Effect of Handover on the Performance of Scheduling Algorithms in LTE Networks

Narges Shojaedin
University of Calgary
nshojaed@ucalgary.ca

Majid Ghaderi
University of Calgary
mghaderi@ucalgary.ca

Ashwin Sridharan
AT&T
asridharan@research.att.com

***Abstract*—In this work, using *ns*-3 and different mobility models, we simulate realistic LTE network scenarios to study the effect of handover on two popular schedulers, namely the Proportional Fair (PF) and Max Weight (MW) scheduler. The performances of these schedulers are widely studied in the literature via simulation and mathematical analysis in the absence of handovers. Specifically, it has been shown that MW is throughput optimal among all scheduling policies that stabilize the system in the sense of bounding the user queues. In our experiments, however, we observe that such general conclusions may not be accurate in the presence of mobile users that hand over across multiple cells. To this end, we show that: i) MW achieves higher throughput than PF when users are confined to a single cell, but ii) when there is handover in the network across multiple cells, PF achieves a throughput similar to that of MW, and in some cases even slightly outperforms MW. Furthermore, these observations are consistent across a wide range of network scenarios in terms of round-trip delay, buffer size and channel fading.**

I. Introduction

To meet the growing demand for cellular services, cellular operators around the world are deploying 4G networks based on the Long Term Evolution (LTE) standard. The promise of LTE is to provide high data rate and low latency by providing a packet-optimized wireless access and core network. The underlying physical layer technology in LTE is Orthogonal Frequency Division Multiple Access (OFDMA) in which the radio frequency is divided into many orthogonal subcarriers. A multi-user scheduler at the evolved NodeB (eNB) assigns subsets of subcarriers to individual users allowing for flexible bandwidth sharing in the system. There is a large body of work on developing scheduling algorithms for wireless networks in general [1], and LTE networks [2] in particular. Specifically, two schedulers, namely, the Proportional Fair (PF) [3] and Max Weight (MW) [4] schedulers, have received significant attention in the literature. The PF scheduler is widely deployed in modern day wireless networks owing to its simplicity and ability to achieve a good balance between system throughput and fairness. The MW scheduler, on the other hand, is specifically designed to maximize the system throughput, while stabilizing the user queues at eNBs.

A defining feature of cellular networks is the ability of users to roam in the coverage area of the network without loosing their end-to-end connectivity. Indeed, achieving seamless user mobility even at a high speed is one of the prominent goals of LTE design. As a user moves from the coverage area of one cell to the coverage area of another cell, the old cell “hands over” the user to the new cell. While the handover procedure in LTE is designed to have a low latency, there is still some time during which the user is disconnected from the network. To alleviate the service disruption during this time, the source eNB temporarily forwards the incoming data and the data that is already in buffer for the user to the target eNB. However, the forwarding of the user data may cause problems of its own when TCP is involved due to increased delay of the forwarded data. There can be a time interval immediately after the handover when packets on both the direct path and the forwarding path arrive in parallel, albeit with different delays, at the target eNB. This may give rise to the problem of out of order packets and unnecessary reduction in TCP throughput due to the ensuing duplicate ACKs and spurious time-outs.

The impact of handovers on TCP performance has been extensively studied in the literature over the past several years [5]. There has also been several recent works specifically on this subject in LTE networks (*e.g.*, [6]). In LTE networks, in order to increase the spatial reuse of the system and consequently increase the wireless bandwidth, *small cells* in the form of micro, pico and femto cells are being deployed [7]. The reduction of cell size in LTE leads to an increased rate of handover for mobile users, which consequently exacerbates the impact of handovers on TCP.

Different from the existing work on TCP, in this paper, we argue that it is not just TCP that is affected by handovers, rather *lower layer network mechanisms such as radio resource schedulers are also affected by handovers*. Specifically, the scheduling mechanisms that take into consideration *queue* backlogs when scheduling users, such as the Max Weight (MW) algorithm, exhibit different performance with handovers which consequently affects not only TCP but UDP traffic as well. This aspect of handover has not been previously studied in the literature.

