

Efficient Delivery of Frequent Small Data for U-healthcare Applications Over LTE-Advanced Networks

Revak R. Tyagi
SenSIP Center & Industry
Consortium
School of ECEE
Arizona State University
Tempe, AZ 85287
rtyagi@ieee.org

Ki-Dong Lee
Research & Standards Dept.
LG Electronics Mobile
Research
San Diego, CA 92131
kdlee@ieee.org

Frank Aurzada
Institute of Mathematics
Technische Universität
Berlin, Germany
aurzada@math.tu-berlin.de

Sang Kim
Research & Standards Team
LGE Mobile Research
San Diego, CA 92131
sanggook.kim@lge.com

Martin Reisslein
SenSIP Center & Industry
Consortium
School of ECEE
Arizona State University
Tempe, AZ 85287
reisslein@asu.edu

ABSTRACT

Ubiquitous healthcare (u-healthcare) applications typically require the frequent transmission of small data sets, e.g., from patient monitors, over wireless networks. We consider the transmissions of such u-healthcare data over an LTE-Advanced network, where each small data set must complete the standardized random access (RA) procedure. We mathematically analyze the delay of the RA procedure and verify our analysis with simulations. We find that our delay analysis, which is the first of its kind, gives reasonably accurate delay characterization. Thus, the presented delay characterization may form the basis for network management mechanisms that ensure reliable delivery of small frequent u-healthcare data sets within small delays.

Categories and Subject Descriptors

C.2.1 [Network Architecture and Design]: Wireless communications; C.2.5 [Local and Wide-Area Networks]: Access schemes

General Terms

Performance, Reliability

Keywords

Delay analysis, LTE-Advanced, Preamble transmissions, Random access, Ubiquitous healthcare

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MobileHealth '12, June 11, 2012, Hilton Head, South Carolina, USA.

Copyright 2012 ACM 978-1-4503-1292-9/12/06...\$10.00.

1. INTRODUCTION

It is expected that the elderly population will continuously grow worldwide for the next few decades; for example, by 2030, more than 22% of the population of European countries is projected to be over 65 years old [7]. This increase is expected to necessitate more public and private expenditures on healthcare; further, this increase in expenditures can encourage the development and application of smart information and communication technology. Due to such a global trend, “ubiquitous healthcare” (or u-healthcare) services have attracted considerable interest, particularly in the field of wireless networking and its applications. U-healthcare involves the use of information and communication technology to facilitate the exchange of health and medical information and services between patients and healthcare service providers. One of the most common applications of u-healthcare is “remote patient monitoring” to carry out various measurements, such as those of blood pressure and heart rate. Remote patient monitoring enables the measurements to be carried out wherever a communication link established and thus, reduces the need for patients to visit health clinics; as a result, the quality of life of the patients, particularly the elderly, is improved.

Driven by such needs, a wide spectrum of research has been performed in the information and communication research community for possible realization of u-healthcare service in a ‘reliable’, ‘secure’, and ‘efficient’ manner in the near future. In this paper, we are driven by the need for improving ‘reliability’ of wireless connection establishment under one of the fourth generation (4G) cellular architectures. When it comes to u-healthcare, there are components of interests that should be considered from networking perspectives: (1) a cluster of sensor nodes logically and/or physically attached to an individual patient, (2) a group of patients receiving u-healthcare service, and (3) a u-healthcare provider. A cluster of sensors belonging to a patient can form a Body Area Network (BAN) with one communication node acting as a Body Control Unit (BCU) transmitting the

gathered information (after pre-processing, if needed) to the u-healthcare provider or server.

