This regular paper was presented as part of the main technical program at IFIP WMNC'2011

An analytical model to study the impact of time-varying cell capacity in LTE networks

Bart Sas
PATS Research Group
University of Antwerp/IBBT
2020 Antwerp, Belgium
Email: bart.sas@ua.ac.be

Elena Bernal-Mor
GIRBA Research Group
Universitat Politècnica de València/ITACA
46022 Valencia, Spain
Email: elbermo@upvnet.upv.es

Kathleen Spaey
PATS Research Group
University of Antwerp/IBBT
2020 Antwerp, Belgium
Email: kathleen.spaey@ua.ac.be

Vicent Pla
GIRBA Research Group
Universitat Politècnica de València/ITACA
46022 Valencia, Spain
Email: vpla@dcom.upv.es

Chris Blondia
PATS Research Group
University of Antwerp/IBBT
2020 Antwerp, Belgium
Email: chris.blondia@ua.ac.be

Jorge Martinez-Bauset
GIRBA Research Group
Universitat Politècnica de València/ITACA
46022 Valencia, Spain
Email: jmartinez@upvnet.upv.es

***Abstract*—Contemporary wireless networks like Long Term Evolution (LTE) employ a technique called adaptive modulation and coding (AMC) to enhance the throughput of the users in the system. Applying AMC however causes the total cell capacity to vary over time as sessions are started and stopped and users move around. The varying cell capacity has an impact on the quality of service (QoS) experienced by the users and also on the admission control (AC) algorithms used for such system as the variation of the cell capacity can cause the cell capacity to drop below the required amount that is needed to service all users in a cell.**

In this paper we present an analytical model that models this time-varying cell capacity and compare the results obtained with it to results obtained from more realistic simulations in order to verify the modelling assumptions made in the analytical model. We then use both the analytical model and the simulations to study the impact of the time-varying cell capacity on a simple AC scheme. Scenarios in which various parameters are varied are simulated and the results of both models are compared to each other.

The results obtained from the analytical model and the simulations show that the analytical model is very accurate. The differences between the results only differ up to a couple of tenths of a percent. Only in extreme conditions both models differ. We also identify the cases and the reason why both models differ.

***Index Terms*—LTE, Time-varying cell capacity, Quality of service, Admission control.**

I. Introduction

Contemporary wireless networks like Long Term Evolution (LTE), the next generation of cellular networks, or Worldwide Interoperability for Microwave Access (WiMAX), a standard for wireless broadband internet access, try to optimise performance by employing adaptive modulation and coding (AMC) [1]. AMC allows different modulation and coding schemes (MCSs) to be used at different points in time depending on the signal quality. By using different MCSs data can be sent to the users at higher or lower bit rates as they move around in a cell. Since the capacity of a cell is determined by the number of active users and the throughput that can be achieved by each of these users, changing the MCS will cause the cell capacity to vary as the number of active users changes (calls are started and stopped) and as users move around. This varying of the cell capacity might influence the decisions taken by admission control (AC), which is the algorithm that decides if a session is admitted to the cell or not. It bases its decisions on the availability of the resources that are needed to guarantee the quality of service (QoS) of the new session and the already accepted sessions. Due to the time-varying cell capacity, it is for instance possible that at one point in time there is sufficient capacity to provide the desired QoS to all users while at another point in time variations in the cell capacity cause the cell capacity to drop below the required amount. When designing AC algorithms it is important to take this possibility into account.

In this paper we will create an analytical model for studying the effects of varying cell capacity on AC. We consider a single cell in which users move around. When a new session arrives it is subjected to AC which will decide if the user is admitted to the cell or not. Only one, real-time, service is considered. Although for our purposes it is not of great importance which traffic direction is considered, we consider only the downlink traffic direction. Users that are accepted remain active for a certain amount of time. In order to maintain a service with a fixed bitrate, active users require a varying number of resources depending on their location in the cell. The modelling assumptions made in the model will then be verified using simulations that model the mobility of the users and the session duration more accurately.

Analytical models for studying AC in cellular networks have already been studied in several papers, but most of them do not consider the varying of the cell capacity due to user movement. For example, in [2] and [3] the efficiency of several AC policies in cellular networks is studied. An analytical model considering varying cell capacity is described in [4] and an AC algorithm that takes into account the mobility of the users is

978-1-4577-1193-0/11/$26.00

proposed in [5]. Studying the effects of varying cell capacity on AC using simulations has already been done in a number of papers. [6] proposes an AC scheme for WiMAX and evaluates its performance and [7] studies self-optimisation of an AC scheme for LTE.

This paper is structured as follows: Section II explains how AMC influences the cell capacity in wireless networks. In Section III the general modelling assumptions are presented. Section IV discusses the analytical model. The simulation model that is used to verify the mobility and session duration assumptions made in the analytical model is described in Section V. In Section VI numerical results obtained with both the analytical model and the simulations are discussed and compared. Section VII concludes this paper and mentions some future work.

II. Varying cell capacity

A. Adaptive modulation and coding

In this paper we will consider AMC as it is implemented in LTE, although the same principles also apply to AMC in other network technologies. LTE defines 15 different MCSs, which are obtained by combining 3 modulation schemes: quadrature phase-shift keying (QPSK), and quadrature amplitude modulation (QAM) with 16 and 64 states per symbol, with a number of code rates. The modulation scheme determines how binary data (bits) is converted to a signal. The code rate represents the ratio of information bits to the total amount of bits (information bits + forward error correction (FEC) bits) that is used to transmit the data. The lower the signal quality, the lower the code rate has to be, (i.e. more FEC bits are needed per information bit) in order to tackle the high number of erroneous bits. An overview of the MCSs that are used in LTE can be found in Table I.

