A Comprehensive Review of Congestion Control Techniques in M2M and Cellular Networks

Chidanandan V. and Arun Raj Kumar P.
Department of Computer Science & Engineering; National Institute of Technology (NIT) Puducherry, Karaikal, India
E-mail: chidanandanv@gmail.com, p.arunrajkumar@nitpy.ac.in

Abstract—Long Term Evolution (LTE) emerges as the promising communication technology for data and voice service in Human to Human Communication (H2H). Currently, LTE network also includes Machine to Machine (M2M) communication. Combining the legacy systems and evolving Machine type devices, increases the access rate to a particular base station leading to congestion in Access network. Congestion in Random Access Network (RAN) that occurs during the Random Access Process affects the performance of H2H communication. The objective of the paper is to analyze the existing congestion control techniques and to provide a comprehensive summary on classified poll and event based mechanisms.

Index Terms—Long Term Evolution, Human to Human, Machine to Machine, Random Area Network, Random Access, Random Access Channel.

I. INTRODUCTION

Enabling the vision of Internet of Things (IoT), Machine-to-Machine (M2M) communication enables the end-to-end communication between autonomous devices without much human intervention [35]. According to CISCO [31], the numbers of machine to machine communication devices are expected to cross 50 billion by 2020 and this was predicted by Ericson and Cisco in 2011 and 2013 respectively by considering an average of 6 devices per person [31]. This promises new revenue opportunities and tremendous growth in the market share with a large set of the application domain. M2M communication differs from the existing sensor network by communication type included with intelligent analytics. The communication may be wired (LAN, Telephone line, etc.) or wireless (Zigbee, IEEE 802.15.11, IEEE 802.15.4, Bluetooth, etc.,). The wired communication is more reliable but difficult to adapt to mobile applications. The wireless communication will be most suitable choice for mobile application, but due to the convergence of various wireless technologies it suffers from the interference problem. Therefore, to achieve the M2M communication of the various applications (irrespective of communication type) there is a need for the broad communication coverage and connectivity. The Cellular networks are the most widely accepted wireless network technology and have undergone various generations to standardize updates for better performance towards Human to Human (H2H) Communication. However, merging the existing cellular networks with M2M networks face challenges such as congestion, channel availability, etc. Due to various traffic characteristics and features such as high device density and small amount of payload and uplink dominant transmissions, simultaneous access to multiple users in Random Access Network and Core Network remains challenging. Transmission in LTE consists of two phases, viz., uplink and downlink. The downlink is for the MTC devices to send the data to the servers and uplink is used for the signaling and access requesting. Large number of devices is accommodated in particular base station and the network is accessed through Random Access mechanism. Random Access Channel (RACH) will be overloaded due to limited uplink resources and situation will worsen if H2H also accesses the network at the same time. The telecommunications standard development organization such as Third Generation Partnership Project (3GPP), Institute of Electrical and Electronics Engineers (IEEE), European Telecommunications Standards Institute (ETSI), etc., have recently initiated developing global standards for supporting M2M over a cellular network. The Technical Report TR 37.868 summarized by 3GPP identified the importance of a cellular uplink signaling channel called the Random Access Channel (RACH). The report also suggests considering congestion control as a first priority area for improvement. The M2M architecture (Fig.1) is classified as follows [1]:

i) **Device domain**: the user elements such as sensors, smart devices, etc., are used to monitor the behavior and activities assigned for that device. In 3GPP, the device domain can be a three type scenario such as;

  i. the devices which establish the connection to the eNodeB,
  ii. The devices which works under the femto cell, gateway assisted network connected devices and
  iii. The device which directly communicates with other device

ii) **Network Domain**: it is the core network that

depends on the need and coverage of the application. Due to Large Coverage, Scalability and Reliability, most of the application prefers the LTE for communication. The Evolved Packet Core (EPC) is IP based core networks which rely on the radio access. This provides connectivity between UE and the Packet Data Network (PDN). The main components of the LTE EPC are as follows:
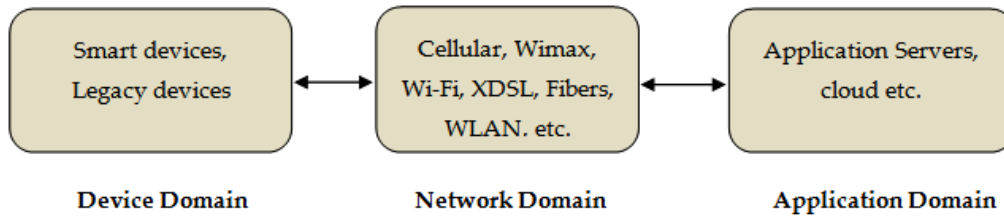

Fig.1. Architecture of M2M.

- Mobility Management Entity (MME): it is the control node which processes the signaling between the UE and the core network. This provides the functionality like mobility management such as Roaming and Handover. Upon initial access, it selects the serving and packet data network gateway. S1 interface connects the EPC with eNodeB
- Serving – Gateway (S-GW): It acts as the IP router between the eNodeB and the Packet Data Network gateway and takes care of user IP packet. It is responsible for handoff between eNodeB
- Packet Data Network–Gateway (P-GW): packet data gate way is responsible for providing connectivity to the external packet data network. It is responsible for user equipment IP allocation; also it acts as an anchor of mobility between 3GPP and non 3GPP technologies.
- Home Subscriber Server (HSS): It is referred to as the centralized database that contains information of all the network subscribers.
- Policy Control and Charging Rules Function (PCRF): resides in P-GW, helps in decision making and charging policy. It also helps in QoS setting information for user session.
- eNodeB: it is known as base station in LTE and it is responsible for allocating the radio access for communication. X2 interface connects the UE with the eNodeB.