For most routine measurement applications, the u-healthcare provider or server is located somewhere else over the Internet cloud and, therefore, it is important to study how to provide wireless connectivity for the BAN nodes toward the u-healthcare provider or server. In this paper, we consider a cellular network as a means to connect BAN nodes to the u-healthcare provider or server. Among several cellular network standards, we specifically consider the Long Term Evolution (LTE) Advanced (LTE-Advanced) standard, which is the most *popular*¹ Radio Access Network (RAN) protocol standard for the 4G cellular network and is known to meet the ITU-R performance requirements for the 4G cellular technology.² Furthermore, the LTE-Advanced standards encompass a wide range of advanced RAN, System Architecture (SA) and Core Network (CN) technologies, such as those for Machine-Type Communication (MTC, or Machine-to-Machine, M2M), enabling many advanced applications, with limited or no human interactions, to be more efficiently realized in our daily life. Due to such reasons, we have chosen to consider the cellular network standards as a means to deliver u-healthcare information to the u-healthcare provider or server. Thus, we believe that our study provides practically relevant insights for implementing low-delay u-healthcare transmissions.

Typically, the measurement information of u-healthcare subscribers (or patients) is small in size and generated somewhat frequently. If the u-healthcare provider requires the measurement information to be updated in a real-time or semi-real time manner, the ‘frequency’ of each individual’s data transmissions can increase accordingly. Thus, one of the key characteristics of u-healthcare information to be considered as a design factor is “frequently transmitted” “small” data. For cellular networks, this characteristic is a challenging hurdle to overcome because the radio resource is scarcely available and because each transmission of small u-healthcare data unavoidably has to go through the so-called Random Access (RA) procedure (see Fig. 1). If a communication node fails at RA, it cannot establish a Radio Resource Control (RRC) connection for its data transmission and reception. Thus, failure at RA may cause urgent data to be delayed (in the blocked-customer-delayed (BCD) queue cases) or cause measurement reports to fail (in the blocked-customer-cleared (BCC) queue cases).

Motivated by this reason, we mathematically model and study the RA behavior over LTE-Advanced when small u-healthcare measurement data are frequently generated and transmitted. Our mathematical delay analysis is in good accordance with simulations, predicting the traffic load bound-

¹In terms of the number of service providers worldwide that have chosen their evolution path to be LTE: Most European service providers, including Vodafone and Orange; most major North American service providers, such as Verizon Wireless, AT&T, and T-Mobile.

²This paper follows the *formal definition* of 4G technology specified by the International Telecommunication Union Radio-communication Sector (ITU-R). As of 2011, many cellular service providers, such as Verizon Wireless, AT&T and T-Mobile, have announced that they have deployed 4G networks. However, those underlying standards are all LTE or HSPA+ (but not LTE-Advanced and the like), which cannot officially meet the ITU-R performance requirements to be 4G technology.

aries within which low queueing delays occur. Our analysis also finds the traffic load regions of where the chances of the data being blocked increase and large delays can occur. We believe that our study can be utilized, after proper enhancement, for careful optimization of RA system parameters in 4G networks that are highly loaded by small frequent data, as generated by u-healthcare applications.

The remainder of this paper is organized as follows. We review the related literature in Section 2. The system model under consideration is discussed in Section 3. The delay characteristics are analytically derived in Section 4. In Section 5, the analytical results are compared with the observations from an independent simulation model. Section 6 presents the conclusions and future work.

2. RELATED WORK

The RA model in LTE (and later Releases) is a type of slotted ALOHA. Early work on slotted ALOHA was done by Kleinrock and Lam [8]. It used a Markovian model to analyze the system and introduced primary trade-off between throughput and stability of the system. Carleial and Hellman [3] studied the system in terms of drift of the system state. Ferguson [6] conducted preliminary work on delay analysis and established that the service time is independent of queue parameters.

Lüders and Haferbeck [10] analyzed the performance of the GSM version of slotted ALOHA using a circuit switching based model. Sarker and Halme [14] used a similar model and analyzed the steady state throughput of the system and impacts of retransmission cut-off. Sakakibara *et al.* [13] analyzed a similar system and introduced several bounds on the number of retransmissions using formulations from cusp theory. They also considered some delay aspects but did not formulate closed-form analytical results. They primarily focused on the stability criterion for the system.

Yang and Yum [17] considered the delay distributions for slotted ALOHA systems. They analyzed different back-off strategies, such as uniform back-off, binary exponential back-off, and geometric back-off and derived the respective distributions. In this work, we study the system without back-off. Naware *et al.* [11] discussed the stability and delay for multipacket reception in slotted ALOHA. They have derived transmission probabilities and related them to delays. Their model is based on a capture channel model and model with finite number of users. Our model considers an unlimited user population without capture effect.