To this end, we study the performance of PF and MW with and without handovers in LTE networks. We measure the performance in terms of the average TCP throughput achieved under each scheduler. While there is significant work on the performance of PF and MW in the absence of handovers, it is extremely hard to capture the effect of handover on the scheduling algorithms using analytical models. We are not aware of any work that has tackled this problem analytically due to many interactions involved in the system. Thus, in this work, we use *ns*-3 and different mobility scenarios to simulate realistic LTE networks across a wide range of network conditions in terms of round-trip delay, buffer size and channel fading. We use the LENA module [8] to create an end-to-end LTE network which has all the major elements of a real LTE system including the air interface Evolved UMTS Terrestrial Radio Access (E-UTRA) and Evolved Packet Core (EPC). The LTE model in *ns*-3 provides a detailed implementation of various aspects of the LTE standard [8] such as OFDMA, hybrid ARQ, adaptive modulation and coding, and handover management. The *ns*-3 implementation follows detailed specification of TCP and 3GPP LTE. Hence, the results provided should be representative of what happens in real systems.

Our results show that that:

1) MW achieves higher throughput than PF when users are confined to a single cell, but,
2) when there is handover in the network across multiple cells, PF achieves a throughput similar to that of MW, and in some cases even slightly outperforms MW.

We note that there exist analytical results on the performance of radio resource schedulers that rely on simplifying assumptions about the network dynamics and ignore handovers. Those models conclude that MW is throughput optimal [9], which is not consistent with the results we obtained when considering a realistic network with user handovers.

The rest of the paper is organized as follows. Section II provides a brief introduction to LTE. In Section III, we describe the implementation of PF and MW in *ns*-3. Our simulation results are discussed in Section IV, while Section V concludes the paper.

II. LTE Primer

This section presents a brief introduction to LTE with specific focus on LTE resource allocation and handovers.

A. LTE Architecture

LTE is a fourth generation high-speed wireless network that evolved from the Universal Mobile Telecommunication System (UMTS), which in turn evolved from the Global System for Mobile Communications (GSM). The main goals of LTE are spectral efficiency, high data rate for different services (such as VOIP, streaming multimedia and video conferencing), flexible carrier bandwidth, and QoS support.

The high-level network architecture of LTE is depicted in Fig. 1. The network has three main components:

- User Equipment (UE). Such as a smartphone or tablet which communicates with the LTE network over an OFDMA radio interface.
- Evolved UMTS Terrestrial Radio Access Network (E-UTRAN). The E-UTRAN consists of eNBs and handles the radio communications between UEs and the evolved packet core. Each eNB is a base station that controls UEs in one or more cells and is capable of fast resource allocation over time slots of 0.5 milli-seconds.
- Evolved Packet Core (EPC). The EPC acts as the intermediate gateway between the radio network (eNB) and the Internet and performs functions such as QoS control, charging, anchor point, *etc*.

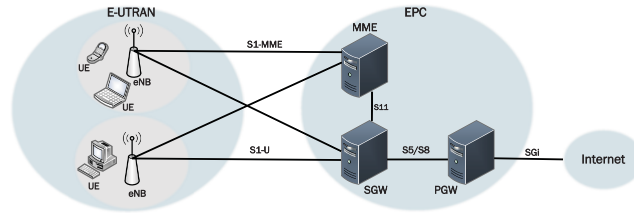


Fig. 1. High-level LTE network architecture.

A key feature of the LTE network is that the radio interface uses OFDMA to significantly enhance transmission speeds above that of 3G technologies and has features such as high robustness against frequency selective fading and high spectral efficiency, and allows flexible bandwidth sharing which we elaborate on further below.

B. LTE Resource Allocation

In LTE, the radio resources are allocated in both time and frequency domains. The smallest radio resource unit that can be assigned to a UE is called a physical Resource Block (RB). The scheduler at the eNB assigns subsets of RBs (not necessarily adjacent ones) to individual users. In the frequency domain, the system bandwidth is divided into multiple sub-channels of 180 kHz, where each RB spans over one sub-channel. An RB in the time domain is one Transmission Time Interval (TTI) which lasts for two time slots each of length 0.5 ms. The time is divided into frames, each frame is made of 10 consecutive TTIs (or sub-frames). In each TTI, UE measures the pilot signal from the serving eNB and periodically reports the Channel Quality Indicator (CQI) to the eNB. The scheduler uses this indicator along with queue information to determine how many RBs allocate to each UE.