Table 1. M2M Applications and Their Characteristics

Applications (features)	Characteristics								
	Mobility	Time Controlled	Time Tolerant	Payload	Online small Data Transmissions	Group based policy	Priority	Connectivity	Location Specific Trigger
Digital Home	fixed	✓	✓	small	-	✓	-	Secured	-
Smart City	fixed	✓	-	small	✓	✓	✓	Secured	✓
Surveillance & security	fixed	✓	-	large	✓	✓	✓	Secured	✓
eHealth	Low	✓	-	moderate	✓	-	✓	Secured	-
Industrial Automation	fixed / Low	✓	✓	large	✓	✓	✓	Secured	✓
Smart Logistics & Transportation	high	✓	✓	small	✓	-	✓	-	✓
Agriculture	fixed / Low	✓	✓	large	-	-	-	-	✓
Smart Metering	fixed	✓	✓	small	✓	-	-	-	✓
Retails & utilities	fixed	-	-	small	✓	-	✓	Secured	
Smart Grid	fixed	✓	✓	small	-	✓	✓	-	✓

iii) Application domain: it may be the cloud / server that holds the entire data sent by the devices and it will be accessed by the end-user for monitoring. The applications with their characteristics of M2M are shown in Table1. For example, in agriculture application, information providing regarding the right time to harvest, weather forecast, etc., are vital to farmers for the increased production of crops. Therefore, characteristics such as time tolerant, delay sensitive, priority access, etc. are significant to achieve the successful communication among the devices or between the devices. Under Third Generation Partnership Project (3GPP), a wide range of M2M applications such as Smart Home to Smart City, Smart Vehicles to Smart Logistics, Health Application to Energy Generation, etc., are supported. In Table 1, all the applications have small payload as characteristics. But this small payload in each application pass through the LTE resulting in congestion and affecting the QoS. Thus

there is a need for congestion control in LTE networks.

II. RANDOM ACCESS PROCEDURE

Every machine in LTE network has to follow the Random Access Channel (RACH) process for communication. Random Access procedure basically is the key for accessing the channel and it allocates the Resource Block (RB) for the devices which was successful in RACH process [2-3]. Resource Block is distributed in both frequency and time domain. Each resource block consists of 12 subcarrier frequency of 15 Khz of total 180 Khz and holds 6 to 7 slots of 1 ms each in time domain. This slot also known as resource elements is the smallest unit that is used for the resource allocation as illustrated in Fig.2. One radio frame of 10 ms consists of 10 frames of 1 ms each and each 1 ms frame is split into two equal sub frames. The physical channel is responsible for assigning resource elements and actual data transmission. This will aggregate the Resource Blocks to form frames used in Uplink and Downlink. Orthogonal Frequency Division Multiple (OFDM) access is used for the downlink transmission in LTE. LTE uses different physical channels. The type of Downlink channels are as follows.

- Physical Broadcast Channel (PBCH): it is responsible to carry the Master Information Block (MBI). It will assist in cell identification and initial access specific parameter.
- Physical Downlink Control Channel (PDCCH): it is a downlink control channel used to allocate resources for efficient data transmission and also informs the user element about transport format and uplink shared channel resource allocation.
- Physical Downlink Shared Channel (PDSCH): it carries all user data and system broadcast information which is not carried in PBCH. PDSCH uses the unused resource in the PDCCH for the data transmission.

- Physical Control Format Indicator Channel (PCFICH): it carries a Control Format Indicator (CFI) which indicates the number of OFDM symbols used in each frame involved in transmission in PDCCH.
- Physical Hybrid ARQ Indicator Channel (PHICH): it carries the ACK and NACK indicating to UE in response to the transmission in uplink data.

Single carrier – Frequency Division Multiple Access (SC-FDMA) is used for uplink transmission in LTE. The Channels involved in uplink are as follows.

- Physical Uplink Control Channel (PDCCH): it is used to carry uplink control information such as Channel Quality Indicator (CQI), ACK and NACK response for HARQ, in response of downlink and uplink scheduling request.
- Physical Uplink Shared Channel (PUSCH): corresponding to uplink shared channel, it carries the user data and reference signal for higher layer.
- Physical Random Access Channel (PRACH): it is used to get access into the LTE by using the Preamble sent by the UE. This channel is always Non-Synchronized.

A device initiates the RACH process due to any one of the following reasons [27]:

a) Cell search for access to the network.
b) Device state transfer from sleep to active state,
c) During Handover
d) Reconnection / exponential after back off.

The RACH process establishes the connection by one of the following access mechanisms:

1. **Contention Based RA Process** - The Accessing Request originated from the User Element (UE) and
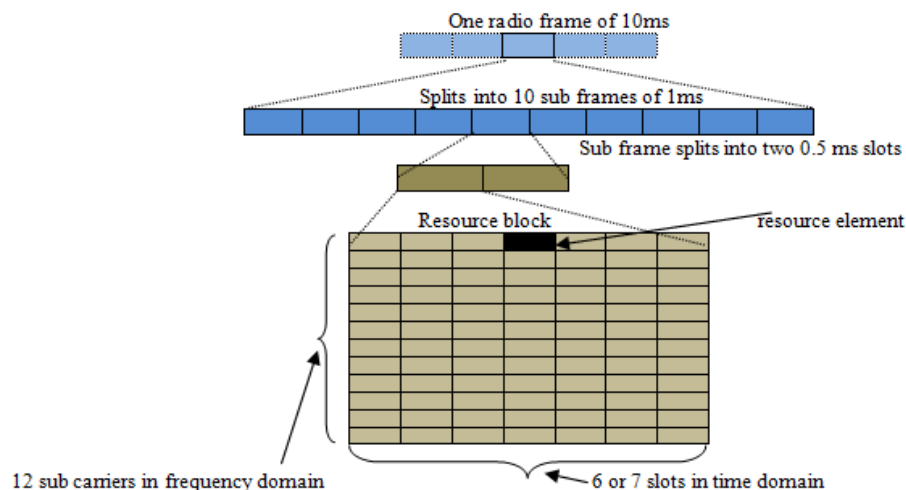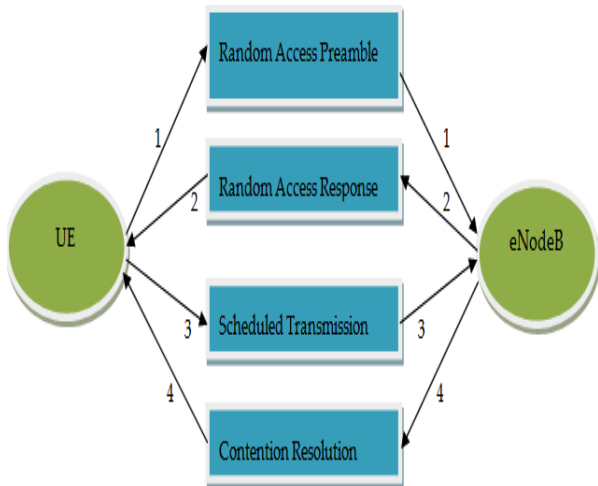

Fig.2. Frame Structure in LTE

Fig.3. Contention Based Random Access

2. **Non Contention Based RA Process** – The Accessing Request originated from the Server or other Communication Agents in network.

