x and standard deviation y.

According to Shannon's capacity equation, the maximum data rate, $r_{i\to ue_k}$, is given as

$$r_{i\to ue_k} = \text{BW}\log_2(1+\gamma^r_{i\to ue_k}), \tag{3}$$

where BW is the carrier bandwidth, and $\gamma^r_{i\to ue_k}$ is the SINR received at user k from base station i. The SINR from the SC i and MC received at user can be respectively written as

$$\gamma^r_{sc_i\to ue_k} = \frac{P^{(d)}_{t,sc_i\to ue_k} \mid h^{(d)}_{sc_i\to ue_k} \mid^2}{P^M_{I\to ue_k} + P^{sc}_{I\to ue_k} + \sigma}, \tag{4}$$

$$\gamma^r_{m\to ue_k} = \frac{P^{(d)}_{t,m\to ue_k} \mid h^{(d)}_{m\to ue_k} \mid^2}{P^{sc}_{I\to ue_k} + \sigma}, \tag{5}$$

where $\gamma^r_{m\to ue_k}$ is the SINR received from MC at the user k, $\gamma^r_{sc_i\to ue_k}$ is the SINR received from SC i at the user k,
$P^{sc}_{I\to ue_k}$ is the total interference power from the neighbouring SCs,

$$P^{sc}_{I\to ue_k} = \sum_{j=1, j\neq d}^{N_{sc}} P^{(int)}_{t,sc_j\to ue_k} \mid h^{(int)}_{sc_j\to ue_k} \mid^2 \tag{6}$$

$P^M_{I\to ue_k}$ is the total interference power from the MC,

$$P^M_{I\to ue_k} = P^{(int)}_{t,m\to ue_k} \mid h^{(int)}_{m\to ue_k} \mid^2 \tag{7}$$

$P^{(.)}_{t,m\to ue_k}$ and $P^{(.)}_{t,sc_i\to ue_k}$ are respectively the transmitting power of MC and SC i on a user k, $h^{(.)}_{m\to ue_k}$ is the channel gain between the UE and MC, $h^{(.)}_{sc_i\to ue_k}$ is the channel gain between the UE and SC i, σ is the noise power, $^{(d)}$ denotes a desired link, $^{(int)}$ denotes an interferer link, and finally N_{sc} is a set representing the total number of SCs in the network.

Taking into account the heterogeneous network architecture, the propagation model between the MC and the user is defined as in [11] by

$$\delta_{m\to ue_k} = 128.1 + 37.6\ \log_{10}(d_{m\to ue_k}) + \xi_{m\to ue_k}, \tag{8}$$

where $d_{m\to ue_k}$ is the distance between the user and the MC base station in kilometres, and ξ is a Gaussian distribution random variable with zero mean and 12 dB standard deviation [12].
For outdoor SC, the path loss is defined as in [13] by

$$\delta_{sc_i\to ue_k} = 38 + 30\ \log_{10}(d_{sc_i\to ue_k}) + \xi_{sc_i\to ue_k}, \tag{9}$$

where $d_{sc_i\to ue_k}$ is the distance between the user and SC i in metres.

Given that the UE time of stay can be expressed using the velocity, V_{ue}, and the expected distance that the user will spend inside the base station coverage area as shown in Fig.2. The angle β_{sc} can be represented as a random variable which is uniformly distributed and restricted to interval $[\frac{-\pi}{2}, \frac{\pi}{2}]$. This random variable has a constant density over the interval i.e. has a probability density function (PDF) $f_{\beta_{sc}}(\beta_{sc})$. Thus, we can define the mean ToS a user will stay in SC as

$$\begin{aligned} \text{E}\Big[ToS_{ue\to sc_i}\Big] &= E\left[\frac{2R_{sc_i}\cos(\beta_{sc})}{V_{ue}}\right] \\ &= 2\int_0^{\frac{\pi}{2}} \frac{2R_{sc_i}\cos(\beta_{sc})}{V_{ue}}\frac{1}{\pi}d\beta_{sc} \\ &= \frac{4R_{sc_i}}{\pi V_{ue}}, \end{aligned} \tag{10}$$

where R_{sc_i}, and β_{sc}, are respectively the SC radius, and the UE angle of entry to the SC. We can have an estimate of the SC radius when the user begins to receive a minimum required signal power for service continuity, denoted as P^{th}_{min}, from the SC [14] as

$$R_{sc_i} = \Big(\frac{P^{(d)}_{t,sc_i\to ue_k}\ 10^{\xi/10}}{P^{th}_{min}}\Big)^{\frac{1}{\zeta}}, \tag{11}$$

where ξ is a Gaussian distribution random variable with zero mean and 12 dB standard deviation [12], and ζ is the path loss exponent.