The choice of which MCS is used is based on measurements of the signal quality that are provided by the user equipment (UE) to the eNodeB (eNB). Based on the modulation scheme and the code rate, the theoretical maximum bit rate for each MCS can be calculated. Using the attenuated Shannon bound [1] the minimum signal to interference ratio (SIR) that is needed to be able to use a certain MCS can be determined. The SIR is the ratio (in the linear domain) of the received signal strength (S) at the UE of the source eNodeB (SeNB) to the total strength of the signals of the interfering neighbouring eNodeBs (NeNBs) (I_i): $SIR = \frac{S}{\sum_i I_i}$. The higher the SIR, the better the quality of the signal. The strengths of the received signals S and I_i in turn depend on a number of factors. First of all there is the power at which the signals are transmitted. The most dominant factor in the reduction of the transmitted signal power is the pathloss. Pathloss is the decreasing of the signal strength because of spatial dispersion and is linked to the distance between the transmitter and the receiver. The received signal strength is also influenced by other factors like shadow fading and multi-path fading. As the pathloss has the biggest impact on the received signal strength and thus also on the SIR, the SIR will especially depend on the relative distance from the SeNB and the interfering NeNBs. If we consider a site that is surrounded by a number of interfering sites, the lines with equal SIR will be cycles around the SeNB (see Fig. 1).

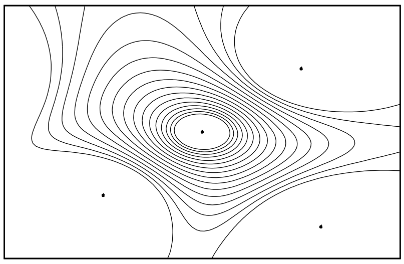

Figure 1. Lines of equal interference surrounding a site.

Table I
THE DIFFERENT MCSS USED BY LTE [8].

#	Modulation	Bits/symbol	Code rate
1	QPSK	2	$78/1024$
2	QPSK	2	$120/1024$
3	QPSK	2	$193/1024$
4	QPSK	2	$308/1024$
5	QPSK	2	$449/1024$
6	QPSK	2	$602/1024$
7	16-QAM	4	$378/1024$
8	16-QAM	4	$490/1024$
9	16-QAM	4	$616/1024$
10	64-QAM	6	$466/1024$
11	64-QAM	6	$567/1024$
12	64-QAM	6	$666/1024$
13	64-QAM	6	$772/1024$
14	64-QAM	6	$873/1024$
15	64-QAM	6	$948/1024$

B. Determining the cell capacity

LTE uses orthogonal frequency-division multiple access (OFDMA). The frequency domain is divided into non-overlapping subchannels measuring 180 kHz and the time domain in slots with a duration of 1 ms. These subdivisions of time and frequency are referred to as scheduling resources (SRs) in this paper. The SRs are distributed among all ongoing sessions by a scheduling algorithm which is implemented at the eNB. In this paper we assume a simple scheduler that makes optimal use of the available resources (i.e. SRs will only be left unused if there is no data to send). In case there are less SRs available than there are required, the scheduler divides the available SRs among the users, proportional to the amount of SRs they require. The number of SRs a user needs depends on its location. This means that when a user moves

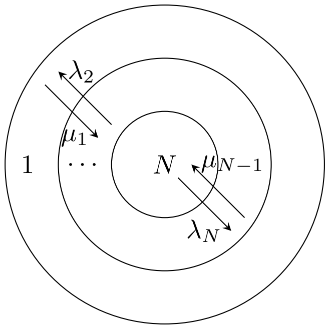

Figure 2. Different zones with transition rates.

around in a cell, the number of SRs that are needed to send a certain amount of bits towards that user will change.

The total amount of SRs that are needed for all active users can be calculated as follows. Suppose r_k is the required bitrate by user k and b_k is the amount of bits that can be sent towards that user in a single SR. The average number of SRs per second (n_k) that are needed by user k is given by: $n_k = \frac{r_k}{b_k}$. The total amount of SRs that are needed in order to service all users is: $R = \sum_k n_k = \sum_k \frac{r_k}{b_k}$. Note that, using the scheduler described above, the cell capacity can be estimated as follows. In case the required number of SRs per second is smaller than the available amount of SRs per second ($R < R_A$), the unused SRs are compensated for by multipglying with the term $\frac{R_A}{R}$. In this case each user will be given the bit rate it requires and the cell capacity is defined as $C = \frac{R_A}{R} \sum_k r_k$. In case there are insufficient SRs the scheduler will divide the available amount of resources among the users, proportionally to the number of SRs they require. I.e. the amount of SRs a user k receives will be $\frac{n_k}{R} R_A$ and therefore the cell capacity is given by $C = \sum_k \frac{R_A}{R} b_k n_k = \frac{R_A}{R} \sum_k r_k$. So, in either case, the cell capacity is given by: $C = \frac{R_A}{R} \sum_k r_k$.

III. General Model

In order to model that the cell capacity varies depending on the location of the users we consider a single cell that is divided into N concentric rings called zones, where zone 1 is the outermost zone (see Fig. 2). Denote by ρ_i the bits transmitted per SR to a user in zone i (with $\rho_1 < \rho_2 < \cdots < \rho_N$). The cell has R_A SRs available per second and an active user requires a bitrate of r bit per second to fulfil its QoS requirements. Therefore, the number of SRs per second that a user needs to achieve its required bitrate r in zone i is: $s_i = \frac{r}{\rho_i}$.