A. Contention Based Access

In Machine Type Communication (MTC), the User element initiates communication through contention based [2] with the help of Preamble. The word "contention" is because the device is still not synchronized and likely to collide with each other. The preamble is the pseudo random number generated by the device based on the information sent by eNodeB through System Information Broadcast (SIB) message. The total number of preambles available for the contention based access is shown in equation 1.

$$Preamble_{CB} = Total\ Preamble - Preamble_{NCB} \qquad (1)$$

Total number of Preamble in the cell is 64 and preambles reserved for non contention based access

Preamble $_{NCB}$ is 10. Therefore, total preamble available for contention based access is 54. During the Process, the devices establish the connection as mentioned in Fig.3 and flow of messages is explained in Table 2. The four steps of connection establishment is as follows;

1) **Random Access Preamble:** The user element selects one preamble out of 54 and sends it to the eNodeB for connection establishment as message 1. Since the RACH process is not synchronized, the other device selects the preamble resulting in collision. Due to collision, the eNodeB is not able to decode the preamble and the preambles are dropped.

2) **Random Access Response**: On successful reception and decode of preamble by the eNodeB, the Temporary Cell Radio Network Temporary Identification number (C-RNTI), time, and frequency slot are sent to UE as message 2. If the user element has not received any message from the eNodeB, it will wait for Random backoff time. As time elapses, the UE restarts the Random Access Preamble process and a new preamble is sent to eNodeB.

3) **Scheduled Transmission**: It is the first scheduled uplink transmission of actual random access request from the UE upon successful recipient of message 2. using Uplink shared channel, the UE sends the request for the connection setup which is identified by the temporary identification number C-RNTI. If the UE is successfully decoded, the contention remains unresolved for the other UEs.

4) **Contention resolution:** Upon successful decoding of message 3 by the eNodeB, it will authorize the connection with permanent identification (i.e., C-RNTI).

Table 2. Details of CB – RACH Process

Msg. no.	Msg. type	Link used	Direction	Process
1	RA Preamble	Uplink shared channel	UE ->eNB	Sends Preamble
2	RA Response	Downlink shared channel	eNodeB->UE	Timing advance, UL grant, TC – RNTI
3	Scheduled transmission	Uplink shared channel	UE ->eNodeB	Scheduled transmission
4	Contention Resolution	Downlink shared channel	eNodeB->UE	Contention Resolution with C – RNTI

Table 3. Details of Non CB – RACH Process

Msg. no.	Msg. type	Link used	Direction	Process
1	RA Preamble	Downlink shared channel	UE ->eNodeB	RA Preamble assignment
2	RA Response	Uplink shared channel	eNodeB->UE	RA Preamble
3	Scheduled transmission	Downlink shared channel	UE ->eNodeB	Random Access Response

B. Non - Contention Based Access

The Non Contention Based (NCB) access is well synchronized and the request is issued by the server for carrying out the specified task. Upon receiving the request by the server for the particular UE device, the eNodeB will allot the reserved preamble to the UE device and allows the UE for communication as shown in table 3.

Preamble $_{NCB}$ = 10 preamble is reserved in every cycle of RA.

NCB process is carried out by the RAN in three steps as follows:

- Random Access Preamble: eNodeB assigns a 6 bit reserved preamble for non contention based access
- Random Access Response: User Equipment sends the assigned symbol to the eNodeB.
- Scheduled Transmission: upon receiving the preamble in RAR. The eNodeB assigns the downlink timing alignment information. Initial uplink link grant and other activities similar to contention based access as shown in Fig.4. Due to individual assignment of timing slot for smooth data communication for the device, there may be a chance of empty slots remaining in every frame resulting in performance degradation.

III. CONGESTION IN RACH PROCESS

Congestion in MTC exists either in Random Area Network (RAN) or Core Network. Due to opportunistic nature of intelligent and embedded smart devices, it is easier to connect the available network for data communication. Due to less channel availability for data transmission, the core network overloads resulting in congestion. Thus, Congestion is a severe problem for the RAN that eventually degrades the performance in the network and also affects the Quality of Service. In 3GPP Long Term Evolution [4][5], accessing Network is broadly classified as follows:.

i. Event Based Access
ii. Polling Based Access

i. **Event-based Access:** The events / Access request is triggered either by the Network / Application domain.

A. **Contention - Free RACH (CF - RACH):** Channel accessing is made available by the eNodeB using the reserved preambles [5][7]. CFR has minimum message exchange between the eNodeB. CFR helps in fast and collision free connection setup

resulting in better performance. The information / calls which started the access from generated network are transferred to another network due to cell switching. This process is known as Handover and it occurs without any loss in information.

B. **Paging message by the server:** The paging message from the Server end to eNodeB is to allow the specific UE to initiate the network access. The application server reports instantly of a particular task.

ii. **Polling Based Access:** A request and response communication is as follows [6]; Devices start accessing the network with the help of preambles. Devices poll the eNodeB for the Downlink and uplink access time alignment grant. Upon resolving the preamble contention and devices those who sent the unique symbol, the eNodeB authenticates the devices and grant the access successfully. Introducing the MTC devices in the LTE network is not optimized and the integration is still in the initial stage due to the characteristic difference from the Legacy system. The Bulk activation / accessing nature of MTC introduces the congestion problem in RAN.

