User Specific Cell Clustering to Improve Mobility Robustness in 5G Ultra-dense Cellular Networks

Mohammad Joud, Mario García-Lozano, Silvia Ruiz
Department of Signal Theory and Communications
Universitat Politècnica de Catalunya-BarcelonaTech, Spain
Email: {m.joud, mariogarcia, silvia}@tsc.upc.edu

Abstract—One of the key challenges in very dense cellular network deployments is keeping mobility robustness. Connection dropping and throughput degradation appear significantly when considering vehicle to infrastructure communications. In order to cope with this problem, this paper proposes the use of User Specific and Adaptive Cell Clustering. Each device is served by a different set of cooperative cells that depends on its mobility state. That is to say, the network customizes the cell cluster size separately for each user. This is accomplished by dual connectivity and splitting the data and control planes. The methodology for evaluation is system level simulations considering a realistic mobility model in which user’s velocity varies based on the traffic load. Results show network performance enhancements in terms of mobility and throughput. Thus, this technique is an interesting mobility management option in such ultra-dense deployments.

Index Terms—5G, Small Cells, Ultra-dense, Mobility, Clustering, Cooperative Multi Points (CoMP)

I. INTRODUCTION

One of the next big leaps in the performance improvement of cellular networks is coming from the reduction of the distance between transmitter and receiver. This implies completing the existing macrocells by deploying low power nodes, i.e. small cells. Small cells improve spectrum efficiency per unit area by bringing the wireless cellular network closer to User Equipments (UEs). This is especially important to redefine the cell edge and homogenize the system Quality of Service (QoS). Small cells offload part of the traffic from macrocells, yield a more effective utilization of radio resources and a significant increase in the system capacity. Moreover, they provide a flexible way to remove coverage holes.

Future wireless networks will need to deploy massive number of small cells to cope with the tremendous increase of data demand. Indeed, Ultra Dense Networks (UDNs) have been identified as one of the pillar technologies in the Fifth Generation of cellular networks (5G) at the International Mobile Telecommunication system (IMT) 2020 [1]. However, network densification is not a straightforward option. High interference levels among small cells (and also from macrocells to small cells if both tiers are co-channel) yield to a degradation in mobility performance, especially for high mobility users.

Nowadays, users demand uninterrupted connectivity, and consume large amounts of data and media content while commuting. Many new use cases related to Internet of Things (IoT) and smart city applications have been addressed and tested which include itinerant devices and equipment, such as autonomous vehicles. Moreover, cellular Vehicle-to-Anything (V2X) communication concept has been introduced and discussed in Long Term Evolution (LTE) Release 14, and the initial standard for this feature was completed recently. Forecasts indicate that the number of connected vehicles to wireless networks will exceed a quarter billion by 2020 [2]. Therefore, new mobility enhancements for high mobility users within small cells networks are required. This has been stated clearly by the 3rd Generation Partnership Project (3GPP) in the latest version of TR 36.932 “Scenarios and requirements for small cell enhancements for E-UTRA and E-UTRA” [3].

Mobility management in dense deployments arises as a serious challenge and many research works are putting in a great deal of effort. Initial works have mainly dealt with the optimization of classic Handover (HO) design parameters [4]. Other contributions have reused interference coordination schemes, such as Almost Blank Subframes (ABS), to decrease Radio Link Failure (RLF) ratios, thus reducing Handover Failure (HOF) occurrence rates [5]. The feature of Dual Connectivity has also brought some enhancements to mobility robustness since it allows the UEs to establish a connection with more than one base station [6]. Moreover, it allows the Control/Data Plane Separation Architecture (CDSA), which has been evaluated as a solution for mobility performance degradation in dense deployments [7]. It has shown better performance since small cells can provide user-specific data only, while macrocells provide a more stable link for control signaling. The majority of these works are basically restricted to UEs up to 30 km/h and, when high mobility UEs are studied, dense deployments are not typically considered [8]. In the most recent works, ultra-dense deployments of small cells are adopted and some of the previously offered solutions were combined to produce new schemes. Examples are the integration of CoMP and CDSA [9], and CDSA plus HO parameters adaption [10]. Backhaul connectivity enhancements were also considered to support mobility with flexible deployment of small cells [11]. The rest employ prediction algorithms based on mobility context awareness to foresee UE-cell transitions [12], or the volume of signaling exchange [13], in order to reinforce mobility robustness. However, high mobility users were not the main interest in these works. On the other hand, cell clustering has been studied in the literature as a means to improve spectral efficiency, energy efficiency and load balancing [14] but mobility issues are hardly addressed.
This paper deals with UEs mobility in the context of small cells deployed in ultra-dense manner. We propose the use of a novel adaptive per-UE cell clustering scheme based on UE mobility state estimation combined with a coherrent CoMP Joint Transmission (JT). This way UEs at moderate speeds can also benefit from the new cell layer and improve their performance with respect to a pure macro-cellular operation.