For evaluating the performance of the models, two performance measurements are defined. These are the total blocking probability (P_T) and the low QoS probability (P_{QoS}). The total blocking probability is the probability that a session that arrives is blocked by the AC policy. The low QoS probability is the fraction of time that active users experience a low QoS, i.e. the bitrate they receive is lower than r.

A. Admission Control Policy

In order to guarantee a minimum QoS (i.e. a minimum bitrate of r) for the users that are in the system, the acceptance of new arrivals is controlled by an AC policy. Let $f \in [0, 1]$ denote the AC threshold that determines which fraction of resources can be used for new users. A new session is accepted if after accepting the session there would still be more than $(1 - f)R_A$ SRs available.

Therefore, to decide on the acceptance of a new session in zone i, the following decisions are taken:

$$R + s_i \begin{cases} \leq fR_A & \text{session accepted} \\ > fR_A & \text{session rejected} \end{cases} \tag{1}$$

Once a user is admitted to the system we assume that it cannot be removed before its session ends. If a user makes a transition to an exterior zone, it can happen that more SRs than there are available are needed ($R > R_A$) since users need more SRs in the destination zone than in the zone where they came from to maintain their bitrate. In this case, the scheduler will give all the users a share of the available SRs that is proportional to the amount they requested, as is mentioned in Section II-B. So all users will receive a lower bitrate than the required r, hence all users will be served with a lower QoS.

IV. Analytical Model

We model the proposed system using a multidimensional continuous-time Markov chain (CTMC), where the system state vector is described by the N-tuple $\boldsymbol{x} = (x_1, \ldots, x_N)$, where x_i represents the number of users in zone i. Since the number of bits that can be transmitted per SR is the highest for a user in zone N, the maximum number of sessions M that can be present in the system is determined by the maximum number of sessions accepted in this zone when there are no active users in the other zones. The set of feasible states is thus given by:

$$W := \left\{ \boldsymbol{x} : x_i \in \mathbb{N}; \quad \sum_{i=1}^{N} x_i s_N \leq fR_A \right\}.$$

The total number of SRs that are needed per second to serve all the users at the required bitrate r when the system is in state $\boldsymbol{x}$ is represented by $R(\boldsymbol{x}) = \sum_{i=1}^{N} x_i s_i$.

For the sake of mathematical tractability we make the common assumptions that new calls arrive according to a Poisson arrival process with arrival rate ϵ and exponentially distributed session durations with rate γ. Assuming uniformly distributed traffic, if A_T is the total area of the cell and A_i is the area corresponding to zone i (i.e. $A_T = \sum_i A_i$), we can consider that the arrival rate for new sessions in zone i is $\epsilon_i = \frac{A_i}{A_T}\epsilon$. We checked this assumption with simulations for the random walk mobility model described in Section V.

The zone residence time, i.e the time that a user stays in a certain zone before entering another one, is also assumed to be exponentially distributed with rate λ_i for transitions from zone i to zone $i-1$ where $i = 2, \ldots, N$ and μ_i for transitions from zone i to zone $i+1$ where $i = 1, \ldots, N-1$ (see Fig. 2).

We consider that the user starts and finishes its session inside the cell.

The function $a_i(\boldsymbol{x})$ denotes whether a session that arrives in zone i when the system is in state $\boldsymbol{x}$ is accepted by the AC algorithm ($a_i(\boldsymbol{x}) = 1$ means that the session is accepted and $a_i(\boldsymbol{x}) = 0$ means that the session is blocked).

Fig. 3 shows a CTMC for $N = 2$ as an example. The notation has been simplified as $a_i(\boldsymbol{x}) = a_i$. This system has only two zones $i = 1, 2$ and therefore one incoming rate μ_1 and one outgoing rate λ_2. Recall that M is the maximum number of users that can be active in the system. If we define phases as the number of users in zone $i = 2$ and levels as the number of users in zone $i = 1$, we can study the model as a finite level-dependent quasi-birth-and-death (QBD) process [9] with $M + 1$ levels where level h ($h = 0, \ldots, M$) has $M + 1 - h$ phases. Therefore we can construct the transition rate matrix with a block design, see (2). The first row of blocks corresponds to level $h = 0$, the second row of blocks to level $h = 1$, etc., where blocks $\boldsymbol{Q_1^h}$ correspond to transitions among phases in level h, blocks $\boldsymbol{Q_0^h}$ to transitions from level h to level $h+1$ and blocks $\boldsymbol{Q_2^h}$ to transitions from level h to level $h-1$. Note that blocks have different sizes for different levels h. The total size of the transition rate matrix for $N = 2$ is $\frac{M^2+3M+2}{2} \times \frac{M^2+3M+2}{2}$.

$$\boldsymbol{Q} = \begin{bmatrix} \boldsymbol{Q_1^0} & \boldsymbol{Q_0^0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \ldots \\ \boldsymbol{Q_2^1} & \boldsymbol{Q_1^1} & \boldsymbol{Q_0^1} & \boldsymbol{0} & \boldsymbol{0} & \ldots \\ \boldsymbol{0} & \boldsymbol{Q_2^2} & \boldsymbol{Q_1^2} & \boldsymbol{Q_0^2} & \boldsymbol{0} & \ldots \\ & & \ddots & \ddots & \ddots & \\ \ldots & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{Q_2^{M-1}} & \boldsymbol{Q_1^{M-1}} & \boldsymbol{Q_0^{M-1}} \\ \ldots & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{Q_2^M} & \boldsymbol{Q_1^M} \end{bmatrix} \quad (2)$$