Altrnan *et al.* [2] used game theory as a means to minimize delay in slotted ALOHA system and found that the use of the game theoretic algorithm led to increased delays. They also varied transmission probabilities in the system. Cunningham [4] presented comparison of delays with various retransmission schemes using simulations. They compared immediate transmission and deferred transmission schemes. Liew *et al.* [9] used a model similar to ours, but consider only a binary exponential back-off system model which does not generalize to our system.

3. THE SYSTEM MODEL

We consider a cellular system comprising of multiple User Equipments (UEs) located in a cell. Considering a single cell is not a limitation of this work since the RA procedure in LTE is an interaction between a certain UE and it’s most rel-

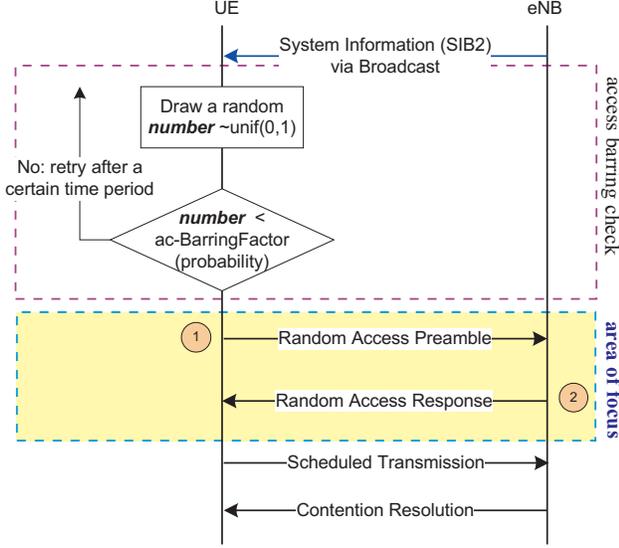


Figure 1: The holistic procedure of Random Access and the Area of Focus in this work: The UEs passing through Access Barring Check (ABC) will start transmitting a preamble, and the contention-based transmissions from multiple UEs will generate RA load in Random Access Channel (RACH).

evant cell chosen while staying in RRC_IDLE mode or while staying in RRC_CONNECTED mode with time synchronization lost (e.g., when timing alignment timer expires).

UEs shall use one of the Random Access Channel (RACH) opportunities configured by the physical (PHY) layer. The RACH is a set of logical resources defined in the 3-dimensional domain of time-frequency-preamble. Each u-healthcare data arrival at a UEs uplink (UL) transmission queue will trigger an RA procedure. First, the UE randomly chooses a preamble from O allowable preambles.

The eNodeB receives RA requests from UEs during an epoch of duration T_s . If multiple UEs transmitted their requests using the same preamble in the same epoch, then those RA requests are considered to have collided.³ When a collision occurs, contention (for RA) is not considered to be resolved (i.e., contention resolution failed). UEs are able to identify the contention resolution result at the fourth step of the RA of LTE [1]. If contention is resolved, the UE will get into RRC_CONNECTED mode to transmit the u-healthcare data to the eNB in the RAN domain.

When contention is not resolved, the UE may need to repeat preamble transmission. Before it's re-transmission, UE shall wait for a certain period of time signaled from eNB, if any, which is referred to as "backoff interval", ranging from 0 ms to 960 ms. For simplicity, we consider a case that backoff interval is 0 ms, that is, UEs whose preamble transmissions collide in a given epoch may re-transmit in the next epoch. In the case of re-transmission, if the same

³It is interesting to take all the physical layer considerations, such as different levels of transmission power among UEs, into account for RA modeling. However, the major focus of this paper lies in capturing MAC layer behavior and thus those considerations are beyond the scope of this paper.

preamble were used it would increase the chance of a collision manifold, resulting in a domino-like effect leading to high losses. In our model, the re-transmitting UE thus, randomly chooses a new preamble from O preambles. Choosing the same preamble again is allowed.

If the W th preamble transmission fails, then that UE drops the request from the system. For the delay considerations, such cases are not included. We only consider the delay of requests which eventually succeed.