C. Handover in LTE

In LTE, the handover (HO) procedure is controlled by the network but assisted by the User Equipment (UE) [10]. The procedure is triggered by the measurement reports sent by the UE to its corresponding eNB. The measurement reports, among other parameters, specify the target eNB to which the UE has to be handed over. The source eNB and target eNB then use the X2 interface between them to directly exchange the necessary handover information without the involvement of EPC. Once a new connection is established, the target eNB informs the Mobility Management Entity (MME), which in turn, informs the Serving Gateway (S-GW) to switch the downlink data path for the UE to the target eNB. This form of handover is referred to as *hard handover* as the UE is disconnected from the source eNB before connecting to the target eNB.

D. ns-*3 Simulator and LTE*

In this work, we use the LTE component of the *ns*-3 simulator (LENA)[1] to accurately simulate dynamics of the LTE radio interface [8]. The LENA module closely emulates 3GPP standards for the data plane, faithfully reproducing interactions of the EPC as well as the various stacks of the LTE radio layer such as the PDCP, RRC, MAC and PHY. It incorporates the Femto Cell scheduling framework as well as accurate models for emulating transmission, fading and decoding on the radio interface. As such, the simulator provides a fairly accurate proxy for the actual network.

III. System Model

In this section, we describe the network configuration and parameters that have been considered in our work. We also describe the Max Weight (MW) scheduling policy that we have implemented in *ns*-3.

A. Network Topology

We have implemented two LTE-EPC networks in *ns*-3; the single-cell and multi-cell network. In the single-cell network, the UEs are distributed randomly in the cell within a distance of 500 m to 5000 m from the eNB and a remote host sends packets to the UEs using TCP. We use channel fading traces for pedestrian, urban and vehicular mobilities that come with the LENA package [11]. In the multi-cell network, 3 tri-sectored macrocell sites where each site has 3 cells, *i.e.*, 9 cells in total, are deployed in a hexagonal layout with 500 m inter-site distance. UEs are randomly distributed around the sites and roam the simulation environment with different movement speeds depending on the selected trace-based channel fading model.

In both topologies, the remote host is connected to the gateway node of the LTE network with a high-speed link (100 Gb/s) in order to avoid any bottleneck effects outside the LTE network. All UEs are connected to the remote host, while multiple TCP servers run on the remote host, each server is dedicated to one UE. In fact, there is one TCP flow from the remote host to each UE. Each flow of traffic is generated by remote host and passes through the gateway to the eNB. The eNB maintains a queue for each flow where traffic flow awaits transmission to associated UE. A scheduler at eNB allocates radio resources to flows by following a specific priority metric (*i.e.*, scheduling policy).

B. System Parameters

Various system parameters are summarized in Table I. Those parameters that are not listed in the table are used with their *ns*-3 default values. As can be seen, some of the parameters that deal with the LTE network setup are fixed during the simulations while those that represent network properties, *e.g.*, Internet delay, vary over a wide range of different values. In this work, we are specifically interested in studying the impact of handover on the performance of various scheduling algorithms across a wide range of network scenarios in terms of the following parameters:

[1]We use ns-3.19 which at the time of writing is the latest release.

TABLE I
SYSTEM PARAMETERS

Parameter	Value
AMC mode	PiroEW2010 [11]
Fading model	Pedestrian, Urban, Vehicular
Mobility model	Steady-State Random Waypoint
Packet size	1024 bytes
Internet delay	10 ms, 20 ms, 50 ms
Number of UEs	5 in single-cell and 19 in multi-cell
Buffer size at eNB	10, 50, 100 packet
Simulation time	15 seconds

- *Internet delay:* This is the one-way propagation delay between the remote host and the LTE gateway. We set the default Internet delay in our experiments to 10 ms, but will run our experiments with a range of Internet delays. The *round-trip time* for a TCP packet is given by twice the "Internet Delay" plus the transmission time of the packet inside the LTE network. In addition to the Internet delay, the handover process also introduces some delay due to packet forwarding from the old eNB to the new eNB, and the signalling involved in initiating the handover and switching the UE.
- *eNB buffer size:* Packet loss and delay that affect TCP throughput are highly dependent on the queue size. We consider a range of queue sizes covering small and large queues.
- *Fading model:* Traces for three different mobility scenarios are provided in *ns*-3: Pedestrian with mobility speed of 3 kmph, Vehicular with mobility speed of 60 kmph and Urban with mobility speed of 3 kmph.