As shown in Fig.3, the HO is taken place at point HO^*_i, which is the point at which the UE is receiving $r_{sc_i\to ue_k} > r_{m\to ue_k}$. The aim is to find $HO^*_i\ \forall i = 1, ..., N_{sc}$, to maximize the throughput and load balance the traffic between MC and SC. Therefore, we can write the formula to find the HO point, HO^*_i, which will offer the maximum downlink throughput, as

$$HO^*_i = \arg\max_{i\in n}\ r_{i\to ue_k}, \tag{12}$$

where n is a set representing the total base stations in the HetNet n = $\{0,1,\ldots,N_{sc}\}$ where 0 represents the MC and N_{sc} is the number of SCs.

In the following subsections we explain the analysis and calculations of the resource assignment and loads, the equivalent SINR required to perform inbound HO to SC and the interference-based load-dependent margin.

A. Resource Assignment and Load Calculations

The influence of the received interference at the user side in HetNet is largely affected by the amount of resources used by the HetNet base stations i.e. SCs and MCs. Also, the power of these resources has a significant influence on interference level. Cell load factor is the amount of resource usage with respect to the available resources in the cell [13]. In other word, a UE receives varied interference levels according to a varied base station load levels. For the i^{th} MC sector, the load L_{m_i} is defined as number of physical resource blocks (PRBs) being used by all mobile users connected to the aforementioned sector divided by the total MC PRBs, that is

$$L_{m_i} = \frac{RB_{m_i}^{ue}}{RB_{tm}}, \tag{13}$$

where $RB_{m_i}^{ue}$ is the number of PRBs used by all active mobile users connected to the MC sector i, and RB_{tm} is the total number of PRBs in the MC.

The number of PRBs used by all active mobile users connected to the MC sector i, i.e. $RB_{m_i}^{ue}$, can be expressed as

$$RB_{m_i}^{ue} = \sum_{k=1}^{N_{ue}^{sec}} RB_{m_i,k}, \tag{14}$$

where N_{ue}^{sec} is the number of UEs in the sector, and $RB_{m_i,k}$ is the number of PRBs used by user k.

B. Equivalent SINR Analysis

The HO_i^*, see Fig.3, is the point at which the data rate of the SC is greater than that of the MC. In other word, it is the point at which $\gamma_{sc_i \to ue_k}^r > \gamma_{m \to ue_k}^r$ for inbound HO. For a given SC i and MC, recall equations (4) and (5). We first apply the condition ($\gamma_{sc_i \to ue_k}^r > \gamma_{m \to ue_k}^r$) to the two equations

$$\frac{P_{t,sc_i \to ue_k}^{(d)} \mid h_{sc_i \to ue_k}^{(d)} \mid^2}{P_{I \to ue_k}^{M} + P_{I \to ue_k}^{sc} + \sigma} > \frac{P_{t,m \to ue_k}^{(d)} \mid h_{m \to ue_k}^{(d)} \mid^2}{P_{I \to ue_k}^{sc} + \sigma}. \tag{15}$$

Substituting (6) and (7) in equation (15) and after some simplifications we get

$$P_{t,sc_i \to ue_k}^{(d)} \mid h_{sc_i \to ue_k}^{(d)} \mid^2 > \frac{A}{B}, \tag{16}$$

where A and B are

$$A = P_{t,m \to ue_k}^{(d)} \mid h_{m \to ue_k}^{(d)} \mid^2 \big(P_{t,m \to ue_k}^{(int)} \mid h_{m \to ue_k}^{(int)} \mid^2 + \sum_{j=1, j \neq d}^{N_{sc}^{**}} P_{t,sc_j \to ue_k}^{(int)} \mid h_{sc_j \to ue_k}^{(int)} \mid^2 + \sigma\big), \tag{17}$$

$$B = \sum_{j=1, j \neq d}^{N_{sc}^{**}} P_{t,sc_j \to ue_k}^{(int)} \mid h_{sc_j \to ue_k}^{(int)} \mid^2 + \sigma. \tag{18}$$

