American Journal of Networks and Communications
2016; 5(1-1): 1-5
Published online June 9, 2015 (http://www.sciencepublishinggroup.com/j/ajnc)
doi: 10.11648/j.sjbm.s.2016050101.11
ISSN: 2326-893X (Print); ISSN: 2326-8964 (Online)

Aloha Based Resource Allocation Scheme to Efficient Call Holding Times in Public Safety Network

Tuyatsetseg Badarch[1], Otgonbayar Bataa[2], Bat-Enkh Oyunbileg[2]

[1]Department of Information Technology, Mongolian National University, Ulaanbaatar, Mongolia
[2]Department of Information Technology, Mongolian University of Science and Technology,Ulaanbaatar, Mongolia

Email address:

ba.tuyatsetseg@mnu.edu.mn (T. Badarch), b_otgonbayar2002@yahoo.com (O. Bataa), o_bat_enkh@yahoo.com (Bat-Enkh Oyunbileg)

To cite this article:

Tuyatsetseg Badarch, Otgonbayar Bataa, Bat-Enkh Oyunbileg. Aloha Based Resource Allocation Scheme to Efficient Call Holding Times in Public Safety Network. *American Journal of Networks and Communications.* Special Issue: Traffic and Performance Engineering for Networks. Vol. 5, No. 1-1, 2016, pp. 1-5. doi: 10.11648/j.sjbm.s.2016050101.11

Abstract: The allocation of limited resources such as links, servers, and agents to support as many as emergency call responses as possible while maintaining reasonable quality of service is a principal requirement in public safety networks. It is difficult to handle the highest priority emergency call responses for a congested period with a limited number of dedicated incoming links from commercial networks. Therefore, in this paper, we present the adaptive scheme based on a different call conversation time for an achieving efficient resource scheduling during a congested period without investing extra resources. The proposed adaptive scheduling approach improves accuracy of the network performance with the dynamic call holding time approach.

Keywords: Aloha Protocol, Resource Allocation, Public Safety Network

1. Introduction

The ability to access emergency services by dialing a nationally designed fixed numbers from the Public Switched Telephone Networks (PSTNs) is a vital component of public safety and emergency preparedness [1]. The "real" emergency call handling is a specific service, since all calls have the same highest priority service. Therefore, it is difficult to handle all these highest priority call responses for a congested period, with a limited number of dedicated incoming links from commercial networks.

Therefore, one main criteria of the emergency call handling network performance to enhance SLA (Service Level Agreement) requirements is the limited call holding time, which a customer may use as the common resource within a time after it has been allocated to a system resource. There are standards, for instance, the document serves as a standard for answering 9-1-1 calls operating procedure for the call taking function within Public Safety Answering Points (PSAPs) presents the standard that 90% of all 9-1-1 calls arriving at the PSAP shall be answered within 10 seconds during the busy hour and 95% of all 9-1-1 calls should be answered within 20 seconds [2], [3]. This shows that limited conversation time is critical function for the emergency system SLA.

In general, the recent study shows that it is common in the SLA for workloads whose intensity varies widely and unpredictably that the dynamic resource allocation maximizes performance of a system [4]-[5]. It is a well known that the dynamic allocation is the more efficient solution than the static allocation scheme. In traditional static resource allocation approaches, when network resources are depleted, more resources (links, trunks, operators, agents, servers) are added to the network. These additions may resolve congestion conditions temporarily, but it is not a long term cost effective solution.

In the literature, a resource allocation discipline, such as time dependent priority, selfish scheduling, shortest to longest remaining time scheduling, and processor sharing scheduling, may prove to be optimal for certain application requirements [6] - [7]. A dynamic holding time allocation scheme that uses adaptive quota ensures an efficiency approaching that of First Come First Serve (FCFS), especially under heavy load conditions [8]. Moreover, the priority resource allocation is proposed to support ongoing activities in the ITU and IETF for developing the International Emergency Preparedness Scheme (IEPS) [9]-[11].

Emergency voice incoming calls for specific emergency services such as ambulance, police, fire, and hazard which are directed to emergency – specific PSAPs transferred through the bulk of the existing E1 trunks across PSTNs are the main part of the network used for the peak period performance evaluation. The reason is that the dedicated limited number of inbound trunks is built under the government regulation.

