Cellular-Based Machine-to-Machine: Overload Control

Adlen Ksentini and Yassine Hadjadj-Aoul, University of Rennes1
Tarik Taleb, NEC Europe

Abstract

One of the most important problems posed by cellular-based machine type communications is congestion. Congestion concerns all the parts of the network, both the radio and the core network impacting both the user data plane and the control plane. In this article, we address the problem of congestion in machine type communications. We propose a congestion-aware admission control solution that selectively rejects signaling messages from MTC devices at the radio access network following a probability that is set based on a proportional integrative derivative controller reflecting the congestion level of a relevant core network node. We evaluate the performance of our proposed solution using computer simulations. The obtained results are encouraging. In fact, we succeed in reducing the amount of signaling while maintaining a target utilization ratio of resources in the core network.

Cellular-based machine-to-machine (M2M) or machine type communications (MTC) are about enabling automated applications that involve machine or device communication without any human intervention over cellular networks. MTC will enable an endless number of applications in a wide range of domains impacting different environments and markets. It will connect a huge number of MTC devices to the Internet and the networks, forming the so-called Internet of Things (IoT). Depending on the use case, an MTC device transmits or receives a determined amount of data at a determined frequency (e.g., a smart meter sending measurement results every day at 23:00h). MTC devices can be either fix installed (implemented in a factory's machine, gas meters, etc.) or mobile (e.g., fleet management devices in trucks). The heterogeneity of the features of MTC devices (i.e., in terms of mobility, amount of transmitted data, security, etc.) enables mobile operators to make different optimizations for grouping MTC devices, charging MTC applications, and controlling network resources used by the MTC devices. The communication network domain can be a wired or wireless network. In this article, our focus is on the case when the communication network is a mobile/cellular network. Indeed, mobile cellular communications have the advantage of easier installation and provisioning, especially for short-term deployments. Besides, cellular mobile networks offer different network technologies and can support a wide range of MTC devices, including those with mobility features. Recent predictions indicate that MTC over cellular mobile networks could be a leverage to provide much needed mobile operator revenue. Indeed, several forecasts state a significant market growth, over the next years, for both the MTC device and the MTC connectivity segments. The growth is expected at a compound annual growth rate (CAGR) exceeding 25 percent [1–3]. According to these forecasts, billions of machines or industrial devices will be potentially able to benefit from MTC.

In order to take advantage of the potential opportunities raised by a global MTC market over cellular networks, the third generation partnership project (3GPP) SA2 group is defining 3GPP network and system improvements that support MTC in the evolved packet system (EPS) [4]. The objective of these studies is to identify 3GPP network enhancements required to support a large number of MTC devices in the 3GPP network domain and to provide necessary network enablers for MTC services. Specifically, transport services for MTC as provided by the 3GPP system and the related optimizations are being considered as well as aspects needed to ensure that MTC devices and/or MTC servers and/or MTC applications do not cause network congestion or system overload.

In this article, we focus on the problem of congestion and system overload when enabling MTC in cellular networks. Cellular networks are mainly designed to handle human-to-human (H2H), machine-to-human (M2H), and human-to-machine (H2M) communications, whereby the proportion of the uplink (UL) traffic is lower in comparison to the downlink (DL) traffic. In contrast, MTC applications communicate without any external intervention, and involve a potential number of cost-effective and low-power devices that generate more signaling than the effective data over the UL channel. Congestion that may occur due to simultaneous signaling messages from MTC devices can be significant as they may lead to peak load situations and may have a tremendous impact on the operations of a mobile network. In the context of MTC, signaling congestion may happen due to a malfunction in the MTC server (e.g., MTC devices rapidly retrying to connect to the remote server which is down) or application (e.g., synchronized recurrences of a particular MTC application) and/or due to massive attempts from a potential number of MTC devices to attach/connect to the network all at once [5]. A straightforward solution to signaling congestion can be in the form of designing MTC applications that are friendly to mobile network operators. However, mobile operators cannot risk the operations of their networks and the quality of the provided services by leaving the whole signaling congestion problem to MTC application developers.

To address the problem of system congestion in cellular-based MTC communication, we propose a novel solution, congestion-aware admission control (CAAC). In CAAC, MTC signaling traffic is rejected at the radio access network (RAN), that is, at eNodeBs, with a probability p that represents the level of congestion of the relevant EPS nodes such as a mobility management entity (MME) as well as the serving gateway (S-GW). The more severe the congestion at the EPS nodes, the higher the reject probability p becomes. An important feature of the proposed scheme consists of the fact that it is based on control theory. In fact, control theory is used:

- To model the impact of the MTC signaling traffic on the queue length at the MME/S-GW
- To use the proportional integrative derivative (PID) controller [6] to avoid MME/S-GW overload by deriving the reject probability p to be used by the RAN part (eNodeB) for admission control of MTC traffic

The remainder of this article is organized as follows. We provide some background on MTC congestion management. We present a general and detailed description of our proposed solution, CAAC. The proposed solution is then evaluated through simulations. Finally, we highlight some future research work and conclude this article.

