Proceedings of the 2016 Winter Simulation Conference
T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, eds.

USING SIMULATION TO EVALUATE LTE LOAD CONTROL FOR PRIORITY USERS

Brittany L. Biagi
Nassissie Fekadu
David A. Garbin
Steven P. Gordon
Denise M. Masi

Noblis, Inc.
3150 Fairview Park Drive South
Falls Church, VA 22042, USA

ABSTRACT

Disasters can cause extraordinary service demand by the public. It is imperative that services supporting disaster response perform with minimal degradation during such events. In order to provide adequate service to special users such as first responders, priority treatment mechanisms have to be developed. Priority treatments have been incorporated for earlier wireless technologies, but have to be established on Long-term Evolution (LTE) / 4G. One of the proposed priority-treatment concepts is Access Class Barring (ACB), which will shed traffic from public users in response to extreme overloads, resulting in priority for special users. However, the degree to which ACB would improve voice call completion is unknown. A discrete-event simulation was performed to model extreme overload situations and predict the performance of ACB under various configurations. The simulation study found that ACB could drastically improve the priority call completion probability in the most extreme overloads while maintaining performance for public traffic.

1 INTRODUCTION

There is a need to provide priority access to commercial wireless networks users, particularly Government representatives and first responders, during local, regional and national emergencies. These events often cause congestion from excessive loads on wireless networks while concurrently causing outages that reduce network capacity to serve the surging demand. This prevents both public and priority users from reliably making voice calls. In order to provide adequate service to special users like first responders, priority treatment mechanisms have to be developed. Priority treatments have been incorporated for special users for earlier wireless technologies such as 2G and the Universal Mobile Telecommunications System (UMTS) / 3G. However, priority mechanisms have to be established on Long-term Evolution (LTE) / 4G.

Previous studies such as Lien et al. (2012), Phuyal et al. (2012), and Jian et al. (2013) have addressed a concept called Access Class Barring (ACB) in LTE which could be a potential priority mechanism. However, the scope of these studies is limited to machine-to-machine (M2M) communications and does not address priority users. Gordon et al. (2014) developed models to evaluate the use of ACB to provide priority treatment for 3G cellular networks, namely UMTS, and found that ACB could drastically improve the priority call completion probability. However, there are significant differences between the 3G and 4G technologies, and there is a need for modeling of LTE access networks with ACB to evaluate the ability of this treatment to enable call success for priority users.

This paper proceeds to describe the LTE system of interest for modeling, the approaches to providing priority treatment on LTE (including ACB), the structure of our simulation model, modeling scenarios of interest, and modeling results.

1.1 System of Interest

Figure 1 shows a generic LTE architecture. The two main components of LTE networks are the Evolved UMTS Terrestrial Radio Access Network (E-UTRAN) and the Evolved Packet Core Network (EPC). The E-UTRAN has two main elements – the User Equipment (UE) (mobile user's smart phone) and the eNodeB. The eNodeB is a logical node with antennas that are often mounted on towers and may be in plain view of mobile users. The eNodeB provides all radio resource management functions supporting the UE over the radio link for the Radio Admission Control for admission of voice call, data session, and video session admittance decisions, the Radio Bearer Control, the Dynamic Resource Allocation and the Packet Scheduling. The EPC contains the Mobility Management Entity (MME) for control plane signaling and the Serving Gateway (S-GW) for user plane data; this is the gateway to the public switched telephone network (PSTN). Our focus for modeling is from the UE to the eNodeB. We model the uplink, as we assume that it is the constrained channel. We assume that the Random Access Channel (RACH) is not a potential congestion point.

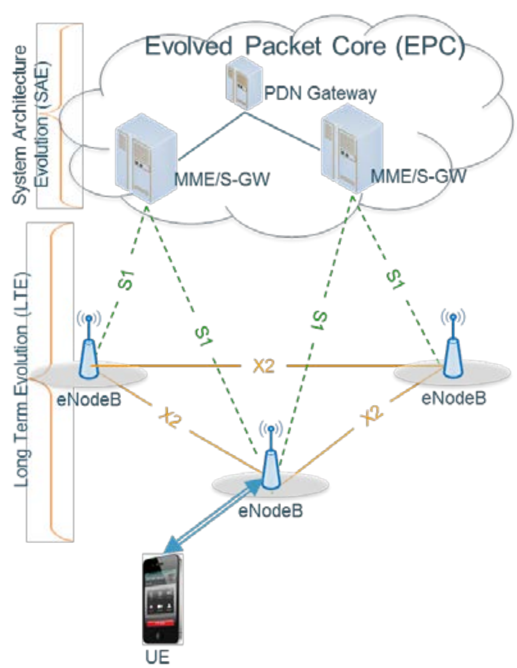

Figure 1: LTE architecture.