The inbound HO will be initiated to the SC with the highest data rate i.e. at HO point HO_i^*. In other word, we can say that the inbound HO is triggered when the downlink received power from the SC satisfy (16). Without loss of generality, we substitute (16) in (4) to obtain the equivalent SINR $\gamma_{sc_i \to ue_k}^{req}$, for inbound HO from MC to SC i, that provides at least the same data rate as the current serving MC, that is

$$\gamma_{sc_i \to ue_k}^{req} > \frac{P_{t,sc_i \to ue_k}^{(d)} \mid h_{sc_i \to ue_k}^{(d)} \mid^2}{P_{I \to ue_k}^{M} + P_{I \to ue_k}^{sc} + \sigma}$$
$$\therefore \gamma_{sc_i \to ue_k}^{req} > \frac{A/B}{P_{I \to ue_k}^{M} + P_{I \to ue_k}^{sc} + \sigma}, \tag{19}$$

It is worth noting that the summation of interference term, in equations (15) to (18), considers only the SCs in a reduced NCL set N_{sc}^{**} (will be given in section III as defined in (25)) which will, in turn, reduce the computation complexity since we only have a reduced number of SCs in this set.

C. Inbound Handover Interference-Based Margin

To maintain mobility load balancing in general, if the serving MC is overloaded then it increases the HO margin so as to trigger the inbound HO early to the SC. However, this unplanned increase may cause radio link failure and ping-pong HO issues, and hence, poor QoS is delivered to the UE. Therefore, these parameters should be adjusted dynamically according to the cell load to maintain the mobility robustness. For this reason, we aim to force the UE to perform inbound HO to SC i, which has lower load and hence lower resource utilization. The proposed method will bias the HO point between the congested MC and SC i.

Given the conventional A3 HO triggering condition, which depends on a power-based margin, when the power of the neighbour SC i is offset greater than that of the serving MC for a period of TTT [3], that is

$$P_{sc_i \to ue_k}^r > P_{m \to ue_k}^r + HM_m - HM_{m,sc_i}, \tag{20}$$

where HM_m is the hysteresis parameter of MC and HM_{m,sc_i} is the SC i specific offset with respect to the MC (i.e. the hysteresis set by MC to HO to the SC).

Inspired by (20), we proposed to modify this criteria to facilitate an interference-based load-dependent hysteresis margin. We will consider the SINR instead of downlink received power and replace the power margin HM_{m,sc_i} with interference-based load-dependent margin, denoted $\gamma_{m \to sc_i}^{pro}$, namely the proposed interference-based load-dependent margin to control the HO point between MC and SC i. For inbound HO, in order to balance the load, the HO point HO_i^* must be moved closer to the serving MC rather than being closer to the target SC i. To adjust the HO point for a UE trying to perform inbound HO from MC to SC i, we must shift HO_i^* point to the left as shown in Fig.4 i.e. the HO point will be changed from the intersection point of the two curves $\gamma_{sc_i \to ue_k}^r$ and ($\gamma_{m \to ue_k}^r + \gamma_{th} - \gamma_{m \to sc_i}$) to the intersection point of the two curves $\gamma_{sc_i \to ue_k}^r$ and ($\gamma_{sc_i \to ue_k}^{req} - \gamma_{m \to sc_i}^{pro}$), note that $\gamma_{sc_i \to ue_k}^{req}$ is taken from (19). In other word, the congested MC adjusts the HO parameter $\gamma_{m \to sc_i}^{pro}$ to allow the UE to perform early HO to SC i. For Fig.4, γ_{th} is the outage threshold and is set to 5 dB [15], $\gamma_{sc_i}^{max}$ is the SINR from SC i when $\gamma_{m \to ue_k}^r$ is equal to γ_{th}. Based on [3] and so as to maintain the radio link failures to a lower level, the hysteresis can be assigned according to UE speed. Therefore, we adjust the value of $\gamma_{m \to sc_i}$ to 4 dB for low speed UE ($V_{ue} \leq$ 20km/h), 3 dB for medium speed UE (20km/h $< V_{ue} \leq$ 50km/h) and 2 dB for high speed UE ($V_{ue} >$ 50km/h).

