SindhUniv. Res. Jour. (Sci. Ser.) Vol. 52 (03) 273-278 (2020)

http://doi.org/10.26692/sujo/2020.09.40

SINDHUNIVERSITYRESEARCHJOURNAL(SCIENCE SERIES)

Dynamic Length Based Data Aggregation in Machine Type Communications

L. ASLAM, S. SARWAR[++], A. AKRAM, U. SAEED*, Z. IRFAN**

Punjab University College of Information Technology (PUCIT), University of the Punjab, Lahore.

Received 11th February 2020 and Revised 8th July 2020

Abstract: For machine type communications, the use of aggregation has got a lot of attention in the recent past. In this article a dynamic length based aggregation technique is presented. Performance of proposed technique is evaluated with respect to three key aspects, namely, number of random access procedures, utilization of radio resources, and transmission latency. It is shown that like other techniques in literature our technique also reduces number of RA procedures, improves utilization of radio resources, at the cost of extra transmission latency. The demonstrated distinguishing feature of our technique is that it does not fail for very low traffic loads. The proposed technique works equally well for very low traffic toads to very high traffic loads.

Keywords: Data Aggregation; IoT; Machine type Communication;

I. INTRODUCTION

Recently, the internet of things (IoT) is becoming a fundamental concept to interconnect numerous devices due to its ubiquitous presence and surge influence in our society. IoT application opportunities are endless and will profoundly change existing infrastructure such as health, agriculture, logistics, and rescue system in the coming days. IoT applications are extending with the support of sensor networks, radio-frequency identification (RFID), machine to machine (M2M) communication, mobile internet, and wireless internet communication network. These machine type communication (MTC) devices cover a wide variety of applications ranging from real-time monitoring, industrial automation, and control of physical environment. MTC devices are applicable in smart cities, smart metering, smart animal farming, connected appliances, public safety, healthcare, and structure monitoring.

These MTC devices employ a diverse group of new wireless technologies (e.g., Bluetooth Low Energy (https://www.bluetooth.org/enus/specificationladopted-specifications.), (IEEE802.11ah, IEEE 802.15.4). These diversified connectivity technologies require supporting infrastructure, which open new research challenges to develop integrated IoT platforms into the current internet infrastructure.

Cellular networks can play an important role in the development and expansion of IoT networks. The cellular network can leverage their ubiquity, security, network management and advanced connectivity for backhaul capabilities into the IOT networks. In this regard, capillary networks aim to provide connectivity between wireless sensor networks and cellular networks. The capillary network consists of a set of devices connected using short-range radio-access technologies while it is connected to the backhaul cellular network using a node called Capillary Gateway (CG). In this way, capillary networks provide connectivity to a large number of devices.

Capillary networks consist of local area networks, these networks enable end-devices to off-load their data on the cellular network. Massive MTC devices access gateway to transmit data which causes congestion and signaling over head at the network.

Several techniques have been proposed to alleviate this problem including pull-based schemes, MTC clustering (Derya *et al.*, 2016) and data aggregation (Zaher *et al.*, 20176) Data aggregation gets more attention among all of these approaches.

All MTC devices are transmitting a small amount of data to a device over the capillary network. The device

---

[++] Corresponding author: shahzad.sarwar@gmail.com

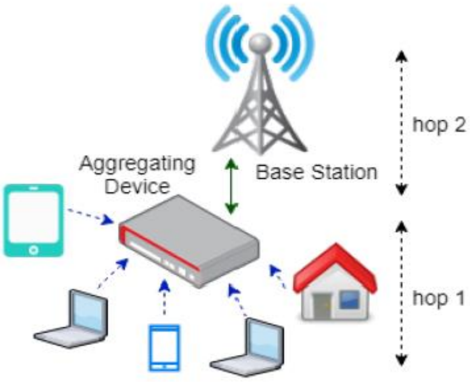
*Department of Computer Science, Virtual University of Pakistan.

