Energy-Efficient Clustering Design for M2M Communications

Peng Zhang and Guowang Miao
Royal Institute of Technology
Stockholm, Sweden
Email: {pzh,guowang}@kth.se

Abstract—Machine-to-machine (M2M) communications in cellular networks enables Internet of Things in the macro geographical regions and is a promising technique in next-generation wireless networks. In this paper, we study energy-efficient clustering design in cellular networks to maximize the overall network battery life time. We propose a new clustering design which finds the optimal number of clusters in a cellular network, and maximizes the network energy efficiency, meanwhile rules of selecting and rotating cluster heads. Different from existing clustering designs, we consider both transmission and circuit energy consumption in this paper. The network performance will be both rigorously analyzed and evaluated using simulations with respect to dead device ratio, residual energy, and network life. The simulation results show that by using the proposed clustering approach, the network battery life can be improved significantly.

Index Terms—M2M, energy efficiency, clustering.

I. INTRODUCTION

Machine-to-machine (M2M) communications in cellular networks enables direct communications between machine devices without human intervention. The objective of M2M communications is to provide comprehensive connections to a large number of devices in a wide geographical area. M2M communications are characterized by small payload, little traffic per device, and a massive number of devices [1, 2]. Cellular networks, which are mainly designed for mobile phone communications, have already been used for some M2M applications nowadays. In comparison to Human-to-Human communications, there are many incompatible factors that we need to consider to enable M2M communications, in which energy efficiency (EE) is one of the main issues.

To make the whole system last as long as possible, a system-level energy-efficient design for M2M communications is necessary. In a large M2M communications system, if every device communicates to the base station (BS) directly, there will be a lot of packet collisions because of the large number of simultaneous access requests. Clustering have been proven to be an effective way to reduce network-level energy consumptions [3]. Some devices are selected as the cluster heads (CH), which are denoted by triangles in the figure, and some as the group members (GM), which are denoted by the circles. GMs transmit their data to the respective CH, which aggregates the data and forwards it to the BS. Clustering reduces the number of devices attempting to access the BS and leads to reduced signaling overhead and saves energy of machine devices. The transmission energy may also be saved as the link distances of the machine devices to the BS are also reduced. Considering a large number of devices, a protocol to find the optimal number of CHs has been proposed for sensor networks in [4]. But it assumes all devices consume a fixed amount of energy that is independent of channel propagation and there is no CH reselection. With this design, the battery of CHs will be exhausted very fast. A simple free space propagation model are mentioned in [5] and [6], which, however, are not applicable to networks deployed in a large region. A comprehensive survey about existing clustering techniques can be found in [7]. However literatures above do not consider massive devices deployed in the big area and pay less attention to CH reselection. In addition, most of the approaches are based on heuristic designs. In our paper, we take an analytical approach and address the clustering design for energy-efficient M2M communications in cellular systems. The proposed clustering design uses a new metric, bits per Joule, in which the energy consumption consists of both transmission and circuit energy consumption. It has been shown recently [8, 9] that a tradeoff exists between transmission and circuit energy consumption in wireless communications and an optimal selection of link adaptation is necessary to minimize the total energy consumption. In this paper, we will show that a similar phenomenon also exists for the clustering design of M2M communications. To be more specific, we will address the following issues:

- Develop rules of selecting CHs and the objective is to maximize the network energy efficiency that is measured by bits per Joule, where the energy consumption consists of both transmission and circuit energy.
- Find the optimal number of clusters in each cell so that the overall network energy efficiency is maximized.
- Develop criteria to rotate the roles of CHs so that the overall network battery life can be maximized.

The rest of the paper is organized as follows. The system model is described in Section II. Section III derives the optimal cluster size. Section IV presents how to form clusters. Numerical and graphic results are discussed in Section V. Section VI concludes the paper.

II. SYSTEM MODEL

We consider energy-efficient communications of a single-cell cellular network to maximize the amount of data sent with a given amount of energy by the machine devices. We will formulate the problem of energy-efficient communications and obtain the optimal number of CHs to make the system energy efficiency (EE), \( U \), maximum. Assume there are \( n \) M2M devices, and let \( p \) denote the probability for a device becoming a CH, the number of CHs will be \( np \). We will find the optimal \( p \) to maximize \( U \). Consider devices deployed in a single cell with a radius \( R_c \). The BS is located at the cell center. The devices are homogeneously distributed within the cell according to a spatial Poisson process of intensity \( \lambda \) in a 2-dimensional space. Each device chooses the CH that is the closest.