To incorporate the impact of the UE velocity on the proposed margin, we proposed to incorporate the margin $\gamma_{m \to sc_i}$ into equation (21) to find the load-dependent parameter, denoted as L_m^{mr}, which will be used later to calculate the proposed margin

$$L_m^{mr} = (1 - L_{m_i}) \cdot \gamma_{m \to sc_i} \tag{21}$$

where L_m^{mr} is the load-dependent parameter for inbound HO. Finally, the proposed interference-based load-dependent margin is calculated as

$$\gamma_{m \to sc_i}^{pro} = \gamma_{m \to sc_i} - L_m^{mr}, \tag{22}$$

The parameter L_m^{mr} depends on L_{m_i}: the higher the value of L_{m_i} the smaller the value of L_m^{mr}, the higher the proposed margin, and eventually the closer the HO_i^* point to SC i. Lower the value of L_m^{mr} means that the serving MC is heavily loaded, hence, the HO point is moved closer to the MC so as to speed up the HO triggering which will balance the traffic load by offloading it from the congested MC to SC i.

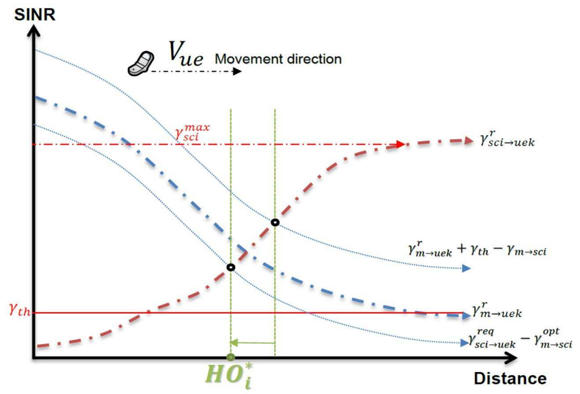

Figure 4: Handover point for inbound HO

Now we have $\gamma^{req}_{sc_i \to ue_k}$, and $\gamma^{pro}_{m \to sc_i}$, then the modified A3 HO triggering event can be rewritten for inbound HO as

$$\gamma^{r}_{sc_i \to ue_k} \geq \gamma^{req}_{sc_i \to ue_k} - \gamma^{pro}_{m \to sc_i}. \tag{23}$$

The above condition in (23) should hold for a period of TTT according to the UE speed [16] as depicted in table I.

Table I: TTT according to UE speed

UE speed (km/h)	$V_{ue} \leq 20$	$20 < V_{ue} \leq 50$	$V_{ue} > 50$
TTT (ms)	1280	512	256

III. Proposed Method Process

Algorithm 1 illustrates the proposed method procedures where $ToS_{ue \to sc_i}$ is the expected time of stay of the user in SC i coverage area, T_{th} is the time threshold for ToS, and N^*_{sc} is a set which represents the total number of SCs with an SINR greater than the outage threshold.

Algorithm 1 Proposed Method

1: **Procedure Starts**
2: User **moves** to SC coverage area
3: **if** $\gamma^{r}_{sc_i \to ue_k} \leq \gamma_{th}$ **then**
4: **Exclude** this SC from Handover target cell list N^*_{sc}
5: **end if**
6: **for** $i \leftarrow 1, N^*_{sc}$ **do**
7: **Estimate** $ToS_{ue \to sc_i}$
8: **if** $\mathrm{E}\big[ToS_{ue \to sc_i}\big] > T_{th}$ **then**
9: **Keep** SC i in the new Handover list N^{**}_{sc}
10: **end if**
11: **end for**
12: **Convert** $\gamma^{r}_{m \to ue_k}$ to its equivalent $\gamma^{req}_{sc_i \to ue_k}$
13: **Calculate** $\gamma^{pro}_{m \to sc_i}$
14: **Select** the SC with the maximum $\gamma^{r}_{sc_i \to ue_k}$ from N^{**}_{sc}
15: **if** $\gamma^{r}_{sc_i \to ue_k} \geq \gamma^{req}_{sc_i \to ue_k}$ - $\gamma^{pro}_{m \to sc_i}$ for TTT **then**
16: **if** $RB^{ue}_{sc_i} < 1$ **then**
17: **Handover** the user to sc_i
18: **end if**
19: **end if**
20: **end procedure**

The proposed algorithm begins by eliminating the SCs that could cause degradation in the user QoS, i.e. SCs with SINR less than the outage threshold γ_{th}, resulting in a candidate list N^*_{sc} which is written as

$$N^*_{sc} = \{sc_i \in N_{sc} \mid \gamma^{r}_{sc_i \to ue_k} > \gamma_{th}\}. \tag{24}$$