The rest of the paper is organized as follows. Section II describes the system model. The new clustering scheme is presented in section III. Section IV contains evaluating assumptions and study cases. Section V includes results and discussion. Finally, section VI concludes the paper and outlines possible future research directions.

II. SYSTEM MODEL

A. Network Scenario

In our study, a two-tier downlink LTE cellular network is considered. The macrocell tier operates at 2 GHz, and its cells are deployed following a tri-sectorial regular layout. The ultra-dense small cell tier operates at 3.5 GHz. Given the fact that small cell locations are restricted by backhaul and street furniture availability, the deployment is considered to be random but with a limited inter-small cell distance. Thus, clusters of small cells with overlapped coverage areas appear in a natural manner. Dual connectivity with CDSA is assumed but utilized in a selective manner. That is to say, not all UEs will simultaneously connect to the umbrella macro-cell and a small cell as explained afterwards.

B. UEs Mobility Model

UEs are uniformly distributed and initially associate with the cell providing the strongest Reference Signal Received Power (RSRP), whether it be a macro- or small cell. Non-pedestrian UEs move in a grid of horizontal and vertical streets (Manhattan Model) with four lanes each and having an inter-street distance of 50 m. UEs route may randomly change at each intersection. When a UE hits a border of the studied area, it bounces back in the opposite direction. Pedestrians can freely move around all the area.

A total of 120 full-buffer UEs are dropped per macrocell sector. One third of them is pedestrian (3 km/h). Another third moves at a fixed velocity in the set of 10, 20, and 30 km/h with equal proportions, note that UEs at 10 km/h are bicycle riders on their own roads. The rest forms the set of moderate speed UEs and move at a variable velocity in the range from 40 km/h to 60 km/h. Their velocity increases or decreases depending on the vehicular traffic load, and it is updated every 3 s. This is done in a per UE basis and by calculating the distance to the other UEs that are traveling in the same direction. If the separation distance reduces (or increases), UE’s velocity is updated by a random value in the range [0, 5] (or [-5, 0]) km/h. Moreover, these UEs decrease their velocities while crossing intersections, whether they keep moving on the same direction or not. The purpose of this realistic traffic model is to assess the effectiveness of the proposed scheme on mobility performance in urban scenarios.

C. Handover Model

UEs are accurately modeled by means of a finite-state machine with 11 states and their corresponding latency times in each transition, as indicated in Fig. 1 [15]. The HO process is event triggered, UEs send a Measurement Report (MR) once the A3 condition has been met during a certain Time to Trigger (TTT) [16]. Intra-frequency HOs are based on RSRP comparisons, whereas inter-frequency HOs rely on the Reference Signal Received Quality (RSRQ) [16]. The different HO types also imply different measurement periods, 200 ms for intra-frequency HO, and 400 ms in inter-frequency HO cases. This causes different delays in detecting the event. Cell scanning is continuously performed on both frequencies, at a fixed interval of 240 ms and with measurements gaps of 6 ms. Based on those MRs, the serving cell triggers the HO procedure. At the radio level, note that HO related messages are presumed to be transmitted with the lowest Modulation and Coding Scheme (MCS).

UEs detect RLF when they are out of synchronization during a time equal to Timer T310. This situation is determined when the average wideband Signal to Interference and Noise Ratio (SINR) is below the one required to decode the minimum MCS. If a RLF happens during a HO, it is considered as HOF. This corresponds to the following situations [16]:

- RLF happens after satisfying the event A3 condition and before receiving the HO command.
- The UE is considered out of synchronization when the HO command is sent or the HO complete message is sent, even if the T310 timer is still running.

