Investigating Mobility Robustness in 5G Networks Using User-Adaptive Handoff Strategies

Masoto Chiputa[1](✉), Peter Han Joo Chong[1], Saeed Ur Rehman[1], and Arun Kumar[2]

[1] Department of Electrical and Electronic Engineering, Auckland University of Technology, Auckland, New Zealand
{masoto.chiputa, peter.chong, saeed.rehman}@aut.ac.nz
[2] Department of Computer Science and Engineering, NIT, Rourkela, India
kumararun@nitrkl.ac.in

Abstract. Millimetre Wave (mmWaves) communication is a major capacity booster to the fifth generation (5G) mobile network. However, challenges of significant attenuation and high propagation losses lead to intermittent user connectivity. This limit mmWave applicability in the 5G mobile network. The dual connectivity (DC) architecture with its split control and data plane functionality has proven to be more effective. The DC model uses the less attenuated Long Term evolution (LTE) bands to coordinate mmWave cells while they provide on demand high capacity rates. This guarantees prolonged network association for mmWave links with minimal to no signalling cost at high user data rates. Minimizing signalling cost in LTE bands considering their scarcity is vital too. To that effect, intelligent Handoff (HO) strategies that prolong mmWave link association using minimal signalling cost are vital for a robust DC system. This paper investigates the performances of adaptive HO strategies given a highway multiuser-type mobility scenario on DC systems. Results show that adaptive HO solutions enhance link reliability in mmWaves with minimal signalling cost.

Keywords: 5G · Millimeter wave · Performance evaluation · Dual Connectivity · 3GPP · NR · Adaptive handoffs

1 Introduction

The next generation of cellular networks will need to use macro frequency diversity solutions to overcome the current resource scarcity and meet new capacity demands [1]. Macro frequency diversity allows usage of multiple frequencies with varying propagation capabilities to reduce the effects of user data distortions in wireless transmission [2]. In 5G, the macro frequency diversity is used to mitigate data distortions and boost capacities for mobile network. This is particularly, due to poor propagation characteristics of mmWaves and the scarcity of idle bands in lower frequencies currently deployed for cellular network communication [3].

J. Zheng et al. (Eds.): ADHOCNETS 2019, LNICST 306, pp. 99–112, 2019.
https://doi.org/10.1007/978-3-030-37262-0_8

The Dual Connectivity (DC) model is one example of a macro frequency diversity solution. It has thus far been adopted by 3rd Generation Partnership Project (3GPP) as a 5G mobile network model where it is technically as New Radio (NR) [1] and [5]. Its goal is to provide continuous service guarantees to mmWave mobile network users. The DC system allows non-co-located deployment of PHY and MAC layer of LTE and mmWave networks and a system level 3GPP LTE protocol stack integration [2, 3, 7] for NR1. This NR1 model paves a clean slate to addressing mmWave propagation challenges. Technically, in NR1 also known as 5G, the lower frequencies, such as the Long Term Evolution (LTE) bands, provide oversight links for mmWave link recovery and cell co-ordination. The mmWave links provide on-demand multi-gigabit user rates. To that effect, it is believed that NR will use multi-connectivity solutions with a sub-6 GHz radio overlay for coverage, possibly based on Long Term Evolution (LTE) thanks to a tight internetworking [6], and mmWave links for capacity.

To optimize performance in NR1, longevity of user association to mmWave eNBs overlaid by LTE eNBs as shown by Fig. 1 is vital [3]. The prolonged association reduces HO signalling costs and improves the effective data rates. Thus, intelligent HO strategies are needed to (1) facilitate both intra and inter frequency user data flow without any distortion, (2) minimize resource usage in LTE cells given the scarcity and Low data rates of in lower frequencies plus the additional functionality of LTE bands to provide both fall back [2] and control-plane links [2, 3, 10].

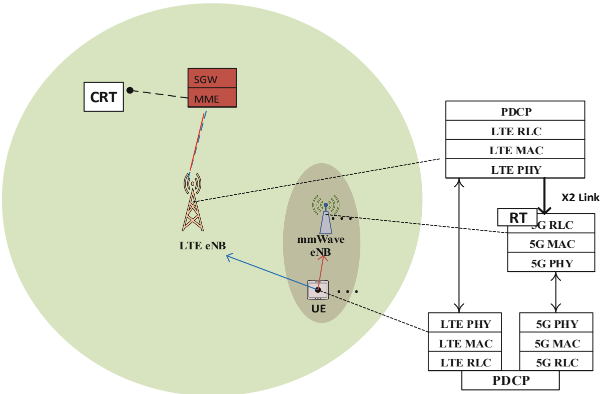

Fig. 1. LTE-mmWave 5G tight integration architecture.