In the case of $N = 3$ zones, we can define phases as the number of users in zone $i = 3$, low-levels as the number of users in zone $i = 2$ and high-levels as the number of users in zone $i = 1$. The model is a three dimensional finite QBD process where the transition rate matrix $\boldsymbol{Q}$ again follows the structure of (2). Moreover the block matrices $\boldsymbol{Q}_0^h$, $\boldsymbol{Q}_1^h$ and $\boldsymbol{Q}_2^h$ are also constructed with a block design, see (3), (4) and (5). The total size of the transition rate matrix is $\frac{2M^3+12M^2+22M+12}{12} \times \frac{2M^3+12M^2+22M+12}{12}$.

$$\boldsymbol{Q}_1^h = \begin{bmatrix} \boldsymbol{A_1^{h,0}} & \boldsymbol{A_0^{h,0}} & \boldsymbol{0} & \ldots \\ \boldsymbol{A_2^{h,1}} & \boldsymbol{A_1^{h,1}} & \boldsymbol{A_0^{h,1}} & \ldots \\ \ddots & \ddots & \ddots & \\ \ldots & \boldsymbol{0} & \boldsymbol{A_2^{h,M-h}} & \boldsymbol{A_1^{h,M-h}} \end{bmatrix} \quad (3)$$

$$\boldsymbol{Q}_0^h = \begin{bmatrix} \boldsymbol{B_1^{h,0}} & \boldsymbol{0} & \boldsymbol{0} & \ldots \\ \boldsymbol{B_2^{h,1}} & \boldsymbol{B_1^{h,1}} & \boldsymbol{0} & \ldots \\ \ddots & \ddots & & \\ \ldots & \boldsymbol{0} & \boldsymbol{B_2^{h,M-h}} & \boldsymbol{B_1^{h,M-h}} \end{bmatrix} \quad (4)$$

$$\boldsymbol{Q}_2^h = \begin{bmatrix} \boldsymbol{C_1^{h,0}} & \boldsymbol{C_0^{h,0}} & \boldsymbol{0} & \ldots \\ \boldsymbol{0} & \boldsymbol{C_1^{h,1}} & \boldsymbol{C_0^{h,1}} & \ldots \\ & \ddots & \ddots & \\ \ldots & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{C_1^{h,M-h}} \end{bmatrix} \quad (5)$$

Blocks $\boldsymbol{A_1^{h,l}}$ correspond to transitions among phases inside high-level h and low-level l, blocks $\boldsymbol{A_0^{h,l}}$ correspond to transitions from low-level l to low-level $l+1$ inside high-level h and blocks $\boldsymbol{A_2^{h,l}}$ correspond to transitions from low-level l to low-level $l-1$ inside high-level h. Blocks $\boldsymbol{B_1^{h,l}}$ correspond to transitions from high-level h to high-level $h+1$ with low-level l, blocks $\boldsymbol{B_2^{h,l}}$ correspond to transitions from high-level h to high-level $h+1$ and from low-level l to low-level $l-1$. Blocks $\boldsymbol{C_1^{h,l}}$ correspond to transitions from high-level h to high-level $h-1$ with the same low-level l and blocks $\boldsymbol{C_0^{h,l}}$ correspond to transitions from high-level h to high-level $h-1$ and from low-level l to low-level $l+1$.

This block design can be generalised to any number of zones N by constructing matrix blocks inside matrix blocks with $N-1$ different levels.

To solve the finite QBD Markov process and obtain the steady state vector ($\pi(\boldsymbol{x})$) we use the linear level reduction (LLC) algorithm [10]. Basically, the algorithm has two stages. First, the state space is reduced by removing one high-level at each step until there is a Markov process on the last high-level left. That Markov process is solved and the stationary vector is constructed in the second stage by adding back one high-level at each step. Note that despite the rate matrices being large, their sparseness makes the computations feasible. The complexity of the algorithm grows with N as the size of the transition matrix is $O(M^N)$.

Let us denote by P_i the blocking probability for new arrivals in zone i and the total blocking probability in the system by P_T. Then:

$$P_i = \sum_{\boldsymbol{x} \in W} (1 - a_i(\boldsymbol{x}))\pi(\boldsymbol{x}); \quad P_T = \frac{\sum_{i=1}^{N} \epsilon_i P_i}{\epsilon}.$$

Let $I(\boldsymbol{x})$ denote the indicator function which takes the value 1 when $R(\boldsymbol{x}) > R_A$ and otherwise it takes the value 0. The low QoS probability is then given by:

$$P_{QoS} = \sum_{\boldsymbol{x} \in W} I(\boldsymbol{x})\pi(\boldsymbol{x}).$$

V. Simulation model

In order to verify the mobility and session duration modelling assumptions made in the analytical model and to verify the results obtained with the analytical model, simulations that model the mobility and the session duration of the users more realistically are performed. In the simulations, sessions are generated according to a Poisson process with arrival rate ϵ. The duration of the session is, in contrary to the analytical model, chosen from a log-normal distribution as this more realistically models the duration of sessions [11], the mean of