Two more factors have been provided to this model, the HO interruption time and the delay of Radio Resource Control (RRC) connection re-establishment [15]. During both periods of time, the UE does not receive any data, thus frequent HOs and HOFs also lead to throughput degradation.
D. Mobility State Estimation

In this study, Mobility State Estimation (MSE) has been modeled as defined in the 3GPP LTE standards [17]. UE mobility state is evaluated by computing the ratio $r_{MSE}$:

$$r_{MSE} = \frac{N_{HO} + N_{reselect}}{T_{MSE}}$$

(1)

which is the number of HOs, $N_{HO}$, and cell re-selections, $N_{reselect}$, over a specified period of time, $T_{MSE}$.

Traditionally, $r_{MSE}$ is compared with two thresholds to determine one out of three possible mobility states: low, medium or high. This estimation is usually done for TTT scaling purposes. In the context of this work, $r_{MSE}$ is used for two purposes: First, to define whether a UE should be served by a cooperative cluster of transmission points and second, to compute (and update) the number of cooperating transmission points. In this respect, further details are provided in the following section.

MSE is a key element for a correct function of the per-UE clustering approach. Standard values for $T_{MSE}$ are 30 s and above. However, a network with a dense deployment of small cells allows having more resolution in the mobility state estimation, i.e. beyond the three basic states previously mentioned. Thus, a second timer can be introduced to track short term speed variations, in particular we assume $T_{MSE} = 5$ s. On the other hand, small cells are irregularly deployed and their density may change in different parts of the network. This requires a proper calibration and matching between the values of $r_{MSE}$ and the corresponding mobility states for different parts of the network. The operator can do this task by means of simulations; analytic estimations are also possible [18].

III. CELL CLUSTERING BASED ON MOBILITY STATE ESTIMATION SCHEME

The core idea is to enable the wireless cellular network to respond to each UE in an independent and different way based on its current mobility state. In other words, the network customizes the cell size separately for each user, which is served by a different cluster of cooperative cells that depends on its mobility state. This scheme aims at reducing the frequent small cell to small cell HOs, and consequently it minimizes HOFs occurrence for moderate speed UEs when connected in the UDN layer. This is achieved through CoMP-JT clustering plus dual connectivity with CDSA. Note that, user specific small cell clustering is only applied for moderate speed UEs and locally among small cells under the same macro site.

Moderate speed UEs are initially connected to the best macrocell. This happens in a natural manner if the operator decides to broadcast scaling factors for TTT and hysteresis margin in the cell (re-)selection procedure. MSE starts to estimate the mobility state of those UEs, the stored history from idle mode cell re-selections is also used. When the cell receives a report from a UE indicating an A3 event towards a small cell, the system assesses the current UE mobility state. If the UE surpasses the velocity threshold, dual connectivity with CDSA is activated. The UE keeps receiving the control plane from the macrocell and the data plane is transferred to a cluster of cooperative small cells that will schedule the UE in a coordinated manner, as illustrated in Fig. 2. The scheme will determine and update the required number of small cells in the serving CoMP cluster. Hence, conventional HO is no longer happening to these UEs while commuting within the same macro site. The process turns into cluster configuration by adding or removing small cells to CoMP clusters based on mobility state assessing. Two configuration steps are required before enabling this scheme in the network:

- Set the activation condition, utilized by the network to identify the targeted UEs (moderate speed UEs) by assessing their current $r_{MSE}$ value. The mobility based clustering scheme is limited to UEs between two figures, named low and high selection thresholds. Such thresholds are chosen by means of simulations and might be tuned by using MSE statistics and HO history once the mechanism is in exploitation, for example, if HOF are larger than expected. Due to the non-regular deployment of small cells, the selection thresholds may vary from one part of the network to another.

- Set the list of sub-conditions that are used to determine the size of CoMP clusters, i.e. number of cooperative small cells. This involves the definition of a second set of thresholds or *interim thresholds*. This step also requires pre-operation simulations with possible post-tuning.