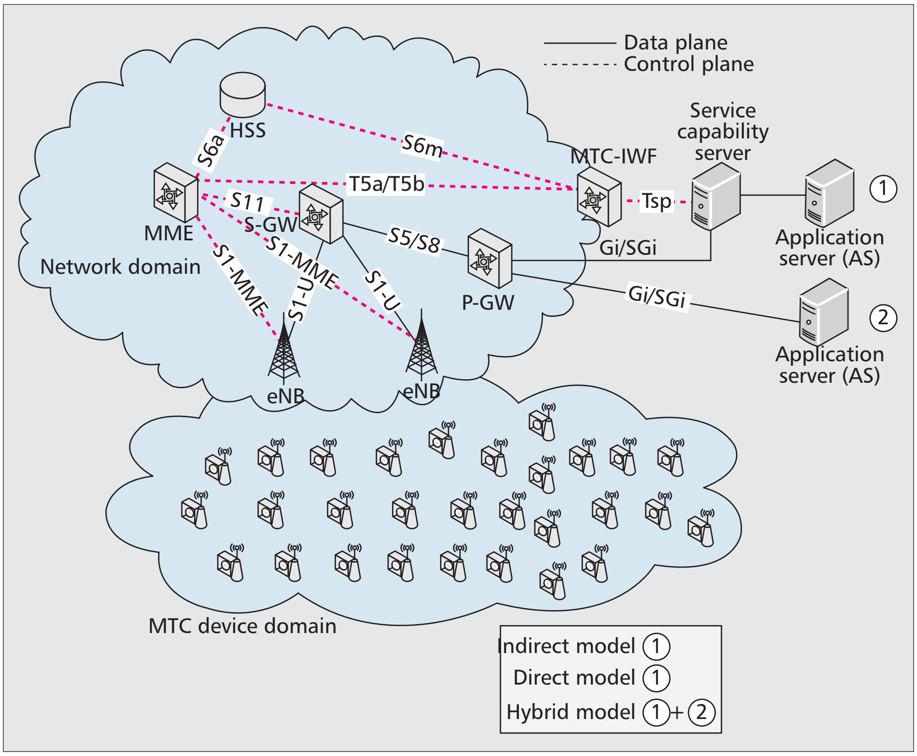

Figure 1. *A typical MTC network architecture.*

Background and Related Work

Figure 1 depicts a typical MTC network architecture, as currently envisioned by 3GPP. It consists of three main domains: the MTC device domain, the communication network domain, and the MTC application domain. In the network domain, most important nodes of a 3GPP EPS network are shown. Table 1 provides a brief description of the most important EPS nodes shown in Fig. 1. The MTC application domain consists of MTC servers, under the control of the mobile network operator or a third party. As per recent agreements within the 3GPP SA group, a new entity, called the MTC interworking function (MTC-IWF), is introduced to interwork between the mobile operator network and the MTC servers via a services capability server (SCS) using a new control interface, Tsp.

The vast majority of signaling congestion avoidance and overload control build on two main assumptions:

- The grouping of MTC devices based on different metrics/features (e.g., low-priority access, low mobility, online/offline small data transmission)
- The allocation, for each MTC device group, of forbidden/grant times, based on the subscription features of its MTC devices in HSS

As detailed by the authors in [7], congestion due to MTC can be handled by selectively rejecting connection/attach requests from MTC devices at specific network nodes. This operation targets only MTC traffic, particularly related to MTC applications that are causing congestion, and shall have no impact on non-MTC traffic. Rejection of connection/attach requests shall be done while ensuring that a rejected MTC device does not immediately reinitiate the same request (e.g., only until after a predetermined backoff time) and that the rejected MTC devices do not attempt connecting to the network all at the same time, but rather at randomized times. In other words, the network indicates to the concerned MTC devices a backoff time and ensures an even distribution of future incoming attach requests. Such randomization of MTC accesses can be triggered at an MTC device or by the network including the MTC server. When it is triggered by the network, the operation of an MTC device starts after receiving paging from the network or application level data from the server. In the case of the former, the operation of the MTC device can start immediately after being paged or by adding some delay since being paged. The rejection can be done based on specific MTC group identifiers and/or by looking at MTC traffic dedicated for a specific access point network (APN).