C. Scheduling Algorithms

There is a build-in PF scheduler in *ns*-3, but we had to implement our own MW as it is not yet included in *ns*-3. The priority metric for MW is the product of queue length and achievable data rate of UEs. In our implementation, we update the queue length of the UE that gets allocated a new RBG during a TTI and do not allocate any RBG to a UE with an empty queue. Let $R_j(k,t)$ denote the maximum rate achievable by UE j on RBG k at TTI t. The scheduling decision of MW and PF are described below.

- The scheduling decisions for MW are made as follows:

$$\hat{i}_k(t) = \arg\max_{1\leq j\leq N} \left\{Q_j(t)\cdot R_j(k,t)\right\},$$

  where, $\hat{i}_k(t)$ is the UE chosen for transmission on RBG k at TTI t, N denotes the number of UEs in the cell, and $Q_j(t)$ is the queue size of UE j before allocating RBG k at TTI t. Clearly, as the queue backlog of a UE grows, its priority for scheduling increases. This guarantees that every UE eventually gets some resources. It has been shown that such a policy is throughput optimal when queue size is unlimited.

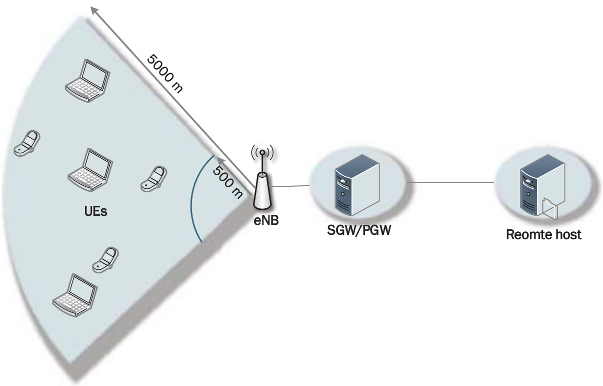


Fig. 2. Single-cell LTE network topology. There is no handover in this topology as all UEs are connected to a single eNB.

- The PF scheduler uses the following scheduling policy to schedule users:

$$\hat{i}_k(t) = \arg \max_{1 \le j \le N} \{R_j(k,t)/T_j(t)\},$$

where, $T_j(t)$ is the moving average throughput achieved by UE j until TTI t, which is computed as follows:

$$T_j(t) = \beta \cdot T_j(t-1) + (1-\beta) \cdot R_j(k,t),$$

for some $0 < \beta < 1$. Similar to MW, PF guarantees that every user eventually is granted access to radio resources regardless of its achievable rate. It has been shown that such a policy achieves proportional fairness among UEs. A scheduling policy is proportionally fair if it maximizes $\sum_{1 \le j \le N} \log(T_j)$, where T_j is the average rate allocated to UE j.

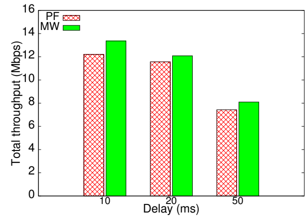
IV. Results and Discussion

In this section, we study the performance of PF and MW scheduling algorithms in terms of their achieved throughput. Throughput plots in this section represent the total throughput achieved in the network by all UEs. Each individual experiment result is the average of 10 independent simulation runs, each lasting for 15 seconds. The default values for Internet delay and buffer size are 10 ms and 100 packets, respectively.

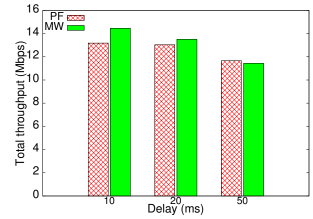
A. *Single-Cell Network*

To isolate the effect of handover on the performance of PF and MW, in the first set of results, we investigate the performance of both algorithms in a single cell network where there is no handover. This is the scenario that is commonly considered in the literature. The remote host sends TCP packets to 5 UEs located in a single cell, in a back-to-back fashion, for the entire duration of the simulation where all UEs are connected to a single eNB node, as depicted in Fig. 2. Note that, we allow UEs to move around in the coverage area of the cell but there is no handover.