This paper investigates mobility robustness optimization in NR1 given a multi-user mobility perspective in a highway mobility scenario. The objective of mobility robustness is to dynamically improve the network performance using adaptive HO strategies. Ultimately, this enhances Quality of Experience (QoE) for the users and as well as increases the network capacity [9]. This is basically done by automatically adapting appropriate cell parameters to adjust thresholds based on feedback of performance indicators. Thus, we adopt and analyse performance of various adaptive HO strategies from a multi-user type perspective on a highway scenario.

Given the lack of a real end-to-end DC system for system-level analysis in NR1, we use the LTE-mmWave DC module in the ns-3 simulator described in [2] and [3]. The DC framework in ns-3 allows designing network procedures faster than standard standalone hard handover (HH) to improve the mobility management [3]. Besides, it uses a 3GPP antenna array model for mmWave links [1]. This gives it a more practical two-dimensional directional antenna pattern at PHY and MAC layer in mmWaves. Unlike in "flat top" antenna model [1], the main-lobe gain and side-lobe gain of mmWave channels vary with signal changes in the receiving angles.

The principle operations of the ns-3 DC module involves mmWave users directionally broadcasting beaconing messages known as sounding reference signals (SRSs) in time-varying angular spaces. Each potential mmWave eNB in response scans all its angular directions, as it monitors the strength of the received SRSs. It then records the wideband SINR values in the Report Table (RT). The SINR estimates are thereafter transmitted via X2 links to LTE eNB to build a Complete RT (CRT) in the mobile management entity (MME). This data is later used to coordinate mmWave Cells in HO procedures. Figure 1 give a glimpse of the DC module in ns-3, detailed operations can be found in [2, 3] and [10].

2 Related Research

Mobility robustness in the 5G wireless communication is vital to mitigate channel instabilities due to poor propagation characteristics exhibited by mmWaves. For example, the Doppler spread at 60 km/h for 60 GHz band is over 3 kHz [4]. Thus, the mmWave channel changes in orders of hundreds of microseconds faster than today's cellular systems within a small time. Tracking mmWave signal at this rate is a challenge. Moreover, mmWaves are easily blocked by brick, mortar, or even by the human bodies [1]. This makes the quality of the mmWave SINR to exhibit high variability with variations in the order of 30 dB for transitions between Line of Sight (LOS) and Non-Line of Sight (NLOS) [6]. Therefore, mmWave Channel links can turn from being excellent to being unusable within milliseconds

Reducing SINR prediction inconsistences is vital to eliminate short network association and wasteful HOs for mobile users. This will enhance not only link reliability but also reduce signaling. In [2] and [4] where the DC model is used, for example, raw SINR values from SRS are filtered to produce time-averaged SINR trace for HO decisions. Additionally, a much higher averaged SINR threshold value is used to avoid unnecessary HOs and beam switch due to SINR fluctuations. The main objective is to eliminate short-term parameter variation from instigating long HO and beam switching processes. Thus, by time averaging raw SINR values and setting higher threshold values DC systems automatically adjust and adapt practical thresholds that will improve performance.

Additionally, proposed HO strategies show that efficient adjustment of the Time-to-Trigger (TTT) intervals [3], channel and RT updates can enhance effective data rates. In fact, more advanced adaptive strategies HO combining most of the alluded strategies above show significantly prolonged network association and high effective data rate [1, 9]. Given outage is environmentally specific in mmWaves [4], if more network link assessment and mobility support strategies are designed in an intelligent way, robust and efficient network performance can be attained. To that effect, other 5G studies have reinverted interference coordination schemes in 4G, e.g. Almost Blank Sub frames (ABS) to mitigate Radio Link Failure (RLF) and Handover Failure (HOF) occurrence rates [8]. They take advantage of the flexible nature of DC systems by focusing only on mmWave systems. Thus adjust range for non-line of sight (NLOS) and Line of sight (LOS) scenarios. Also, others have integrated DC with of CoMP strategies [10]. They employ prediction algorithms based on user mobility context awareness to predict channel state, or the volume of signaling exchange before reinforcing mobility robustness mechanisms.

Low mobility users are the main interest in most of the works above. Additionally, the focus is on improving load balancing, spectral and energy efficiencies on mmWave cells. Thus, the majority of these works neglect the impact of DC models on LTE cells responsible for coordinating mmWave cells and providing links for fall back and higher-speed users. Thus, a multi-user type mobility analysis is vital because user diversity will stretch both LTE and mmWave networks resource. Ultimately, results obtained from a multi user type gives a truer performance of 5G DC systems.