A. Existing Congestion Control Techniques

The Taxonomy of existing congestion control technique is shown in Fig.5. In [7], 3GPP explains a method to control the RA overload of delay sensitive devices. The LTE frame structure is used for the system access. The access to the network works similar to the paging system. All the devices in the cell must page for their access grant. The access cycle consists of consecutive access slots in a cyclic fashion. The access slot is synchronized with the system frame (10 ms) and identified by SFN mod K. The devices are permitted to send the request in their respective time slots. The access request congestion is reduced by delaying the device to enter into Random access process with a backoff interval [4].
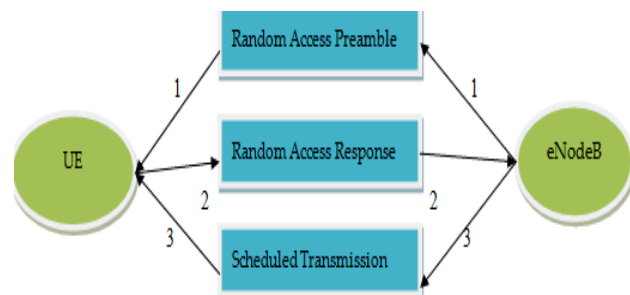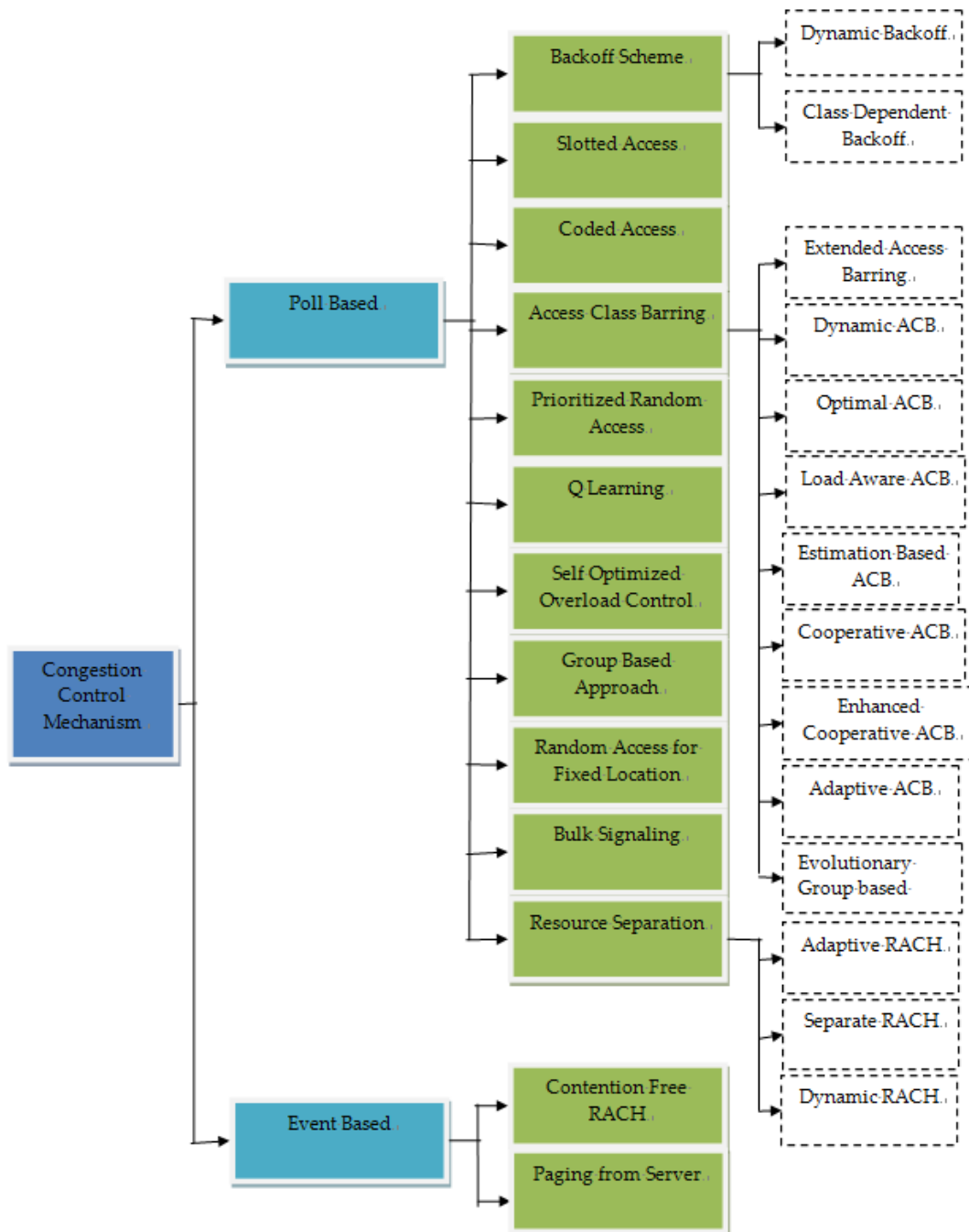

Fig.4. Non Contention Based Access

Fig.5. Taxonomy of Congestion Control Techniques

If the device fails to get random access response message during the contention access, the device is set with a longer back off interval time and starts access request on timer elapse. By backoff scheme, the performance of the network is boosted. The drawback of the technique is longer backoff which degrades the performance when both H2H and M2M access the network. In Dynamic Backoff approach [8], the backoff time divided between the M2M and H2H and modeled as the bankruptcy problem. Shapley value method is used for the backoff time split proposed in [8] is based on game theory framework. A novel Class Dependent Backoff is proposed in [9] to overcome the drawback of slotted access. Instead of applying a single backoff to

multiple classes with different arrival rate, it uses the system current load information and classification information sent by eNodeB to generate the class dependent backoff. Access class barring [4], is an improvised version of backoff scheme that comes along with the random access mechanism in the LTE/ LTE-A. ACB has two parameters such as barring factor and backoff interval. ACB rejects the connection with the class barring factor. The devices randomly pick from 0 to 1 and matches with Class Barring (CB) factor. Access is granted only if the number randomly picked is less than CB factor. Otherwise, enters into the backoff state. In Extended Access Barring [10], the probability of success is increased by blocking the low priority devices from