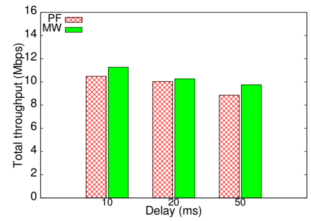
Effect of Internet Delay. Internet delay has a significant effect on TCP performance as it affects the round-trip delay of TCP packets. We change the Internet delay from the default value of 10 ms to 20 and 50 ms and measure the total throughput. This range of delays represents scenarios with low, medium and high delay. As shown in Fig. 3, increasing the Internet delay dramatically reduces TCP throughput. This is, of course, a well-known behavior of TCP. What we are looking for is the relative performance of MW and PF. The expectation is that MW should outperform PF. This is indeed confirmed by our simulations as MW achieves higher throughput in most experiments for the range of Internet delays considered.



(a) Pedestrian mobility.

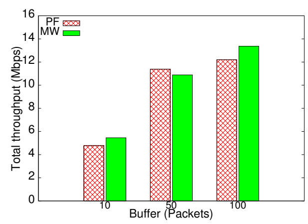


(b) Urban mobility.

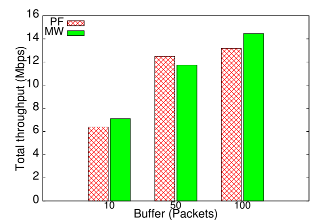


(c) Vehicular mobility.

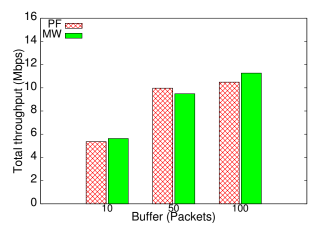
Fig. 3. Effect of Internet delay in a single-cell network: MW achieves higher throughput in most experiments, as expected.



(a) Pedestrian mobility.



(b) Urban mobility.



(c) Vehicular mobility.

Fig. 4. Effect of buffer size in a single-cell network: MW achieves higher throughput in most experiments.

Effect of Buffer Size. Fig. 4 shows the performances of the schedulers for a range of buffer sizes. It is well-known that TCP performance is sensitive to the amount of buffer available at the eNB. It can be seen that as the buffer size increases so does the TCP throughput. The reason is that if there is a drop

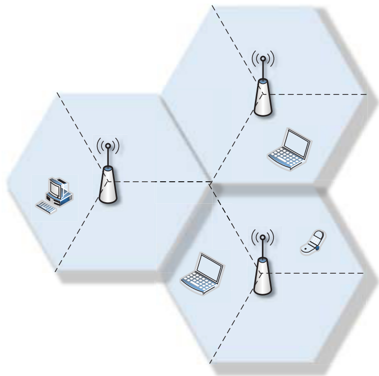
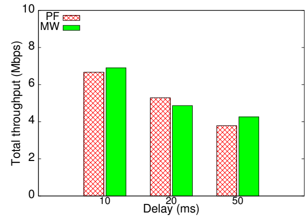
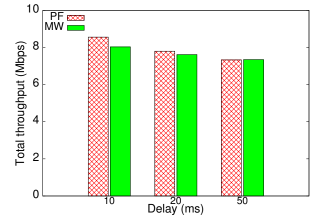


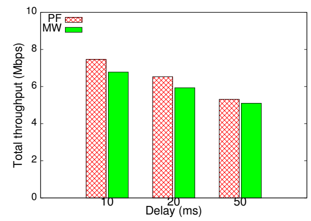
Fig. 5. Multi-cell LTE network topology. UEs are connected to any of the three available eNBs. The handover algorithm selects the target eNB based on the best possible Reference Signal Received Quality (RSRQ).



(a) Pedestrian mobility.

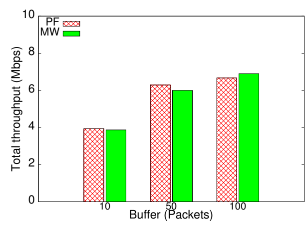


(b) Urban mobility.

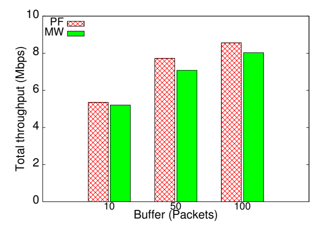


(c) Vehicular mobility.

