Energy Efficient MAC for Cellular-Based M2M Communications
Amin Azari and Guowang Miao
KTH Royal Institute of Technology

Abstract—In Machine-to-Machine (M2M) networks, an energy efficient scalable medium access control (MAC) is crucial for supporting massive access requests in the network. Aiming to address the challenge of efficiently managing machine access in cellular networks, we investigate the energy efficient MAC design to minimize battery power consumption in cellular-based M2M communications. We present an energy efficient MAC protocol that not only adapts contention and reservation-based protocols for M2M communications in cellular networks, but also benefits from partial clustering to handle the massive access problem. We then study the energy efficiency and access capacity of contention-based protocols and present an energy efficient contention-based protocol for intra-cluster communication of the proposed MAC, which results in huge power saving. The simulation results show that the proposed MAC protocol outperforms the others in energy saving without sacrificing much delay or throughput. Also, the lifetimes of both individual nodes and the whole M2M network are significantly extended.


I. INTRODUCTION

INTERNET of Things (IoT) enables smart devices to participate more actively in everyday life, business, industry, and health care. Among large-scale applications, cheap and widely spread machine-to-machine (M2M) communications supported by cellular networks will be one of the most important approaches for the success of IoT [1]. M2M communications, also known as machine-type communication (MTC), means the communications of machine devices without human intervention [2], which is applicable to health monitoring, smart metering, remote security, and so on [3]. Smart devices are battery driven and long battery life is crucial for them, especially for devices in remote areas, as there would be a huge amount of maintenance effort if their battery lives are short. Medium access control (MAC) design for short-range wireless networks, like wireless sensor networks, is extensively studied in literature [4]-[5]. However, regarding the particular characteristics of the M2M communications such as the massive access request, energy efficiency, and fairness, these MAC protocols are failed to address large-scale concurrent channel access in an M2M network [6]. Also it is raised by the 3GPP that such a high efficiency in current cellular network infrastructure, which is designed for human-to-human communication, is not possible [7]-[8].

Regarding the fundamental differences between M2M and human-to-human communications, many research works have been launched to understand how current infrastructure need to change to be able to provide large-scale massive access [9]-[10]. 3GPP LTE have defined research projects to support massive machine access [8]. Some other challenges in LTE networks for supporting M2M communication have been investigated in [9]-[10]. In [11], massive access management in cellular networks for satisfying delay requirements of machine nodes is considered. Also they proposed to divide machine nodes into the clusters based on the different QoS requirements. Power-efficient multiple access access protocols for a limited number of machine devices with reliability constraints in cellular networks is considered in [12]. The study of these research works has been focused on improving network performance for supporting massive access, but the energy efficiency in massive machine access has not been considered. Notice that contention-based MAC protocols for wireless sensor networks, e.g. IEEE 802.15.4, can not be used here as they are designed for short-range ad hoc or mesh type of networks [13]. In addition, their designs are not designed for enabling a massive number of devices accessing the BS at the same time. Addressing the numerous concurrent machine access with current cellular network infrastructure is still an open problem. This is the focus of the paper.

In this paper, energy efficient MAC design for M2M communications is considered. We present a large-scale energy efficient MAC protocol for cellular-based M2M communications. We will investigate the energy efficiency of contention-based protocols and devise a multi-phase protocol for intra-cluster communications in the proposed MAC. The simulation results show that the proposed MAC protocol outperforms the others in energy saving without sacrificing much delay or throughput.

The remainder of this article is organized as follows: In the next section, system model and related works are introduced. In section III, MAC design for cellular-based M2M is presented. Performance evaluation and improvement for contention-based protocols are presented in section IV. In section VI, we present the simulation results. Concluding remarks are presented in section V.

II. SYSTEM MODEL AND MEDIUM ACCESS

Consider a single cell with one base station (BS) and \( N \) static machine nodes that are uniformly distributed in the cell. The packet arrival at the machine nodes follows a Poisson distribution [14]. The packet arrival rate for each node is very low and the packet size is small, but the overall network load is dependent upon the number of active machine nodes in the cell. When a packet is generated at a machine node, it tries to access the base station. The machine nodes are battery driven and desire long battery life. The objective is to extend the battery life of the whole network, while minimizing the implementation and maintenance costs.