**Department of Computer Science, Bahria University, Lahore Campus, Lahore.

that combines these data messages is normally the gateway and can also be called an aggregator. Data aggregation is the process in which the gateway aggregates arriving packets for a limited time and then forwards all the aggregated packets at once to the base station (BS). This leads to lesser number of BS random access (RA) and thus aggregation alleviates the need to access BS for every small packet. Furthermore, it also reduces signaling overhead, RA congestion, and resource wastage of cellular network (Nour *et al.*, 20114). In **(Fig. 1),** we can see capillary network with and without aggregation devices.

M2M devices are transmitting small amounts of data to an aggregator over a capillary network. The aggregator combines the received messages in a large message and transmits it over a cellular access network. In this paper, our goal is to provide a dynamic length based aggregation technique which can work equally well for low and high data rates. We provide a detailed discussion about different data aggregation methods, identified their strengths, and limitations. In the next section literature review is presented. In the Section III dynamic length based aggregation technique is analyzed. In the end the Section IV concludes the article by presenting conclusions and future research directions on this topic.

(a)



(b)

Fig.1: Machine type communication (a) without aggregation (b) with aggregation scenario.

2. LITERATURE REVIEW

(Shariatmadari *et al.* 2015) proposed time based aggregation scheme for the capillary networks. The gateway aggregates packets coming from different devices for a fixed period of time, and then forward this aggregated data to the BS. Aggregation of data packets for the fixed period of time decreases the RA interactions and radio resources needed to deliver the data to the network. However, aggregation for a fixed period of time may lead to resource wastage in case of low traffic. Furthermore, if any data packet has crucial information to offload and it's been waiting for that fixed period of time. There are chances that this information expires after that time.

(Kouzayha *et al.* 2014), employs two aggregator queues for periodic and alarm reporting. Periodic reporting consists of measurement information while alarm reporting contains information about specific events.

(Kim *et al.* 2016), discussed different aggregation scenarios such as single and multiple aggregators. They put forward multiple aggregators to avoid the link congestion problem at the single aggregator. They employ the nearest neighbor rule to forward data packets to the aggregator. However, it is a costly solution in respect of excessive hardware, resource allocation, and maintenance.

(Derya *et al.*, 2016) have proposed an energy efficient hierarchical data aggregation scheme using methods from stochastic geometry. Each level in this hierarchy consists of a transmitter and aggregator. At each level, transmitters transmit their payload to the nearest aggregator and transmitters are excluded after they complete all the transmissions. In the next level, aggregators from the previous level become the transmitters and the aggregation process is repeated over multiple stages.

(Salaam *et al.* 2018) have proposed a cooperative data aggregation scheme for massive MTC networks by employing the MTC gateway and multiple user equipment (UE). MTC gateway aggregates delay intolerant MTC data packets while UE collects delay tolerant MTC data packets. They also provide numerical and simulation results which conforms to their claim that the proposed cooperative algorithm outperforms as compared to single and multiple aggregator schemes.

(Zhou and Nikaein 2016) have proposed an adaptive packet aggregation method to alleviate the problems of collision and power consumption for RA in LTE/LTE-A. A UE sends preamble for multiple aggregated packets instead of every single packet. However, this reduction of collisions by considering aggregation of packets causes an increase in waiting time. However, as the collision rate reduces, the latency is increased which is not desirable for delay-sensitive applications.

To minimize the network latency and to enhance the system reliability of IoT applications (Bhandari *et al.*, 2017), proposed cloud-assisted priority-based channel access and data aggregation schemes. It can be seen from the numerical results that priority-based data aggregation schemes provide substantial improvements as compared to non-prioritized schemes.

(Matmoros and Anton-Haro 2013) proposed a direct access method in capillary networks to monitor environmental applications. They proposed a data compression strategy at the aggregator to alleviate the problem of network overloading and to incorporate massive MTC devices.