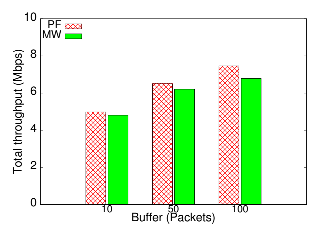
Fig. 6. Effect of Internet delay in a multi-cell network: PF outperforms MW in most experiments.



(a) Pedestrian mobility.



(b) Urban mobility.



(c) Vehicular mobility.

Fig. 7. Effect of buffer size a multi-cell network: PF outperforms MW in most experiments.

in TCP rate (*e.g.*, due to a packet loss), with a small buffer at eNB, there is not enough data at the buffer to saturate the wireless link. What we are interested in, however, is to find out how different schedulers perform under different amount of eNB buffer size. As can be seen from the figures, for small and large buffer sizes, MW outperform PF, as expected. However, for a medium buffer size of 50 packets, PF outperforms MW, which is in contrast to the throughput optimality of MW. This observation can be justified as follows. For small buffer sizes (*e.g.*, 10 packets), the congestion control algorithm of TCP is dominant so that there is little room for the schedulers to have any significant effect on TCP throughput. Between PF and MW, MW is more sensitive to buffer size and requires a larger buffer size to realize its optimality. Indeed, as our simulation results show, for medium size buffers (*e.g.*, 50 packets), MW is not able to achieve its potential, while for larger buffer sizes (*e.g.*, 100 packets) it outperforms PF.

B. Multi-Cell Network

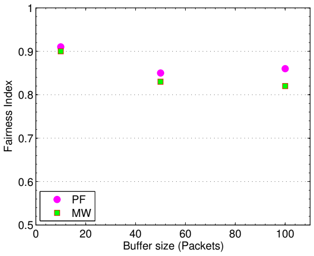
After observing the effect of PF and MW on TCP throughput in the absence of handovers, we turn our attention to network scenarios with multiple cells and user handovers. The LTE network scenario simulated in this section is depicted in Fig. 5. There are 19 UEs in the network that are randomly distributed in the coverage area of 9 cells. The handover algorithm used in the simulations is the algorithm “A2-A4-RSRQ” implemented in *ns*-3, which selects the target eNB based on the best possible Reference Signal Received Quality (RSRQ).

Effect of Internet Delay. For the sake of comparison, we have conducted a set of experiments similar to those conducted for the single-cell network. The results are depicted in Fig. 6. Interestingly, unlike the single-cell network, PF achieves higher throughput in most experiments. Note that the only difference between single- and multi-cell scenario is the existence of handovers in multi-cell scenario. In both scenarios users are mobile. Thus, any difference in the results is contributed to handovers and their effect on the scheduling algorithms.

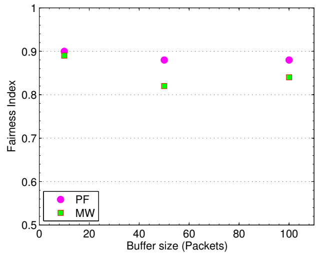
Effect of Buffer Size. The simulation results for various buffer sizes are depicted in Fig. 7. Recall that in the single-cell network, MW outperformed PF when the buffer size was large. In the multi-cell scenario, on the contrary, PF outperforms MW in almost all experiments. Again, the only difference between multi- and single-cell scenario is the presence of handovers. To understand the surprising behavior of MW with and without handovers, we have studied the buffer occupancy in single- and multi-cell scenario. The results are presented later in subsection IV-D along with a justification for the observed behavior.

C. Fairness Comparison

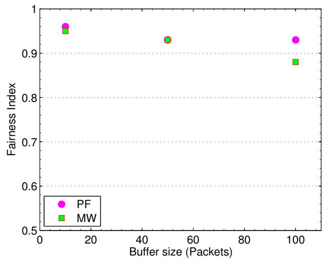
An important measure of performance in multi-user networks is the fairness among users. There are several notions of fairness and fairness measures such as α*-fairness* [12] and



(a) Pedestrian mobility.



(b) Urban mobility.



(c) Vehicular mobility.

Fig. 8. Jain's fairness index comparison of PF and MW schedulers for different fading models and buffer sizes. In all experiments, PF and MW indices are very close.