The eNB shall periodically broadcast the following parameters values via System Information Block Type 2 (SIB2) (first three parameters) or via MAC Protocol Data Unit (PDU) (the last parameter),

1. Blocking probability P_b , any request has P_b probability of being transmitted immediately.
2. ac-BarringTime, if a request is not transmitted immediately then it waits for a random duration dependent on ac-BarringTime.
3. W , the number of times a request is made to eNodeB before the requesting UE drops the request.
4. T_o^{\max} , if the contention is not resolved then it will be made again after a random duration between 0 and T_o^{\max} .

For the model, under the stated assumptions these values are set as, $P_b = 1$, ac-BarringTime = 0, and $T_o^{\max} = 0$. Thus, with a prescribed value of W , the system is static with no change in these parameters, i.e., the parameters are specified at the beginning and then there is no change due to periodic broadcast.

4. DELAY CHARACTERIZATION

In this section the expected delay for the system described in Section 3 is analytically determined by first analyzing the request population in the random access system, followed by the delay characterization.

4.1 Steady-State System Population

Let us consider that λ new requests arrive in an epoch, and that x is the population of new plus re-transmitted requests in steady state. Let us denote f for the fraction of successful transmissions by the system population and denote δ for the fraction of dropped requests to incoming traffic. In the steady state, the rate λ of new requests balances the rate of dropped requests $\lambda \cdot \delta$ plus the rate of successful requests $x \cdot f$, i.e.,

$$\lambda = \lambda \cdot \delta + x \cdot f, \quad (1)$$

i.e.,

$$x \cdot f = \lambda - \lambda \cdot \delta. \quad (2)$$

Note that the probability of success is given by,

$$f = e^{-x/O} \quad (3)$$

and the probability of drop is given by,

$$\delta = (1 - f)^W = (1 - e^{-x/O})^W. \quad (4)$$

We arrive at the steady state equation,

$$xe^{-x/O} = \lambda \left(1 - (1 - e^{-x/O})^W \right) \quad (5)$$

where

- λ is the number of new requests received in an epoch.
- x is the total number of requests received at the eNodeB in one epoch. It includes new requests and re-transmitted requests.
- O is the number of preambles available for UEs.
- W is the number of times a request is made to eNodeB before it is dropped by the UE.

Eqn. (5) is similar to the state equation described by Sarker *et al.* [15] and by Sakakibara *et al.* [13]. The solution(s) x to the equation (5) give the expected value of incoming requests per epoch in the system model. From this we can compute success and drop probabilities using equations (3) and (4).

4.2 Expected Delay

Given the preamble transmission success probability f , the probability of exactly n collisions before a success, can be modeled as,

$$P[n \text{ collisions}] = (1 - f)^n \cdot f. \quad (6)$$

Hence, the probability of a UE to experience n collisions, given that it sends is

$$P[n \text{ collisions} | \text{sends}] = \frac{(1 - f)^n f}{\sum_{k=0}^{W-1} (1 - f)^k f}, \quad (7)$$

where $n = 0, 1, 2, \dots, (W - 1)$. Thus, the expected number of collisions $\mathbb{E}[C]$ is given by

$$\mathbb{E}[C] = \sum_{n=0}^{W-1} n \cdot \frac{(1 - f)^n f}{\sum_{k=0}^{W-1} (1 - f)^k f}. \quad (8)$$

Since an epoch has duration T_s , the expected delay $\mathbb{E}[D]$ is

$$\mathbb{E}[D] = T_s \cdot \sum_{n=0}^{W-1} n \cdot \frac{(1 - f)^n f}{\sum_{k=0}^{W-1} (1 - f)^k f}. \quad (9)$$

Equation (9) gives the expected delay for the system model described in Section 3. To account for the delay with no collisions we introduce an addend of $T_s/2$, which is the expected value for uniform delay in $[0, T_s]$. Thus, the expected delay for a given solution f to Equation (5) is given by,

$$\mathbb{E}[D] = \frac{T_s}{2} + T_s \cdot \sum_{n=0}^{W-1} n \cdot \frac{(1 - f)^n f}{\sum_{k=0}^{W-1} (1 - f)^k f} \quad (10)$$

$$= \frac{T_s}{2} + \frac{T_s}{\sum_{k=0}^{W-1} (1 - f)^k} \cdot \sum_{n=0}^{W-1} n \cdot (1 - f)^n \quad (11)$$