2 PRIORITY TREATMENT

2.1 Call Admission Control Description

The Call Admission Controller (CAC) is the mechanism in the eNodeB that determines if a call request can be admitted to the network or not, depending on whether bandwidth resources are available to provide the required call quality of service. The admission control algorithm implemented in this analysis is based on the work in Spaey et al. (2010). We assume that there is a five percent margin of bandwidth that is not utilized to provide a noise margin. A buffer provides bandwidth for priority calls, 911 calls, and calls handing over from another cell sector, sufficient for ten calls ("ThHO" threshold). If a priority, 911, or handover call arrives, and finds that the total bandwidth utilized (including the new arrival) exceeds ThHO, then the call is blocked. If a Public voice or video Guaranteed Bit Rate (GBR) call arrives, and finds that the total bandwidth utilized (including the new arrival) exceeds the ThGBRu threshold, the call is blocked. The ThHO threshold is higher than the ThGBRu threshold, providing some priority for the handover, priority, and 911 calls upon call admission. Lastly, if a Public data (Non-GBR) call arrives, and ThGBR_inc kbps is not available, then the data call is blocked. In addition to these bandwidth-based thresholds for call admission control decisions, there are also limitations on the number of connected UEs using the eNodeB, the number of bearers in use in the eNodeB, and the number of concurrent voice calls associated with the eNodeB. The use of the higher ThHO threshold to provide bandwidth for priority, 911, and handover calls does result in performance benefits for these calls, but additional priority treatments such as ACB may be needed at extreme overload levels.

2.2 Access Class Barring Description

Access Class Barring selectively limits the load offered by the public user equipment (UEs) while preferentially allowing priority users. The ACB mechanism is specified in the LTE standards (3GPP). However, ACB is a highly configurable feature, and a simulation analysis is required to analyze the expected performance under various configuration options and parameter settings, and provide recommendations for implementing and allowing for dynamic updates of ACB parameters in response to changing levels of system congestion. ACB is generally configured such that (1) it engages only during extreme overloads, and (2) it limits public traffic only. In order to detect and react to the traffic level where the cell is experiencing an overload, ACB is often configured to trigger off failures from the Call Admission Controller.

We have implemented a dynamic ACB algorithm in our simulation model, whereby the degree of barring that is imposed is modified dynamically based on system congestion levels (Figure 2). The eNodeB broadcasts ACB parameters which include the degree of barring (ac_barringfactor) and a wait time parameter (ac_waittime) which is used to compute the amount of time before a barred UE can retry. The eNodeB regularly checks the trigger versus the threshold values to determine if ACB should be imposed or decreased, and will adjust the degree of barring upon seeing multiple consecutive positive triggers. After adjusting the barring factor a hold period of 30 seconds is imposed on the ACB system, preventing further barring during that time, which allows the traffic conditions to settle and prevents excess barring. As each UE attempts arrival, a determination is made as to whether the UE should be barred based on the UE access class and the access class barring factor parameter (ac_barringfactor). If the UE is barred, the wait time parameter (ac_waittime) is used to compute the time the UE must wait until it can re-attempt arrival. While ACB is engaged, each UE may attempt arrival up to 3 times before the call is considered barred and is ultimately dropped.