We want to find which value of \( p \) maximizes EE:

\[
U(p) = \frac{R(p)}{P(p)}, \quad (1)
\]
where $R$ and $P$ are the overall system data rate and power consumption respectively. We assume the communications within clusters experience a short distance propagation model $PL(a)$ and the corresponding shadow effect $Z_a$. Meanwhile, for the communications between the CH and the BS, we use a long distance propagation model $PL(e)$ and the corresponding shadow effect $Z_e$. $PL$ and $Z$ are in dB values. Let $P_t(GM)$ denote the transmission power of GM and $P_t(CH)$ denote the transmission power of CH. The received power for each device is

$$P_j(GM) = P_t(GM) \times 10^{-\frac{PL(a)_j + Z_a}{10}}, \quad (2)$$

$$P_K(CH) = P_t(CH) \times 10^{-\frac{PL(e)_K + Z_e}{10}}, \quad (3)$$

where $J$ and $K$ indicate the $J^{th}$ GM and the $K^{th}$ CH respectively. The data rate for each CH and GM can be written as:

$$R_j(GM) = \frac{w_2}{M_K + 1} \log(1 + \frac{P_j(GM)}{N_0w_2}), \quad (4)$$

$$R_K(CH) = w_{2} \log(1 + \frac{P_K(CH)}{N_0w_2}), \quad (5)$$

where $w_2$ is the bandwidth for the communications from the CHs to the BS and $w_2$ is the bandwidth for intra-cluster communications. To simplify the analysis, the bandwidth is assumed to be equally divided among all accessing devices. $M_K$ is the number of GMs for the $K^{th}$ CH. $\frac{w_2}{M_K + 1}$ is the bandwidth which is allocated to each member in the $K^{th}$ cluster. $N_0$ is thermal noise power. Moreover let the circuit power $P_c$ represent the average consumption of electronics, the overall power consumption for GM and CH are

$$P_o(GM) = P_c + P_t(GM), \quad (6)$$

$$P_o(CH) = P_c + P_t(CH). \quad (7)$$

The number of CHs and GMs in the whole system are $np$ and $n(1-p)$ respectively. Therefore the overall system power consumption is

$$P(p) = nP_c + pnP_t(CH) + (1-p)nP_t(GM). \quad (8)$$

Meanwhile, the throughput for the whole system is defined as

$$R(p) = \sum_{K=1}^{np} R_K(CH) + \sum_{J=1}^{(1-p)n} R_j(GM). \quad (9)$$

### III. Optimal Cluster Size

Based on (8) and (9), the system power consumption depends on different $p$. The data rate of each device is based on its location and bandwidth. If we want to compute the system throughput by using (4) (5) and (9), we need to figure out the path loss and the bandwidth for each device, which equals to knowing the value of $M_K$. In the following, we derive a reasonable expected value for the path loss of each device and the size of each cluster.

We have assumed the devices are distributed according to a homogeneous spatial Poisson process. The number of devices in a circle area with radius $R_c$ is a Poisson random variable, $N_c$ with mean $\lambda s$ where $S = \pi \times R_c^2$ and $\lambda$ is intensity of this Poisson process. Let’s assume that for a particular realization of the process, there are $n$ devices in this area. As mentioned before, the probability of one device becoming a CH is $p$. Therefore, on average, there are $np$ CHs. Also, the CHs and GMs are distributed as per independent homogeneous spatial Poisson processes. For the communications between the CH and the BS, we use a long distance propagation model $PL(e)$ for the communications between the CH and the BS, we use a long distance propagation model $PL(e)$ and the corresponding shadow effect $Z_e$. $PL$ and $Z$ are in dB values. Let $P_t(GM)$ denote the transmission power of GM and $P_t(CH)$ denote the transmission power of CH. The received power for each device is