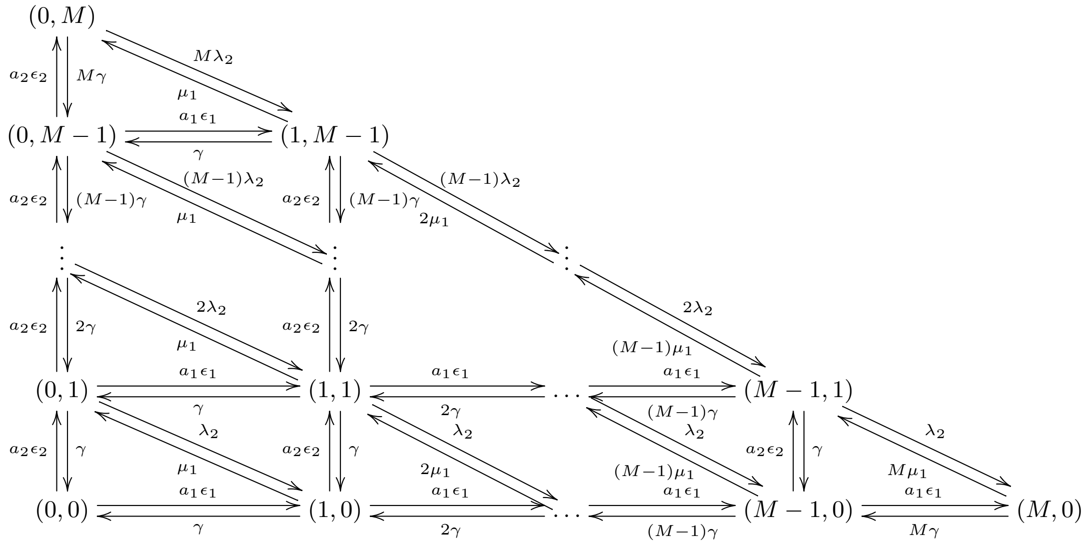

Figure 3. Transition diagram of the bi-dimensional model.

this log-normal distribution is $1/\gamma$ and the variance is $1/\gamma^2$, i.e. the log-normal parameters are chosen considering that this log-normal distribution has the same mean and variance as the exponential distribution considered in the analytical model. When a session is generated, it is placed uniformly in the cell and is subjected to AC. If it is admitted to the cell, it starts moving around. Users move around according to a random walk mobility model [12]. This means that when a session is started, it chooses a direction ϕ (in radians) uniformly distributed in the interval $[0, 2\pi[$ and starts moving in the chosen direction at a fixed velocity v. After the user has travelled over a fixed distance d, it again chooses a direction and starts moving in the newly chosen direction. This is repeated until the session finishes and the user is removed from the system. When a user reaches the boundary of the cell, it bounces against the circular edge and continues its path in the reflection direction. The simulator does not model the transitions between zones using exponential distrubutions. Instead, transitions between zones occur when a user crosses the border of a zone.

The call blocking probability (P_T) is calculated by counting the total number of generated sessions and the number of sessions that are rejected by AC and dividing the latter by the former. The low QoS probability is calculated by recording the time that the system has a low QoS and dividing it by the total simulation time.

VI. RESULTS

In this section we present the results of a comparative study between the analytical an the simulation models. Different scenarios were simulated for various settings of the parameters. In these simulations the the distance travelled by the users in a single leg of the mobility model (d), mean session inter-arrival time ($1/\epsilon$), mean session duration ($1/\gamma$) and AC threshold (f) were varied.

In the simulation users move around with a certain velocity and distance travelled in a single leg of the mobility model. The rates λ_i and μ_i that are used in the analytical model only depend on those parameters and the size of the zones. In order to determine these transition rates, simulations with only one user were executed. In these simulations a single user walks around without starting any sessions. Each time the user crosses the border of a zone, the event is recorded and at the end of the simulation, the mean transition rates are calculated.

A. Model parameters

The parameters that are fed into the analytical and simulation models are based on the evaluation scenarios described in [13]. The carrier frequency f_c is chosen to be 2 GHz, the pathloss model that is associated with this frequency is $L = 37.6 \log_{10}(D) + 128.1$ where D is the distance between the eNB and the UE. The operating bandwidth is 5 MHz which means that there are 25 subchannels of 180 kHz (plus guard band), resulting in 25000 SRs per second. We model the traffic as a fluid flow with a bitrate of 128 kbit/s. Unless otherwise indicated, we consider $v = 30$ km/h, $d = 30$ m $1/\gamma = 300$ s, $1/\epsilon = 3$ s and $f = 1$. The radii of the zones can be determined from the maximum bitrates in the zones. Using the attenuated Shannon bound [1], the minimum SIR that is needed to achieve the bitrate corresponding to a MCS can be calculated according to (6) where β_i is the bitrate per Hertz in zone i and α is the attenuation factor which is 0.6 [1]:

$$\mathrm{SIR}_i = 2^{\frac{\beta_i}{\alpha}} - 1. \quad (6)$$

Combining this equation with the pathloss model from [13], assuming no noise and assuming that for every direction the interference comes from a single source at a distance D_{s2s} from the SeNB with the same transmit power (P_{Tx}) as the SeNB, an expression for the SIR (in the logarithmic domain) at a given distance from the SeNB (D) can be constructed:

$$\begin{aligned}\mathrm{SIR} &= P_{\mathrm{Tx}} - L_S - (P_{\mathrm{Tx}} - L_I) = L_I - L_S \\ &= 37.6\log_{10}(D_{\mathrm{s2s}} - D) + 128.1 \\ &\quad - (37.6\log_{10}(D) + 128.1) \\ &= 37.6\log_{10}\left(\frac{D_{\mathrm{s2s}}}{D} - 1\right)\end{aligned} \tag{7}$$

where L_S is the pathloss of the signal and L_I the pathloss of the interference. The site-to-site distance D_{s2s} is set to 500 m, this inter-site distance is related to cells in urban environments. As can be seen from (7) the SIR at a given distance from the SeNB does not depend on the transmit power of the eNBs. From (7) the radius D_i of zone i, given the SIR can be calculated:

$$D_i = \frac{D_{\mathrm{s2s}}}{10^{\frac{\mathrm{SIR}_i}{37.6}} + 1}. \tag{8}$$

By combining (6) and (8) the radius of a zone given the bitrate of the MCS can be calculated.