3 System Model

3.1 Network Model

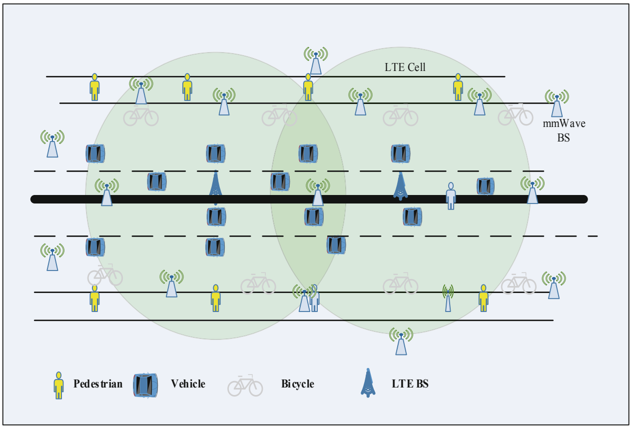

Fig. 2. LTE-5G general user mobility model.

In our scenario, the DC model is assumed and utilized strategically. This is to say, different user type initially connect to which ever cell has the highest Signal to Interference plus Noise Ratio (SINR), before simultaneously connecting to either overlaid LTE-cells or under laid mmWave small cell based on the adaptive HO as explained in Sect. 3.3. LTE operates at 2 GHz while all mmWave cells operate at

28 GHz and connect to LTE's macro cells via the X2 link as illustrated in Fig. 1 based on the NYU and University of Padova DC implementation model [3]. Three LTE eNBs are deployed in the middle of the road stretching over a total length of 1500 m. The coverage radius of the LTE eNB is 250 m and the distance between two LTE eNBs is 500 m. Initially, 20 mmWave primary cells operating at 28 GHz with DC connection are placed over the road as shown in Fig. 2. The source code of the LTE-mmWave DC framework is publicly available in [2] with the ns-3 script (source: mc-example-udp.cc) and is thus used for the simulation scenario here. It features the implementation of the 3GPP channel model for frequencies above 6 GHz and a 3GPP-like cellular protocol stack.

3.2 User Mobility Model

In Users are uniformly distributed and initially associated with the cell providing the highest SINR, that is LTE or mmWave cell. This enables faster initial user access in DC [3]. Three types of users are assumed defined based on their speeds, these include; cars, cyclers and pedestrians. Cars move in four lanes with a length of 1.5 km and each having a width of 3 m. Cycler's paths stretch a width of 1 m from the outermost lanes of the cars on each side of the highway while pedestrians randomly move and change directions on the edges of the cycler's paths also on both sides of the highway. Pedestrian paths have each a width of 1 m. When pedestrians hit the edges of the highway, they bounce back in opposite directions. One third of the total users are pedestrian moving at 3 km/h. They are randomly placed in both walking paths. Another one third are cyclers moving constantly at 10 km/h. The other one third are cars equally distributed in each of the 4 lanes and initially interspaced equally over the length of the highway. They move at a fixed speed 50 km/h. Our assumption on cars is that those in the same lane i.e., in a busy highway are likely to move at similar velocities to avoid crushes. Hence, in each lane, cars' speeds and interspaces are assumed fixed to avoid crushes. The multi-user type mobility model unlike single user type mobility models in [2] and [3], can simultaneously test both LTE and mmWave transmission resources. This is so, given the fact that high speed users in will mostly likely use LTE links due to the band's low Doppler spread. Hence will leave the preferred mmWave links to low speed users despite the higher data rates. This indirectly balances load. Further it gives a fairer assessment of how far the DC network can support users given various adaptive HO strategies. A snap shot of the view of the multi-user type mobility model is as shown in Fig. 2.

3.3 Adaptive Handoff Strategies

The DC framework in ns-3 allows designing network procedures that are fast and dynamic, to improve mobility management in mmWave networks. Particularly in ns-3, it allows designing fast switching mechanisms between the LTE and primary mmWave cells via X2 links at access level as shown in Fig. 1. It further allows HO to Secondary Cells whereby the user does not need to resend SRS signals and incur additional delays before another HO so long as the TTT has not expired [2]. Thus, in our simulation, we investigate three adaptive event triggered fasting switching HO strategies in the DC model for NR1 based on distance, speed, and load, and are as follows:

A. (HO1) Load and SINR only Based HO: In this scenario [3], the, HO decision is based on the data rate, *R,* and key input information for the HO decision includes (i) instantaneous channel quality based on SINR, (ii) channel robustness based on rate, and (iii) cell occupancy or load [3]. Thus, the rate, *R*, experienced by the user connected to eNB_m is approximated using the Shannon capacity [2]:

$$R(m) = \frac{W_{mmW}}{N_m} \log_2(1 + SINR(m)), \quad (1)$$

where N_m is the number of users that are currently being served by eNB_m and W_{mmW} is the available total bandwidth. Unlike traditional procedures where users are not aware of the surrounding cells' current states, the user chooses to connect to a Cell providing the maximum rate, *R*, taking into account the load of the cell. It is a three phased HO process for DC and is as follows:

First Phase: involves uplink (UL) measurements at mmWave cells, user transmits SRSs to the surrounding mmWave Cells through different antennae directions. mmWave cells performs an exhaustive search to collect the SRSs using different angular directions. Thereafter, based on SRS value, they estimate mmWave cell SINR Once the RTs are filled with the SINR metrics, the second phase is invoked.