accessing the network. The EAB is activated by the eNodeB (eNB) with the help of metric congestion coefficient, is the probability of successful access over the contending devices. By receiving the barring page consists of barring bitmap, the user element does not contend the RACH process. Adaptive ACB [11], enhances the ACB by the providing probability factor which is efficient to bar the devices from accessing the network with the help of optimal barring factor called congestion coefficient. If the Congestion Coefficient is greater, the device will be blocked. In Cooperative Access Class Barring scheme [12], a set of barring is associated to *multiple* eNodeB and is calculated despite calculating to a particular eNodeB based on congestion level. As Devices are distributed throughout the network of various cells, the optimization problem as the min-max problem is formulated to balance the devices attached to eNodeB. The parameter like access probability inside ACB is hardwired [26]. This is not suitable in heavy load situations. This will create more retransmission which results in poor scalability. Thus, the barring factor is selected based on the number of device arrival on current slot [28]. Markov chain estimation is used to evaluate the load, based on total requests in the current slot. This helps in changing the barring factor inside eNodeB. When access requests by devices are more, the low barring probability factor helps in controlling the network overload. In Load aware ACB [34], parameters such as path loss, fading, inter – cell interface, etc., are considered while predicting the load. Optimal ACB (OACB) [25], stationary devices has an identical timing advancement which helps in reducing the packet collision. It is calculated based on propagation delay. There might be a scenario in which more devices having the same timing advancements will ultimately mess in the network. OACB is a combination of ACB and Timing advancement information. Here the barring factor is determined in each random access slot to maximize the expected number of successfully accessed devices. T Jiang et al [13] proposed the Evolutionary Game based ACB for effective improvement in minimizing the access delay. Selection decision depends on the evolutionary game of phases such as replicator dynamics and evolutionary equilibrium. Until equilibrium is achieved, the macro cell players play the game. Once achieved, the eNodeB will be ready for random access. The MTC devices are divided into different buffer zones of each base station, which helps to find the optimal ACB factor for the corresponding cell. Duan et al [14] proposed a Dynamic ACB (DACB) scheme to reduce the congestion in RACH. The goal is to manage the access attempts of the M2M devices instead of dismissing the access requests. They propose a novel algorithm to reduce both the congestion and access delay by changing the barring factor adaptively. As expected, the access delay is reduced but at the price of increased complexity since continuous monitoring is required. In Self Optimize *O*verload Control mechanism (SOOC) [15], the base station adds or reduces RACH resources when it detects an increase or decrease in RACH load by implementing

intelligent control loop. SOOC includes a composite of overload mechanism, which comprises dynamic PRACH resource allocation, PRACH resource separation, the ACB scheme, the slotted-access scheme, and the *p*-persistent scheme. Prior allocation of random access channel for different classes with the class dependent backoff [16] was examined by J.P.Cheng [16]. The method is combined with Dynamic Access Baring (DAB) to prevent the huge simultaneous access request to network. The five types of categories and they are pre-assigned certain amount of virtual resources as shown in the Fig.7. According to the statistical information about the MTC, the eNodeB will allocate the virtual resources to the corresponding class.

IN *PRACH resource separation scheme:* In this scheme, PRACH resources such as preamble sequences and RA slots are separately allocated to M2M terminals and H2H terminals. If the M2M terminals share the same PRACH resources as H2H terminals, then PRACH overload condition could significantly affect Random Access quality of H2H terminals. 3GPP standard proposed a solution of splitting the preamble resource into two distinct set pools as shown in Fig.6.
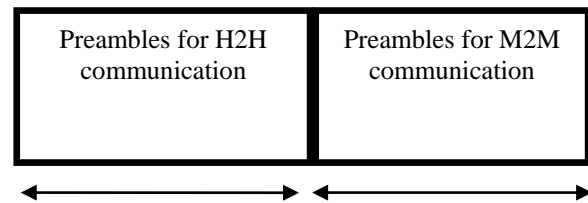

Fig.6. Preamble Splitting Into Equal Space from Available Resource
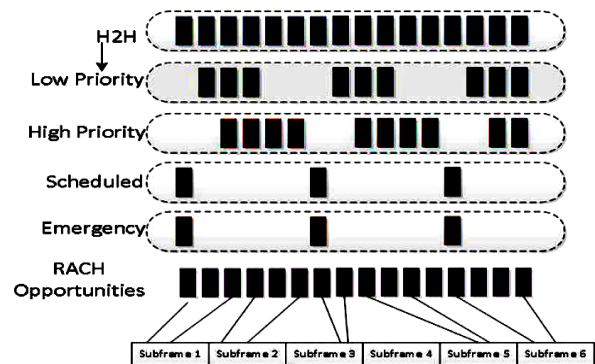

Fig.7. Virtual resource Allocation for Different Classes

One is dedicated to M2M communication and another ID dedicated to H2H communication. In Dynamic RACH partition [17], the preamble for different service requirement clusters such as Delay Tolerant and Delay Sensitive devices along with the ACB mechanism is split dynamically. The main task is to maximize the access control of devices in the cluster. For every time slots eNodeB will obtain the β value and allocate corresponding preamble for the non delay sensitive devices. In Adaptive radio resource allocation Method [18], Resource blocks (RB's) adjusted dynamically based on the availability and reserved signaling. Reserved set of RB is used for rejection message signaling purpose. The

objective of the adaptive mechanism is to maximize the available RB for MTC without affecting the H2H RB so that H2H user gets enough RB. It also mitigates RAN level congestion, which occurs between MTC device and eNB. Handling bulk messages is discussed in [6]. A group of MTC devices sends the bulk signaling messages simultaneously. Fig.8 explains a scenario of handling a common tracking area update message from a group of different MTC devices. Based on the feedback from the PDN / gateways, the connection is rejected in RAN / MME due to congestion. NAS messages are generated simultaneously by a group of MTC devices and corresponding eNodeB processes these signaling messages in bulk and forwards them to the MMEs. In Q Learning-RACH [19], transmission history is used to identify the suitable slot in a transmission frame. Transmission history consists of user list waiting for slot in the repeating frame. The weights are self organized by learning the unique slots during the contention process. The Q – Frame is updated by equation 3.

