ABSTRACT

A three-layer framework is proposed for mobile data collection in wireless sensor networks, which contains the sensor layer, cluster head layer, and mobile collector (called SenCar) layer. The framework works distributed load balanced clustering and multiple data uploading, which is called as LBC-MIMO. The objective is to achieve good scalability, long network lifetime and low data collection latency. At the sensor layer, a distributed load balanced clustering (LBC) algorithm is proposed for sensors to self-organize themselves into clusters. In contrast to existing clustering methods, our scheme generates multiple cluster heads in each cluster to balance the work load and facilitate dual data uploading. At the cluster head layer, the inter-cluster transmission range is carefully chosen to guarantee the connectivity among the clusters. Multiple cluster heads within a cluster cooperate with each other to perform energy-saving inter-cluster communications. Through inter-cluster transmissions, cluster head information is forwarded to SenCar for its moving trajectory planning. At the mobile collector layer, SenCar is equipped with two antennas, which enables two cluster heads to concurrently upload data to SenCar in each time by using multi-user multiple-input and multiple-output (MU-MIMO) technique. The trajectory planning for SenCar is optimized to fully use dual data uploading capability by properly choosing polling points in each cluster. By visiting each selected polling point, SenCar can efficiently gather data from cluster heads and transport the data to the static data sink. Extensive simulations are conducted to evaluate the effectiveness of the proposed LBC-MIMO scheme. The results show that when each cluster has at most two cluster heads, LBC-MIMO achieves over 50 percent energy-saving per node and 60 percent energy saving on cluster heads comparing with data collection through multi-hop relay to the static data sink, and 20 percent shorter data collection time compared to modern mobile data gathering.

KEYWORDS:
Wireless sensor network (WSNs), data collection, load balanced multi-head clustering, multiple data uploading, multi-user multi-input multi-output (MU-MIMO), polling points, mobility of control.

INTRODUCTION
The proliferation of the implementation for low-cost, low-power, multifunctional sensors has made wireless sensor networks (WSNs) a prominent data gathering paradigm for extracting local measures of interest. In such application, sensor are generally densely deployed and randomly scattered over a sensing field and left unattended after being deployed, which makes it difficult to recharge or replace their batteries. After sensor form into autonomous organizations, those sensors near the data sink typically deplete their
batteries much faster than others due to more relaying traffic. When sensors around the data sink deplete their energy, networks connectivity and coverage may not be guaranteed. Due to these limitations, it is crucial to design an energy-efficient data gathering scheme that consumes energy uniformly across the sensing field to achieve long network lifetime. Furthermore, as sensing data in some application are time-sensitive, data gathering scheme may be required to be performed within a specified time frame?

Fig.1. Structure of mobile computing

DIFFERENT TYPES OF DEVICES USED FOR THE MOBILE COMPUTING:
1. Personal digital assistant/enterprise digital assistant
2. Smartphones
3. Tablet computers
4. Netbooks
5. Ultra-mobile PCs
6. Wearable computers

Applications of Wireless sensor networks:
Sensor nodes are used in a variety applications which required constant monitoring and detection of specific events.

1. Military applications:
Military applications of sensors nodes include battlefield surveillance and monitoring guidance system of intelligent missiles and detection of attacks by weapons of mass destruction such as chemical, biological or nuclear.

2. Vehicles:
Tomorrow’s cars will comprise many wireless communication systems and mobility aware applications. Music, news, road conditions, weather reports, and other broadcast information are received via digital audio broadcasting (DAB) with 1.5M-bits/s. For personal communication, a global system for mobile communications (GSM) phone might be available offering voice and data connectivity with 384 k-bits/s. For remote areas satellite communication can be used, while the current position of the car is determined via global positioning system (GPS).