Second Phase: involve cell Coordination, at this stage, the mmWave Cell sends its RT to the LTE Cell, via the X2 link in Fig. 1. The overlaid LTE Cell builds a CRT by collecting all the received RTs. The LTE Cell makes HO decisions by selecting optimal mmWave Cell for each user to connect to given (1) before the third phase is evoked.

Third Phase: is decision time, at this stage the LTE Cell forwards the best decision for the transceiver. And the mmWave Cell is informed through the backhaul X2 link.

With respect to the existing algorithms in [2], the use of both the sub-6 GHz and the mmWave control planes in DC is a key functionality to the HO technique. Especially for highly unstable and scarcely dense scenarios, the LTE connectivity ensures a ready backup in case of a failure in mmWave links else will suffer an outage. Furthermore, the handover/beam switch decision is forwarded to the user through the macro Cell, whose legacy link is much more robust and less volatile than its mmWave counterpart, thereby removing possible points of radio link failure (RLF) in the control signaling path [2, 3, 10]. Detailed process is explained in [2] where intra (secondary) cell HO occur if TTT has not expired yet the channel state has changed. Summarily, Inter and intra frequency HO rely on the SINR and the load variation to trigger a handover within TTT. Further, for secondary HOs, they are only processed if the Target cell becomes 3 dB better than that of a serving cell, the HO is cancelled [3]. Secondary HOs are HO that happen between TTT intervals [2, 3, 10].

B. (HO2) Distance Based HO: The second HO algorithm is based HO1 but considers the effect of user distance relative to the target [6] and serving cell besides load and SINR variation in Phase 2 as shown in Fig. 3. The addition of distance is to avoid wasteful HO requests that may arise even when the user is still within the good transmission range of the serving cell. By factoring distance in process HO1, we minimize HO frequencies and prolong network association to one stable cell to enhance the effective rate.

The hypothesis is that if the user displacement relative to the serving eNB is closer than that of the target eNB. The network association with that target cell will be longer so long the correct prediction of a user's direction relative to the target eNB is made. While the initial rate, *R,* in (1) may be low, so long the SINR with a probability, *P,* is above a given threshold, the network association to that target cell is likely to be longer. Subsequently, the rate, *R,* in (1) is likely to improve considering the fact that the user will move closer to the target eNB for the next HO. Thus, we define SINR probability, *P*, according to dynamic, f_{TTT}, TTT equation in DC model for ns-3 simulator [3] such that:

$$P(SINR_s > t) = \frac{\Delta - \Delta_{min}}{\Delta_{max} - \Delta_{min}}, \tag{2}$$

where Δ is the instantaneous SINR, Δ_{min} is the minimum threshold SINR, *t*, and Δ_{max} is maximum given the maximum transmission power.

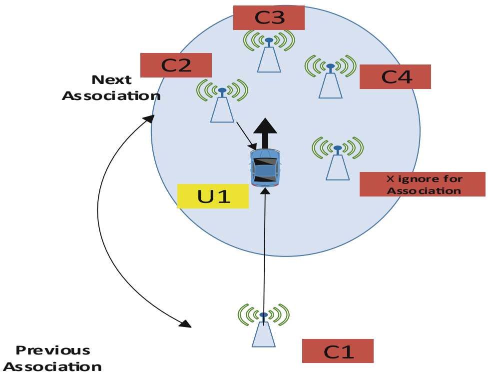

Fig. 3. DC mmWave-LTE-5G distance based HO mechanism.

The concept aims at mitigating radio link failure (RLF) rate and signaling cost due to frequent HOs. It prolongs TTT than in traditional cellular networks. The minimum HO displacement D_{min} is dynamic in our setup. It varies according to user's current displacement relative to the serving eNB at the point the SINR of serving cell falls below 3 dBm. Principally, the HO is evoked, if, within the TTT, the SINR (m) in (1) of another cell becomes 3 dB better than that of the serving mmWave cell and meets the minimum current HO distance at the time the channels update. The user maintains connectivity with serving cell if the serving cell, before the TTT expires, retains the highest SINR again, i.e., the HO process is canceled. Otherwise, the user maintains the HO process to the target cell if the D_{min} condition is met else user connects to the LTE cell before the next mmWave HO is initiate.