Network attachment requests can be rejected either at (or near) the RAN (e.g., eNodeB) or by the MME. Where the MTC access is controlled by the RAN, whenever the network is about to get congested by MTC applications (i.e., following congestion status feedback from PDN GWs/SGWs), the MME sends a notification message to RAN nodes triggering MTC access control indicating MTC barring information (barring factor, MTC group to block, barring time, etc.). In order to avoid random access channel (RACH) overload, work in [8] introduced different directions to control MTC access at the RAN level. In fact, a RACH random access request is issued when user equipment (UE) turns on or performs handover between eNodeBs or loose synchronization with the UL frame. Usually, RACH resources are dimensioned for low and

Node	Description
eNB	Evolved Node B, the LTE's base station.
MME	Mobility management entity, a control plane entity for all mobility related functions, paging, authentication, bearer management in the EPS.
MTC-IWF	MTC interworking function hides the internal public land mobile network (PLMN) topology and relays or translates signaling protocols used over Tsp to invoke specific functionality in the PLMN.
HSS	Home subscriber server, main database containing subscription-related information.
S-GW	Local mobility anchor for intra-3GPP handoffs.
P-GW	Packet data network gateway, interfaces with the packet data network (e.g., Internet).

Table 1. *EPS's most important nodes.*

medium traffic load, which is not always the case for MTC. Therefore, the collision probability increases for both MTC and non-MTC traffic. The main propositions introduced in [8] are:

Separate RACH resources for MTC and non-MTC devices: In case of congestion due to MTC traffic, the separation of RACH resources between MTC and non-MTC devices allows the limitation of the number of MTC devices with the possibility to enter the network, while maintaining normal network access for non-MTC traffic.

Dynamic allocation of RACH resources: This solution could be complementary to the precedent one in case the network is aware of the period of time the MTC devices have to transmit. In fact, the network can dynamically increase the RACH resources dedicated to MTC devices if it is aware of the scheduling of MTC traffic. In some scenarios the network can predict when access load will surge due to MTC devices.

MTC-specific backoff scheme: Since the non-MTC traffic has higher priority than MTC traffic, giving a large backoff window for MTC devices increases the access probability for non-MTC devices. With this mechanism, the access attempts from MTC devices could be dispersed in a large time interval to prevent contending for the RACH resources with UEs. Hence, this solution guarantees network availability for UEs.

Access class barring (ACB): Another solution to reduce RACH overload is to use the access class barring concept. ACB is a solution that effectively reduces the collision probability of transmitting the bulk of preambles at the same RACH resource. Based on the broadcasted parameters, the UE determines whether it is temporarily barred from accessing the cell. An ACB factor or access probability (p) determines the probability that access is allowed. If a random number n generated by the UE is equal to or greater than p, access is barred for a mean access barring time duration. In the legacy ACB scheme, there are 16 access classes. AC 0–9 represents normal UEs, AC 10 represents an emergency call, and AC 11–15 represents specific high-priority services, such as security services and public utilities (e.g., water/gas suppliers). User equipment may be assigned one or more access classes depending on the particular cell access restriction scheme. In [8], the ACB is introduced for MTC communication, where a higher value of p and *access class barring duration* could be assigned to MTC devices in order to reduce the contention on RACH resources, since the MTC devices will likely be blocked by the small probability p.

Using network simulation, the work in [9] evaluated and compared the performance of ACB and the backoff procedure for reducing RACH contention. The findings of this work showed that using fixed ACB parameters for MTC devices and UEs is not optimal, as high values of p and access barring duration increase the access delay of MTC devices. In contrast, a low access probability p and a short barring duration increase the contention on RACH resources under heavy traffic load. Accordingly, the authors proposed using adaptive ACB parameters, by allowing eNodeBs to periodically adjust these values based on the system load. In case of high loads, higher values of p and the access barring duration are assigned to the MTC devices, whereas under low loads these values are decreased. The obtained results clearly show that adaptive A-ACB can increase the system performances when it operates under high load while reducing the MTC access latency in case of low loads. In addition, the work shows that A-ACB exhibits better performance compared to the Adaptive Backoff procedure and ACB with fixed values. The authors also point out that ACB parameters can be updated by sending SIB (system information block) that carries physical channel information (such as Random Access Channel information, Random Access parameters Hybrid Adaptive ReQuest (ARQ)). However, SIB broadcasting cycle is around a second to several minutes. Therefore, the adaptive ACB parameter assignment can not immediately reflect the system load and may cause MTC access delay. Another work that addressed MTC congestion at the RAN level is presented in [10]. Based on QoS characteristics and requirements, in terms of the cluster packet arrival rate and the maximum tolerable jitter by the MTC composing the cluster, the authors proposed the grouping of MTC devices into M clusters. When a MTC device attempts attaching to the network, it sends its QoS characteristics and requirements to the current eNodeB. If there are enough resources to satisfy the MTC requirements, the MTC device is definitely accepted and attached to an existing MTC cluster having the same constraints, or a new cluster is created. Otherwise, the attach request of the MTC device is rejected.