Additionally, cars driving in the same area build a local ad-hoc network for fast information exchange in emergency situations or to help each other keeping a safe distance. In case of an accident, not only will the airbag be triggered, but also an emergency call to a service provider informing ambulance and police. Cars with this technology are already available. Future cars will also inform other cars about accidents via the ad hoc network to help them slowdown in time, even before a driver can recognize the accident. Buses, trucks, and train are already transmitting maintenance and logistic information to their home base, which helps to improve organization (fleet management), and thus save time and money.

3. Emergency:
Just imagine the possibilities of an ambulance with a high quality wireless connection to a hospital. After an accident, vital information about damaged persons can be sent to the hospital immediately. There, all necessary steps for this particular type of accident can be prepared or further specialists can be consulted for an early diagnosis. Furthermore, wireless networks are the only means of communication in the case of natural disasters such as hurricanes or earthquakes.

4. Business:
Today’s typical traveling salesman needs direct access to the company’s database: to ensure that the files on his or her laptop reflect the actual state, to support the company to keep track of all activities of their
traveling employees, to keep databases consistent etc., with wireless access, the laptop can be turned into a true mobile office.

**BENEFITS OF MOBILE COMPUTING:**
- Improve business productivity by streamlining interaction and taking advantage of immediate access
- Reduce business operations costs by increasing supply chain visibility, optimizing logistics and accelerating processes
- Strengthen customer relationships by creating more opportunities to connect, providing information at their fingertips when they need it most
- Gain competitive advantage by creating brand differentiation and expanding customer experience
- Increase work force effectiveness and capability by providing on-the-go access
- Improve business cycle processes by redesigning work flow to utilize mobile devices that interface with legacy applications

**EXISTING SYSTEM:**
- Several approaches have been proposed for efficient data collection in the literature. Based on the focus of these works, we can roughly divide them into three categories.
- The first category is the enhanced relay routing, in which data are relayed among sensors. Besides relaying, some other factors, such as load balance, schedule pattern and data redundancy, are also considered.
- The second category organizes sensors into clusters and allows cluster heads to take the responsibility for forwarding data to the data sink. Clustering is mainly useful for applications with scalability requirement and is very effective in local data aggregation since it can reduce the collisions and balance load among sensors.
- The third category is to make use of mobile collectors to take the burden of data routing from sensors.

**DISADVANTAGES OF EXISTING SYSTEM:**
- In relay routing schemes, minimizing energy consumption on the forwarding path does not necessarily prolong network lifetime, since some critical sensors on the path may run out of energy faster than others.
- In cluster-based schemes, cluster heads will inevitably consume much more energy than other sensors due to handling intra-cluster aggregation and inter-cluster data forwarding.
- Though using mobile collectors may alleviate non-uniform energy consumption, it may result in unsatisfactory data collection latency.

Based on these observations, in this we propose a three-layer mobile data collection framework, named as **Load Balanced Clustering with MIMO Uploading Techniques** (LBC-MIMO) the main motivation is to utilize distributed clustering for scalability, to employ mobility for energy saving and uniform energy consumption, and to exploit Multiple-User Multiple-Input and Multiple-Output (MU-MIMO) technique for concurrent data uploading to shorten latency.

**PROPOSED SYSTEM:**
- We propose a three-layer mobile data collection framework, named **Load Balanced Clustering and Multiple Data Uploading** (LBC-MIMO).
- The main motivation is to utilize distributed clustering for scalability, to employ mobility for energy saving and uniform energy consumption, and to exploit Multi-User Multiple-Input and Multiple-Output (MU-MIMO) technique for concurrent data uploading to shorten latency. The main contributions of this work can be summarized as follows.
First, we propose a distributed algorithm to organize sensors into clusters, where each cluster has multiple cluster heads.

Second, multiple cluster heads within a cluster can collaborate with each other to perform energy efficient inter-cluster transmissions.

Third, we deploy a mobile collector with two antennas (called SenCar in this paper) to allow concurrent uploading from two cluster heads by using MU-MIMO communication. The SenCar collects data from the cluster heads by visiting each cluster. It chooses the stop locations inside each cluster and determines the sequence to visit them, such that data collection can be done in minimum time.