Jain's fairness [13]. Since our focus in this work is on the throughput performance of TCP under MW and PF, we only show the results of fairness using the Jain's fairness index for both algorithms. Jain's fairness index is measured by the following metric:

$$\text{Jain's fairness index} = \frac{(\sum_{1\leq i\leq N} x_i)^2}{N \cdot \sum_{1\leq i\leq N} x_i^2}.$$

The maximum value for Jain's index is 1, which is attained when all UEs achieve equal throughput.

The results are depicted in Fig. 8. As can be seen, both algorithms achieve fairness indices that are very close to each other. In other words, there is negligible difference between the two algorithms in terms of achieved fairness among UEs in the scenarios we have simulated.

D. Discussion

The only change from the single-cell network to the multi-cell network is the introduction of handovers. As alluded to earlier, it is well-known that TCP throughput suffers from the extra latency and packet duplicates caused by handovers. Thus, the reduction in the total throughput in the multi-cell network compared to the single-cell network can be attributed to TCP. However, the inferior performance of MW compared to PF in the multi-cell network is orthogonal to this and can not be explained by TCP behavior.

An interesting observation in both single-cell and multi-cell experiments is that the MW throughput increases as the buffer size increases (see Figs. 4 and 7). Indeed, the MW algorithm is designed to stabilize user buffers assuming that the buffer occupancy can grow to infinity. If the buffer size is limited (which is the case in any LTE network) then MW may not performed as expected, as can be observed in our experiments. The MW performance is very sensitive to the buffer size and as the buffer size reduces so does the MW performance. To correlate this sensitivity to the behavior we observed in our experiments, we have reported the average measured buffer occupancy in the single-cell and multi-cell network in Table II. As can be seen, the average buffer occupancy with handovers is much lower than that without handovers. While PF is insensitive to buffer occupancy, MW achieves a lower throughput due to lower buffer occupancy.

TABLE II
AVERAGE BUFFER OCCUPANCY.

	Pedestrian	Urban	Vehicular
Single-Cell	780	3344.67	13884
Multi-Cell	111.84	716.6	9482

V. CONCLUSION

In this work, we studied the throughput performance of PF and MW schedulers in LTE networks using *ns*-3. We found that while MW generally outperforms PF in a single-cell network, in a multi-cell network, PF actually achieves a higher throughput across a diverse set of network configurations in terms of round-trip delay, buffer size and fading model. Our results show that this behavior can be attributed to the sudden drops in the occupancy of user buffers at the eNBs due to the hard handover mechanism of LTE, which negatively affects MW as it is sensitive to the amount of backlog in user buffers.

REFERENCES

[1] A. Pantelidou and A. Ephremides, "Scheduling in wireless networks," *Foundations and Trends in Networking*, vol. 4, no. 4, 2009.

[2] F. Capozzi, G. Piro, and G. Boggia, "Downlink packet scheduling in LTE cellular networks: Key design issues and a survey," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 2, 2013.

[3] F. Kelly, A. Maulloo, and D. Tan, "Rate control in communication networks: Shadow prices, proportional fairness and stability," *J. Oper. Res. Soc.*, vol. 49, no. 3, 1998.

[4] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," *IEEE Trans. Autom. Control*, vol. 37, no. 12, 1992.

[5] Y. Tian, K. Xu, and N. Ansari, "TCP in wireless environments: Problems and solutions," *IEEE Commun. Mag.*, vol. 43, no. 3, 2005.

[6] D. Pacifico *et al.*, "Improving TCP performance during the intra LTE handover," in *Proc. IEEE Globecom*, Dec. 2009.

[7] Small Cell Forum. [Online]. Available: http://www.smallcellforum.org/

[8] G. Piro, N. Baldo, and M. Miozzo, "An LTE module for the ns-3 network simulator," in *Proc. SIMUTools*, Mar. 2011.

[9] K. Jagannathan, M. Markakis, E. Modiano, and J. N. Tsitsiklis, "Throughput optimal scheduling over time-varying channels in the presence of heavy-tailed traffic," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, 2014.

[10] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2, Rel 12," Jan. 2014.

[11] C. T. de Telecomunicacions de Catalunya (CTTC), "LTE simulator documentation release M6," Apr. 2013.

[12] M. Uchida and J. Kurose, "An information-theoretic characterization of weighted α-proportional fairness," in *Proc. IEEE Infocom*, Apr. 2009.

[13] R. K. Jain, *The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling*. Wiley, 1991.