The summation in the equation (11) can be simplified using, for $0 < y < 1$, the identities,

$$\sum_{k=0}^{W-1} y^k = \frac{1 - y^W}{1 - y}, \quad (12)$$

and

$$\sum_{k=0}^{W-1} k \cdot y^k = y \cdot \frac{(1 + (W - 1)y^W - Wy^{W-1})}{(1 - y)^2}. \quad (13)$$

Substituting results from equations (12) and (13) in equation (11) and substituting $\gamma = (1 - f)$, we get the closed form solution for the delay as,

$$\mathbb{E}[D] = \frac{T_s}{2} + \frac{T_s \cdot \gamma \cdot (1 + (W - 1)\gamma^W - W\gamma^{W-1})}{(1 - \gamma) \cdot (1 - \gamma^W)}. \quad (14)$$

4.3 Multiple Solutions

In some cases, Equation (5) has more than one solution. In such cases, we consider two of the roots (the smallest and the largest values) to represent stable low delay/low drop state and unstable high delay/high drop states. The third solution indicates an unstable state in which the system does not stay for long. If Equation (5) has more than one solution, we compute the expected delay for the f value found using the following interpolation in Equation (14),

$$f = \alpha \cdot f_{\text{stable}} + (1 - \alpha) \cdot f_{\text{unstable}} \quad (15)$$

where f_{stable} and f_{unstable} are the probabilities of success for the stable and unstable solutions. The parameter α is given by

$$\alpha = \frac{\lambda - \lambda_{\text{stable}}}{\lambda_{\text{unstable}} - \lambda_{\text{stable}}}, \quad (16)$$

whereby λ_{stable} is the minimum arrival rate at which more than one solutions exist. $\lambda_{\text{unstable}}$ is the maximum arrival rate at which more than one solution. λ is the arrival rate, between λ_{stable} and $\lambda_{\text{unstable}}$, at which the expected delay is being computed.

5. NUMERICAL EVALUATION AND SIMULATIONS

In this section, we examine the expected delay of u-healthcare related small data from many patients. As previously noted, each small data is supposed to go through RA so that the UE can obtain a UL grant, and ultimately establish a RRC connection, for its data delivery. We compare the expected delay obtained from Equation (14) with verifying simulations. The simulation we performed is based on OMNeT++ [16] in C++. The simulation model has a generator of RA requests with a Poisson arrival rate of λ . The statistics collection is done using Akaroa2 [5]. The observations for $W = 10$ and $W = 12$ are based on 95% confidence intervals with a relative error of 5%. The simulation observations for $W = 15$ are based on 90% confidence interval with a relative error of 10%. The value of O is set to 54, and the arrival rate is varied from $\lambda = 0.01$ to $\lambda = 25$ in steps of 0.01. The values of other system parameters are $P_b = 1$, $T_s = 0.5$ ms, ac-BarringTime = 0 and $T_o^{\text{max}} = 0$.

For comparison with analytical results, all the roots to the equation (5) are found using exhaustive search within the range $[0, 10000]$, since values below zero are meaningless and those above 10000 are not expected. If a single solution is found for Equation (5), then the delay is computed using Equations (3) and (14). Otherwise, Equations (15) and (14) are used.