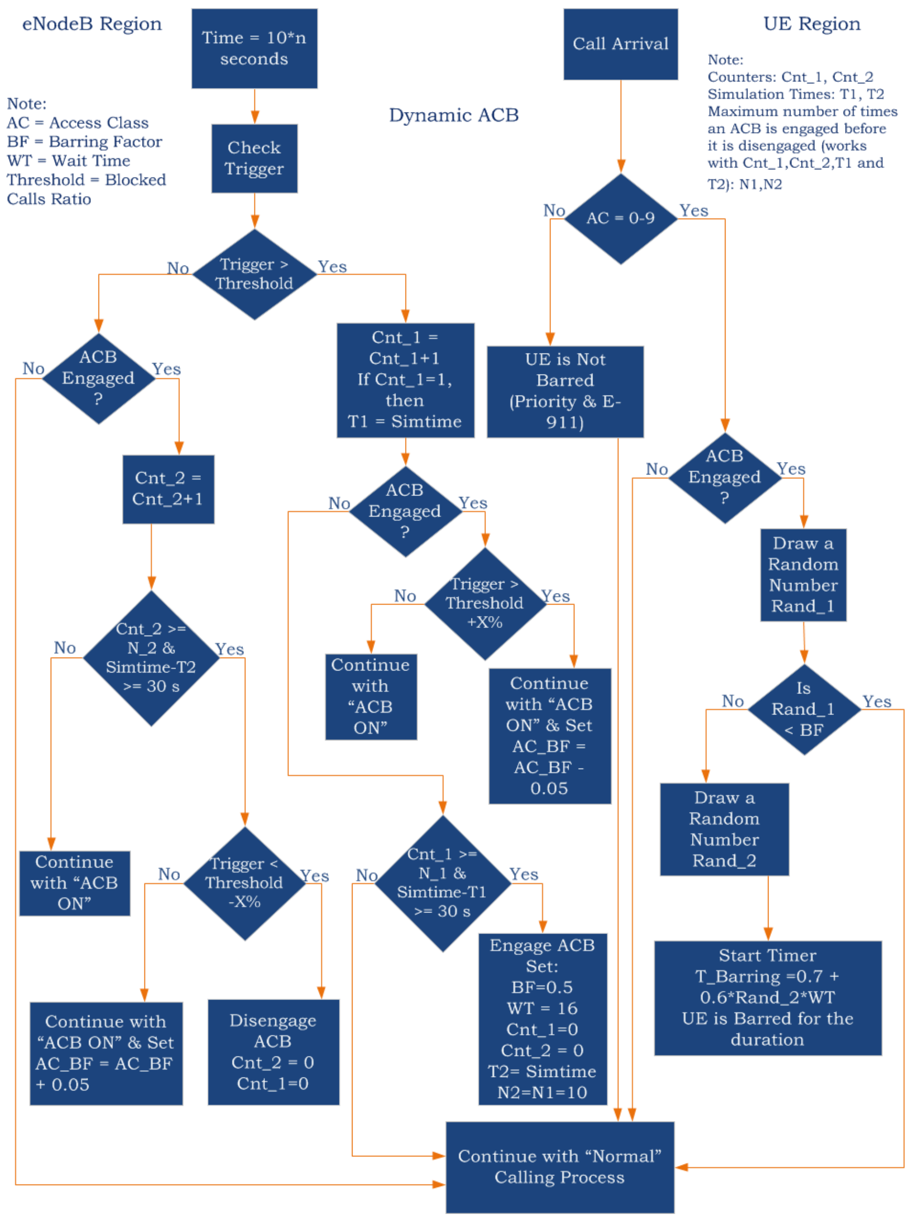

Figure 2: Access Class Barring Algorithm.

3 SIMULATION TOOL STRUCTURE

A customized discrete event simulator was created for this study using the object-oriented Xojo programming environment. Figure 3 illustrates the simulator input screens. Six traffic streams, which can be independently specified, are modeled on a call-by-call basis. The interarrival and service times are exponentially distributed. The model assumes that each UE can handle only one call type at a time. The user is able to specify numerous input parameters in the tool GUI, including parameters related to the run execution and output analysis, management of the UEs including limitations imposed by the eNodeB and UE registration, traffic parameters for each stream, specification of scale factors for the traffic surges in run scenarios, parameters on the Call Admission Controller module and Access Class Barring module. The simulator provides both steady-state and transient results for each traffic stream, in terms of the call served ratio and other output measures.

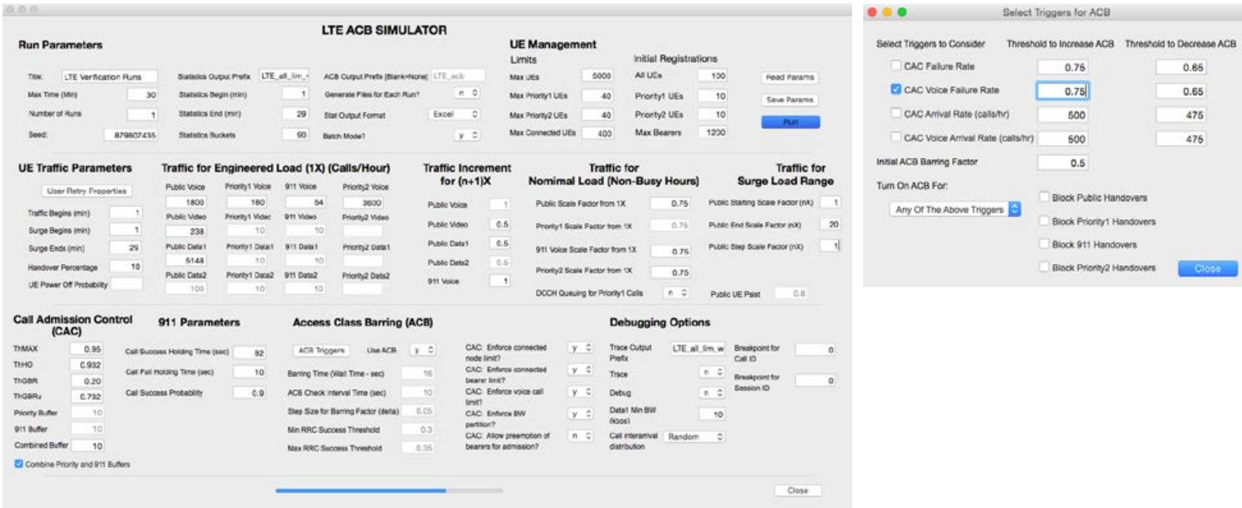

Figure 3: Simulation GUI Input Screens.