In case the MTC access is controlled by MME, HSS may provision MME with information on grant time and forbidden time intervals as part of MTC subscription. Based on this feedback and also on local operator policies, MME then determines authorized times for each MTC device and communicates them to the respective MTC devices via MTC server or by NAS (Non-Access-Stratum) signaling directly from MME [11]. In case of congestion occurrence during the authorized times, MME may reject connections from concerned MTC devices and provide them with back off times for later accesses, or simply send them a congestion notification message triggering them to reduce their data transmission rate. It should be noted that the latter incurs major impact on the MME implementation. Intuitively, MME-based access control works under the assumption that there are still some signaling resources available to receive and reject attach/connect requests. The RAN-based access control is more efficient as it bars the specific congesting MTC group/APN from attempting access and thus prevents MTC devices from congesting the NAS signaling further.

In summary, one of the advantages of the RAN-based solution consists in the fact that there is no wastage in signaling from MTC devices that need to be blocked at first place. It also assists in controlling overload of both RAN nodes and core network nodes. These benefits come at the price of additional implementations in terms of broadcasting barring information to RAN nodes from the network. One aspect, which was missing so far in MTC signaling congestion control consists in the lack of a mechanism that handles a bulk of similar

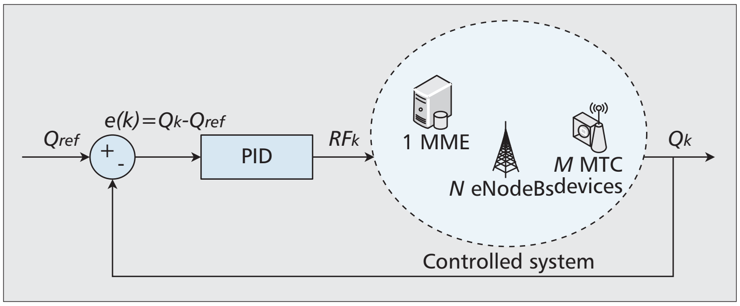

Figure 2. *The envisioned PID-controller based system.*

signaling messages from MTCs in a single shot (i.e., bulk MTC signaling handling). In [7], the authors show the potential of handling signaling messages common to a group of MTC devices in bulk, as a complementary or alternative approach to the above-mentioned solutions. Indeed, overload/congestion can also be mitigated by handling signaling messages, common to a number of MTC devices, by means of bulk processing. This signaling bulk can be handled by different protocols (e.g., S1 Application Part "S1-AP," Diameter, and GPRS Tunneling Protocol for the Control Plane "GTP- C").

Proposed Solution: Congestion-Aware Admission Control

Similar in spirit to the access class barring solution, we propose the congestion-aware admission control (CAAC) scheme. In CAAC, MTC devices are grouped according to their priority classes. In case of congestion, each class is blocked with a probability p as in ACB. In contrast to ACB, where there is no indication on how this probability is computed, in CAAC, the reject probability p is derived for each class by the relevant EPS nodes (e.g., MME, S-GW, and P-GW) being under congestion. Based on this probability, each eNodeB (at the RAN level) accepts/rejects MTC traffic belonging to a specific class.

In this article, we particularly focus on congestion that may happen at an MME. Such congestion can be directly induced from the delay in processing incoming packets at the application layer. This delay may lead to high and variable latencies at the input buffer, which may induce buffer overflow (i.e., packet loss). In contrast, congestion at EPS nodes, particularly P-GWs, concerns the link layer, which may affect the outgoing buffer length. In CAAC, MMEs control their level of congestion by adjusting (through the reject probability p) the amount of incoming signaling traffic from the MTC devices. If the congestion is detected, the probability p is set to a high value. Otherwise, this probability remains low. This probability is then communicated through dedicated signaling packets, or incorporated in existing ones, to the relevant eNodeBs that use it as part of their admission control operation.

To achieve the above-mentioned goal, CAAC uses PID controller, a well-known controller in the field of classical control theory [6]. Each MME implements an independent PID controller that uses only locally available information to derive the probability p in order to maintain the MME's queue around the optimal value that avoids both system overloads and underutilization at the same time. It shall be noted that lots of research work in the literature has indicated that the queue length and queue length variation represent a good indication to quantify the severity of congestion [12]; thus, our usage of queue length for congestion assessment in CAAC.