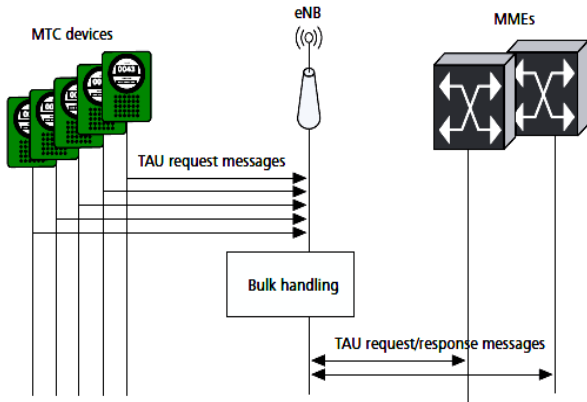$$Q^{'} = (1-\alpha)\,Q + \alpha\,r \qquad (2)$$

Fig.8. NAS Signaling Messages Can be aggregated for Bulk Handling in the NW

Where, Q is the weight, Q' is the updated weight, α is the learning rate and r is the reward. The reward is positive (+1) if it is successful and negative (-1) if it is unsuccessful. In Game theory scheme [20], three groups such as H2H, M2M and hybrid are considered for resource distribution. The strategy of the players (σ) is shown in equation 4.

$$\sigma i\,(ai) = \begin{array}{l} Pi,M \\ Pi,B \\ 1 - Pi,M - Pi,B \end{array} \left\{ \begin{array}{l} \text{if } ai = M, \\ \text{if } ai = B, \\ \text{if } ai = S \end{array} \right. \qquad (3)$$

And gain of each player is shown in equation 5.

$$gi(t) \left\{ \begin{array}{ll} 1 - \lambda & \text{if transmission succeeds,} \\ -\lambda & \text{if transmission fails,} \\ 0 & \text{no transmission.} \end{array} \right. \qquad (4)$$

And expected reward will be equation 6.

$$E[gi] = \sum_{ai=M,B} \sigma i(ai) * (PS(ai) - \lambda) \qquad (5)$$

If the preamble is selected from the pool, the successful probability is greater than the threshold λ. In [32], Q-learning is applied for eNodeB selection based on channel availability and signal strength. Reinforcement learning based eNodeB selection method balances the network load and ensures QoS by learning from previous outcome. In [21], the collision probability is reducing by the Time Alignment (TA) Information. Since the distance between machine device and eNodeB are unchanged, the Timing Alignment is fixed for this access. If the TA is matched with the eNodeB then machine device starts accessing the network. Otherwise it will backoff with the backoff Indicator (BI) received by the eNodeB as shown in Fig.9.
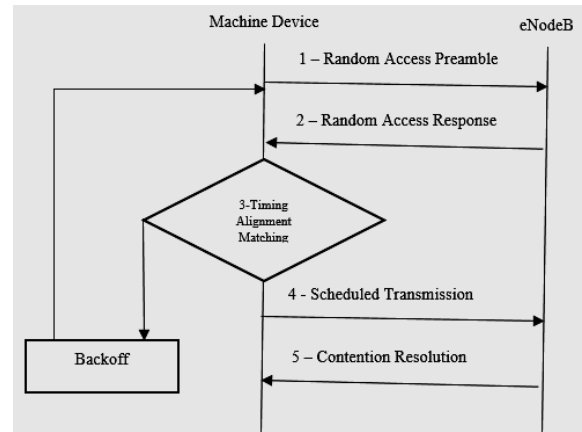

Fig.9. Random Access Procedure for Fixed Location M2M Communication

Group Based approach: grouping of machine devices according to the service requirement or geographic location are needed to solve the signaling overload and resource congestion. Instead of individual UE accessing directly the eNodeB for resource, a cluster is formed and communication happens through the cluster head leading to reduction in collision probability. K Lee et al [22] proposed the Group-based Communication mechanism based on the Location Information.

All the requests sent by the multiple user equipment in a particular window are grouped. The group formation is divided into 4 phases as follows:

i.   Phase 1: Pre Connection
ii. Phase 2: Connection Management.
iii. Phase 3: Group Establishment.
iv. Phase 4: Group Communication.

Yunhun chang et al [23] identified that recurring transmissions of redundant messages in the group leads to degradation of the network performance due to congestion in eNodeB. The author uses the coordinated random access mechanism to enhance the success probability.

Table 4. Advantages and Disadvantages of Random Access Methods

Method Name	Functions	Advantages	Disadvantages
Slotted Access	Specific RA slot	Access collision B/w M2M & H2H will reduce	Slot utilization unevenly spreaded.
Backoff Scheme	Separate Backoff	It controls the congestion in low congestion state	It severely degrades the performance of the system in peak utilization state.
	Dynamic Backoff		
	Class Dependent Backoff		
Code–Expanded RA	Code wise access	Increases RA opportunity without increasing physical resource	Associated with additional energy consumption
Access Class Barring	Individual ACB	eNB reduces load by blocking one or more user classes	Sacrifice the RACH throughput
	Extended Access Barring	Blocking of low priority UE leads to high access success probability for High priority UE	It causes long RA delay for low priority devices
	Dynamic ACB	It will Manage access attempt during overload condition instead of rejecting it.	Not compatible with delay constrained application
	Optimal ACB	Reduce the load by estimating access request in current slot	Reduce the overload but increases the access delay
	Estimation Based ACB	Improves Scalability	Cannot change the hardwired parameter of eNodeB in current slot.
	Cooperative ACB	Distributes the load between the eNB	Lack of knowledge about the number of device may attach to eNB
	Enhanced Cooperative ACB	Load will be distributed between the eNB leading to reduction in access delay	Delay in connecting to eNB remains there. Not suitable for critical applications
	Adaptive ACB	Barring factor adaption based on access intensity prediction increases the success probability rate	Access delay may get increased and difficult to predict in each RA slot
	Evolutionary Group based ACB	It shares the traffic load equally with the available eNodeBs	Initial access delay to achieve equilibriums
Prioritized Random Access	RACH Split, ACB	Combination of virtual resources with dynamic ACB alleviates overload	Loss in RACH throughput
Q Learning	Learning based eNB selection	Delay will get reduced by quick learning rate	Performance is dependent on the density of H2H
Self Optimized Overload Control	RA Separation, ACB & slotted access	Responds to varying load in real time	Increases access delay for low priority devices if the slots reaches maximum
Group Based Approach	Preamble reuse	It will reduce collision rate	Preamble mis-detection may happen
Random Access for Fixed Location	TA consideration	It will reduce collision rate considering fixed TA.	is not preferable to mobility application
Bulk Signaling	Bulk Messaging	Reduces overhead by bulk processing	batch transmission on access network will not be suitable for priority access
Resource Separation	Adaptive RACH	It will maintain the average access delay under the threshold	overhead and access delay will be more when both CBF & N-CBF operated
	Separate RACH	It will reduce congestion and interference B/w H2H	Negatively affects the M2M during high traffic load
	Dynamic RACH	Considering the system load, dynamic adjustment of resources increases the Access probability.	adjustment decision making for limited available resources
Non Contention Based Access	Individual paging	Easy access to the network during congested scenario.	Wastage of the Preamble due to reservation