$$P_j(GM) = P_t(GM) \times 10^{-\frac{PL(a)_j + Z_a}{10}}, \quad (2)$$

$$P_K(CH) = P_t(CH) \times 10^{-\frac{PL(e)_K + Z_e}{10}}, \quad (3)$$

where $J$ and $K$ indicate the $J^{th}$ GM and the $K^{th}$ CH respectively. The data rate for each CH and GM can be written as:

$$R_j(GM) = \frac{w_2}{M_K + 1} \log(1 + \frac{P_j(GM)}{N_0w_2}), \quad (4)$$

$$R_K(CH) = w_{2} \log(1 + \frac{P_K(CH)}{N_0w_2}), \quad (5)$$

where $w_2$ is the bandwidth for the communications from the CHs to the BS and $w_2$ is the bandwidth for intra-cluster communications. To simplify the analysis, the bandwidth is assumed to be equally divided among all accessing devices. $M_K$ is the number of GMs for the $K^{th}$ CH. $\frac{w_2}{M_K + 1}$ is the bandwidth which is allocated to each member in the $K^{th}$ cluster. $N_0$ is thermal noise power. Moreover let the circuit power $P_c$ represent the average consumption of electronics, the overall power consumption for GM and CH are

$$P_o(GM) = P_c + P_t(GM), \quad (6)$$

$$P_o(CH) = P_c + P_t(CH). \quad (7)$$

The number of CHs and GMs in the whole system are $np$ and $n(1-p)$ respectively. Therefore the overall system power consumption is

$$P(p) = nP_c + pnP_t(CH) + (1-p)nP_t(GM). \quad (8)$$

Meanwhile, the throughput for the whole system is defined as

$$R(p) = \sum_{K=1}^{np} R_K(CH) + \sum_{J=1}^{(1-p)n} R_j(GM). \quad (9)$$

Let $D_1$ be a random variable that denotes the length of the segment from a device located at $(x, y)$ to the BS. A device could be located at any $(x, y)$ point on the terrain with uniform intensity, the probability density function of its location is constant $\frac{1}{s}$. So the average distance to the BS is

$$E[D_1|N = n] = \frac{1}{2\sqrt{\lambda p}}. \quad (12)$$

Based on the path loss model in Table I, the expectation for $PL(e)$ and $PL(a)$ will be

$$E[PL(e)|N = n] = 128.1 + 37.6 \log(\frac{R_c}{1500}), \quad (14)$$

$$E[PL(a)|N = n] = 38.5 + 20 \log(\frac{1}{2\sqrt{\lambda p}}). \quad (15)$$

So far, we can derive the expectation for the data rate of each CH using (5) and (14)

$$E[R(CH)|N = n] = \frac{w_2}{np} \log(1 + \frac{P_t(CH)np}{N_0w_2} \frac{PL(e)|N = n + 1}{10}). \quad (16)$$

The average data rate of each GM is

$$E[R(GM)|N = n] = \frac{w_2}{E[M_K|N = n] + 1} \log(1 + \frac{P_t(GM)E[M_K|N = n] + 1}{N_0w_2} \frac{PL(a)|N = n + 1}{10}). \quad (17)$$

Combining the equations (8) (9) (16) and (17), meanwhile calculating the value of $PL(a)$ and $PL(e)$ with average value
of $D_1$ and $D_2$, moreover including the value of $Z_a$ and $Z_e$, the expected network EE can be derived as

$$E[U(p)|N=n] = E[R(p)|N=n]$$

$$= \sum_{p} w_1 \log(1 + \frac{P_i(CH)}{R}) + C$$

$$= \sum_{p} w_1 \log(1 + \frac{P_i(CH)}{R}) + C + D_p$$

where $A = \frac{P_i(CH)}{R}$, $B = \frac{P_i(GM)}{R}$, $C = \lambda R_0^2 (P_c + P_l(GM))$, $D = \lambda R_0^2 (P_c(CH) - P_l(GM))$ and $E = w_2 \lambda R_0^2$. Similar to EE function in [8], it can be easily proven that $E[U(p)]$ is strictly quasi-concave in $p$ and a local maximum is also globally optimal. We can let $E[U(p)] = 0$ and obtain the optimal value of $p$, $p^*$, and the corresponding $E[U(p^*)]$. Fig. 1 demonstrates the relationship between $p$ and $E$ function.