Figure 2 depicts the envisioned control system using a PID controller. The system involves all elements impacting its length, including MTC devices, eNodeBs attached to an MME and the MME itself. The PID controller takes as input the error signal $e(k)$ that represents the difference between the current queue length and the queue reference (Q_{ref}) value (i.e., the threshold above which the MME is considered overloaded), and gives as output the reject factor (RF), which represents the amount of traffic to reject in order to reduce the error signal. The principal objective is to retrieve the optimal value of the RF (reject probabilities) considering two metrics in conflict, for instance, the minimization of the queuing delay (i.e., small buffers) and the maximization of the throughput (i.e., avoiding underflows). Note that the loop represents the action of rejecting MTC traffic through the admission control operation enforced at eNodeBs. The envisioned system is based on a PID controller. We therefore use PID control function [6] in order to express the relationship between RF_i for each class of traffic C_i and $e(k)$ as follows:

$$RF_i(k) = RF_i(k-1) + k_p\left(1 + \frac{T}{T_i} + k_p\frac{T_d}{T}\right)e(k)$$
$$-k_p\left(1 + 2\frac{T_d}{T}\right)e(k-1) + k_p\frac{T_d}{T}e(k-2)$$

where T, $e(k)$, $e(k-1)$, and $e(k-2)$ denote the sampling period, the error signal at the kth sampling instant ($t = k^*T$), the error signal at the $(k-1)$th sampling instant, and the error at the $(k-2)$th sampling period, respectively. The parameters T_i and T_d depend on the proportional gain k_p, the integral gain k_i, and the derivative gain k_d. They are equal to (k_p/k_i) and (k_d/k_p), respectively. The three parameters k_p, k_i, and k_d represent the positive definite gains of the PID controller (i.e., proportional, integrative, and derivative gains, respectively), which have an important impact on the stability and speed of convergence of the system to the reference value. The proportional action (with k_p gain) is used to reduce the tracking error but does not eliminate it; the integrative action (with k_i gain) allows an asymptotic convergence and error rejection, and the derivative action (with k_d gain) improves the stability and the transient response. It should be noted that these values are obtained from empirical results and do not change the fundamental performance of the controller under variable network conditions. Thanks to the developed controller, the following actions are achieved:

- When the congestion is high, the tracking error is high (positive values): the system triggers a reaction by increasing the RF_i value for each class, which increases the reject probability $p_{i,j}$ to be sent to eNodeBs.
- When the system workload is low, the tracking error takes a negative value (also the RF_i values), that is, no need for rejecting MTC signaling traffic at eNodeBs.

Now we need to translate the RF_i values into a reject probability to be sent to eNodeBs participating in the congestion alleviation (i.e., eNodeBs that forward MTC signaling to the MME). We defined the following formula in order to derive the value of $p_{i,j}$ for each traffic class i at an eNodeB j:

$$P_{i,j} = \begin{cases} \min\left(\dfrac{traffic_j}{\sum_{l=1}^{n} traffic_l}.RF_i(k), 1\right) \\ 0 \end{cases}$$

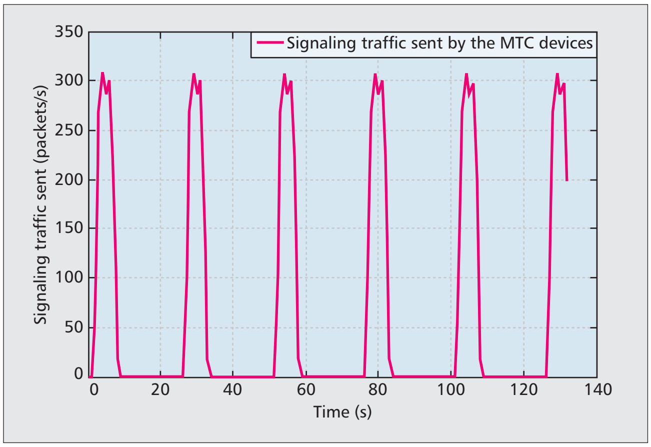

Figure 3. *The signaling traffic generated by MTC devices.*

Parameter	Value	Details
nMME	1	The number of MMEs
nP/SGW	1	The number of P/S-GWs
neNodeB	10	The number of eNodeBs
nMTCperCell	150	Number of MTC devices per cell
k_p	0.032	Proportional gain
k_i	0.00051	Integrative gain
k_d	0.000312	Derivative gain
T	0.01 s	Sampling period
Buffer size	100	Buffer size (packets)
Reference queue length	50	MME's reference length (packets)
S	276	Service rate in the MME (request/s)

Table 2. *Simulation parameters.*

where $traffic_l$ denotes the amount of traffic coming from eNodeB l. Indeed, the probabilities $p_{i,j}$ represent the proportion of MTC signaling traffic that need to be rejected, in order that the system reaches the reference value of the queue length at MME. Aiming at ensuring fairness among eNodeBs, this probability was carefully dimensioned to depend mostly on the amount of traffic generated by an eNodeB, which participates in the congestion alleviation. Thus, the higher the traffic forwarded by an eNodeB, the higher the value of $p_{i,j}$ (to be applied by this eNodeB to reduce the MTC traffic) will be.