Then, for an active mobile user k, a SC NCL is formed, denoted as N^{**}_{sc} set, containing all SCs whose predicted mean UE ToS is greater than the time threshold T_{th}. Thus, we can re-write the candidate SC list as

$$N^{**}_{sc} = \{sc_i \in N^*_{sc} \mid \mathrm{E}\big[ToS_{ue \to sc_i}\big] > T_{th}\}. \tag{25}$$

The UE performs inbound HO from MC to SC i if the data rate from SC i is higher than that of the MC. This is done by converting the SINR received from the MC, i.e. $\gamma^{r}_{m \to ue_k}$, to its equivalent SINR received from the SC, i.e. $\gamma^{req}_{sc_i \to ue_k}$, which gives a higher data rate. Then, the resulted SINR from the previous step is compared with the actual SINR received from the SC i considering the interference-based load-dependent margin $\gamma^{pro}_{m \to sc_i}$ (i.e., applying the proposed modified A3 event). The inbound HO is performed to SC i providing that the PRBs of this SC is sufficient to provide resources to the user. In line (16) in Algorithm 1, the condition $RB^{ue}_{sc_i} < 1$, the value 1 here means that the SC resources are all occupied by other users and it is not possible to perform the inbound HO to this SC.

IV. Performance Analysis and Results

The performance of our proposed method is compared against the performance of two competitive methods, namely the conventional method and the estimated time-of-stay-based cell selection (ETCS) method presented in [7]. In the conventional method, the UE periodically performs neighbourhood scanning, based on the downlink received power, to form the HO target cell list. This means that the UE will spend a significant time period to select the proper target. Then, the UE performs the inbound HO to the SC with the strongest downlink received power without considering the interference and load balancing scenario which means that the HO point HO^*_i for this method is downlink power constrained. This will cause an UE throughput reduction and wasting the battery power of the UE due to the frequent scanning measurement especially in dense SC environment. Therefore, the HO target SC for the conventional method, denoted as sc^t_{conv}, can be expressed as

$$sc^t_{conv} = \Big\{sc_i \in N_{sc} \mid \; P^{(d)}_{t, sc_i \to ue_k} \mid h^{(d)}_{sc_i \to ue_k} \mid^2 > P^{(d)}_{t, m \to ue_k} \mid h^{(d)}_{m \to ue_k} \mid^2 \Big\}. \tag{26}$$

Whereas the ETCS method in [7], forms the HO target cell list based on the downlink received power and ToS criteria which means to avoid the SC that could cause short time of stay phenomena. Then, the UE performs the inbound HO to the cell with the strongest power from the list. Also the interference scenario and cell load balance are not considered in this method and the HO point HO^*_i is based on the power difference between the serving the target cells. We can write the HO target SC, sc^t_{etcs}, for this method as

$$sc^t_{etcs} = \Big\{sc_i \in N_{sc} \mid \; (\mathrm{E}\big[ToS_{ue \to sc_i}\big] > T_{th}) \wedge (P^{(d)}_{t, sc_i \to ue_k} \mid h^{(d)}_{sc_i \to ue_k} \mid^2 > \; P^{(d)}_{t, m \to ue_k} \mid h^{(d)}_{m \to ue_k} \mid^2) \Big\}. \tag{27}$$

On the contrast, our proposed method forms the HO target SC list based on the ToS criteria and interference constraint. Then, the UE performs the inbound HO to the cell that gives a better data rate with load balancing considerations, providing that the PRBs are available, considering a modified A3 HO triggering condition to ensure high QoS which means that the HO point is interference based as given in (23).

The outage probability or the probability of transmission failure happens either when the inbound HO is initiates but an interruption stops the process before completion (before the HO execution time expires) due to the degrades SINR from the serving MC and the target SC, or when the SINR of the serving MC is degraded, the SINR of SC is sufficient enough and the SC has lack of resources.