If the number of calls addressed to the system is too large during the peak period, the incoming trunks will be overloaded and operators cannot handle the volume of emergency calls. Hence, our aspect of the peak period performance is to characterize the system capacity using the model of the dynamic call holding time scheduling approach handling the overload of incoming calls.

The rest of the paper is organized as follows: Section 2 presents the dynamic call holding time scheduling scheme based on the Bernoulli theory, Section 3 describes the analytical model formulation of this dynamic scheme. Section 4 and Section 5 present the results based on the proposed method.

2. Proposed Model Formulation

The ALOHA protocol is an interesting example of a MAC protocol of the contention-type. It is the precursor of Ethernet and its subsequent standardization as IEEE 802.3[12].

Based on the protocol, this model considers the vulnerable (non retrial) and non vulnerable (retrial) period of call holding time of the PSN [13]. Under this scheme, a dynamic quota with a vulnerability coefficient k is used to determine a threshold level to allow a number of customers to occupy a resource (agent) within a call holding time limitation, especially under heavy load conditions. Applying the Bernoulli theory under the assumption of multiple Poisson arrival process is the fundamental analytics used.

We consider a customer who is granted permission to use a resource without waiting, is allowed to use a resource for the total CHT with the following condition:

$$T = kT_v \tag{1}$$

where T is the proposed total Call Holding Time (CHT), T_v is the threshold value of the CHT, and k is the vulnerability coefficient.

When CHT is exceeded the threshold T_v, the system may initiate retrial calls or drop-off calls. The threshold can be measured by k. When $k < 1$, it describes a call success in the threshold value and the condition $k > 1$.

May describe retrials or drop off CHT. When $k = 1$, the CHT equals to threshold value of the time. For the idle system, we assume $k = 0$.

In this scheme, we consider the condition that the T_v which may be equal to the proposed CHT (1), it may be computed within the limited time: $T_{\min} < T_v < T_{\max}$.

In general, to compute the theoretical throughput in the proposed model, first we assume the arrival process is Poisson, which means the probability P, of getting no call at T time interval $P = [0, t]$ is:

$$P(T) = e^{-\lambda T} \tag{2}$$

where $\lambda = 1/T$ is the mean call arrival rate. Then, the probability P_T, of getting some calls at T time interval:

$$P_T = P(T) = 1 - e^{-\lambda T} \tag{3}$$

We assume the condition that the mean service time equals to the mean call holding time plus a constant small amount of the processing time. This small processing time is negligible in this study. Then,

$$T_v = T_{\max} - T_{\min} = T / k \tag{4}$$

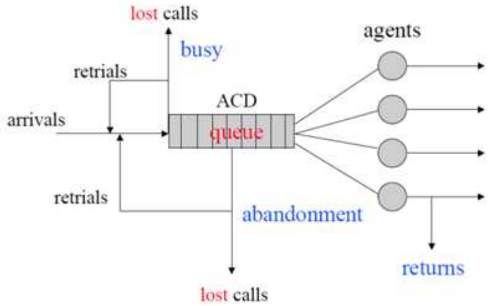


Fig. 1. *Feature of emergency call operational scheme.*

Offered load λ^* is described by the total call arrival rate divided by the total service rate (inverse of call holding time in the case of the negligible small amount processing time) as follows:

$$\lambda^* = \lambda \times T_v \tag{5}$$

Under the condition, using the Bernoulli theorem, we can get the probability of the holding time of retrials, P_{T_v}, of the network getting at least one call within the range of kT_v. The call departs from the system with probability $1 - P_{T_v}$. If calls passed within the threshold time, P_{T_v} is zero at an agent tier level.