As mentioned before and following 3GPP standards, admission control is implemented at the RAN level. Upon receiving an attach request from a MTC device, an $eNodeB_j$ applies admission control based on the reject probability $p_{i,j}$ received from the MME. Requests from MTC devices belonging to a class C_i is, thus, rejected with the probability $p_{i,j}$, by randomly choosing a uniform value between zero and one. If this value is greater than $p_{i,j}$, the MTC attach request is accepted. Otherwise, the request is rejected, and a backoff time value is indicated to the MTC device in order to ensure a good distribution of future incoming attach requests over time.

Performance Evaluation

We implemented the CAAC solution using the NS3 simulator [13]. We simulated a system of 1 MME, 1 P-GW, 1 S-GW and n eNodeBs. The MTC traffic is modeled as a bursty traffic, as shown in Fig. 3, which represents the connection requests originating from the MTC devices. In fact, bursts are what characterize MTC applications since MTC devices are more likely to send data at synchronized periods due to time expiration or the occurrence of an event [14]. MTC devices are evenly distributed among the simulated n eNodeBs. For the sake of simplicity, in the simulations, we considered only one class of MTC devices. The PID gains were empirically obtained. They do not change any of the fundamental observations made about the proposed scheme. Other simulation parameters are shown in Table 2.

Figure 4 illustrates the relation existing between the RF and the reject probability p used by the eNodeBs for admission control. The figures clearly indicate that the system reacts by increasing the RF value (i.e., the amount of MTC traffic to

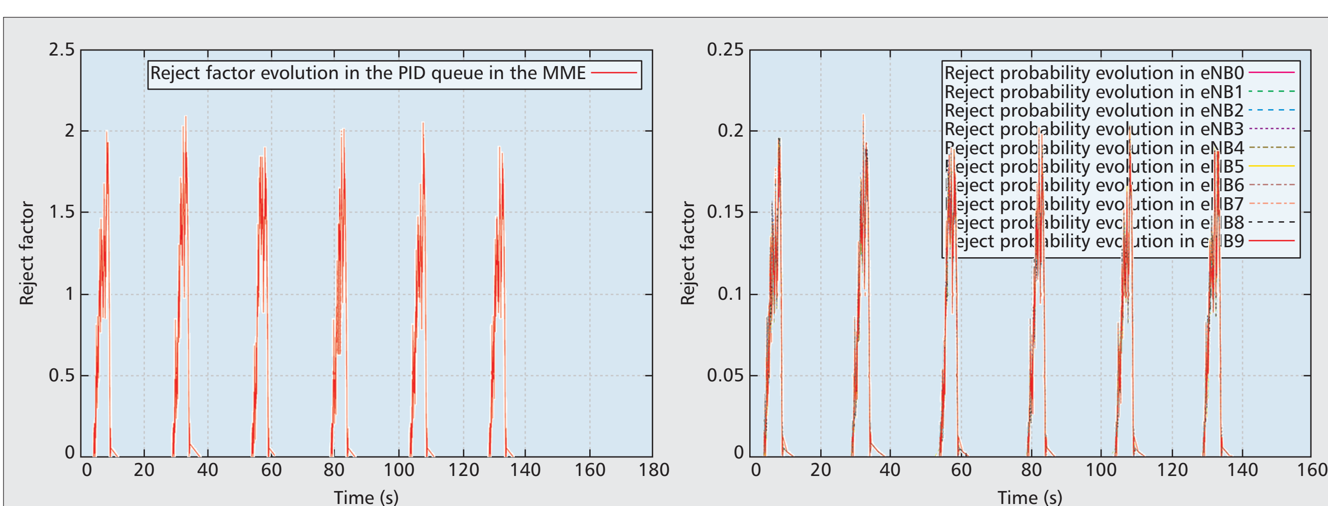

Figure 4. *Reject factor and reject probability evolution, respectively at the MME and eNodeBs.*