ADVANTAGES OF PROPOSED SYSTEM:

- In contrast to clustering techniques proposed in previous works, our algorithm balances the load of intra-cluster aggregation and enables dual data uploading between multiple cluster heads and the mobile collector.
- Different from other hierarchical schemes, in our algorithm, cluster heads do not relay data packets from other clusters, which effectively alleviates the burden of each cluster head. Instead, forwarding paths among clusters are only used to route small-sized identification (ID) information of cluster heads to the mobile collector for optimizing the data collection tour.
- Our work mainly distinguishes from other mobile collection schemes in the utilization of MUMIMO technique, which enables dual data uploading to shorten data transmission latency. We coordinate the mobility of SenCar to fully enjoy the benefits of dual data uploading, which ultimately leads to a data collection tour with both short moving trajectory and short data uploading time.

SYSTEM ARCHITECTURE:

An overview of LBC-MIMO framework is depicted in Fig.2, which consist of three layers: sensor layer, cluster head layer and Sencar layer. The basic layer is the bottom and basic layer. For generality, we do not make any assumption on sensor distribution or node capability, such as local awareness. Each sensor communicated with only its neighbor, sensor are self-organized themselves into clusters. Each sensor decides to be either a cluster head or cluster member in a distributed manner. Sensors with high residual energy would become cluster heads or each cluster has at most \( M \) cluster heads, where \( M \) says number of cluster heads in each cluster. The multiple cluster heads within a cluster are called cluster head Group (CHG) with each cluster head being the peer of other others. The algorithm constructs clusters such that each sensor in a cluster is one hop a way from at least one cluster head. In the case that a sensor may covered by multiple cluster heads in CHG for load balancing .to avoid collisions during data aggregation the CHG adopts time-division-multiple-access (TDMA) based technique to coordinate communications between sensor nodes. Right after the cluster heads are elected, the nodes synchronize their local clocks via beacon messages. For example, all the nodes in a CHG could adjust their local clocks based on that of the node with the highest residual energy. After local synchronization is done, an existing scheduling scheme can be adopted to gather data from cluster...
members. Note that only intra-cluster synchronization is needed here because data are collected via SenCar. In the case of imperfect synchronization, some hybrid techniques to combine TDMA with contention-based access protocols (Carrier Sense Multiple Access (CSMA)) that listen to the medium before transmitting are required. For example, hybrid protocols like Z-MAC can be utilized to enhance the strengths and offset the weaknesses of TDMA and CSMA. Upon the arrival of SenCar, each CHG uploads buffered data via MU-MIMO communications and synchronizes its local clock with the global clock on SenCar via acknowledgement messages. Finally, periodical re-clustering is performed to rotate cluster heads among sensors with higher residual energy to avoid draining energy from cluster heads.

The cluster head layer consists of all the cluster heads. Inter-cluster forwarding is only used to send the CHG information of each cluster to SenCar, which contains an identification list of multiple cluster heads in a CHG. Such information must be sent before SenCar departs for its data collection tour. Upon receiving this information, SenCar utilizes it to determine where to stop within each cluster to collect data from its CHG. To guarantee the connectivity for inter-cluster communication, the cluster heads in a CHG can cooperatively send out duplicated information to achieve spatial diversity, which provides reliable transmissions and energy saving. Moreover, cluster heads can also adjust their output power for a desirable transmission range to ensure a certain degree of connectivity among clusters.

The top layer is the SenCar layer, which mainly manages mobility of SenCar. There are two issues to be addressed at this layer. First, we need to determine the positions where SenCar would stop to communicate with cluster heads when it arrives at a cluster. In LBC-MIMO, SenCar communicates with cluster heads via single-hop transmissions. It is equipped with two antennas while each sensor has a single antenna and is kept as simple as possible. The traffic pattern of data uploading in a cluster is many-to-one, where data from multiple cluster heads converge to SenCar. Equipped with two receiving antennas, each time SenCar makes dual data uploading whenever possible, in which two cluster heads can upload data simultaneously. By processing the received signals with filters based on channel state information, SenCar can successfully separate and decode the information from distinct cluster heads.