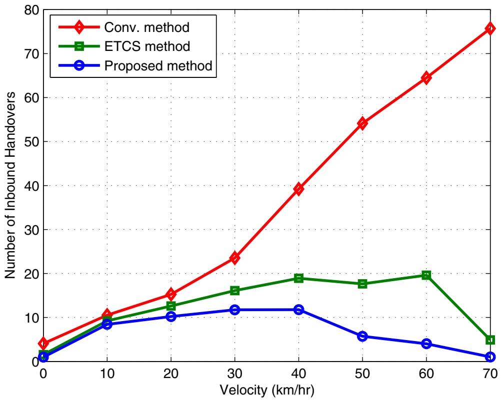

Figure 5: Total number of inbound Handovers

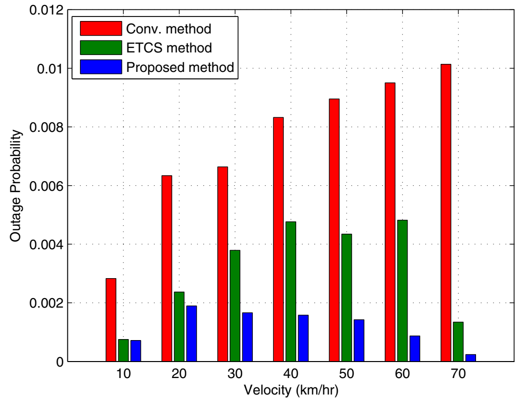

Figure 6: Outage Probability

Therefore, we can define the outage probability as

$$P_{out} = \mathbb{P}\Big[(\gamma^r_{sc_i \rightarrow ue_k} < \gamma_{th} \ \wedge \ \gamma^r_{m \rightarrow ue_k} < \gamma_{th}) \ \text{ for } t < T^{exe}_{ho} \\ \vee \\ (\gamma^r_{m \rightarrow ue_k} < \gamma_{th} \ \wedge \ \gamma^r_{sc_i \rightarrow ue_k} > \gamma_{th} \ \wedge \ RB^{ue}_{sc_i} = 1) \\ \text{for } t \ < \ T^{exe}_{ho}\Big], \tag{28}$$

where T^{exe}_{ho} is the time required to complete the HO process. The simulation parameters are listed in table II [7].

Table II: Simulation Parameters

Bandwidth	10 MHz
Carrier Frequency	2.5 GHz
Macrocell Transmit power	43 dBm
Macrocell Radius	800 m
Maximum Small cell Transmit power	23 dBm
Number of Small cell within Macrocell sector	10
Number of UEs within Macrocell sector	40
Maximum number of UEs per Small cell	5
Minimum required signal for service continuity	-70 dBm
Outage threshold	5 dB
Handover completion time	1 sec
Mean velocity of the UE	3 km/h
Standard deviation for UE velocity	1 km/h
Period between two updates of the mobility model	1 sec
Path loss exponent	3.5
Time threshold for ToS	5 sec

The total number of inbound HOs is depicted in Fig.5. The conventional method has the higher rate of increase in the number of inbound HOs. In fact, for both ETCS method and our proposed method, the number of inbound HOs to SCs is highly minimized because of the reduction in the number of target SCs in the NCL owing to the ToS condition. Our proposed method outperformed both the conventional and ETCS methods by reducing the unnecessary HOs for different user velocities since our method initiates the inbound HO at a point when the data rate from the target SC is good enough with the consideration of interference-based load-dependent modified A3 HO condition, unlike the conventional and ETCS methods which depend on the downlink received power to initiate the HO to the SC via traditional A3 HO condition. The proposed method reduces the unnecessary inbound HOs as V_{ue} increases because the final HO candidate NCL only contains few number of SCs as the velocity increases, hence, the reduction happens.

Fig.6 shows the outage probability for all methods. The proposed method yields lower link failure compared to the other two methods because the proposed method only initiates the inbound HOs when there is a sufficient data rate received from the target SC