The eNodeB continuously broadcasts the merged data which is received from the cell to all the members. The UE undergoes redundancy check every time. If data is available in the eNodeB, the UE will be in silent mode. The assigned signaling resources may be shared and successfully utilized within the formulated groups. Code-expanded random access [24] is based on the mechanism to increase the access resource instead of increasing the system resources such as preambles by using code – expanded random access. The virtual frame with multiple

sub frames and exchange of preamble from UE forms a codeword. Instead of sending single preamble, this will help in sending multiple preambles in each slot. In Virtual Frame, if 'L' RA slots are formed and 'M' preambles can accommodate, then the available number of code words are $[((M+1)^L) - 1]$. This increase in the contention resource gives more access to the machine type devices with less delay and more energy constraints. Congestion Aware Admission Control [30] works similar to ACB. Congestion in RAN is always high due to randomly setting up of access probability 'P', without considering the feedback of core network. CAAC considers core elements information to set probability. With this, the eNodeB rejects the new access request. It uses the proportional integrative derivative (PID) controller to reveal the congestion level of core element, such as MME by measuring the variations in the queue. Limited number of preambles is the main bottleneck of slotted aloha based RA for massive MTC in LTE networks. In [29], Distributed Queuing (DQ) is used for improving the random access during high load. It is based on M-Ary tree splitting for the control mini slots [33]. This helps in allowing the optimum number of simultaneous access and data transmissions in each slot. It maintains two virtual queues, one for collision resolution for resolving the previous collision and another for data transmission that carries the data which they contended successfully in control mini slots. This helps in the full channel utilization. The existing congestion control techniques are summarized in Table 4.

IV. CONCLUSION

M2M Communication enables Internet of Things by the autonomous communication without human meddling. As per 3GPP technical reports, contention based random access channel is one of the crucial problem in LTE supported M2M communication. The challenges such as service blockage and network overload arose due to massive access requests from active devices within the coverage area. This led to a significant overload in LTE RACH that degrades the performance of the system. From the comprehensive review, it is found that the existing congestion control techniques are not capable of managing burst of access requests due to dynamic nature of traffic load. The access opportunities are reduced and the delay is increased leading to congestion. Hence, it is difficult to adopt solutions such as game based, code expanded RA, reinforcement learning, etc., since some modification is required in existing MAC to function smoothly. The M2M is anticipated to offer advanced connectivity. Therefore, an efficient congestion control technique is required to achieve good Quality of Service for the growing number of devices/users.

REFERENCES

[1]   TS 102 689 V1.1.1 (2010), Machine-to-Machine communications (M2M); M2M service requirements, ETSI Technical Specification, FRANCE

[2]   TS 36.321 V12.5.0 (2015 - 04), Medium Access Control

(MAC), protocol specification, 3GPP Technical Specification, FRANCE

[3]   TS 36.331 V13.0.0 (2016 - 01), Radio Resource Control (RRC), Protocol Specification, 3GPP Technical Specification, FRANCE

[4]   TR 37.868 V11.0.0 (2011-09), Study on RAN Improvements for Machine-type Communications, 3GPP Technical Report, FRANCE

[5]   R2 - 104007 (2010), Pull vs Push approach for MTC, 3GPP TSG RAN WG2 #70bis, Sweden

[6]   T. Taleb and A. Kunz, "Machine type communications in 3GPP networks: potential, challenges, and solutions," in *IEEE Communications Magazine*, vol. 50, no. 3, pp. 178-184, March 2012.

[7]   TS 36.300 (2010-07) Evolved universal terrestrial radio access (E-UTRA) and Evolved universal terrestrial radio access network (E-UTRAN); Overall description, 3GPP Technical Specification, FRANCE

[8]   D. Aragao, D. Vieira, M. F. de Castro and Y. Ghamri-Doudane, "A dynamic backoff approach to control the congestion in M2M communication over LTE networks," *2015 Global Information Infrastructure and Networking Symposium (GIIS)*, Guadalajara, 2015, pp. 1-6.

[9]   X. Jian, Y. Jia, X. Zeng and J. Yang, "A novel class-dependent back-off scheme for Machine Type Communication in LTE systems," *2013 22nd Wireless and Optical Communication Conference*, Chongqing, 2013, pp. 135-140

[10]   R. G. Cheng, J. Chen, D. W. Chen and C. H. Wei, "Modeling and Analysis of an Extended Access Barring Algorithm for Machine-Type Communications in LTE-A Networks," in *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 2956-2968, June 2015.

[11]   Jihun Moon and Yujin Lim, "Adaptive Access Class Barring for Machine-Type Communications in LTE-A," *2016 Eighth International Conference on Ubiquitous and Future Networks (ICUFN)*, Vienna, 2016, pp. 398-402.

[12]   S. Y. Lien, T. H. Liau, C. Y. Kao and K. C. Chen, "Cooperative Access Class Barring for Machine-to-Machine Communications," in *IEEE Transactions on Wireless Communications*, vol. 11, no. 1, pp. 27-32, January 2012.

[13]   T. Jiang, X. Tan, X. Luan, X. Zhang and J. Wu, "Evolutionary game based access class barring for machine-to-machine communications," *16th International Conference on Advanced Communication Technology*, Pyeongchang, 2014, pp. 832-835.

[14]   S. Duan, V. Shah-Mansouri and V. W. S. Wong, "Dynamic access class barring for M2M communications in LTE networks," *2013 IEEE Global Communications Conference (GLOBECOM)*, Atlanta, GA, 2013, pp. 4747-4752.

[15]   A. Lo, Y. Law, M. Jacobsson, M. Kucharzak "Enhanced LTE-Advanced Random-Access Mechanism for Massive Machine-to-Machine (M2M) Communications", *Proc. 27th Meeting of Wireless World Research Form*, Düsseldorf, Germany, Oct 2011, pp. 1-5.