We have the following formula of the carried load considering the probability of retrials P_T under the model in Fig.1.

$$L_c = \lambda + P_T L_c \tag{6}$$

where L_c is the carried load, the arrival rate λ describes the offered load, it shows that the carried load may be depend on strongly the probability of retrials time P_T. Hence, the probability P_T, of getting some calls within at kT_v time interval (Fig. 1):

$$P_T = P(T \le T_v) = 1 - e^{-L_c T} \tag{7}$$

Using (7) substitute P_T in (6) and rewriting, the call arrival rate λ has the following relationship within the call holding time T considering k coefficient.

$$\lambda = L_c e^{-kL_cT_v} \quad (8)$$

From (5), we easily update the offered load :

$$\lambda^* = \lambda \times kT_v = (kL_cT_v)e^{-kL_cT_v} \quad (9)$$

Where the normalized total carried load is described as follows:

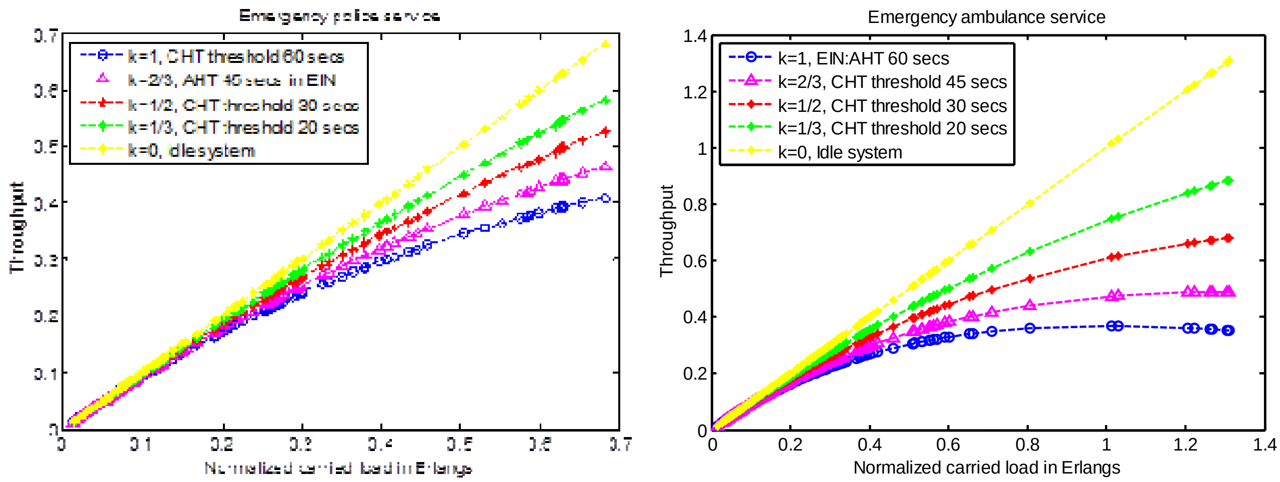
$$L_{c.n} = kL_cT_v \quad (10)$$

We often consider that the call holding time T equals to the number of calls N in unit time divided by the call arrival rate capacity μ . If all offered traffic can be serviced, so that the offered load λ^* may be described by the throughput. If the offered load is greater than this value, the system becomes unstable.

We sample throughput data points by feeding calls. We get the throughput vs. total carried load including retrials. It proves that CHT has the best performance in terms of the maximum throughput. The maximum achievable throughput is presented by the differentiation of the right side of equation (8).