(Salman 2017), proposed a priority-based data aggregation scheme for MTC. He classifies data from MTC devices into three different priorities based on their applications. The author follows the M/G/1 queuing model to perform aggregation for these priorities. The data from a device with higher priority is served faster as compared to the lower priority device. Furthermore, he also analyzed the system performance concerning mean waiting time, system delay, and power consumption. The numerical results indicate that, as parameters such as arrival rate and aggregation level increases the energy consumption also increases.

(Liang *et al.* 2018) proposed cluster-based congestion mitigation access schemes (CCAS) to minimize the severe collision of MTC devices. They designed a modified version of the spectral clustering algorithm to assign same MTC devices into one cluster based on service requirements and their locations. An aggregator is selected to aggregate data packets from MTC devices in a cluster. Then, it forwards these packets to the BS after some time. They also employ a queuing model at aggregator to analyze system performance with respect to system delays and the probability of collision. Simulation results reveal that CCAS significantly increases the number of received packets while decreases the collision probability also.

3. DYNAMIC LENGTH BASED AGGREGATION

The main objectives of introducing aggregation are to reduce the accumulative link establishment cost and to reduce the required radio resources for data delivery to the network. By employing aggregation these objectives are achieved to a certain extent at the cost of increasing the overall transmission latency. The performance of any aggregation technique should be evaluated simultaneously with respect to three different aspects, namely, reduction in RA interactions, utility of radio resources, and overhead of excessive latency. Another aspect which should be considered is the robustness of the benefits offered by an aggregation technique with respect to varying traffic loads.

In (Kaijie and Nikaein, 2016) packets are aggregated up to a certain threshold. This fixed threshold if kept very small would benefit low traffic loads while compromise on aggregation gains for high traffic loads. Similarly if this threshold value is kept very high, it would be a disaster for low traffic loads.

In (Hamidreza, *et al.*, 2015) packets are aggregated for a certain time called as aggregation period. This technique would again face the same problem. If traffic load is high, then this technique would perform better but for very low traffic load scenarios this technique can altogether fail.

We are proposing a closed form formula for dynamic length based aggregation which would be able to cater different traffic load scenarios. The main idea is to use length based aggregation and change the aggregation length dynamically based on the traffic load.

A. Assumptions and Variales

We assume that N devices are connected to the network through a single aggregator. Every device when has some data it forwards it to the aggregator and aggregator responds with an acknowledgment. It is assumed that an acknowledgment is never dropped. To model the link quality between any device and aggregator a probability p_1 is taken by which a packet fails to reach to the aggregator. It is further assumed that this probability is same for the link of all devices to the aggregator. If a packet fails to reach the aggregator, such packet is resent by the device and we assume that m such tries are taken before dropping a packet. Aggregator uses a scheduling request to acquire the required resources for transmitting the data to the network through BS. The aggregator is usually connected to the network through a shared link and hence it may face a failure for sending its request. We assume that the probability of failure of an RA process isp_2. Further it is assumed that in case of failure of an RA process, the retries are limited to a maximum n times.

It is assumed that each device produces packets of approximately same size following a Poisson process. The average arrival rate from this Poisson process is taken to be λ packets per second per device.

B. *Performance Analysis of RA Processes*

As there are N devices generating packets with an arrival rate λ, the total arrival rate from all devices to the gateway (aggregator) is $N\lambda$. The number of packets per second received by the gateway are $N\lambda(1-p_1^m)$. The time, in seconds, to receive one packet by the gateway can be represented as $1/N\lambda(1-p_1^m)$.