As M2M usually has small size packets, it make no sense to perform Ping-Pong authentication and transmit some reservation packets, while the actual data packet size is comparable with or even smaller than the reservation packet size [8]. Therefore it wastes a lot of energy if M2M devices send data directly to the BS, which implements reservation-based protocols. On the other hand, contention-based protocols are not energy efficient for massive machine access because of collisions in massive access and idle listening. To address the energy efficient massive access problem in cellular networks, we introduce partial clustering with hybrid MAC protocol in this section.
A. Partial Clustering

With clustering, the cluster head (CH) relays the messages from the cluster members to the base station. This reduces the contention for channel access between nodes and saves energy. Clustering always decreases the number of direct access requests to the base station and makes the protocol scalable; however, it may not always be energy efficient in data transmission, owing to relatively the same distance from BS to both cluster head and member. To achieve the highest energy efficiency, we propose partial clustering, in which only machine nodes far from the BS are grouped in clusters. In other words, the communication from the machine nodes to the BS might be in one hop. Simulation results will show that partial clustering outperforms full clustering and non-clustering MAC protocols in energy efficiency.

B. Frame Formation

The frame is divided into two parts, one for the communication from the cluster members to the cluster heads and the other one for sending data from the cluster heads to the BS. We treat unclustered nodes as cluster heads, where their cluster has no member. To be scalable and decrease the deployment costs, we propose to use contention-based protocols for intra-cluster communication. It has been shown that contention-based protocols outperform the others in time efficiency, i.e., delay, while they are not energy efficient, due to the collisions [12]. With clustering, the number of nodes in each cluster is relatively small and then, traffic load within each cluster is too light to cause idle listening or collisions. Therefore we can use carrier sense multiple access-collision avoidance (CSMA/CA) protocol for intra-cluster communications. In the second part of the frame, CHs have different numbers of packets to send to the BS, depending on the cluster sizes. Then to tackle the heterogeneous traffic pattern in CHs and making the communications more energy efficient, we use reservation-based protocols, e.g., dynamic time division multiple access (TDMA) where a short reservation phase is used to schedule the resources for all users, for the communications between cluster heads and the base station. This is also compatible with existing cellular standards.

C. Medium Access Design

To further improve the energy efficiency in intra-cluster communications by reducing collisions and idle listening, the former half of the frame is further split into \(n\) phases, where in each phase, a portion of the cluster members transmit their packets using the CSMA/CA protocol. The complete design works as follows:

- For a given SNR requirement at the BS, the transmission power of each node can be calculated. The nodes whose transmission power is higher than a threshold, \(T\), are grouped into clusters. Inside each cluster, the machine node with the lowest transmission power is selected as the cluster head. By feasible increase in the number of clusters, the traffic load of the clustered nodes for communication to the BS will be decreased. The choice of \(T\) determines the number of clusters in the cell and the traffic loads in each cluster.

- In the intra-cluster communication phase, each cluster head divides its members into \(n\) groups and allocates \(n\) phases to them. Each member node wakes up for data transmission only in its assigned phase.

- In the dynamic TDMA period, the cluster heads and the unclustered nodes communicate directly to the base station. In the notification phase, the base station broadcasts the reservation probability, \(q\), and the number of reservation slots. The reservation is made with probability \(q\), i.e., nodes with probability \(q\) randomly choose a reservation slot to send reservation packets and with \(1-q\) wait for the next beacon.

- In the transmission phase, nodes wake up and send packets into corresponding slots. After this phase, the unclustered nodes switch to the sleep mode and cluster heads start to listen to their cluster members in the next CSMA phase.

III. Energy Efficiency Evaluation and Improvement for Intra-Cluster Communication

When network load increases, collisions and idle listening cannot be avoided in contention-based protocols. In the following, we evaluate the performance of multi-phase CSMA/CA scheme for intra-cluster communication to avoid idle listening and as many collisions as possible. The goal is to realize a close-to-zero power-wasting MAC protocol.