Ideally, 15 zones should be considered, each corresponding to a single MCS. Since using all 15 different MCSs would produce too much computational overhead, the cell was instead divided in 3 zones with equal areas (i.e. the size of each zone is $\frac{1}{3}$ of the size of the cell), where the ρ_i of each zone is calculated considering the zones that correspond to the different MCSs. The cell border itself coincides with the circle on which the SIR is 0 dB. As the 0 dB border is in the middle between the SeNB and a NeNB, this border lies on $\frac{D_{\mathrm{s2s}}}{2} = 250$ m and it is inside the zone that corresponds to MCS 10. The number of bits per SR that is used in each of these zones is the weighted average number of bits per SR of the areas that overlap with the 3 different zones. The three different zones that are used in this paper are listed in Table II.

Table II
THE THREE DIFFERENT ZONES WITH THEIR PARAMETERS THAT ARE CONSIDERED IN THE SCENARIOS.

Zone	Bits/SR	Radius (m)
1	154.81	250
2	349.87	204.12
3	466.59	144.33

B. Numerical results

In this section numerical results obtained with both the analytical and simulation models are presented and compared. All simulation results have been obtained by running 10 simulations per point. Fig. 4 shows the blocking probabilities P_T while varying the session inter-arrival time ($1/\epsilon$). The results of the analytical model are represented using a solid line while the results of the simulations are represented using discrete crosses. Increasing the mean inter-arrival time will cause the blocking probability P_T to decrease. When the mean inter-arrival time is low, more sessions will be started at a shorter amount of time. As there are only a fixed number of resources available, this will cause more sessions to be blocked by the system resulting in a higher blocking probability. As can be seen in Fig. 4, the results obtained with the analytical model and the simulations are very similar.

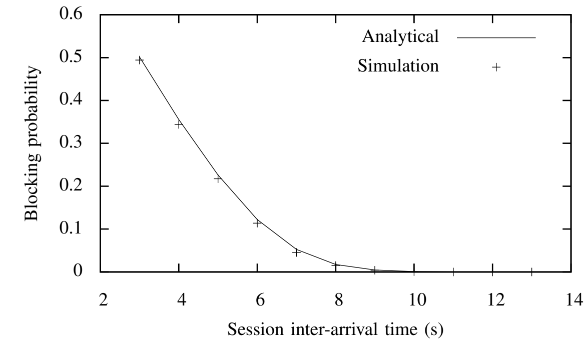

Figure 4. Blocking probabilities for various mean session inter-arrival times.

Also the low QoS probabilities P_{QoS} obtained with the analytical model and with the simulations are very similar as can be seen in Fig. 5. As with the blocking probabilities, the QoS is worse when the session inter-arrival time is low than when the session inter-arrival time is high. The reason that the probability of a low QoS is higher when the session inter-arrival time is low is a consequence of the varying cell capacity: as the AC threshold f is set to 1 in these results, the system will be filled up until all SRs are used, implying that whenever the variations in the cell capacity cause more resources to be needed, the system will no longer be able to fulfil the QoS needs of the users.

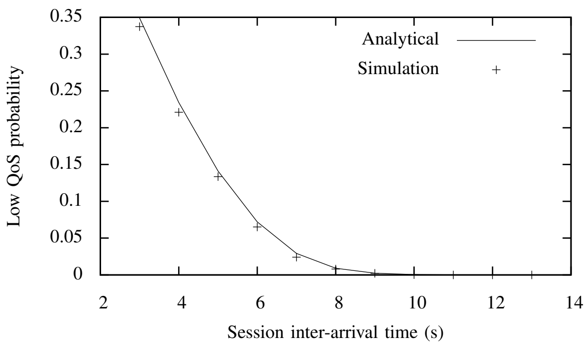

Figure 5. Low QoS probability for various mean session inter-arrival times.

Fig. 6 shows the total blocking probability when, instead of the mean inter-arrival time, the mean session duration is varied. When sessions are longer (high $1/\gamma$) more sessions will be blocked as accepted sessions will be active in the system (and thus require resources) for a longer time. For the same reason the low QoS probabilities are higher when the session duration is longer as can be seen in Fig. 7. As with the case in which the mean inter-arrival time is varied, the results of the analytical model and the simulations are very similar in both figures.

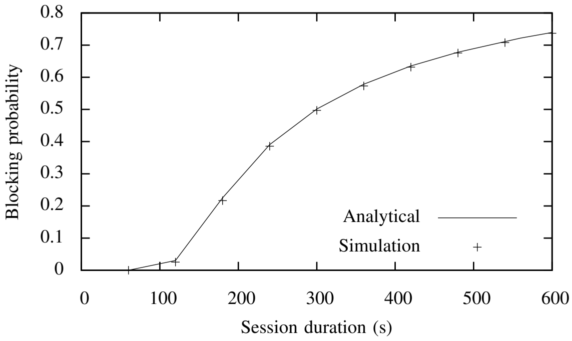

Figure 6. Blocking probabilities for various mean session durations.