If we assume packet size to be s bytes and size of a resource block (RB) to be c bytes then one RB can carry $\lfloor c/s \rfloor$ packets. Assume that, under certain traffic load, aggregation length to be k RBs. In this case, aggregation length would be $\lfloor kc/s \rfloor$ packets. The time required to fill this aggregation length can be expressed by $\lfloor kc/s \rfloor / N\lambda(1-p_1^m)$. If we call the time duration, in seconds, between start of two consecutive aggregation lengths as D, then it can be expressed as following.

$$D = \frac{\left\lfloor \frac{kc}{s} \right\rfloor}{N\lambda(1-p_1^m)} + \frac{1}{N\lambda(1-p_1^m)} = \frac{\left\lfloor \frac{kc}{s} \right\rfloor + 1}{N\lambda(1-p_1^m)}$$

The number of RA processes per second when aggregation is not used is $N\lambda(1-p_1^m)(1-p_2)\sum_{i=1}^{n} i p_2^{i-1}$. We can express expected number of RA processes for one RB aggregation length as $(1-p_2)\sum_{i=1}^{n} i p_2^{i-1}$. If the aggregation length is k RBs, then the expected number of RA process per second are:

$$E\left[\frac{RAs}{sec}\right] = \frac{(1-p_2)\sum_{i=1}^{n} i p_2^{i-1} \times N\lambda(1-p_1^m)}{\left\lfloor \frac{kc}{s} \right\rfloor + 1}$$

To get the normalized value of RA processes, we divide number of RA processes with aggregation to the number of RA processes without aggregation. This normalized gain comes out to be $1/(\lfloor kc/s \rfloor + 1)$. We can express the percentage saving in expected RA processes as:

$$\rho(\lambda) = \frac{100 \times \lfloor kc/s \rfloor}{\lfloor kc/s \rfloor + 1}$$

If we assume that one RB can contain four packets then value of c/s would be 4. To make this aggregation model adaptive we define k as a function of λ. As k is number of RBs in an aggregation length we want to make it dependent on arrival rate such that, if traffic load is very low it takes a value of one and if traffic load is high its integer value increases very slowly. Keeping this in view we define k in terms of λ as:

$$k = \lceil 2 \times \lambda^{0.25} \rceil$$

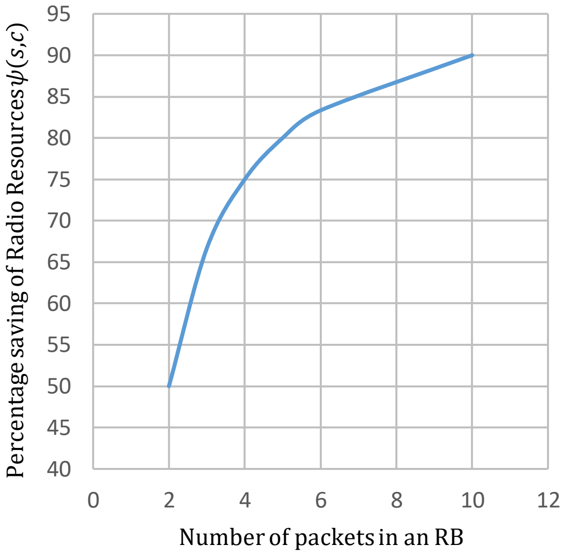


Fig. 2. Percentage saving of RA processes for varying traffic loads.

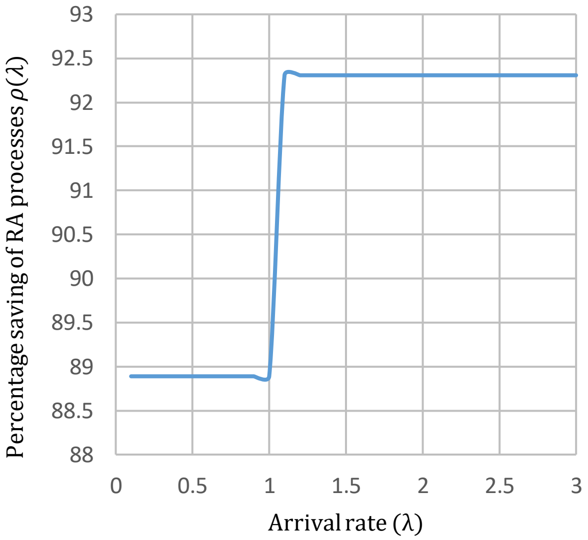


Fig.3. Percentage saving of radio resource against changing packet size.