To collect data as fast as possible, SenCar should stop at positions inside a cluster that can achieve maximum capacity. In theory, since SenCar is mobile, it has the freedom to choose any preferred position. However, this is infeasible in practice, because it is very hard to estimate channel conditions for all possible positions. Thus, we only consider a finite set of locations. To mitigate the impact from dynamic channel conditions, SenCar measures channel state information before each data collection tour to select candidate locations for data collection. We call these possible locations SenCar can stop to perform concurrent data collections polling points. In fact, SenCar does not have to visit all the polling points. Instead, it calculates some polling points which are accessible and call them selected polling points. In addition, we need to determine the sequence for SenCar to visit these selected polling points such that data collection latency is minimized. Since SenCar has pre-knowledge about the locations of polling points, it can find a good trajectory by seeking the shortest route that visits each selected polling point exactly once and then returns to the data sink.

The proposed framework aims to achieve great energy saving and shortened data collection latency, which has the potential for different types of data services. Although traditional designs of WSNs can support low-rate data services, more and more sensing applications nowadays require high-definition pictures and audio/video recording, which has become an overwhelming trend for next generation sensor designs. For example, in the scenario of military defence, sensors deployed in reconnaissance missions need to transmit back high-definition images to identify hostile units. Delays in gathering sensed data
may not only expose sensors or mobile collector to enemy surveillance but also depreciate the time value of gathered intelligence. Using MU-MIMO can greatly speed up data collection time and reduce the overall latency. Another application scenario emerges in disaster rescue. For example, to combat forest fire, sensor nodes are usually deployed densely to monitor the situation. These applications usually involve hundreds of readings in a short period (a large amount of data) and are risky for human being to manually collect sensed data. A mobile collector equipped with multiple antennas overcomes these difficulties by reducing data collection latency and reaching hazard regions not accessible by human being. Although employing mobility may elongate the moving time, data collection time would become dominant or at least comparable to moving time for many high-rate or densely deployed sensing applications. In addition, using the mobile data collector can successfully obtain data even from disconnected regions and guarantee that all of the generated data are collected.

IMPLEMENTATION

MODULES:

- Initialization Phase
- Status Claim
- Cluster Forming
- Synchronization among Cluster Heads

MODULES DESCRIPTION:

Initialization Phase
In the initialization phase, each sensor acquaints itself with all the neighbors in its proximity. If a sensor is an isolated node (i.e., no neighbor exists), it claims itself to be a cluster head and the cluster only contains itself. Otherwise, a sensor, say, $s_i$, first sets its status as “tentative” and its initial priority by the percentage of residual energy. Then, $s_i$ sorts its neighbors by their initial priorities, and picks neighbors with the highest initial priorities, which are temporarily treated as its candidate peers. We denote the set of all the candidate peers of a sensor by $A$. It implies that once $s_i$ successfully claims to be a cluster head, its up-to-date candidate peers would also automatically become the cluster heads, and all of them form the CHG of their cluster. $s_i$ sets its priority by summing up its initial priority with those of its candidate peers. In this way, a sensor can choose its favorable peers along with its status decision.

Status Claim
In the second module, each sensor determines its status by iteratively updating its local information, refraining from promptly claiming to be a cluster head. We use the node degree to control the maximum number of iterations for each sensor. Whether a sensor can finally become a cluster head primarily depends on its priority. Specifically, we partition the priority into three zones by two thresholds, $th$ and $tm$ ($th > tm$), which enable a sensor to declare itself to be a cluster head or member, respectively, before reaching its maximum number of iterations. During the iterations, in some cases, if the priority of a sensor is greater than $th$ or less than $tm$ compared with its neighbors, it can immediately decide its final status and quit from the iteration.