Two main direction of users are used (X, −X) representing North and South. eNBs determined to be in a similar direction of the user have at most a 60-degrees variation to serving eNB. This variation accounts for road way anomalies. Besides, users moving in a similar direction stay connected to the same BS longer than users moving in the opposite direction; hence, the eNB may have a stable load prediction in (1). To that effect, the user movement direction (or vector) is crucial when determining neighbors in target eNBs range.

To include the effect of distance on dynamic TTT, we define the Relative distance probability, $P_{TTT}(t)$, such that:

$$P_{TTT}(t) = P(SINR_s > t | D_{\Delta s} > D_{min}), \quad (3)$$

where $P_{TTT}(t)$ is the coverage approximation [9] given the target eNB distance, $D_{\Delta s}$, is greater than user distance D_{min} relative to serving the cell and the user SINR is greater than the threshold t. The measure of $P_{TTT}(t)$ expresses the network's reliability obtained by improving the reliability of the TTT for handoff timing. In other words, it represents the link reliability loss if the TTT is too unrealistic or short for the user HO, forcing users to engage in wasteful HOs. Thus, the dynamic time, $f_{TTT}(\Delta)$, to trigger HO is defined as

$$f_{TTT}(\Delta) = TTT_{max} - P_{TTT}(t)(TTT_{max} - TTT_{min}), \quad (4)$$

where the minimum SINR instantaneous Δ expected between the best target cell and of the current serving cell is met. TTT_{min} is the minimum of TTT to trigger HO. Thus, we use (4) instead of (5) used in (HO1) for secondary HOs [4]:

$$f_{TTT}(\Delta) = TTT_{max} - P(SINR_s > t)(TTT_{max} - TTT_{min}). \quad (5)$$

C. (HO3) User speed Based HO: In this HO, we use user speed and Mobility State Estimation (MSE) with its scaling factor ratio defined as sf_{MSE} [8] and [9];

$$sf_{MSE} = \frac{N_{cell-Change}}{T_{MSE}}, \quad (6)$$

where $N_{Cell-change}$ is the number of HOs and reselections over a specified period of time, T_{MSE}. MSE give the option of scaling up the TTT and SINR Threshold values of a cell. If the user is moving so fast, the HOs can be triggered earlier to avoid RLF. Three MSE scales are used to categories speed; low, medium and high mobility. For high sf_{MSE}, the recommended $N_{cell-Change}$ is 10 and for medium, $N_{cell-Change}$ is 16 [12] over a period T_{MSE} for LTE links. According to 3GPP TS 36.331 [9], the values that high sf_{MSE} or medium sf_{MSE} can take are 0.25, 0.5, 0.75 or 1.0.

Given the DC model, a user connects to more than one band i.e., LTE uses lower carrier frequencies with high transmission range, and has low Doppler spread for high speed users [4]. In contrast, mmWaves have small coverage [4, 11] and high Doppler spread at high speed due to high carrier frequencies. So, even though the user speed on the ground is the same for both LTE eNB and mmWave eNB, for scaling mobility parameters; low, medium and high must be different for the eNBs. Practically, if (6) is used, the user will have crossed over multiple boundaries of the mmWave cells before crossing over the boundary of the overlaid LTE cells. Thus, given the DC network with a dense deployment of mmWave small cells, to allow more resolution in the mobility state estimation, a, second timer [8] is introduced to track mmWave's short-term speed variations. We assume T_{MSE} = 5 s to give corresponding mobility states for different cell-type of the network. Hence, the time, T_{HO}, taken to complete HO can be defined as [9]

$$T_{HO} = TTT + HO\ execution\ time(T). \tag{7}$$

The principle operation is done in Phase 3 of the HO1, where the LTE cell passes either the sf-High or sf-medium to the mmWave cell i.e. the factor by which the mmWave cell will scale the TTT and SINR threshold given the user speed. If it doesn't pass any sf value, it is regarded low. By doing passing MSE values, LTE trigger HOs earlier. For instance, if the user is fast moving and high sf_{MSE} = 0.3, if normal TTT value is 400 ms and hysteresis value is −2 dB, the new TTT value and the hysteresis will be multiplied by high sf_{MSE}. So, the new value TTT will be 120 ms and −0.6 dB hysteresis. Technically, the MSE procedure just correlates user-type speed defined with a scaling of mobility parameters. It doesn't estimate speed.

Finally, for all HO processes, given the different HO types, channel updates time used, is 200 ms for mmWave HO, and 400 ms for LTE HO. Normal TTT is set at 480 ms. This allows mmWave to initiate a HO event before LTE links can be used as fall backs. RT updates are continuously performed at an interval of 200 ms and with measurements gaps of 6 ms. Based on SRS SINR values, LTE cells trigger the HO procedure.