IV. CLUSTER HEAD SELECTION AND ROTATION

The $np^*$ clusters should be distributed in the cell as evenly as possible. For the devices in each cluster, we select one device as CH. Which one will be CH is based on the proposed Cost Function.

A. Cost Function in the Initial CH Selection

The Cost Function models the overall energy cost if a node is selected as the CH. It takes three factors into account, normalized intra-cluster communication cost ($f_i$), normalized inter-cluster communications cost ($F_i$) and normalized energy having been consumed by each device ($E_i$).

$$COST_i = f_i + F_i + \omega E_i$$

where $\omega$ is a weight coefficient. Base on long distance propagation path loss, the inter-cluster cost, $F_i$, is

$$F_i = \frac{10^{-\frac{p_i(CH)}{10} + Z_e}}{10^{-\frac{P_i(CH)}{10} + Z_e}}$$

$$= \frac{10^{-\frac{p_i(CH)}{10} + Z_e}}{10^{-\frac{P_i(CH)}{10} + Z_e}}$$

where $i$ is the $i^{th}$ device in the cluster, $D_i$ is the distance between the $i^{th}$ device and BS, which can be easily estimated based on cell reference signals. $F_i$ indicates the normalized energy cost (consumed by path loss) for member $i$ to send the data of all cluster members to the BS.

The intra-cluster cost is defined as:

$$f_i = \frac{\sum_{j=1}^{M_k+1} 10^{-\frac{P_i(CH)}{10} + Z_e}}{\sum_{j=1}^{M_k+1} 10^{-\frac{P_i(CH)}{10} + Z_e}}$$

where $D_{ij}$ is the distance between the $i^{th}$ device and others in the same cluster. $f_i$ is the normalized energy cost for cluster members to send the data to member $i$. Initially, every device has full energy, we let $E_i$ equal to zero. In each cluster, the device with the minimum COST means the device using the least consumption for both intra- and inter-cluster communications. Therefore it will be selected as the CH.

B. Cost Function in CH Reselection

The CHs consume much more energy than their GMs, and their energy will be exhausted faster than normal devices. If some devices are always kept as the CHs, they will die much sooner than other devices. To solve the issue, we rotate the role of CHs among all cluster members. After a certain time duration $T$, the amounts of energy consumptions of all cluster members are collected. We normalize each device’s consumption in its cluster group and obtain $E_i$ which indicates the normalized energy having been consumed.

$$E_i = \frac{e_i}{\sum_{j=1}^{M_k+1} e_j}$$

where $e_i$ is the total energy having been consumed of the $i^{th}$ member. In each cluster group, the devices with the minimum COST will be the cluster heads in next round.

V. SIMULATION RESULTS AND DISCUSSION

The system parameters are listed in Table I. 6000 machine devices are randomly dropped in the cell. The optimal $p$ is found to be 0.1577. Therefore, the optimal number of clusters is 947 and the remaining number of devices will be GMs. Fig. 2 presents the energy consumption results for different $p$ between the $250^{th}$ day and the $300^{th}$ day. In the figure, the $p^*$ line always performs the best. Moreover, Non-cluster design (NonCluster) consumes the highest energy. As dead devices increasing, $p_2 = 0.1$ will become more close to $p^*$ so that it’s performance is better than $p_2 = 0.3$. The same situation happen between $p_1 = 0.05$ and $p_2 = 0.5$.

Fig. 3 compares the number of dead devices with the different clustering designs. Non-Cluster design, which is the black dotted line, has the worst performance, almost all of
In this paper, we provide an EE cluster design for M2M communications. The proposed design has the advantages for fast deployment with low computing complexity. Specifically, an appropriate number of clusters can be calculated easily, while offering a significant performance improvement in terms of energy saving. We have proposed simple cluster head selection and reselection rules. Based on simulation results, the proposed clustering technique can extend the overall network battery life by about 50%.

VI. CONCLUSION

In this paper, we provide an EE cluster design for M2M communications. The proposed design has the advantages for fast deployment with low computing complexity. Specifically, an appropriate number of clusters can be calculated easily, while offering a significant performance improvement in terms of energy saving. We have proposed simple cluster head selection and reselection rules. Based on simulation results, the proposed clustering technique can extend the overall network battery life by about 50%.

REFERENCES

[1] 3GPPTR. Service improvement for machine type communication. (22.368), December 2012.