While the targeted UE moves on the grid, its velocity changes and so it does its mobility state. The mobility based clustering scheme tracks these changes by computing $r_{MSE}$ every $T_{MSE}$. If necessary, the serving CoMP cluster size is modified according to the current mobility state.

Three interim thresholds are adopted. Note that this number is directly proportional to the small cell density, a denser network allows a more precise mobility state estimation. Given this, moderate speed UEs may be served by an adaptive CoMP cluster of 2, 3, 4, or 5 small cells, depending on their current $r_{MSE}$. Fig. 3 summarizes the considered configuration of the new scheme proposed in this study.

Adaptive CoMP clusters are not created for low mobility UEs, whose Radio Resource Management (RRM) is normally done by the serving macro or small cell. They also perform classic HO among all cells and without control/user plane split. Normal HO occurs between macrocells no matter the mobility state of UEs. But with the clustering scheme, HO only happens between macros or small cells under the control of different
would accommodate the rest of UEs.

Then, the scheduler can easily account for the rest of UEs in each small cell. On the other hand, if the architecture is distributed, small cells would first receive the allocated resources for JT from the macrocell and then, they

measurements [19] for all the cells in the cluster, and the maximum CQI is used to decide the MCS. If small cells are generated by means of Remote Radio Heads (RRHs), their RRM is centralized in the macrocell (semi-cloudified architecture). Then, the scheduler can easily account for the rest of UEs in each small cell. On the other hand, if the architecture is distributed, small cells would first receive the allocated resources for JT from the macrocell and then, they would accommodate the rest of UEs.

The RRM of moderate speed UEs is centralized at the macrocell, which decides on the radio resources to be used by the cooperative cluster (same resources for all small cells in the group). Allocations are based on time-filtered wideband Channel Quality Indicator (CQI) measurements. In particular, the UE performs parallel Channel State Information (CSI) measurements [19] for all the cells in the cluster, and the maximum CQI is used to decide the MCS. If small cells are generated by means of Remote Radio Heads (RRHs), their RRM is centralized in the macrocell (semi-cloudified architecture). Then, the scheduler can easily account for the rest of UEs in each small cell. On the other hand, if the architecture is distributed, small cells would first receive the allocated resources for JT from the macrocell and then, they would accommodate the rest of UEs.

IV. METHODOLOGY AND ASSUMPTIONS

Dynamic system level simulations have been performed to investigate the performance of the proposed mechanism. Statistics have been collected from simulations of 1000 s with a time resolution of one Transmission Time Interval (TTI) (1 ms). The main simulation parameters not mentioned so far are summarized in Table I. For a more detailed review of the simulator, the reader is referred to [15]. Three baseline scenarios have been used for comparison purposes:

- **Baseline cases:**
  1) **Macro only:** Small cells are not deployed. Thus, all UEs connect only to macrocells.
  2) **SC-all:** Small cells are deployed and UEs can connect to any cell whether it be macrocell or small cell.
  3) **SC<30:** UEs at >30 km/h are able to attach only to macrocells.

- **Cluster+MSE:** New proposed scheme.

V. SYSTEM PERFORMANCE EVALUATION

A. UE SPECIFIC CELL CLUSTERING SCHEME PERFORMANCE

In order to assess the value of using the UE specific clustering mechanism in UDN, we examine its impact on the mobility and throughput performance, HOF rates and average UEs throughput as a function of UE velocity for all studied scenarios are illustrated in Fig. 4 and Fig. 5 respectively.

It is clear that the ‘Macro only’ case has the best mobility performance due to the lack of small cells. This absence leads to the worst throughput result among all cases. Off-loading data into small cells brings benefits to all UEs in terms of throughput, as observed in the ‘SC-all’ case. Though this gain decreases significantly with velocity due to an increase in HOF rates. The reasons behind this poor performance are, the severe inter-small cell interference, frequent small-to-small HOs and the late small cells detection. As a result, UEs get a poor QoSs because of longer out-of-service time. One possible solution to tackle this issue is to prevent moderate speed UEs from connecting to the small cells layer, as in ‘SC<30’ case. This keeps HOF rates similar to the Macro only case, but it deprives those UEs from small cells benefits. Note that the SC<30 line is discontinuous because joining the HO/throughput at 30 km/h with that at 40 km/h does not interpolate the intermediate values, since the strategy is different at the different velocities.