Figures 2, 3, and 4 show comparison of results from our analysis with the values estimated using simulations. It is clear from these figures that the values are in close agreement when the RA request rate, λ , is small. When the system is in the unstable state, then the variability increases. However as the λ increases further the variability reduces and the analysis results are in greater agreement with the simulation

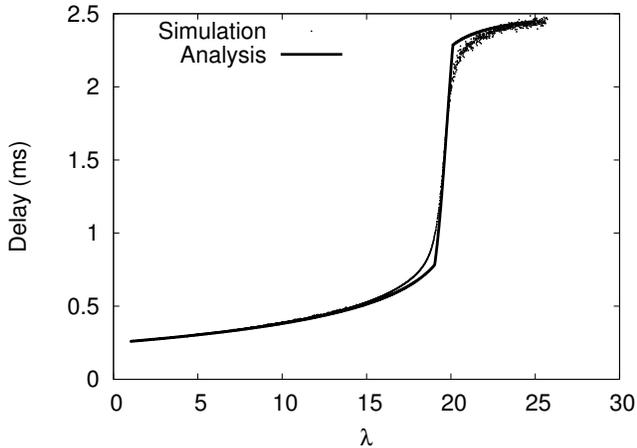


Figure 2: Delay Comparison at $W = 10$.

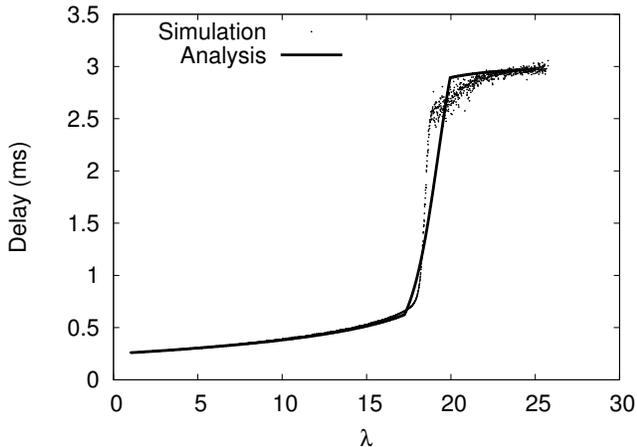


Figure 3: Delay Comparison at $W = 12$.

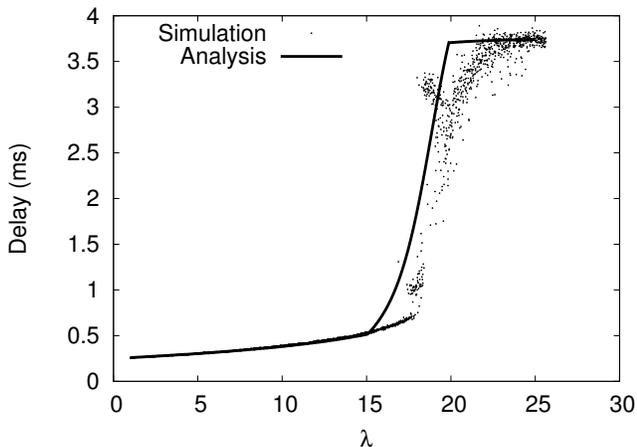


Figure 4: Delay Comparison at $W = 15$.

results. The interpolation adopted in Equation (15) is quite accurate when the attempts are limited to 10. The increase in the number of allowed attempts leads to a slight deviation from the interpolated model though. The deviation is still small and the steep change in the state of the system is within the limits in which the interpolation operation works.

From the simulations it is clear that the system is stable (is in the low delay part in Figures 2, 3 and 4) for values of λ close to 19, this is the well known limit of stability for ALOHA systems [12, 8, 6] (for our model, this value is $e^{-1} \cdot 54 = 19.86$). From Figures 2, 3 and 4 we observe that for stable values of λ , the upper limit on the expected delay is close to one epoch, i.e., T_s .

6. CONCLUSION

Motivated by the global trend of aging population and their potential needs for ubiquitous healthcare services, we have considered a use case of small data applications for u-healthcare, such as the delivery of u-healthcare measurement information over an LTE-Advanced network. Since the small data delivery must complete the Random Access (RA) procedure over the radio access network protocols, we focused on the modeling and analysis of the RA behavior specified in the MAC protocol of LTE-Advanced.

With particular attention to the RA behavior under heavily congested RACH conditions in mind, we have mathematically modeled the RA behavior of User Equipments (UEs) over the LTE-Advanced Radio Access Network protocols and validated the analytical model with simulations. Although our analytical model is simple, it is demonstrated in simulation results that the results obtained from our model show good agreement with those from simulation. Owing to the simple form of our analytical model, we believe that our model can be utilized as a guideline for realtime optimization and control of RA parameters. Furthermore, since our model tends to have a good agreement (i.e., good prediction) when the RA load is heavy, it can be effectively utilized for cellular systems with heavily populated u-healthcare subscribers.