The **(Fig. 2)** shows the percentage saving of RA processes for varying traffic loads. It should be noted that this scheme is not inefficient at very low traffic loads. It keeps on saving in RA processes and does not fail.

C. *Utility of Radio Resources*

On a successful reservation of dedicated radio resources, the gateway gets a minimum unit of one resource block (RB) allocated. In machine type communication normally the size of a packet is small as compared to the size of a RB. We assumed that the size of a packet is s bytes and size of a RB is c bytes. Without aggregation the gateway will try to reserve one RB with a probability of success p_3. The gateway will receive a total of $N\lambda(1-p_1^m)$ packet in one second and would use $N\lambda(1-p_1^m)p_3$ RBs. While in case of aggregation the gateway would reserve the resources once it has aggregation length filled with packets. The gateway has an aggregation length of k RBs. The number of packets which fit into an RB are$c/s$and we assume that this number is an integer hence an RB can be completely filled with whole packets. In one second $N\lambda(1-p_1^m)p_3 s/c$RBs would be filled and for every k of them resources would be reserved with a success probability ofp_3. If we look at the percentage saving of radio resources per second, it comes out to be,

$$\psi(s,c) = \left(1 - \frac{N\lambda(1-p_1^m)p_3 s}{cN\lambda(1-p_1^m)p_3}\right) \times 100 = \left(1 - \frac{s}{c}\right) \times 100.$$

It is observable that the utility of radio resource is dependent on packet and RB sizes. As we are focusing on machine type communication it is reasonable to assume that packet size is less than RB size and traffic load is low making value of k as one. The percentage saving in radio resource for a fixed size of RB with changing size of packet is demonstrated in **(Fig. 3).** For data represented in Fig. 3 it is assumed that the physical RB size is 60 bytes.

While using aggregation, the smaller packet size saves more radio resource. Hence it is more useful for machine type communication from this aspect. It should be noted that if the size of packet and RB are same then there is no saving in radio resource utilization. However, in machine type communication like in the context of IoT, packet sizes are usually small. Further, if the value of k is more than one, it is not hard to see that utilization of radio resources can improve a little and would not decrease.

D. Analysis of Latency

The benefits of aggregation discussed so far are achieved at the cost of extra latency. From the point when a device generates a packet to the point when it is successfully received at the BS, there are mainly three types of delays. First one is from the device to the gateway, second is at gateway waiting for the completion of aggregation length, and the third is from gateway to the BS. It should be noted that first and third type of delay is same in both cases if we use aggregation or not. Only the extra delay incurred due to aggregation is of second type. It is shown in (Hamidreza, *et al.*, 2015) that this extra delay is approximately $D/2$, on average, where D is the time between the start of two consecutive aggregations. The reason behind is that, the packets at the start of aggregation length would face more delay but the packets at the end of aggregation length would face very little delay and on average it comes out to be $D/2$. Hence, in our case, extra transmission latency due to aggregation can be expressed as follows.

$$\eta(\lambda) = \frac{D}{2} = \frac{\left\lfloor \frac{kc}{s} \right\rfloor + 1}{2N\lambda(1-p_1^m)} sec.$$

Let us assume a scenario in which 20 devices (nodes) are connected to a gateway, with$p_1 = 0.1$, $m = 5$, and $c/s = 4$. As k is dependent on λ, we can see a relationship between extra transmission latency and arrival rate (λ). This relationship is shown in **(Fig. 4).** It can be seen in that latency overhead for aggregation is negligible for modest to high traffic load scenarios. For very low traffic loads the latency is proportionally high.