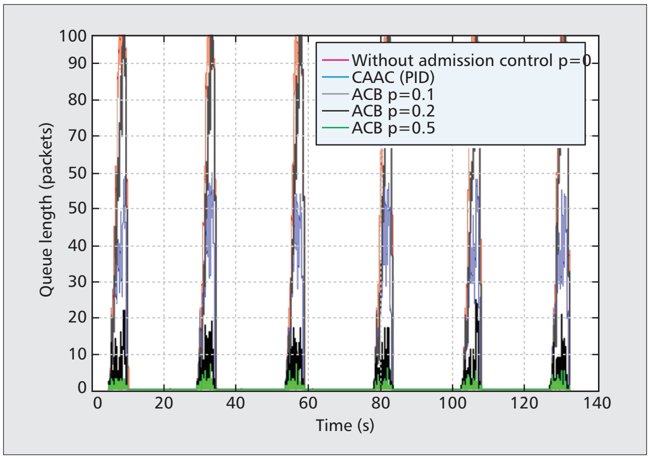

Figure 5. *MME's queue length variation.*

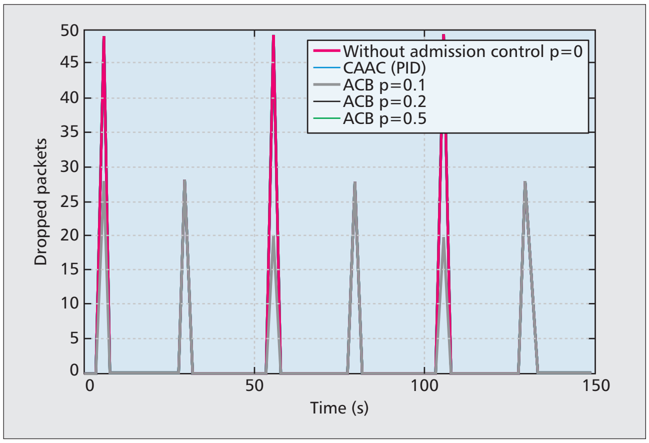

Figure 6. *Packet drops at the MME.*

reject) each time the MME is overloaded by the MTC signaling traffic. In addition, as the eNodeBs forward the same amount of uplink traffic to the MME, the p value is practically the same for all the eNodeBs. Here, the p value is around 0.15, which represents $RF/neNodeB$ (neNodeB=10). Furthermore, the graphs also clearly demonstrate that the reject probability value follows the evolution of the RF, which reflects the need to reduce the UL traffic of the eNodeBs when the MME is overloaded.

Now, we compare the CAAC solution against the conventional approach whereby no admission control is used at eNodeBs for MTC signaling traffic as well as against the ACB solution introduced by the 3GPP standards. We consider three constant values of the reject probabilities p to be used in ACB. For instance, setting p to 0.2 ($p = 0.2$) means that the MTC devices successfully attach to the network with a probability equal to $(1 - p)$, that is, 0.8. Figures 5 and 6 show the MME queue length evolution and the dropped packets at the MME, respectively. The obtained results show that the ACB and CAAC mechanisms can alleviate congestion at the MME level thanks to their admission control feature. The worst performances are observed when no-admission control is used, where bursts affect the queue length (i.e., overutilization) resulting in packet drops. At some points, 49 packets (15 percent of the received traffic in the MME) are dropped. In case of ACB, we remark that using high reject probability ($p = 0.5$) considerably reduces the congestion at the MME, but at the price of underutilizing the network resources (i.e., low throughput). Furthermore, this situation may impact the quality of service (QoS) expected by the MTC provider, as the MTC traffic to be sent to the remote server is severely decreased. In contrast, a low value of the reject probability ($p = 0.1$) increases the network resource utilization, but causes packet losses at the MME. In fact, we notice that packets are dropped when bursts of MTC signaling traffic are received at the MME; a fact which shows that ACB ($p = 0.1$) fails in avoiding congestion. On the other hand, in the case of CAAC, the evolution in time of the queue length at the MME indicates stable behavior of the system (no packet drops due to congestion), despite receiving highly bursty traffic from MTC devices. The queue length mostly oscillates around the reference value (50 packets or attach requests). The demonstrated system stability validates the configuration of the PID gains (k_p, k_i and k_d) used in CAAC.

Conclusion

In this article, we address the problem of congestion and system overload, which occur when MTC communications are deployed in EPS. We introduce a solution that rejects MTC traffic at RAN using feedback from core network nodes regarding their congestion status. The system was modeled with a PID controller that controls the congestion level around an acceptable value by tuning the amount of MTC traffic to be rejected at eNodeBs. Simulation results show clearly that the proposed solution avoids congestion and maintains good performance of the system in comparison to other existing solutions such as ACB.

Finally, it shall be stressed that the proposed solution is compatible with the 3GPP standardization efforts as CAAC follows the concept of ACB. It can easily be deployed and implemented at relevant EPS nodes, including eNodeBs. Furthermore, CAAC can easily be extended to reduce the congestion at the RACH level as well, by applying admission control at the MTC device itself; that is, eNodeBs broadcast the computed reject probabilities to the MTC devices, which in turn enforce them locally.