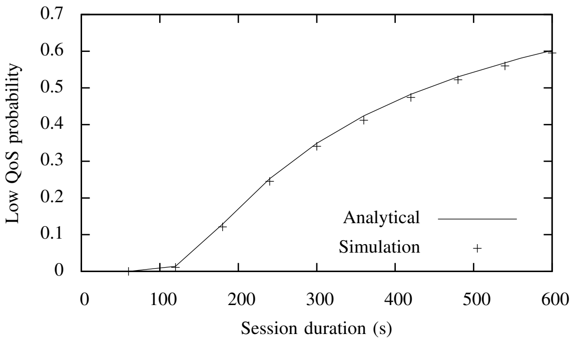

Figure 7. Low QoS probabilities for various mean session durations.

C. Deviations between the analytical model and the simulations

Although the results obtained with the analytical and simulation models fit very well, there are occasions where there are deviations between both models. One of the cases where the results of the analytical model and the simulations differ is when the distances travelled by the users in a single leg of the random walk mobility model (d) are either very short or very long in comparison to the radii of the zones. An example of this can be seen in Fig. 8. In this figure, the distance is varied between 1 m and 512 m (note that the cell radius is 250 m). As can be seen in this figure, the differences between the results obtained from the analytical model and the simulations are similar when the distance d lies between 10 m and 100 m. When d is either shorter or longer the deviations between both models become bigger. In the case of short d, this deviation is because users that are close to the border of a zone will cross the border of the zone many times while users that are further away from the border will most likely already have finished their session before they make a transition because the short distance causes them to remain at the same location. In the case of long d, this means that the distance that is travelled by the users is bigger than the radii of the zones and that the users will cross the zones more than once before choosing a new direction. This in turn has as a consequence that the time between entering a zone and leaving it again is bounded by the minimum and maximum distances that can be travelled in a zone in a straight line, divided by the velocity of the users and can no longer be either very long or very short as users will travel in a straight line to the zone and will not be able to change direction within a zone. This will have an influence on the distribution of the transition times. Fig. 9 shows the distribution of the time users spend in zone 2 before going to zone 1 ($1/\lambda_2$). The times are measured in simulations where the distance travelled in a single leg is 1 m, 32 m and 512 m. The plots also contain the probability density function (PDF) of the exponential distribution that is used in the analytical model to model the time that a user stays in that zone. As can be seen in the figures, the distribution of the times resembles the exponential distribution best when the travelled distance is around 30 m.

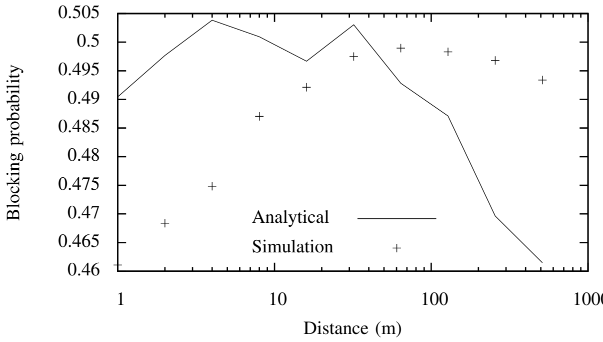

Figure 8. The blocking probabilities for various distances travelled by the users.

D. Analysis of the AC algorithm

In this section we briefly evaluate the AC algorithm that was used in this paper using the developed analytical and simulation models. Fig. 10 shows the blocking probabilities for various values of the AC threshold f. As can be seen in this figure the blocking probability decreases as the AC threshold increases. When the AC threshold increases, more sessions will be allowed to the cell, causing the blocking probability to decrease.

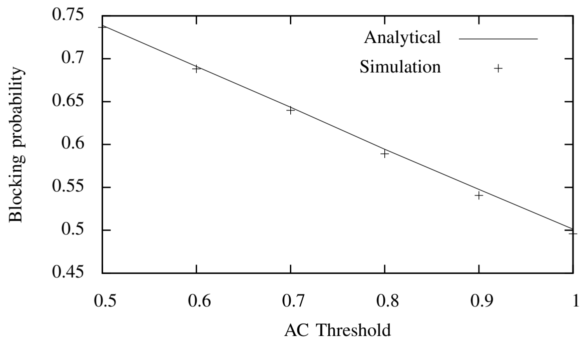

Figure 10. Blocking probabilities for various AC thresholds f.

When looking at the low QoS probability in Fig. 11 the QoS remains good until f reaches a value of more than 80 %. This shows the effects of the time-varying cell capacity on the QoS experienced by the active users.

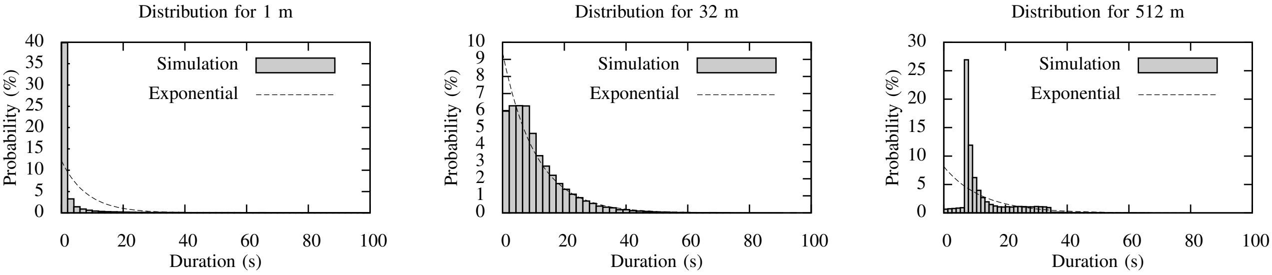

Figure 9. Distribution of the time spent in zone 2.