3. Results of the Proposed Approach

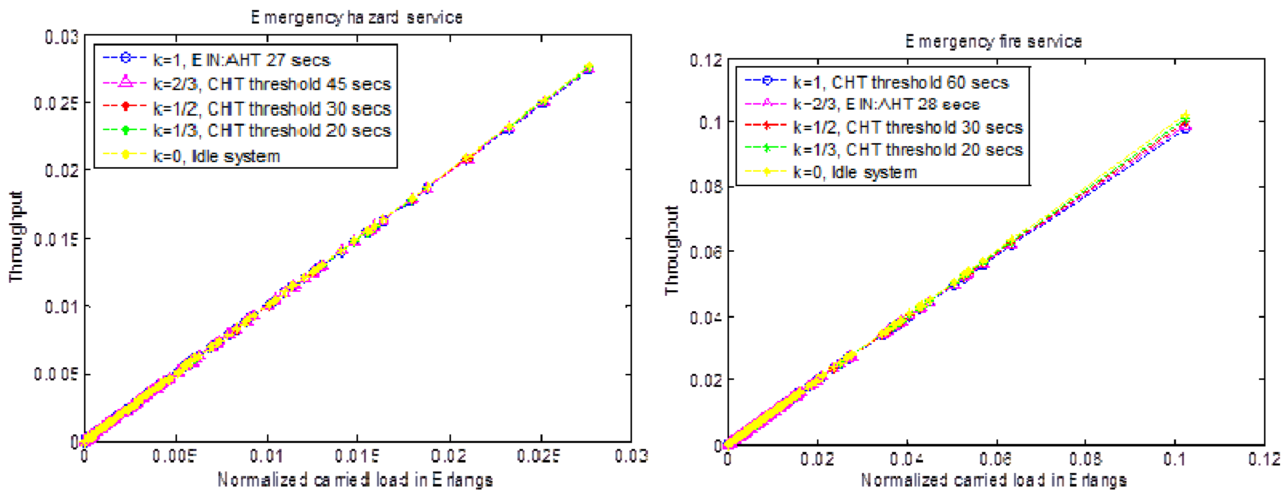
We monitored the incoming trunks at edge port of IP-PBX of the EIN network [14]. The experimental period was one week including week days, weekends, and the biggest national holiday Naadam, therefore, the workload analysis enables us to account for the existing system's peak period performance.



(a) Load vs. Throughput of "102". (b). Load vs. Throughput of "103".

Fig 2. *Throughput performance for a high load.*

In Fig.2a, there are maximum emergency users who generate as high as 1.2 Erlangs of the ambulance class as well as light users generating less than 0.03 Erlangs of the hazards class.



(a). Load vs. Throughput of "105". (b). Load vs. Throughput of "101".

Fig 3. *Throughput performance for a light load.*

During the period (Fig.3a). The figure presents the CHT value of all "103-ambulance" calls arriving at the PSAP is 60 seconds in average during the peak week and 45 seconds show the "102-police" CHT in average. As a result of the proposed approach, the phenomena in the case of "103" is clear the throughput is increased by approximately two times more than the existing throughput when the CHT is reduced to 20 seconds.

It is clear to see there is high demand to use the proposed scheme under the priority scheduling for the heavy load situation. Therefore, even though the arrival rate is increased up to a maximum load level, the calls of each class may be served by Long Holding Time (LHT), Medium Holding Time (MHT), or Short Holding Time (SHT) scheduling approaches. However, the idle system presents linear relationship between the throughput and traffic load, the existing EIN system has shown the exponential relationship between the throughput and traffic load.

Fig. 3 shows the idle case performance of the scheme.

Reducing the CHT, we can keep the system more idle compared to the congested system.

According to the results, in the light load of the emergency fire ("101") and hazard call ("105") classes, there is no difference between the long and short call holding time scheduling advantage. Then, FIFO scheme may be the best approach with a light load, then a dynamic holding policy may not be a main metric in the light load (See Fig. 3).

4. Simulation Verification

We also used Matlab Simulink to support the proposed model. Using the discrete event simulation SimEvents library, we built the model diagram. The diagram is omitted, due to the space limitations. The simulation was the time-based 24 hours period - simulation for both LHT, MHT, and SHT respectively. The block generates Poisson arrival events packed with the random exponential CHT attribute that we extract the data as the experimental data set for the Matlab Simulink simulator from statistics of the EIN by averaging the week into the day.

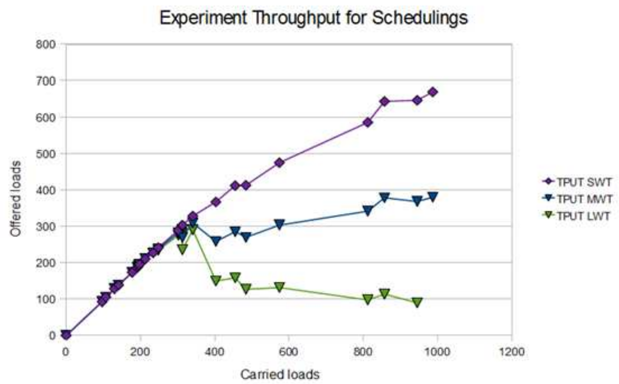


Fig 4. *Carried vs. Offered number of incoming calls depends on a call conversation time.*

The simulator provided output on queue length, throughput, and utilization. Our analysis focuses on the relationship between workload and throughput. The result (Fig. 4) proved that the calls into the network depend on a call conversation time for the long-LWT, medium-MWT, and short-SWT call conversation times, it clearly shows the same pattern as our analytical result this supports (8).