A. Energy Efficient and Network Capacity

Different transmission algorithms can be used in CSMA/CA, for example 1-persistent CSMA (1P-CSMA), \(p\)-persistent CSMA, non-persistent CSMA (0P-CSMA) or RTS/CTS mechanism. As M2M has very small packet sizes, we choose 0P-CSMA. This also achieves the lowest implementation cost. In non-persistent CSMA, the machine node waits for a random amount of time after sensing a busy channel and repeats this algorithm until finding the channel idle, to transmit data. In following, we investigate the energy efficiency and access capacity of non-persistent CSMA protocol. Define the aggregated packet arrival rate of a machine node as \(g\), which includes both new arrivals and retransmitted ones. We assume that the acknowledgment packets are transmitted in a separate collision free channel to simplify the analysis. By long-term observation of the channel, one can see two different periods in channel utilization: idle period and busy period, where the transmission in the later can be either successful or unsuccessful. We consider a 2-state Markov model for idle and busy states of the channel utilization that is shown in Fig. 1. Based on this model, the limiting probabilities of the idle and busy states are the same, i.e., \(\pi_I = \pi_B = 0.5\). Also, the probability of each possible transition between states is 1. Define \(\tau_p = \tau_p + \delta_Y\), where \(\tau_p\) and \(\tau_r\) stand for transmission delay and round trip time delay from successful packet transmission to the acknowledgment packet arrival respectively. The average duration of the idle state is the average time between each pair of consecutive packets, i.e., \(B_I = 1/g\). The average duration of busy period is \(B_B = \tau_p + \delta + \gamma\), where \(\gamma\) denotes the average time at which the last interfering packet is scheduled within a transmission period that started at time 0, and is calculated as follows:

\[
F_Y(y) = pr(\text{no arrival during } \delta_d - y) = e^{-g(\delta_d-y)}
\]

\[
\rightarrow \hat{Y} = \delta - \frac{1 - e^{-g\delta_d}}{g}
\]

where \(\delta_d\) is the detection delay. Packet transmission will be successful if it starts after an idle period and no other node starts transmission after it. Then, the probability of successful packet transmission is the multiplication of time-averaged idle channel probability \((p_I)\) and no collision after that \((p_I)\), as follows:

\[
p_s = p_I \times p_t = \frac{\pi_I B_I}{\pi_I B_I + \pi_B B_B} 
\times \frac{pr(\text{no transmission in } \delta_d)}{g \tau_p + \delta_d + \delta e^{g \delta_d} + 1}
\]

(3)
The average packet delay is derived by considering the average time spent in backoffs and retransmissions before a successful packet transmission, as follows:

$$D^{cs} = \tau_s + \sum_{k=0}^{K} (1 - p_s)^k p_s k \left( \frac{1 - p_c}{1 - p_s} \theta_b + p_s \left( \frac{1 - p_c}{1 - p_s} (\theta + \tau_s) \right) \right)$$

in which, $1 - p_c$ and $p_s 1 - p_c$ are the probability of unsuccessful transmission due to a busy sensed channel and collision respectively. Also $K$ is the maximum number of times that a machine nodes tries to transmit a specific packet, and $\theta$, $\theta_b$ are the average backoff after sensing a busy channel and collision respectively. Define the power consumption in listening and transmitting modes for node $i$ as $P_i$ and $P_s$, respectively. Define the power consumption in that interval for listening to the channel, data transmission (successful or unsuccessful), and idle listening. Then one can derive the energy efficiency of node $i$ as follows:

$$EE^{CS} = \frac{gT_c p_s}{gT_c (P_c + P_t) \tau_p + P_i (\tau_c + (1 - p_i)\theta) + [1 - p_i] P_s \theta_b}$$

Under a delay constraint, one can find the threshold probability of successful transmission, $p_i^{th}$, from (4). Then the network capacity, i.e. the maximum number of sustained machine devices is derived from (3) as

$$N_{max} = p_i^{th} \ln \left( \frac{1 - p_i^{th}}{p_s \theta_c (\tau_c + \delta + \delta_d)} / \delta_d \lambda_0 \right)$$

in which, $\lambda_0$ is the packet arrival rate of each node.