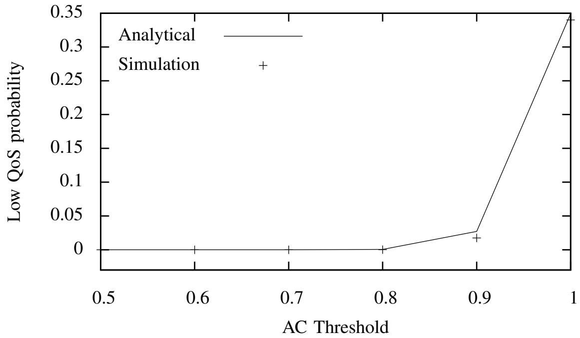

Figure 11. Low QoS probabilities for various AC thresholds f.

VII. Conclusions

In this paper an analytical model that models the time-varying cell capacity of cellular networks in general and LTE networks in particular was described. The cell capacity in these systems changes over time as the MCS that is used to send data to the users is changed depending on the signal quality. User movement causes the signal quality and thus the MCS to change over time. The time-varying cell capacity in this paper is modelled by dividing a cell in multiple concentric zones in which a certain bitrate can be achieved by the users in that zone. By assuming that the times between users changing from one zone to another are distributed according to an exponential distribution with a certain rate and the session duration is exponentially distributed, the system can be modelled using a CTMC. In order to verify these assumptions, results obtained with the analytical model in various scenarios were compared to simulation results that were obtained from a simulator that models the user mobility and the session duration more realistically. Results show that the analytical model models the user mobility very accurately and that the assumption of exponentially distributed session duration is also accurate. The analytical model can be used to study the impact of a varying cell capacity on the QoS experienced by the user and for system design issues such as resource dimensioning.

A. Future work

The goal of this paper was to compare the results of the analytical model that was developed to more realistic simulations in order to verify whether the mobility modelling and the exponentially distributed session duration assumptions that were made in the analytical model are accurate. The developed analytical model can now be used to further investigate AC in cellular networks: other AC algorithms can be implemented in the analytical model and the performance of the implemented algorithms can be compared. The model can also be used to investigate the optimisation of the parameters of the AC algorithm like was done in [7] using simulations. The model can also be extended by introducing handovers.

References

[1] *3GPP TR 36.942: Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Radio Frequency (RF) system scenarios*, 3GPP, September 2010.

[2] R. Ramjee, R. Nagarajan, and D. Towsley, "On optimal call admission control in cellular networks," *Wireless Networks Journal (WINET)*, vol. 3, no. 1, pp. 29–41, 1997.

[3] D. Garcia, J. Martinez, and V. Pla, "Admission control policies in multiservice cellular networks: Optimum configuration and sensitivity," *Lecture Notes in Computer Science*, vol. Wireless Systems and Mobility in Next Generation Internet, Gabriele Kotsis and Otto Spaniol (eds.), no. 3427, pp. 121–135, 2005.

[4] D. K. Kim, D. Griffith, and N. Golmie, "A novel ring-based performance analysis for call admission control in wireless networks," *IEEE Communications Letters*, vol. 14, no. 4, pp. 324–326, 2010.

[5] S. E. Elayoubi and T. Chahed, "Admission control in the downlink of WCDMA/UMTS," *Mobile and Wireless Systems*, vol. LNCS 3427, no. 5, pp. 136–151, 2005.

[6] S. S. Jeong, J. A. Han, and W. S. Jeon, "Adaptive connection admission control scheme for high data rate mobile networks," in *Vehicular Technology Conference, 2005. VTC-2005-Fall. 2005 IEEE 62nd*, vol. 4, 2005, pp. 2607 – 2611.

[7] K. Spaey, B. Sas, and C. Blondia, "Self-optimising call admission control for LTE downlink," in *COST 2100 TD(10)10056, Joint Workshop COST 2100 SWG 3.1 & FP7-ICT-SOCRATES, Athens, Greece*, February 2010.

[8] *3GPP TR 36.213: Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Physical layer procedures*, 3GPP, June 2010.

[9] M. Neuts, *Matrix-geometric Solutions in Stochastic Models: An Algorithmic Approach*. The Johns Hopkins University Press, 1981.

[10] G. Latouche and V. Ramaswami, *Introduction to Matrix Analytic Methods in Stochastic Modeling*. ASA-SIAM, 1999.

[11] J. Guo, F. Liu, and Z. Zhu, "Estimate the call duration distribution parameters in GSM system based on K-L divergence method," in *International Conference on Wireless Communications, Networking and Mobile Computing, Shangai, China*, September 2007, pp. 2988 – 2991.

[12] T. Camp, J. Boleng, and V. Davies, "A survey of mobility models for ad hoc network research," *Wireless Communications and Mobile Computing*, vol. 2, no. 5, pp. 483–502, 2002.

[13] *NGMN Radio Access Performance Evaluation Methodology*, NGMN, January 2008.