The upper limit thresholds have advantages over other resource allocation policies and is efficient to implement, it still must be considered against traditional resource management approaches, e.g., using excess capacity in the network to control a peak load [15].

5. Conclusion

The paper showed the benefits of a limited call holding time through throughput vs. carried load analysis for the dynamic resource scheduling policy in a public safety network. The new method uses the adaptive time quota ensureing an efficiency approaching the dynamic call holding time. We evaluated the performance of the heterogeneous EIN in Mongolia using the proposed model. Under this priority scheduling with the Bernoulli - loop, the theoretical, experimental and simulation results show the larger the throuhgput, the smaller the call holding time that the existing system produces during the peak period. In the light load, there is no difference of the throughput between the long and short call holding time scheme. Therefore, the paper brings the contribution which results in the smaller CHTs for emergency calls such as ambulance, police, fire and hazard calls of the existing network to influence an efficient dynamic resource allocation. The scheme can be a simple and generic approach, it can be applied to similar time - conscious systems such as time dependent call centers.

References

[1] B. Rosen, H. Schulzrinne, U. Columbia, J. Polk, and A. Newton, "Framework for Emergency Calling Using Internet Multimedia", RFC 6443, December 2011.

[2] National Emergency Number Association (NENA) Standard Operating Procedures Committee, Call taking Working Group, "NENA Call Answering Standard/Model Recommendation", Document 56-005, June 10, 2006.

[3] NENA Guidelines for Minimum Response to Wireless 9-1-1 Calls, Document 56-001 (revised November 18, 2004**).**

[4] Bennani, N. Mohamed, A. Daniel Menasce, "Resource Allocation for Autonomic Data Centers Using Analytic Performance Models." *In Proc. Sec.Int.Conference on. IEEE*, 2005.

[5] D. Ardagna, B. Panicucci, M. Trubian, and L. Zhang, "Energy-Aware Autonomic Resource Allocation in Multitier Virtualized Environments", *IEEE Transactions on Services Computing*, pp. 2 – 19, vol. 5, 2012.

[6] L. Kleinrock, P. Roy Finkelstein "Time Dependent Priority Queues Source", Journal of Operations Research, vol. 15, no. 1, pp. 104-116, 1967.

[7] E. Angel, E. Bampis, and F. Pascual, "Truthful Algorithms for Scheduling Selfish Tasks on Parallel Machines", *In Theoretical Computer Science,* 2006.

[8] A. N. Tantawy, A. N. Tantawi, D. N. Serpanos, "An Adaptive Scheduling Scheme for Dynamic Service Time Allocation on a Shared Resource", *Proc. of the 12th International Conference on Distributed Computing Systems*, pp. 294-300, 1992.

[9] H. Schulzrinne, "Requirements for Resource Priority Mechanisms for the Session Initiation Protocol (SIP)", *RFC 3487,* Feb 2003.

[10] K.Carlberg, I. Brown, C. Beard, "General Requirements for Emergency Telecommunications Services (ETS)", *RFC 3689*, Feb 2004.

[11] ITU, Telecommunications Standardization Sector, "International Emergency Preparedness Scheme", E-106 Recommendation, March, 2000.

[12] Norman M. Abramson, "Development of the ALOHANET", IEEE Transactions on Information Theory 31(2), 119-123 (1985).

[13] G. Koole, A. Mandelbaum, "Queueing Models of Call Centers: An Introduction", Annals of Operations Research, vol. 113, pp 41-59, 2002.

[14] B. Tuyatsetseg, "Parametric Modeling Approach for Call Holding Times of IP based Public Safety Networks," WASET journal. issue 73, pp. 568-575. Available: http://www.waset.org/journals/waset/v73.php.

[15] C.Beard, V.Frost, "Prioritized Resource Allocation for Stressed Networks", *IEEE/ACM Trans, Network*, vol. 6, no. 5, pp. 618-633, Oct. 2001.