We denote the potential cluster heads in the neighborhood of a sensor by a set $B$. In each iteration, a sensor, say, $s_i$, first tries to probabilistically include itself into $s_i.B$ as a tentative cluster head if it is not in already. Once successful, a packet includes its node ID and priority will be sent out and the sensors in the proximity will add $s_i$ as their potential cluster heads upon receiving the packet. Then, $s_i$ checks its current potential cluster heads. If they do exist, there are two cases for $s_i$ to make the final status decision, otherwise, $s_i$ would stay in the tentative status for the next round of iteration.

Cluster Forming
The third module is cluster forming that decides which cluster head a sensor should be associated with. The criteria can be described as follows: for a sensor with tentative status or being a cluster member, it would randomly affiliate itself with a cluster head among its candidate peers for load balance purpose. In the rare case that there is no cluster head among the candidate
peers of a sensor with tentative status, the sensor would claim itself and its current candidate peers as the cluster heads.

In case a cluster head is running low on battery energy, re-clustering is needed. This process can be done by sending out a re-clustering message to all the cluster member. Cluster members that receives this message switch to the initialization phase to perform a new round of clustering.

**Synchronization among Cluster Heads**

To perform data collection by TDMA techniques, intracluster time synchronization among established cluster heads should be considered. The fourth phase is to synchronize local clocks among cluster heads in a CHG by beacon messages. First, each cluster head will send out a beacon message with its initial priority and local clock information to other nodes in the CHG. Then it examines the received beacon messages to see if the priority of a beacon message is higher. If yes, it adjusts its local clock according to the timestamp of the beacon message. In our framework, such synchronization among cluster heads is only performed while SenCar is collecting data. Because data collection is not very frequent in most mobile data gathering applications, message overhead is certainly manageable within a cluster.

**EXAMPLES OF LBC-MIMO ALGORITHM:**

Example to show the impact of some parameters on the clustering results in fig.4. Fig.4a shows a random arrangement of 80 sensors on 100×100 area. The connectivity identified link between any two neighboring nodes. Fig. 4b displays the result of LBC with $M$ set 2. Since the priority of a sensor is the sum of its initial priority and those of its $M-1$ candidate peers, the two thresholds, $th$ and $tm$ are proportionally set to $M\times0.9$ and $M\times0.3$, respectively fig.4b sensors self-organized into 6 clusters each having two clusters heads shown blue. The link between each cluster head and its members, which are shown in grey, indicate the final association pattern of the sensors. And link shown in blue, represent the connectivity between two cluster heads in a CHG.

Fig.4c, we keep $M$ unchanged and vary the setting of $th$ and $tm$ to be $M\times0.75$ and $M\times0.4$ respectively. These relaxed thresholds imply that there is more opportunity a sensor to immediately claim itself and its candidate peers to be cluster heads, and there is also more possibility for a sensor to quickly “retire” to be a cluster member. This helps a sensor effectively reduce the computational iterations, however, it leads to more clusters in the networks. For instance. There are seven cluster formed fig.4c one more compared to fig.4b with more tightened thresholds.

**Fig.3. Example of LBC-MIMO algorithm with M=2**

**Fig.4. Example of LBC-MIMO Different Parameter Settings.**

(a) Network Configuration with 80 sensors distributed in a 100 × 100 area. (b) LBC Clustering with $M = 2$, $th = M \times 0.9$, and $tm = M \times 0.3$.
Finally we vary $M$ to investigate its impact on cluster on clustering. The results shown fig.4d further $M$ is reset to 4, and $th$ and $tm$ are still proportionally set to $M \times 0.9$ and $M \times 0.3$ respectively. A totally of five clusters is to is observed when $M=4$, less than the result in fig.4b with $M=2$.