[16]   J. P. Cheng, C. h. Lee and T. M. Lin, "Prioritized Random Access with dynamic access barring for RAN overload in 3GPP LTE-A networks," *2011 IEEE GLOBECOM Workshops (GC Wkshps)*, Houston, TX, 2011, pp. 368-372.

[17]   Du, Q., Li, W., Liu, L., Ren, P., Wang, Y., Sun, L. Dynamic RACH Partition for Massive Access of Differentiated M2M Services. *Sensors* 2016, *16*, 455.

[18]   D. Kim, W. Kim and S. An, "Adaptive random access preamble split in LTE," *2013 9th International Wireless*

Communications and Mobile Computing Conference (IWCMC), Sardinia, 2013, pp. 814-819.

[19] L. M. Bello, P. Mitchell and D. Grace, "Application of Q-Learning for RACH Access to Support M2M Traffic over a Cellular Network," *European Wireless 2014; 20th European Wireless Conference*, Barcelona, Spain, 2014, pp. 1-6.

[20] Y. C. Pang, S. L. Chao, G. Y. Lin and H. Y. Wei, "Network access for M2M/H2H hybrid systems: a game theoretic approach," in *IEEE Communications Letters*, vol. 18, no. 5, pp. 845-848, 2014.

[21] K. S. Ko, M. J. Kim, K. Y. Bae, D. K. Sung, J. H. Kim and J. Y. Ahn, "A Novel Random Access for Fixed-Location Machine-to-Machine Communications in OFDMA Based Systems," in *IEEE Communications Letters*, vol. 16, no. 9, pp. 1428-1431, September 2012.

[22] K. Lee, J. Shin, Y. Cho, K. S. Ko, D. K. Sung and H. Shin, "A group-based communication scheme based on the location information of MTC devices in cellular networks," *2012 IEEE International Conference on Communications (ICC)*, Ottawa, ON, 2012, pp. 4899-4903.

[23] Y. Chang, C. Zhou and O. Bulakci, "Coordinated Random Access Management for Network Overload Avoidance in Cellular Machine-to-Machine Communications," *European Wireless 2014; 20th European Wireless Conference*, Barcelona, Spain, 2014, pp. 1-6.

[24] Thomsen, H., Pratas, N. K., Stefanović, C. and Popovski, P. (2013), Code-expanded radio access protocol for machine-to-machine communications. Trans Emerging Tel Tech, 24: 355–365.

[25] Z. Wang and V. W. S. Wong, "Optimal Access Class Barring for Stationary Machine Type Communication Devices With Timing Advance Information," in *IEEE Transactions on Wireless Communications*, vol. 14, no. 10, pp. 5374-5387, Oct. 2015.

[26] He H., Ren P., Du Q., Sun L. (2015), Estimation Based Adaptive ACB Scheme for M2M Communications. In: Xu K., Zhu H. (eds) Wireless Algorithms, Systems, and Applications. WASA 2015. Lecture Notes in Computer Science, vol 9204. Springer, Cham

[27] sesia, s., Baker, M., Toufik, I., LTE – the UMTS Long Term Evolution: From Theory to Practice, Wiley, Wiltshire., 2011.

[28] H. He, Q. Du, H. Song, W. Li, Y. Wang and P. Ren, "Traffic-aware ACB scheme for massive access in machine-to-machine networks," *2015 IEEE International Conference on Communications (ICC)*, London, 2015, pp. 617-622.

[29] L. Alonso, R. Agusti and O. Sallent, "A near-optimum MAC protocol based on the distributed queueing random access protocol (DQRAP) for a CDMA mobile communication system," in *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 9, pp. 1701-1718, Sept. 2000.

[30] A. Ksentini, Y. Hadjadj-Aoul and T. Taleb, "Cellular-based machine-to-machine: overload control," in *IEEE Network*, vol. 26, no. 6, pp. 54-60, November-December 2012.

[31] Dave Evans, The Internet of Things How the Next Evolution of the Internet Is Changing Everything, CISCO white paper, Cisco Internet Business Solutions Group (IBSG), 2011

[32] M. Hasan, E. Hossain and D. Niyato, "Random access for machine-to-machine communication in LTE-advanced networks: issues and approaches," in *IEEE Communications Magazine*, vol. 51, no. 6, pp. 86-93, June 2013.

[33] A. Laya, L. Alonso and J. Alonso-Zarate, "Is the Random Access Channel of LTE and LTE-A Suitable for M2M Communications? A Survey of Alternatives," in *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 4-16, First Quarter 2014.

[34] Song L., Zhou W., Hou Y., Gao M. (2017) Load-aware ACB Scheme for M2M Traffic in LTE-A Networks. In: Barolli L., Xhafa F., Yim K. (eds) Advances on Broad-Band Wireless Computing, Communication and Applications. BWCCA 2016. Lecture Notes on Data Engineering and Communications Technologies, vol 2. Springer

[35] Jaideep Kaur, Kamaljit Kaur,"Internet of Things: A Review on Technologies, Architecture, Challenges, Applications, Future Trends", International Journal of Computer Network and Information Security (IJCNIS), Vol.9, No.4, pp. 57-70, 2017.

Authors' Profiles

Chidanandan, V. is a Computer Science and Engineering graduate from Coorg Institute of Technlogy (CIT), Ponnampet (2010) and a Post-Graduate from Sri Jayachamarajendra College of Engineering (SJCE), Mysore (2012) from Visvesvaraya Technological University, Belagavi, Karnataka. He is currently pursuing Doctoral degree at National Institute of Technology (NIT), Puducherry, Karaikal, India.

Dr. Arun Raj Kumar, P. is an Assistant Professor in the Computer Science and Engineering Department at National Institute of Technology (NIT) Puducherry, Karaikal. He obtained his Doctoral degree in Computer Science and Engineering from National Institute of Technology (NIT) Tiruchirappalli, India. His research interests include Computer Networks, Network Security, Machine Learning and Wireless Sensor Networks. He has published 12 papers in reputed and refereed International Journals and IEEE Conferences. He is also invited reviewer for publications / Journals at International Conferences and International Journal of Network Security (IJNS), etc.