4 LOAD AND PRIORITY TREATMENT SCENARIOS

4.1 Traffic Profile

The simulation model represented a large population of UEs in a single cell with an eNodeB, running the CAC and the ACB as described above. The traffic that the UEs generate is a mix of public voice calls, two streams of priority voice calls, emergency 911 calls, public video calls and public data sessions (Table 1). Voice (except e911 calls) and video sessions are assumed to have a mean hold time (call length) of 150 seconds; e911 calls are assumed to have a shorter mean hold time. Data sessions are assumed to have a total mean session holding time of 300 seconds. The duty cycle and the hold time specify the on/off nature of the traffic pattern. So, for example, Data1 will be in the active part of the duty cycle for 10 seconds, then inactive for 40 seconds. This active/inactive cycle repeats for a total of 300 seconds. The baseline engineered load (“1X”) traffic arrival rates are also given in Table 1. Engineered load was assumed to be the offered traffic load that produced approximately 2% blocking probability in 10 MHz cell.

Table 1: Baseline (“1X”) Traffic Model.

Parameter	Voice (public + priority + emergency)	Video	Data (web, email, file transfer, video clip)
Engineered Load (1X) (calls/hr.)	Public: 1800; Priority1: 180; Priority2: 3600; E911: 54	238	5148
Hold time (seconds)	Public and Priority: 150 E911: 82	150	300
Duty cycle	1	1	0.2
Rate (kbps)	15.6	128	10 (min for CAC admission)

The traffic is conceptualized as having arrival rates with different means during each of three periods. An initial period of *nominal* load (some fraction of the 1X load) represents the traffic level before the event of interest. A subsequent period of *surge* load represents the traffic during the event of interest. Lastly, a final period of *nominal* load is used. The surge load is an extended traffic surge for 20 minutes of simulated time such as an earthquake or Boston Marathon bombing, when the traffic quickly reaches a high overload and stays high for an extended period of time. This extended traffic surge is modeled at various overload levels, to examine performance with increasing arrival rates from 1X up to 20X. The Public Voice load scales exactly with overload, that is, the 5X load has 5 times the 1X load of Public Voice. We assume that video and data streams have a different (smaller) scaling factors, due to expected user behavior during emergency situations. While e911 voice load scales with overload, the priority voice traffic streams have fixed traffic levels regardless of the overload level.

4.2 Modeling Scenarios

We consider two main scenarios of interest: (1) ACB Off, which is a baseline against which ACB will be compared, and (2) ACB On. For both scenarios, the performance is estimated by separately executing the simulation at increasing load levels as described above. One hundred model replications are performed for each load level. During the ACB Off scenario, the priority voice streams do have some priority over Public Voice calls, due to the CAC buffer which provides reserved bandwidth for up to ten priority, e911, and handover calls. However, during extreme overload events, this CAC buffer may not provide the required performance for priority calls, and thus ACB is evaluated as an additional priority mechanism.

The system performance is evaluated primarily in terms of the call served ratio. Call served ratios of 0.9 - 0.95 for Priority1 voice calls are deemed to be acceptable. The use of ACB is deemed to be beneficial if it increases the Priority1 call served ratio to acceptable levels, while any decreases to the Public Voice call served ratio are minimal. Any decreases in Public Video and Data1 served ratios are deemed acceptable considering the nature of the event being simulated. Speed of response to the priority treatment after the onset of the event is another measure of interest.

4.3 ACB Parameters

As mentioned previously, ACB is highly configurable, with flexibility as to the trigger mechanism, parameter values, and dynamic behavior of these parameters. This analysis examines performance for two triggers which will engage ACB:

- CAC overall failure rate—measures blocking of *all* calls
- CAC voice failure rate—measures blocking of *voice* calls

Parameter values utilized are given in Table 2. Sensitivity of the performance of ACB to the engage threshold value, which is the point at which ACB turns on or increases in severity, and the threshold to decrease ACB severity were examined. We utilize the notation “Voice x% y%” (and similarly for “overall” CAC failures) to indicate:

- Engage/increase barring severity when CAC voice failure rate > y%
- Decrease barring severity when CAC voice failure rate < x%
- ACB continues at current severity when x% < CAC voice failure rate < y%

Table 2: ACB Parameters.

Parameter	Parameter Value
Engage ACB threshold	5%, 10%, 20%, 40%
Threshold to decrease ACB severity	0.5%, 1%, 2%, 4%
Barring Time (Wait Time)	16 seconds
ACB check interval time	10 seconds
Step size of Barring Factor	0.05
Initial ACB Barring Factor	0.5
Number of iterations to decrease ACB severity	N = 3
Hold time at each ACB level before another change	30 seconds

5 RESULTS

Steady-state and transient modeling results are given below.