References

[1] "Worldwide Cellular M2M Modules Forecast Market Brief," *Beecham Research*, Aug. 2010.
[2] S. Lucero, "Maximizing Mobile Operator Opportunities in M2M: The Benefits of an M2M-Optimized Network," *ABI research*, 1Q 2010.
[3] A. Malm and T. Ryberg, "Wireless M2M and Mobile Broadband Services," *Berg Insight*, Feb. 2007.
[4] 3GPP, "System Improvements for Machine Type Communications," TS 22.368 V10.1.0, June 2010.
[5] 3GPP, "Service Requirements for Machine-Type Communications," TS 22.368 V10.1.0, June 2010.
[6] A. Visioli, "Practical PID Control," Springer-Verlag Editor, ISBN 1-84628-585-2, 2006.
[7] T. Taleb and A. Kunz, "Machine Type Communications in 3GPP Networks: Potential, Challenges, and Solutions," *IEEE Commun. Mag.*, vol. 50, no. 3, Mar. 2012, pp. 178–84.
[8] 3GPP, "Study on RAN Improvements for Machine-type Communications (Release 10), Technical Specification Group Radio Access Network" TS 37.868.
[9] CMCC TSG R2-113197, "Performance Comparison of Access Class Barring and MTC Specific Backoff Schemes for MTC," 3GPP Meeting, 23–27 Aug. 2010.
[10] S.-Y. Lien, K.-C. Chen, and Y. Lin, "Towards Ubiquitous Massive Accesses in 3GPP Machine-to-Machine Communications," *IEEE Commun. Mag.*, special issue on Recent Progress in Machine-to-Machine Communications, Apr. 2011.
[11] 3rd Generation Partnership Project, "Non-Access-Stratum (NAS) protocol for Evolved Packet System (EPS); Stage 3," 3GPP TS 24.301 V9.3.0, June 2010.
[12] M. Chen *et al.*, "Normalized Queuing Delay: Congestion Control Jointly Utilizing Delay and Marking," IEEE/ACM Trans. Net., vol. 17, no. 2, Apr. 2009, pp. 618–31.
[13] Network Simulator (NS) 3, http://www.nsnam.org.
[14] KPN TSG SA, "Problem of Overload Control with Machine Type Communications," 3GPP Meeting 10–14 May 2010, Kyoto, Japan.

Biographies

ADLEN KSENTINI (Adlen.Ksentini@irisa.fr) is an associate professor at the University of Rennes 1, France. He is a member of the joint UMR IRISA/INRIA team Dionysos. He received an M.Sc. in telecommunication

and multimedia networking from the University of Versailles. He obtained his Ph.D. degree in computer science from the University of Cergy-Pontoise in 2005, with a dissertation on QoS provisioning in IEEE 802.11-based networks. His other interests include future Internet networks, cellular networks, green networks, QoS, QoE, and multimedia transmission. He is involved in several national and international projects on QoS and QoE support in future Internet networks. He is a co-author of over 40 technical journal and international conference papers. He has been in the technical program committee of major IEEE ComSoc conferences, ICC/GLOBECOM, WCNC, and PIMRC.

YASSINE HADJADJ AOUL (Yassine.Hadjadj-Aoul@irisa.fr) is currently working as an associate professor at the University of Rennes 1, France, where he is also a member of the IRISA Laboratory and the INRIA project team Dionysos. He received a B.Sc. in computer engineering with high honors from Mohamed Boudiaf University, Oran, Algeria, in 1999. He received his Master's and Ph.D. degrees in computer science from the University of Versailles, France, in 2002 and 2007, respectively. He was an assistant professor at the University of Versailles from 2005 to 2007, where he was involved in several national and European projects such as NMS, IST-ATHENA, and IST-IMOSAN. He was also a post-doctoral fellow at the University of Lille 1 and a research fellow, under the EUFP6 EIF Marie Curie Action at the National University of Dublin (UCD), where he was involved in the DOM'COM and IST-CARMEN projects, which aim at developing mixed Wi-Fi/WiMAX wireless mesh networks to support carrier grade services. His main research interests concern the fields of wireless networking, multimedia streaming architectures and protocols, congestion control protocols and QoS provisioning, and satellite and space communications. He has been on the technical program committees of different IEEE conferences, including GLOBECOM, ICC, VTC, PIMRC, and IWCMC. His work on multimedia and wireless communications has led to more than 30 technical papers in journals and international conference proceedings.

TARIK TALEB (tarik.taleb@neclab.eu) is currently working as a senior researcher at NEC Europe Ltd. Prior to his current position, he worked as an assistant professor at Tohoku University, Japan. He received his B. E degree in information engineering with distinction, and M.Sc. and Ph.D. degrees in information sciences from GSIS, Tohoku University, in 2001, 2003, and 2005, respectively. His research interests lie in the field of architectural enhancements to 3GPP core networks, mobile cloud networking, mobile multimedia streaming, intervehicular communications, congestion control protocols, and social media networking. He has also been directly engaged in the development and standardization of the Evolved Packet System. He is on the editorial board of a number of IEEE and Wiley journals. He is serving as secretary of the Wireless Communications Technical Committee. He is the recipient of many awards, such as the 2009 IEEE ComSoc Asia-Pacific Young Researcher award and the 2008 TELECOM System Technology Award.