Acknowledgement

This work was supported in part by LG Electronics, and the SenSIP Center and NSF NCSS I/UCRC Site.

7. REFERENCES

- [1] G. T. 36.321. *Evolved Universal Terrestrial Radio Access (E-UTRA): Medium Access Control (MAC) Protocol Specification*, v.10.4.0. 2011.
- [2] E. Altrnan, D. Barman, R. El Azouzi, and T. Jimenez. A Game Theoretic Approach for Delay Minimization in Slotted ALOHA. In *Proc. of IEEE International Conference on Communications*, volume 7, pages 3999–4003, June 2004.
- [3] A. Carleial and M. Hellman. Bistable Behavior of ALOHA-Type Systems. *IEEE T. Commun.*, 23(4):401–410, Apr. 1975.
- [4] G. A. Cunningham, III. Delay versus Throughput Comparisons for Stabilized Slotted ALOHA. *IEEE T. Commun.*, 38(11):1932–1934, Nov. 1990.
- [5] G. Erwing, K. Pawlikowski, and D. McNickle. Akaroa2: Exploiting Network Computing by

- Distributing Stochastic Simulation. In *Proc. of European Simulation Multiconference (ESM)*, pages 175–181, Warsaw, June 1999. International Society for Computer Simulation.
- [6] M. Ferguson. On the Control, Stability, and Waiting Time in a Slotted ALOHA Random-Access System. *IEEE T. Commun.*, 23(11):1306–1311, Nov. 1975.
- [7] K. G. Kinsella and W. He. *An Aging World: 2008*. U. S. Census Bureau, Washington, DC: U. S. Government Printing Office, 2009.
- [8] L. Kleinrock and S. Lam. Packet Switching in a Multiaccess Broadcast Channel: Performance Evaluation. *IEEE T. Commun.*, 23(4):410–423, Apr. 1975.
- [9] S. C. Liew, Y. J. Zhang, and D. R. Chen. Delay Analysis of Aloha Network. In *Proc. of IEEE Global Telecommunications Conference (GLOBECOM)*, pages 1–6, Dec. 2008.
- [10] C. Lüders and R. Haferbeck. The Performance of the GSM Random Access Procedure. In *Proc. of IEEE Vehicular Technology Conference*, volume 2, pages 1165–1169, June 1994.
- [11] V. Naware, G. Mergen, and L. Tong. Stability and Delay of Finite-User Slotted ALOHA with Multipacket Reception. *IEEE T. Inform. Theory*, 51(7):2636–2656, July 2005.
- [12] L. G. Roberts. Dynamic Allocation of Satellite Capacity through Packet Reservation. In *Proc. of the National Computer Conference and Exposition*, AFIPS '73, pages 711–716, New York, NY, USA, 1973. ACM.
- [13] K. Sakakibara, H. Muta, and Y. Yuba. The Effect of Limiting the Number of Retransmission Trials on the Stability of Slotted ALOHA Systems. *IEEE T. Veh. Technol.*, 49(4):1449–1453, July 2000.
- [14] J. H. Sarker and S. J. Halme. The Prudence Transmission Method I (PTM I): A Retransmission Cut-Off Method for Contention based Multiple-Access Communication Systems. In *Proc. of IEEE Vehicular Technology Conference*, volume 1, pages 397–401, May 1997.
- [15] J. H. Sarker and S. J. Halme. An Optimum Retransmission Cut-Off Scheme for Slotted ALOHA. *Wireless Pers. Commun.*, 13:185–202, 2000. 10.1023/A:1008930002745.
- [16] A. Varga. The OMNeT++ Discrete Event Simulation System. In *Proc. of European Simulation Multiconference (ESM)*, pages 319–324, June 2001.
- [17] Y. Yang and T.-S. P. Yum. Delay Distributions of Slotted ALOHA and CSMA. *IEEE T. Commun.*, 51(11):1846–1857, Nov. 2003.