5.1.1 Steady-State Results

Figure 4 provides a summary of system performance in terms of the average steady-state call served ratio for each traffic stream for the ACB Off scenario as well as six ACB Trigger scenarios. The primary objective for comparison is obtaining the highest served ratio for Priority1 Voice traffic, with minimum impact to the Public Voice traffic. The served ratio displayed in Figure 4 is an average after steady state is achieved (between t=500s and t=1200s) and across traffic levels 1X through 20X. Figure 4 shows that performance with ACB Off in terms of Priority1 served ratio is only about 80 percent. With all of the ACB scenarios in Figure 4, Priority1 served ratio is significantly improved, to between about 0.9 and 0.95 on average. The Voice 2% 20%, Voice 4% 40% and Overall 2% 20% trigger combinations result in the best performance for Priority1, while also providing a slight increase for the Public Voice call served ratio.

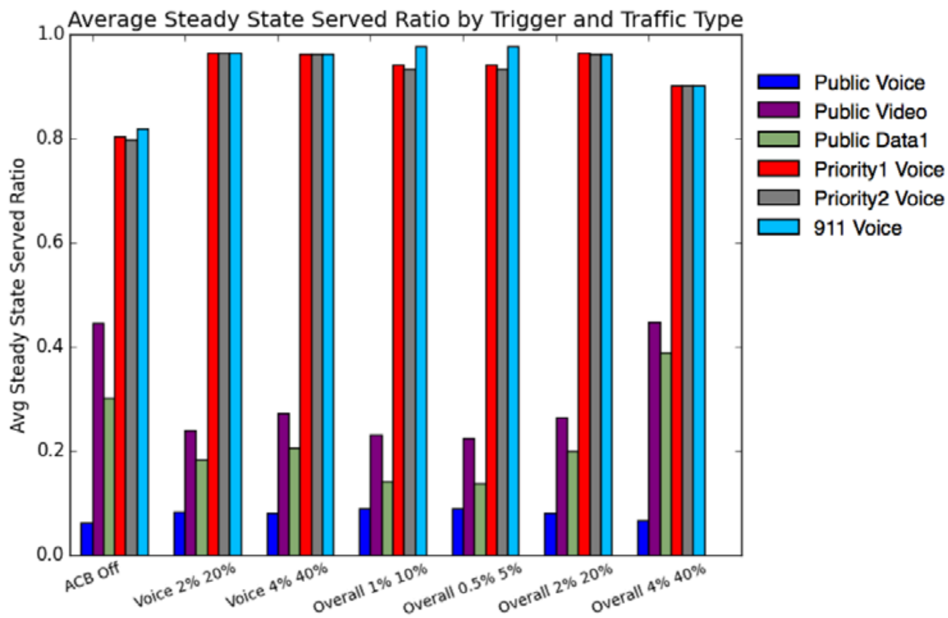

Figure 4: Steady-state ACB Trigger Comparison Summary.

Figure 5 examines the trigger performance of the Priority1 stream for traffic from 1X to 20X, showing the served ratio for each trigger versus the traffic overload factor. The call served ratio is highest for the Voice 2% 20%, Voice 4% 40%, and Overall 2% 20% trigger combinations, which are circled in Figure 5. These trigger combinations yield a Priority1 served ratio of about 0.95 for 10X traffic level. Other trigger combinations and ACB Off have significantly worse performance. Results are similar for the Priority2 traffic class.

Overall, the use of ACB changes the traffic mix being carried by the eNodeB. Figure 6 shows the served ratio at each traffic level for each trigger tested for the Public Voice traffic. The Public Voice call served ratio increases slightly when ACB is used versus ACB Off, which is a good result. The Public video served ratio (not shown) decreases substantially because the higher resources required for video are not available due to the resources being used instead for the Priority1 voice traffic. There is also some residual GBR bandwidth used for Public Voice rather than video, thus resulting in the slight increase in calls served for Public Voice.