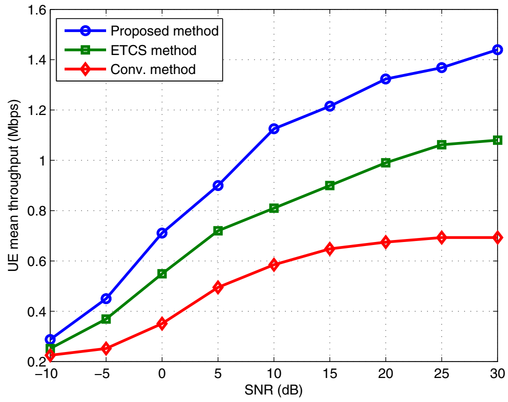

Figure 7: UE mean throughput vs SNR

which means that the HO is initiated with QoS consideration by considering the interference powers from the other neighbouring cells. The conventional method has an instantaneous increase in the link failure owing to the fluctuated downlink received power due to the UE mobility in the HetNet and the level of link failure increases rapidly with the increase in UE velocity. The difference in link failure between ETCS and the proposed method starts to be distinct at a speed of 20km/h and it increases as the speed increases because, in addition to the ToS criteria, the proposed method takes the interference from adjacent cells and the availability of PRBs into account when performing the inbound HO to SC resulting in QoS HO process. This reduction in the outage probability emphasizes that the proposed load-dependent margin, $\gamma^{pro}_{m \rightarrow sc_i}$, has properly managed the load distribution among cells in the HeNet scenario.

The performance of our proposed method, in terms of the maximum throughput a UE can gain while moving through the network, is also compared with the other two methods. Fig.7 illustrates the UE's mean throughput with respect to different signal to noise ratio (SNR) values. The throughput is increased with the increase in SNR accordant with common sense in all methods. The proposed method consistently supplies the UE with the highest throughput compared to the other two methods under different SNR values because the HO point for a UE trying to perform inbound HO from overloaded MC to a target SC is moved closer to the MC (i.e. the HO is triggered earlier), hence, the load is balanced between the two cells resulting in higher throughput.

For a range of network load factor of 5% to 100% with an increment of 10%, Fig.8 shows the UE mean throughput vs different load factors. Our proposed method outperformed the ETCS and the conventional methods in terms of the average UE throughput at

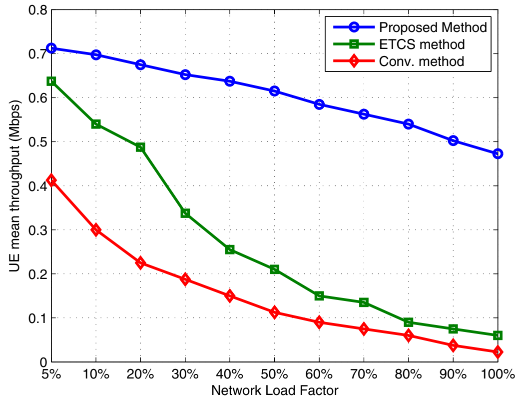

Figure 8: UE mean throughput vs load factor

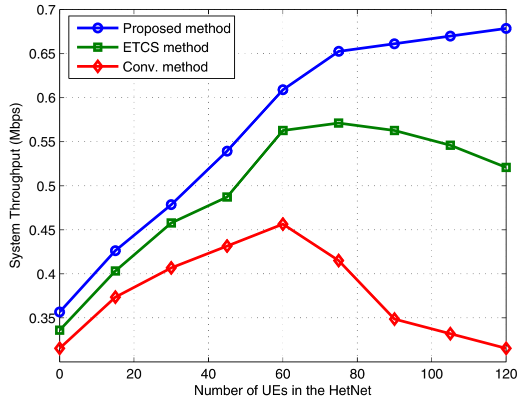

Figure 9: System throughput vs Number of UEs

all load factors. When the load factor increases the cell congests and its radio resources reduce which in turn leads to the drop in the achieved data rate resulting in lower throughput gain. As the load factor goes towards 1 (100% load), the noise will increase which in turn will reduce the SINR resulting in lower UE mean throughput. From Fig.8 we can also notice the sudden drop in the UE mean throughput for ETCS and the conventional methods since they trigger the inbound HO to the target SC based on the downlink received power using A3 event, hence, higher dropping in calls is expected resulting in lower throughput sudden decrease. On the contrary, the drop in the UE mean throughput for our proposed method is less compared to the other two methods because the inbound HO is happened upon the occurrence of our proposed modified interference-based load-dependent A3 event where the UEs are offloaded from the congested MC to the SC by forcing the HO.

Fig.9 depicts the system throughput with respect to the number of UEs moving in the network. It is obvious that the system throughput of both the proposed and the ETCS methods are always greater than that of the conventional method. Below 60 UEs in the network, the throughput of the conventional method keeps going up since the capacity of the MC is still sufficient enough to deliver resources to the incoming UEs but a sudden drop in the throughput happens after that because the MC will be overloaded and its capacity will be limited. When the number of UEs is 60, we can notice that the proposed method has 15% and 4.5% improvement in the throughput compared to the conventional and the ETCS methods respectively, and these percentages increase as the number of UEs increases. Generally, the proposed method's throughput is significantly higher than that of other two methods because of the incorporation of the load-dependent margin which proves the proper distribution of the load between MC and SC tiers.