### B. Multi-Phase CSMA Protocol

Idle listening in CSMA is the time when colliding nodes are back off and keeping sensing the channel, which consumes energy. As the number of nodes increases, the probability of collision increases, which results in more energy consumption. To save energy, we try to reduce the contention between nodes. The proposed multi-phase CSMA divides each contention interval into multi phases (Fig. 2). In each phase, only a portion of the nodes are permitted to compete for channel access. Before the allocated phase starts, each node keeps sleeping instead of listening. Potential arrived packets in inactive phases are buffered. With this multi-phase scheme, the probability of successful packet transmission increases, then the number of collisions and the idle listening time can be shortened significantly.

To design the proposed multi-phase protocol with the optimal number of phases, in the following we derive the energy efficiency and packet delay versus the number of phases.

Because of the page limit, we derive the performance metrics for ALOHA as an extreme case of 0P-CSMA instead of the general case to simply the analysis. The general analysis will be provided in the journal version. The probability of successful packet transmission for slotted ALOHA system with $N$ machine nodes is derived in [15] as

$$N \lambda_0 (1 - \sigma)^{N-1},$$

where $\sigma$ is the probability of packet generation in a time slot for each machine node. Using the proposed $n$-phase protocol, the probability of packet generation in each subslot of the active phase will be $n \sigma$, due to the packet buffering in $n - 1$ inactive phases. Then, the probability of successful packet transmission for the proposed protocol is calculated as follows:

$$p_s = M \sigma (1 - n \sigma)^{N-n}$$

The energy efficiency of the $n$-phase protocol is derived by considering the number of successfully transmitted bits in time interval $T_c$ and the energy consumption in that interval for data transmission (successful or unsuccessful) and backoff after collisions. Then one can derive the energy efficiency for node $i$ as follows:

$$EE^{di} = \frac{gT_c p_s}{gT_c (P_c + P_t) \tau_p + P_i (\tau_c + (1 - p_i)\theta) + [1 - p_i] P_s \theta_b}$$

It is evident that by the feasible increase in the number of phases, $n$, the probability of successful transmission and energy efficiency of the system increase. Also, as machine nodes buffer arrived packets in inactive phases, the packet delay increases in the proposed scheme. The average packet delay for proposed scheme is derived as follows:

$$D_{al} = \sum_{k=0}^{K} (1 - p_s)^k p_s (\tau_s + k \theta + (k + 1) \delta_d)$$

where $\delta_d$ is the access delay caused by dividing the contention window into $n$ phases and is calculated as:

$$\gamma \sum_{i=1}^{(n-1) T_c} \left( (1 - \sigma) (n-1) \frac{T_c}{\gamma} \right)^{-i}$$

in which, $\gamma$ is the length of each time slot and $T_c$ is the length of $n$-phase contention window. Using performance metrics in (6)-(9), one can derive the optimal number of phases for proposed multi-phase protocol.

### IV. Simulation Results

In this section, we evaluate the performance of the proposed medium access design. The parameters for M2M communications are based on [16], where they consider M2M communications for LTE. We compared the proposed MAC (pMAC)
In this paper, energy efficient MAC protocols have been studied to minimize the battery power consumption of M2M communications. We have proposed a new MAC protocol for massive machine access support in cellular networks. This solution not only adapts contention and reservation-based protocols for M2M communications, but also takes the advantage of partial clustering in cellular networks to be scalable. The proposed hybrid protocol combines the advantages of contention-based and -free medium access designs. Also, we have investigated the energy efficiency of the proposed protocols and given the optimal design. Simulation results showed that with the proposed MAC protocol, the lifetimes of both individual nodes and the whole M2M network are significantly extended.

ACKNOWLEDGMENT

The authors would like to thank Xiaohang Chen for helpful investigation of feasibility of the project.

REFERENCES