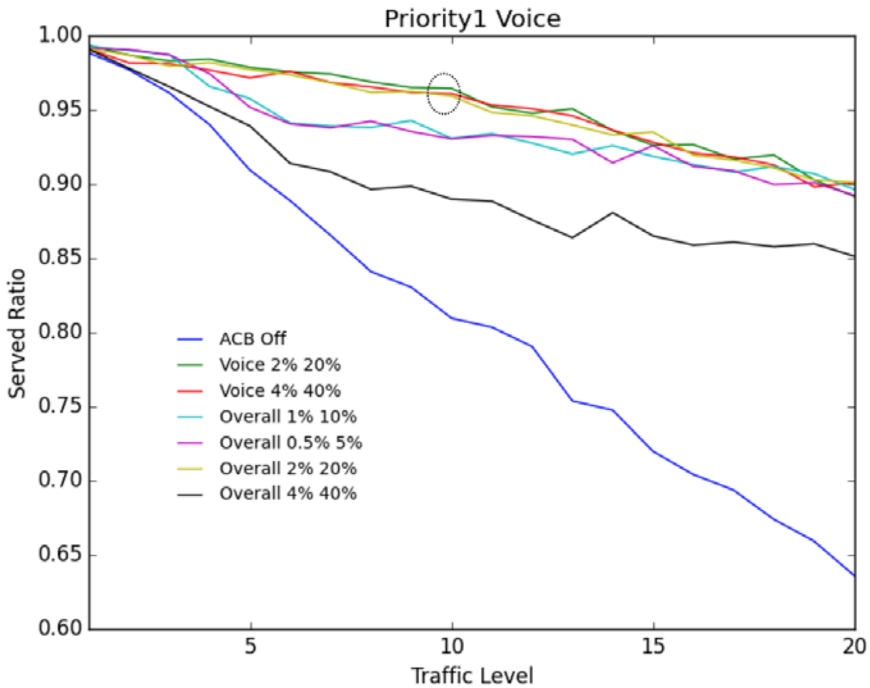

Figure 5: Steady-state ACB Trigger Comparison for Priority1 Traffic.

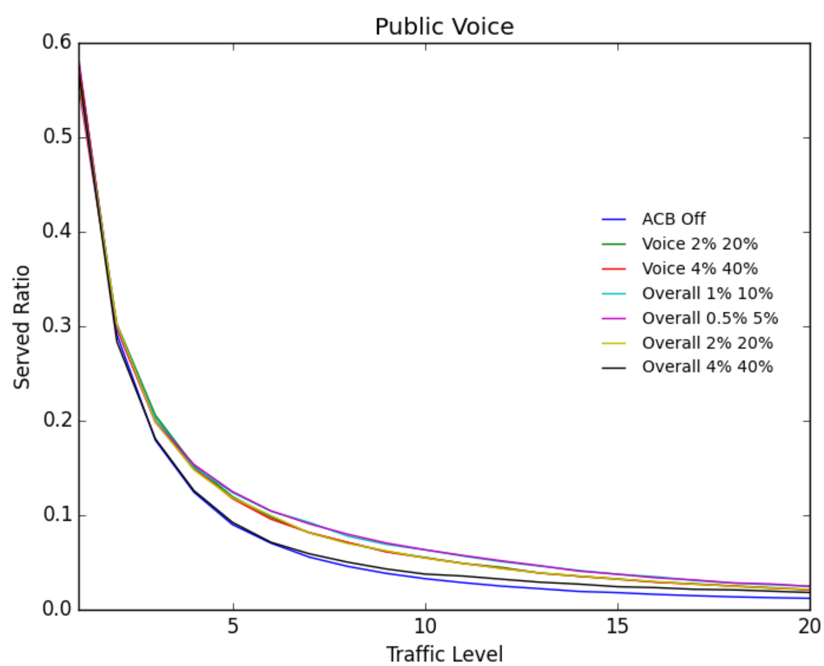

Figure 6: Steady-state ACB Trigger Comparison for Public Voice Traffic.

5.1.2 Transient Results

Transient simulation results immediately after the surge load begins are next discussed. It is desired that the system responds quickly to priority treatments such as ACB, and without oscillatory behavior. Figure 7 shows the call served ratio over time for the Priority1 traffic stream, using the Voice 2% 20% ACB Trigger, for various traffic loads. The traffic surge is imposed at 60 seconds. ACB responds quickly and engages at between 90 and 450 seconds, depending on traffic load. The Priority1 served ratio rises shortly after ACB engages at about 90 seconds, reaching steady state between 140 and 220 seconds after surge begins.

Figure 8 shows the progression of the barring factor over time across multiple traffic levels. The barring factor is adjusted dynamically per the algorithm shown in Figure 2. This factor varies between 1 and 0, and smaller values indicate greater degrees of barring. The initial barring factor after ACB is engaged is 0.5. The degree of barring increases (barring factor decreases) after initially being engaged, and the maximum degree of barring is reached at about 400 seconds.

Figure 9 similarly shows the call served ratio over time for the Public Voice traffic stream, using the Voice 2% 20% ACB Trigger, for various traffic loads. The Priority1 served ratio rises shortly after ACB engages at about 90 seconds, reaching steady state between 140 and 220 seconds after surge begins.