4 Performance Evaluation

4.1 Tools and Techniques Used

In Simulations were done using ns-3 dual connectivity mmWave-LTE. ns-3 is an open source, C++ based discrete event simulator [2] and [3]. For mmWave SRS, the signals are instantaneous wideband with 10 MHz bandwidth, 1 GHz carrier frequency, a time varying sampling periods of 40 ms and a Doppler frequency varying from 1 to 166 Hz (low, medium and high Doppler frequencies). The radio channel follows the 3 GPP Extended Typical Urban (ETU) environment requirement [13–16]. We consider three classes of user speeds; 3 km/h, 10 km/h and 50 km/h. It is an event-based adaptive HO network simulation with a multi-user road traffic model described and shown in Fig. 2.

4.2 Performance Metrics

The metrics used to assess the performance are the following: average handoff failure (HOF) rates, average delivery ratio, average Handoff number and average throughput per user.

4.3 Baseline Considerations

A hard HO mechanism is considered where macro (LTE) and micro (mmWave) cells do not share the X2 link at access level. Small cells are not coordinated by the LTE cells at access level but share the same core network. Thus, mmWave cells and LTE cells must completely disconnect to one cell prior to choosing a better link as explained in detail in [2, 3, 10].

4.4 Simulation Parameters

Table 1. Simulation parameters.

Parameters	LTE		mmWave
3GPP channel scenario		Urban Micro	
Simulation time		200 s	
Cell transmission power	46 dBm		30 dBm
Carrier frequency	2 GHz		28 GHz
Bandwidth	10 MHz		1 GHz
Number of cells	3		20/LTE cell
Inter cell distance	500 m		[0,125 m]
Path loss		3GPP urban model	
X2 link latency		1 ms	
S1 link latency		10 ms	
TTT		160	
Handoff: Intra - cell delay		10 ms	
Inter - cell delay		[60 ms, 100 ms]	
S1-MME link latency	10 ms		
RLC buffer size B_{RLC}		5 MB	
RLC AM reordering timer		1 ms	
mmWave outage threshold			−5 dB
UDP source rate		100 Mbits	
UE velocity		[3, 60] Km/h	
RLF	0 dB		$- \leq$ 6 dB

5 Simulation Results and Discussion

Simulation parameters are shown in Table 1 while results are shown in Figs. 4, 5, 6 and 7 below. To evaluate our adaptive HO model, Fig. 4 shows the HOF rate for hard HO and soft HOs including HO1, HO2 and HO3.

Figure 4 shows the HOF rate for hard HO and soft HOs including HO1, HO2 and HO3. In Fig. 4, results show that the HOF rates are different regardless of user density. The rapid increase of HOF rate in hard HO can be attributed to lack of an access level

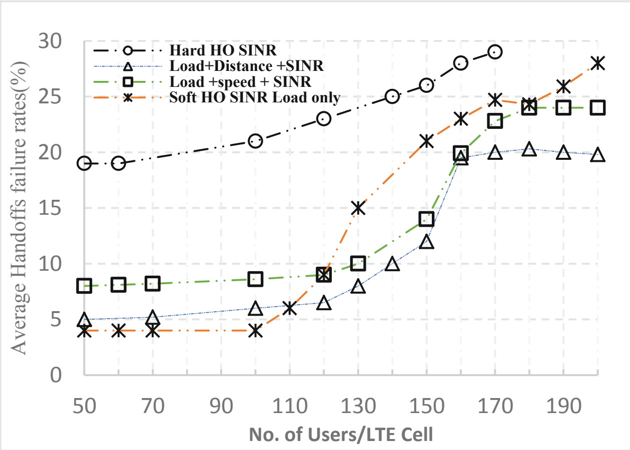

Fig. 4. Comparison of average HOF rate.

coordination resulting into the time increase of a cell selection which ultimately increases the HO latency. Adopting the Soft HO with distance strategy, i.e., (**HO2**), greatly reduces the HOF rate because it corresponds to reduced time needed for HO involvement among involved BSs. Technically, the long network association to one eNB due to distance consideration minimizes the HOF rate. This is proven by the fewer number of HOs when soft HO with distance, **HO2,** consideration is involved as shown in Fig. 5.

It should further be understood that with distance consideration in HO2, a limited number of SRS SINR values is considered at the LTE coordinating cell unlike in hard HO where there is no coordination. Since control plane resources are limited or scarce

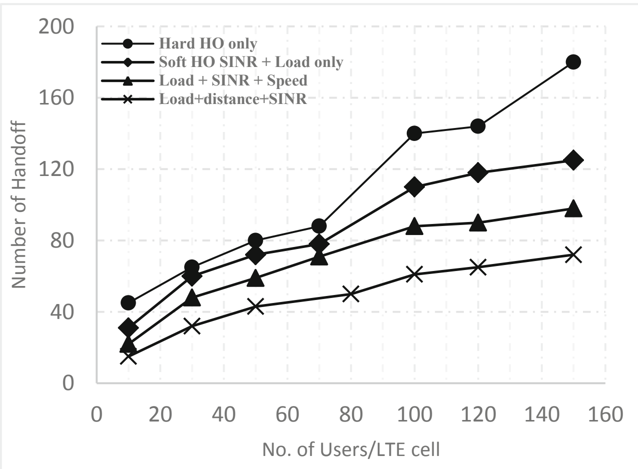

Fig. 5. Comparison of average number of Handoff times.