![Fig.4d. LBC clustering with M×0.9 and M×0.3](image)

**MU-MIMO UPLOADDING:**

We jointly consider the selection of the schedule pattern and selected polling points for the corresponding scheduling pairs, aiming at achieving the maximum sum of MIMO uplink capacity in a cluster.

Once the selected polling points for each cluster are chosen, Sencar can finally determine its trajectory. The moving time on the trajectory can be reduced by a proper visiting sequence of selecting polling points. Since Sencar departs from the data sink and also needs to return the collected data to it, the trajectory of Sencar is a route that visits each selected polling point once. This is the well-known travelling sales person problem (TSP). Since Sencar has knowledge about the location of polling points, it can utilize approximate or heuristic algorithm for the TSP problem to find the shortest moving trajectory among selected polling points, e.g. the nearest neighbor algorithm.

**PERFORMANCE ANALYSIS:**

We evaluate the performance of our framework and compare it with other schemes. Since the main focus of this paper is to explore different choices of data collection schemes, for fair comparison, we assume all the schemes are implemented under the same duty-cycling MAC strategy.

We develop a simulator in MATLAB and discuss the parameter settings in the following. A total of $ns$ sensors are randomly scattered in an $l \times l$ field. The static data sink is located at $(0,0)$. There are a total of $np$ polling points uniformly distributed in the field. The sensor transmission range $Rs$ is 40 m. Each sensor has installed an AA battery of 1,500 mAh.

Fig.5a it is demonstrated that a larger $M$ also leadsto longer data latency. The reason is that more selected polling points need to be visited, which leads to a longer moving trajectory. For instance, when $l = 400 m$, the data latency in the case of $M = 6$ and $M = 4$ is 14 and 7 percent longer than the case of $M = 2$, respectively.

![Fig.5a. Performance Comparison of proposed framework with different M in no of cluster versus l](image)

Fig.5b shows the number of clusters formed with different $M$. It furthersvalidates that a larger $M$ typically leads to fewer clusters. However, it is also noticed that the difference becomes less evident when $l$ becomes larger. This is because that many clusters formed under this condition actually have fewer cluster heads than $M$ since sensors become sparsely distributed such that the controlling impact of $M$ on the cluster size is not fully extracted.

![Fig.5b. Performance Comparison of proposed framework with different M in Data latency versus l](image)
The number of cluster generated by mobile SISO and MIMO. We can see that the mobile MIMO typically yields fewer clusters than mobile SISO. Note that the average energy consumption of mobile SISO and MIMO become indistinguishable as n increase. This is because that although fewer clusters are observed with mobile MIMO, more cluster heads are generated for each cluster. Thus, the total number of cluster heads in the two schemes turns to be comparable, which is actually a dominant factor determining energy consumptions.

Finally, we compare the energy overhead with mobile SISO and MIMO, and illustrate the energy consumption by MIMO uploading itself in Fig5b. For the mobile approaches, overhead is mainly comprised of status and synchronization messages notifying SenCar of cluster locations and circuit energy consumption if MIMO uploading is adopted. First, the number of status messages exchange is proved to be upper bounded by 2n.

By considering these aspects, we plot the energy overhead in Fig5b. First of all, it is observed that energy over the total energy over-head result less than 6 percentage over the total energy consumptions. Then, the gap between mobile SISO and MIMO represents extra energy consumed in the MIMO circuitry is small. It indicates that compared to the energy overhead on clustering and notifying Sencar of cluster of cluster information, the energy consumed by MIMO circuitry itself is negligible.

Finally, we would like to point out that there are some interesting problems that may be studied in our future work. The first problem is how to find polling points and compatible pairs for each cluster. A discretization scheme should be developed to partition the continuous space to locate the optimal polling point for each cluster. Then finding the compatible pairs becomes a matching problem to achieve optimal overall spatial diversity. The second problem is how to schedule MIMO uploading from multiple clusters. An algorithm that adapts to the current MIMO-based transmission scheduling algorithms should be studied in future.

REFERENCES