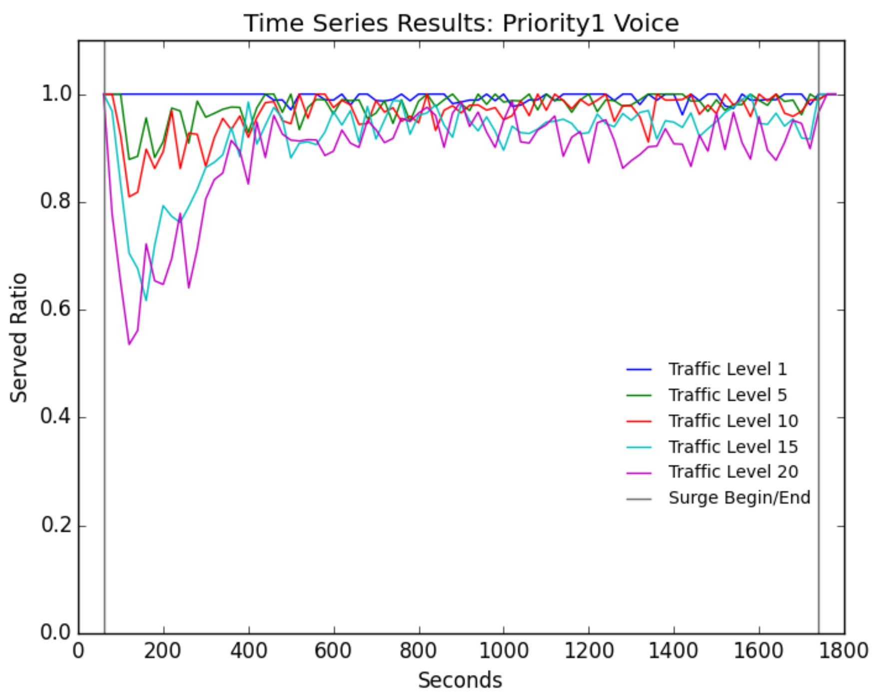

Figure 7: Transient Served Ratio for Priority1 Traffic, for Voice 2% 20% Trigger.

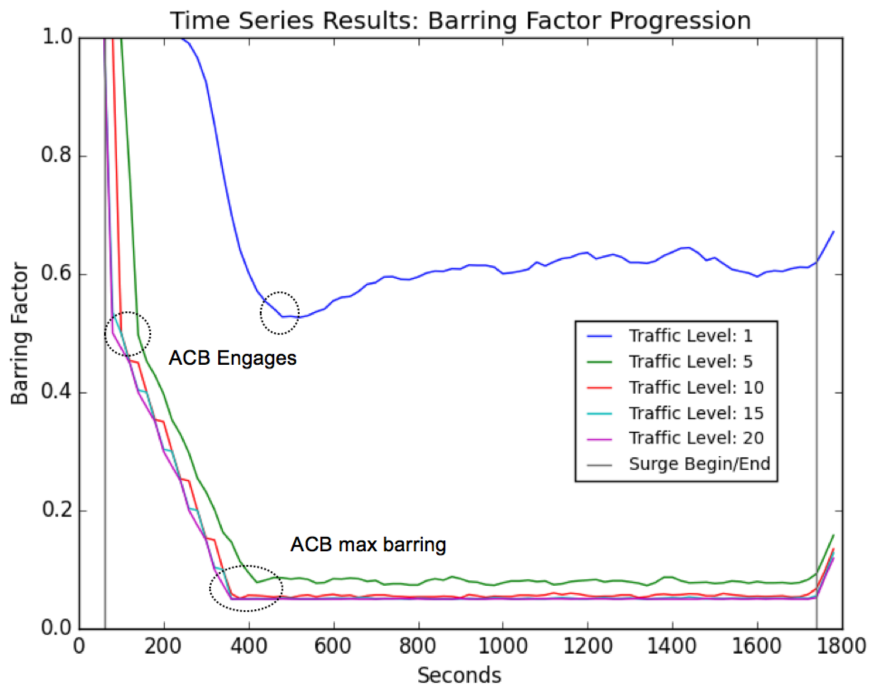

Figure 8: Barring Factor Progression over Time, for Voice 2% 20% Trigger.

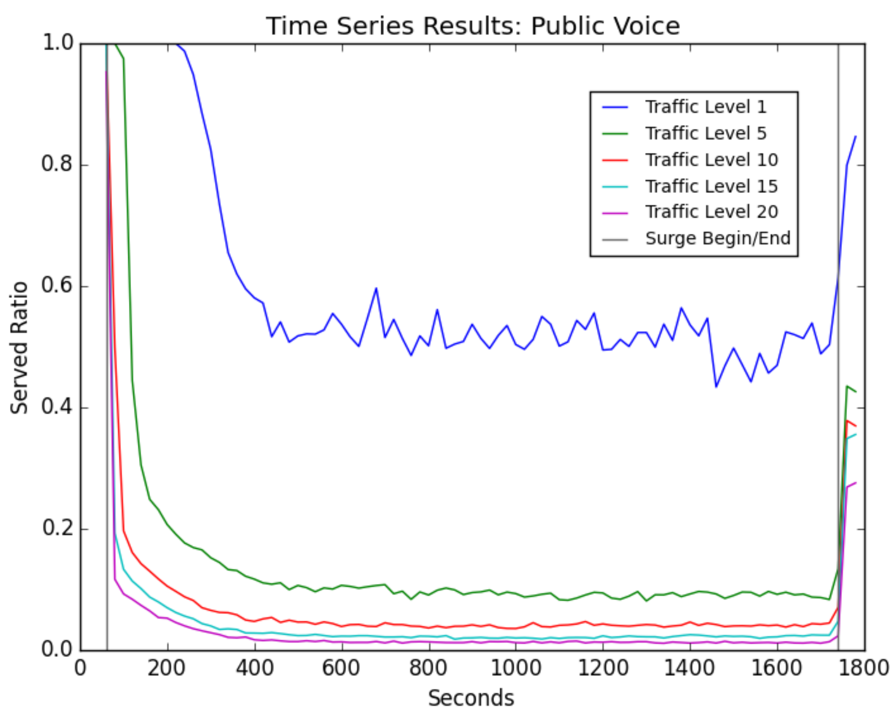

Figure 9: Transient Served Ratio for Public Voice Traffic, for Voice 2% 20% Trigger.