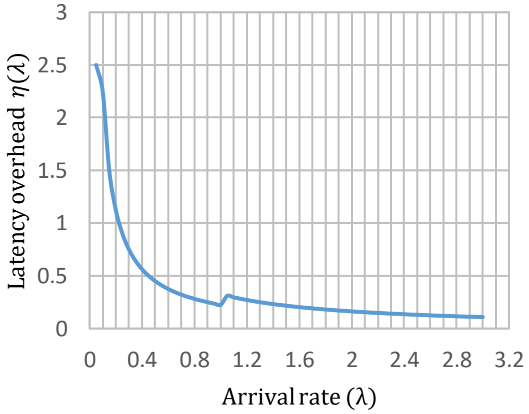


Fig.1: Transmission latency overhead due to aggregation for different arrival rates.

4. CONCLUSIONS AND FUTURE WORK

In this article a dynamic length based aggregation technique is presented. Performance of proposed technique is evaluated with respect to three key aspects, namely, number of RA procedures, utilization of radio resources, and latency. It is shown that like other techniques in literature our technique also reduces number of RA procedures, improves utilization of radio resources, at the cost of extra transmission latency. The demonstrated distinguishing feature of our technique is that it does not fail for low traffic loads like others in the literature. The proposed technique works equally well for very low traffic toads to very high traffic loads. In future we plan to extend our work by supporting our argument with simulative results. Further, we plan to improve our technique by involving both time and length to decide that how many packets should be aggregated for a more efficient aggregation solution.

REFERENCES:

Bluetooth standards. https://www.bluetooth.org/enus/specificationladopted-specifications. Accessed: 22-05-2019.

Dong M. K., R. B. Sorensen, K. Mahmood, O. N. Osterbo, A. Zanella, and P. Popovski, (2017) "Data aggregation and packet bundling of uplink small packets for monitoring applications in LTE", IEEE Network, vol. 31(6), 32–38.

Derya M., H. S Dhillon, and J. G Andrews, (2016) "Optimizing data aggregation for uplink machine-to-machine communication networks", IEEE Transactions on Communications, vol. 64(3), 1274–1290.

Hamidreza S., P. Osti, S. Iraji, and R. J¨antti, (2015) "Data aggregation in capillary networks for machine-to-machine communications", In 26th IEEE Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), 2277–2282.

IEEE 802.11 standard. http://www.ieee802.org/III.

IEEE 802.is standard. http://www.ieee802.org/IS

Javier M. and C. Anton-Haro, (2013) "Traffic aggregation techniques for environmental monitoring in m2m capillary networks", In 77th IEEE Vehicular Technology Conference (VTC Spring), 1–5.

Kaijie Z and N Nikaein, (2016) "Random access with adaptive packet aggregation in LTE/LTE-A", EURASIP journal on wireless communications and networking, vol. 36, 1–15.

Liang L., Lu Xu, B. Cao, and Y. Jia, (2018) "A clusterbased congestion-mitigating access scheme for massive m2m communications in internet of things", IEEE Internet of Things Journal, vol. 5(3), 2200–2211.

Nour K., M. Jaber, and Z. Dawy, (2014). "M2m data aggregation over cellular networks: signaling-delay tradeoffs", In 2014 IEEE Globecom Workshops (GC Wkshps),1155–1160.

Nour K., M. Jaber, and Z. Dawy, (2017) "Measurement-based signaling management strategies for cellular IoT", IEEE Internet of Things Journal, vol. 4 (5), 1434–1444.

Salman A. AlQahtani, (2017), “Analysis and modelling of power consumption-aware priority-based scheduling for m2m data aggregation over long-term-evolution networks”, IET Communications, vol. 11(2), 177–184.

Sabin B., S. K. Sharma, and X. Wang, (2017), “Latency minimization in wireless IoT using prioritized channel access and data aggregation”, In GLOBECOM-2017 IEEE Global Communications Conference, 1–6.

Tabinda S., W. Ur Rehman, and X. Tao, (2018) "Cooperative data aggregation and dynamic resource allocation for massive machine type communication", IEEE Access, vol. 6, 4145–4158.

Zaher D., W. Saad, A. Ghosh, J. G. Andrews, and E. Yaacoub, (2016) "Toward massive machine type cellular communications", IEEE Wireless Communications, vol. 24(1), 120–128.