The CoMP clustering scheme is applied for moderate speed UEs in the ‘Cluster+MSE’ case. The proposed scheme adaptively configures clusters by adding cells according to UEs needs based on their current $r_{MSE}$ values. On the other hand, it does the opposite when UE’s velocity decreases. This
specific per UE cluster updating reduces the small cell HOF occurrence. It completely cancels HO occurrence within the small cell layer, which justifies the gains from the mobility viewpoint, as shown in Fig. 4. This low HOF rate plus downlink reinforcements by means of CoMP-JT with coordinated scheduling translates into throughput improvements for UEs at >30 km/h, as shown in Fig. 5. When compared to the SC-all case, low speed UEs suffer a minor throughput degradation, this is due to the fact that several cells are now coordinating to serve the same UEs. This could translate in a lack of efficiency in the radio resource usage. But, the reduction is marginal thanks indeed to a more efficient use of resources, that are not wasted in unsuccessful transmissions with Cluster+MSE. The strategy performs fair and accurate management, redistribution and releasing of radio resources dynamically.

B. UE Specific Cell Clustering Scheme Tuning

The CoMP clustering scheme creates dynamically and continuously different sets of cooperative cells based on UEs mobility state. The fundamental factor that ensures the proposed scheme effectiveness is the accuracy of the actual $r_{MSE}$ values calculation. This is related directly to $T_{MSE}$. In the previous discussion, 5 s was selected, which guaranteed an accurate tracking for moderate speed UEs. On the other hand, this short period leads to extra overhead and signaling. Fig. 6 shows the effect of tuning $T_{MSE}$ on the average UEs throughput. Increasing $T_{MSE}$ brings benefits to few portion of UEs, but worsens the experience of the majority. The error in the estimation of MSE mechanism makes the scheme blind somehow. This allows a UE at higher velocity to keep the same size of CoMP cluster while its velocity decreases. This UE stays connected to the same number of small cells until the next $r_{MSE}$ value update. Although, its velocity decreases and CoMP cluster size should be reduced. The opposite happens to lower speed UEs while their velocities increase. They reach higher velocities with a cluster of few small cells. Thus, increasing this interval in UDN leads to unfair management and distribution of radio resources among UEs. The high
The presence of vehicular traffic implies that the highest speed UEs tend to decrease their velocities fast and often. That is why UEs at 40 and 45 km/h only get benefits, and of course at the expense of the rest. The reason behind this behavior can be understood well after looking at Fig. 7, which depicts the average number of small cells in adaptive clusters as a function of UE velocity considering the three values of $T_{MSE}$.

Increasing $T_{MSE}$ reduces the average number of small cells in CoMP cluster for the highest mobility UEs and the opposite occurs for the lowest mobility UEs. In general, we can state that, the wrong tuning of this parameter leads the performance of the proposed scheme to an unwanted area with negative effects to the majority of UEs. However, increasing this period reduces the amount of extra required signaling over front-/backhauls. As it is clear in Fig. 8, which represents the average number of cluster configuration and updates per second. This number comprises updating the CoMP cluster by adding and removing small cells according to the received MRs while commuting, plus modifying CoMP cluster size due to $T_{MSE}$ value change. Consequently, a correct tuning of $T_{MSE}$, or managing more than one timer is essential to get the highest gain.

VI. CONCLUSION

This paper addresses mobility degradation in the context of ultra-dense small cell deployments for moderate speed UEs. UE specific cell clustering based on mobility state estimation with non-coherent CoMP-JT is evaluated as a means to overcome this issue. Results indicate that it is a feasible means to allow those UEs to connect to the ultra-dense layer. HOF rates are kept low and their throughput is importantly increased. The fact that several cells serve to the same user does not translate in an ineffective use of radio resources, given the reduction in interruption times that were caused by HOFs. The procedure requires a correct tuning of the MSE. This will be particularly true and complex in highly irregular deployments. Our future work goes in that direction along with the combination of the current coordinated scheduler with coordinated beamforming.

ACKNOWLEDGMENT

This work has been funded by the Spanish National Science Council through the project TEC2014-60258-C2-2-R and European Regional Development Funds (ERDF).

REFERENCES