6 CONCLUSIONS

In summary, Access Class Barring can be an effective method of providing priority access to commercial wireless networks with LTE. It takes a fully-utilized eNodeB and reallocates the way it is using its resources. Conclusions that can be derived from this study are the following:

- ACB gives significant improvement to the priority voice service.
- ACB has minimal impact on the Public Voice traffic that is transported during extreme events, and in fact can slightly improve Public Voice traffic performance, as a result of the reallocation of the resources from the other public streams.
- ACB for priority voice can be improved with configuration changes. This modeling has shown that some ACB trigger combinations provide greater benefit to priority traffic than other triggers. Three of the trigger combinations tested yield similar results and are recommended over the three remaining trigger combinations that were simulated.

Next steps are to explore the sensitivity of performance to additional ranges of ACB trigger combinations and other ACB parameters, such as the barring time parameter and initial ACB barring factor. In addition, the authors hope to have the opportunity to validate the simulation results with laboratory testing.

ACKNOWLEDGMENTS

We would like to thank Richard Kaczmarek and Janet Gunn of CSC Government Solutions who provided valuable technical input for this work.

REFERENCES

3GPP Technical Specification Group Radio Access Network E-UTRA Radio Resource Control (RRC) Protocol specification, http://www.3gpp.org/dynareport/36331.htm.

Gordon, S., D. Garbin, D. Masi and P. McGregor. 2014. "UMTS Load Control with Access Class Barring." In *Proceedings of the 2014 IEEE Military Communications Conference (MILCOM)*, 1009-1014. Baltimore, MD: Institute of Electrical and Electronics Engineers, Inc.

Jian, X., Y. Jia, X. Zeng and J. Yang. 2013. "A Novel Class-dependent Back-off Scheme for Machine Type Communication in LTE systems." *Wireless and Optical Communication Conference (WOCC)*, 135-140. Chongqing.

Lien, S. Y., T. H. Liau, C. Y. Kao and K. C. Chen. 2012. "Cooperative Access Class Barring for Machine-to-Machine Communications." *IEEE Transactions on Wireless Communications* 11(1):27-32.

Phuyal, U., A. T. Koc, M. H. Fong and R. Vannithamby. 2012. "Controlling Access Overload and Signaling Congestion in M2M Networks." *Asilomar Conference on Signals, Systems and Computers*, 591-595. Pacific Grove, CA.

Spaey, K., B. Sas, and C. Blondia. 2010. "Self-Optimising Call Admission Control for LTE Downlink." *Joint Workshop COST 2100 SWG 3.1 & FP7-ICT-SOCRATES*, Athens, Greece.

AUTHOR BIOGRAPHIES

BRITTANY L. BIAGI is a Data Scientist at Noblis. She is a graduate of The College of William and Mary where she received a B.S. in Computer Science and a B.S. in Mathematics, concentrating in Computational Mathematics. Her experience and research interests include modeling and simulation, statistical analysis, and machine learning methods. Her email address is brittany.biagi@noblis.org.

NASSISSIE FEKADU is a Lead wireless engineer at Noblis. Her experience and research interests include antennas, modeling and simulation and RF communications. Ms. Fekadu holds an M.S. in Astronomy from San Diego State University and a B.S. in Electrical Engineering from Virginia Tech. Her email address is nassissie.fekadu@noblis.org.

DAVID A. GARBIN is a Senior Fellow at Noblis. His 40+ years of experience is in the telecommunications and networking field, focusing on the design and economic analysis of large networks, both for carriers and their customers. He holds advanced degrees from MIT and is the co-author of the best-selling New McGraw-Hill *Telecom Factbook*. His email address is david.garbin@noblis.org.

STEVEN P. GORDON is a Principal at Noblis. He has 24 years in the area of Telecommunications, including 9 years with Telecommunications Service Providers. His areas of expertise include network architecture, network design, capacity planning, capacity management and performance analysis of large networks. Mr. Gordon holds a Master of Science in Systems Engineering from Case Western Reserve University. His email address is steven.gordon@noblis.org.

DENISE M. MASI is a Fellow at Noblis. Her experience and research interests include queueing theory and simulation applied to telecommunications networks, along with cybersecurity data analytics. Dr. Masi received her B.S. in Industrial Engineering from Texas A&M University and an M.S. in Industrial Engineering from Purdue University. She received her Ph.D. in operations research at George Mason University. Her email address is denise.masi@noblis.org.