in lower frequencies, LTE resource blocks allocated to mmWave network requests should be limited to optimize resource availability on LTE fall back links. Since the distance consideration in HO can reduce the number of HOs, the LTE with limited resources can still efficiently coordinate and response to the decentralized SRS even under a large number of mmWave eNBs. At the same time, fewer HO sessions allow enough LTE resource blocks for network association to users that cannot connect to mmWave cells, e.g., due to high velocity which will result into high Doppler spread. This justifies the fewer HO failure rate when distance is considered.

When considering hard HO against soft HO with SINR and Load consideration only, **HO1**, the HOF rate curves in Fig. 4 have a radically different trend from **HO2** and **HO3**; they remarkably increase with increase in user numbers per cell. This attributed to the fact that, in principle, different eNBs have different user service queues, populated with different HO commands at different rates. Therefore, the mmWave eNBs involved in the initial association of the same user may start the association at different times when the network is loaded. There is a high chance (probability) that the HO association begins with non-optimal, but available eNB, because the optimal ones are busy for association with other existing users. As soon as optimal eNBs due to either speed or distance effects in HO assessment have less load or empty their queues, they quickly engage the users well positioned users due to better SINR and less load besides distance and speed estimations for HO2 and HO3. This leads to a more of HOs as seen in Fig. 5 for the Hard HO and HO1 strategies. This becomes worse by the rise in users, where the HO rate keeps raising as seen in Fig. 5 because every well position user in terms of SINR in Hard HO or load in HO1 will be queuing to join the best network. The rise in HOs and HOF is even worse for Hard HO where the user load variation is totally ignored as shown in Fig. 4, and HO complexity keeps increasing due to large number of users. Ultimately, a significant reduction in effective data rate is seen as shown in Fig. 6. For distance and speed-based HOs, HO requests and complexity is minimized by fewer numbers of HOs due to distance and

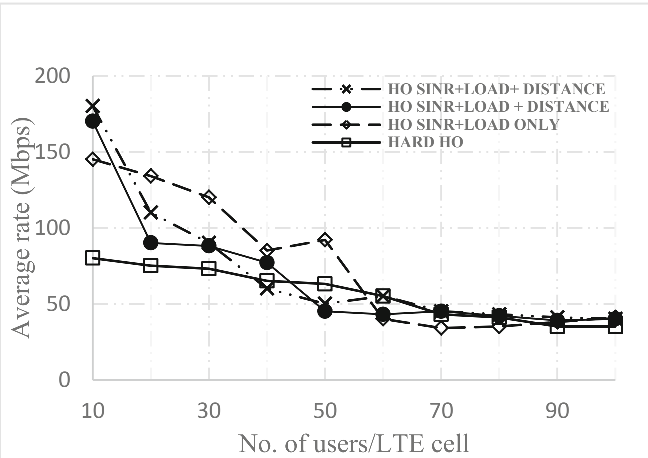

Fig. 6. Comparison of average throughput per user.

speed restrictions. Given a higher number of users, HO2 is able to stabilize HO failure rate better than HO3 as the population grows, this can be attributed to lower number of HOs due to distance consideration unlike speed as shown in Fig. 5.

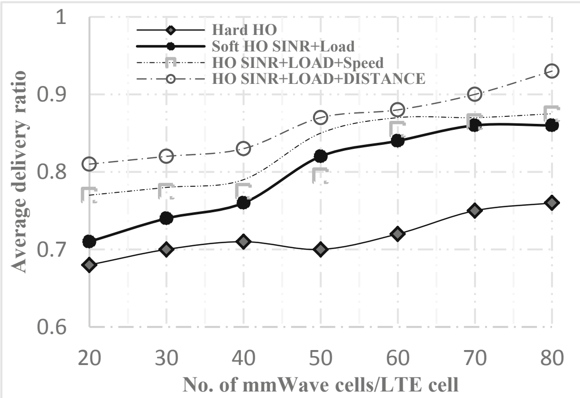

Fig. 7. Comparison of average packet delivery per session.