V. Conclusion

This paper presents an inbound HO method which takes into account the interference and load balancing in HetNet. The effect of interference and short ToS are used to reduce the number of SCs in the candidate NCL so that the user performs an inbound HO to the SC tier that gives a sufficient data rate and has enough resources from a reduced NCL which contains a few and appropriate HO target SCs. We proposed a modified A3 HO triggering condition considering the interference and cell load. Hence, traffic offloading from MC tier to SC tier is accomplished. Results show that our proposed method reduces the unnecessary HO and outage probability compared to the other existing methods. The proposed method has also outperformed the competitive methods by delivering higher throughput as the number of UEs increase in the network. Under different network conditions, including SNR and load factor, we tested and compared the proposed method against the ETCS and the conventional methods. Under all network conditions our proposed method outperformed the other two methods by providing higher throughput.

References

[1] J.-S. Kim and T.-J. Lee, "Handover in umts networks with hybrid access femtocells," in *Advanced Communication Technology (ICACT), 2010 The 12th International Conference on*, vol. 1. IEEE, 2010, pp. 904–908.

[2] A. Ulvan, R. Bestak, and M. Ulvan, "The study of handover procedure in lte-based femtocell network," in *Wireless and Mobile Networking Conference (WMNC), 2010 Third Joint IFIP*. IEEE, 2010, pp. 1–6.

[3] X. Chu, D. López-Pérez, Y. Yang, and F. Gunnarsson, *Heterogeneous Cellular Networks: Theory, Simulation and Deployment*. Cambridge University Press, 2013.

[4] J. Acharya, L. Gao, and S. Gaur, *Heterogeneous Networks in LTE-advanced*. John Wiley & Sons, 2014.

[5] J. G. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. C. Reed, "Femtocells: Past, present, and future," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, pp. 497–508, 2012.

[6] Z. Li, Z. Xiao, D. Wang, and X. Li, "Hysteresis automatic configuration for outbound handover in heterogeneous small cell networks," in *Fuzzy Systems and Knowledge Discovery (FSKD), 2015 12th International Conference on*. IEEE, 2015, pp. 2127–2131.

[7] C.-H. Lee and J.-H. Kim, "Time-of-stay estimation-based cell selection scheme in multitier heterogeneous mobile networks," *IEEE Communications Letters*, vol. 19, no. 9, pp. 1596–1599, 2015.

[8] M. Alhabo and L. Zhang, "Unnecessary handover minimization in two-tier heterogeneous networks," in *Wireless On-demand Network Systems and Services (WONS), 2017 13th Annual Conference on*. IEEE, 2017, pp. 160–164.

[9] M. Alhabo, L. Zhang, and N. Nawaz, "A trade-off between unnecessary handover and handover failure for heterogeneous networks," in *European Wireless 2017; 23th European Wireless Conference; Proceedings of*. VDE, 2017.

[10] J. Zhang and G. De la Roche, *Femtocells: technologies and deployment*. John Wiley & Sons, 2011.

[11] Q. Europe, "Hnb and hnb-macro propagation models," *3GPP R4–071617, Oct*, 2007.

[12] G. L. Stüber, *Principles of mobile communication*. Springer Science & Business Media, 2011.

[13] E. U. T. R. Access, "Mobility enhancements in heterogeneous networks," 3GPP TR 36.839, Tech. Rep., 2012.

[14] X. Yan, N. Mani, and Y. A. Sekercioglu, "A traveling distance prediction based method to minimize unnecessary handovers from cellular networks to wlans," *IEEE communications letters*, vol. 12, no. 1, pp. 14–16, 2008.

[15] H. Kalbkhani, S. Yousefi, and M. G. Shayesteh, "Adaptive handover algorithm in heterogeneous femtocellular networks based on received signal strength and signal-to-interference-plus-noise ratio prediction," *IET Communications*, vol. 8, no. 17, pp. 3061–3071, 2014.

[16] Y. Lee, B. Shin, J. Lim, and D. Hong, "Effects of time-to-trigger parameter on handover performance in son-based lte systems," in *Communications (APCC), 2010 16th Asia-Pacific Conference on*. IEEE, 2010, pp. 492–496.