Finally, results in Fig. 7 clearly show packets generated and successfully received. Thus, the average delivery ratio shows the ability of the HO to transfer data successfully, end-to-end. In Fig. 7, results show that both distance-based, and the speed-based HOs are able to achieve approximately similar average delivery ratio. The average delivery ratios of speed-based and distance-based HO strategies increase by 10% to 15% as compared to that of Load only. Load only HO has also an increased performance compared of about 0.15 to that of Hard HO. Hence, Hard HO has a highest recurring routing overhead while distance-based HO has more reliable dynamic TTT intervals which leads to minimal routing overhead more consistent user rate.

6 Conclusions

We have thoroughly investigated the rich set of network access and HO challenges posed by the introduction of mmWave technologies in the 5G networks with DC capabilities. Results have demonstrated that badly positioned users can severely harm the HO efficiency hence effective user data rates. It requires context aware strategies in HOs and beam switching to sustain user connectivity. Essentially badly positioned users in a network causing multiple handoffs than carefully selected users, thus increase the overhead. While context awareness, such as knowing distance and speed, may increase complexity and signaling, we believe the trade-offs of HO simplicity against HO complexity can be offset by sustained network association or link maintenance. For instance, complexity brought about by considering distance and speed, which requires user and mmWave BS positions to be known by coordinating LTE cells, is offset by fewer HOs and failures. Finally, given the challenges highlighted in HOs, together with the proposed solutions of investigating DC at system level, major contribution that would improve mmWave cell management in 5G networks can be provided.

References

1. Xiao, B.M., et al.: Millimeter wave communications for future mobile networks. IEEE J. Sel. Areas Commun. **35**(9), 1909–1935 (2017)
2. Polese, M., Mezzavilla, M., Zorzi, M.: Performance comparison of dual connectivity and hard handover for LTE-5G tight integration. In: Proceeding of the 9th EAI International Conference on Simulation Tools and Techniques (SIMUTOOLS), pp. 118–123 (2016)
3. Polese, M., Giordani, M., Mezzavilla, M., Rangan, S., Zorzi, M.: Improved handover through dual connectivity in 5G mmWave mobile networks. IEEE J. Sel. Areas Commun. **35** (9), 2069–2084 (2017)
4. Rangan, S., Rappaport, T.S., Erkip, E.: Millimeter-wave cellular wireless networks: potentials and challenges. Proc. IEEE **102**(3), 366–385 (2014)
5. GPP: NR and NG-RAN Overall Description - Rel. 15, TS 38.300 (2018)
6. Park, J., Jung, S.Y., Kim, S., Bennis, M., Debbah, M.: User-centric mobility management in ultra-dense cellular networks under spatio-temporal dynamics. In: Proceeding of IEEE Global Communications Conference (GLOBECOM), Washington, DC, pp. 1–6 (2016)
7. 3GPP: Scenarios and requirements for small cell enhancement for EUTRA and E-UTRAN (Release 14), 3rd Generation Partnership Project (3GPP), TR 36.932, March 2017
8. Joud, M., García-Lozano, M., Ruiz, S.: User specific cell clustering to improve mobility robustness in 5G ultra-dense cellular networks. In: Proceeding of 14th Annual Conference on Wireless On-demand Network Systems and Services (WONS), Isola, pp. 45–50 (2018)
9. Thapliyal, A.: Mobility Robustness in 5G Networks. Ph.D. thesis, Aalto University School of Electrical Engineering, Finland (2016)
10. Gimnez, L.C., Michaelsen, P.H., Pedersen, K.I.: UE autonomous cell management in a high-speed scenario with dual connectivity. In: Proceeding of IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), pp. 1–6, September 2016
11. Lee, S., Jung, J., Moon, J., Nigam, A., Ryoo, S.: Mobility enhancement of dense small-cell network. In: Proceeding of IEEE Consumer Communications and Networking Conference (CCNC), pp. 297–303, January 2015
12. Ibrahim, H., ElSawy, H., Nguyen, U.T., Alouini, M.S.: Mobility aware modeling and analysis of dense cellular networks with CPlane/U-plane split architecture. IEEE Trans. Commun. **64**(11), 4879–4894 (2016)
13. Kuruvatti, N.P., Molano, J.F.S., Schotten, H.D.: Mobility context awareness to improve quality of experience in traffic dense cellular networks. In: Proceeding of International Conference on Telecommunications (ICT), pp. 1–7, May 2017
14. Sun, Y., Chang, Y.: A universal predictive mobility management scheme for urban ultra-dense networks with control/data plane separation. IEEE Access **5**, 6015–6026 (2017)
15. Bassoy, S., Farooq, H., Imran, M.A., Imran, A.: Coordinated multi point clustering schemes: a survey. IEEE Commun. Surv. Tutor. **19**(2), 743–764 (2017)
16. Kurda, R., Boukhatem, L., Kaneko, M., Yahiya, T.A.: Mobility-aware dynamic inter-cell interference coordination in HetNets with cell range expansion. In: Proceeding of IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), pp. 1115–1119, September 2014