

Exploring Group Moving Pattern for an Energy-Constrained Object Tracking Sensor Network

Hsiao-Ping Tsai¹, De-Nian Yang¹, Wen-Chih Peng², and Ming-Syan Chen¹

¹ National Taiwan University
{hptsai@arbor, tony@kiki, mschen@cc}.ee.ntu.edu.tw

² National Chiao Tung University
wcpeng@cs.nctu.edu.tw

Abstract. In this paper, we investigate and utilize the characteristic of the group movement of objects to achieve energy conservation in the inherently resource-constrained wireless object tracking sensor network (OTSN). We propose a novel mining algorithm that consists of a global mining and a local mining to leverage the group moving pattern. We use the VMM model together with Probabilistic Suffix Tree (PST) in learning the moving patterns, as well as Highly Connected Component (HCS) that is a clustering algorithm based on graph connectivity for moving pattern clustering in our mining algorithm. Based on the mined out group relationship and the group moving patterns, a hierarchically prediction-based query algorithm and a group data aggregation algorithm are proposed. Our experiment results show that the energy consumption in terms of the communication cost for our system is better than that of the conventional query/update based OTSN, especially in the case that on-tracking objects have the group moving characteristics.

Key Words: OTSN, Grouping, Data Aggregation, Prediction.

1 Introduction

Energy conservation is paramount among all design issues in the inherently resource-constrained Object Tracking Sensor Network (OTSN). However, the current energy conservation mechanisms and algorithms may not be properly suitable for WSN because WSN differs from other network in many aspects [1][2][3]. Energy conservation in OTSN is even a harder affair because the target objects are moving. Many researches aim for designing an energy efficient OTSN. In [4][5], a cluster-based or tree-based local network collaborates multiple nearby sensors to handle the tracking task. By utilizing the local network, data aggregation and on-off scheduling are applied for reducing communication cost, saving energy, or prolonging the lifetime of the efficient route to the sink. However, while the object moving speed is relatively high, the frequency in rebuilding an energy efficient route and the complexity in maintaining a local network make the route and network structure no longer efficient. Prediction-based schemes

predict the future movement of the on-tracking objects according to their latest detected or average velocity to allow an energy efficient wake-up mechanism [6][7]. Probability-based prediction approaches take advantage of object moving patterns for future location prediction [8][9][10][11]. While investigating the factors that dominate energy cost of OTSN, we observe that many creatures such as animals, birds, or insects have social behavior that is they usually form mass organization and migrate together for food, breeding, wintering or other unknown reasons. The famous annual wildebeest migration is an example. Many water birds such as Black-faced Spoonbill have patterned flight-path. Ultradian rhythm phenomenon can be found in insect world. Each year, millions of monarch butterflies travel from Canada to Mexico and back again. These examples points out that the movement of creatures is correlated in space and dependent in time. Our point is that if the group relationship can be explored and the next location of a group of objects can be predicted together, the long distance network traffic for query can be reduced and the update data can be aggregated. That's the main idea of the paper and to our best knowledge, this is the first study that considers group relationship and object movement simultaneously in OTSN.

In this paper, we propose a group moving pattern mining algorithm, a hierarchically prediction-based query algorithm, and a group data aggregation algorithm. Our contribution has two folds: first, exploring group relationship among object moving patterns. Second, based on the group moving pattern, an efficient energy-constrained OTSN is designed. In our algorithm, a group probabilistic suffix tree (GPST) is used to predict objects' next locations in hierarchically prediction-based query algorithm. The group data aggregation algorithm utilizes group information in local and small scale data aggregation. The rest of the paper is organized as follows. Preliminaries, definitions and problem formulation are presented in Section 2. Algorithms for mining group moving patterns, a hierarchically prediction-based query algorithm, and an efficient group data aggregation algorithm are proposed in Section 3. Experiments are presented in Section 4. This paper concludes with Section 5.

2 Preliminaries, Definitions, and Problem Formulation

2.1 Preliminaries

Hierarchical Sensor Network In a hierarchical WSN, nodes are heterogeneous in energy, computing and storage capacity. Higher-energy node can be used to perform high complexity computing and send data while low-energy node can be used to perform the sensing and low complexity computing. A sensor equipped with higher energy acts as a cluster head (CH) on which high complexity task are assigned. In this paper, we adopt the hierarchical cluster structure. Sensors within a cluster have a locally unique id, and CH logical represents sensors within the cluster and acts as a sensor in view from upper layer. When a sensor detects an on-tracking object, it informs the location information to the CH. The CH aggregates location information then forwards to CH of upper layer. The process repeats until sink receives the location information. Here,

location information of an object is the corresponding sensor id. The movement of an object is represented by a sequence of sensor id visited by the object.

Location Modeling The movement dependency is that the next location in a moving sequence can be predicted from the sequence of preceding locations. We use Variable Markov Model (VMM) for learning statistic of object moving sequences and also a data structure called Probabilistic Suffix Tree (PST) is used together for mining significant moving patterns [12]. PST building algorithm is an efficient lossy compression scheme that converts a large data set into a variable length tree. The tree represents a dictionary of significant sequences that are meaningful for predicting next location.

2.2 Definitions and Problem Formulation

Definition 1. *Group Data Aggregation Radius (GDAR) is the number of hop counts between a sensor and its furthest neighbor in the participation of the group data aggregation.*

Definition 2. *Moving Sequence is a sequence of sensor id visited by one or a group of objects. Moving Speed is the number of sensor that an object crosses in a time unit.*

Given a moving sequence data set, our problem is to find the group relationship and the group moving pattern and obtain an energy efficient hierarchical OTSN.

3 Designs and Algorithms for OTSN

While a sensor detects an object, it invokes a group data aggregation and then informs a list of detected objects and a sensor id to the CH. CHs therefore collect objects' moving sequences within its cluster. In this section, we first present the group moving pattern mining algorithm and then propose an efficient object tracking sensor network.

3.1 Group Moving Pattern Mining Algorithm

The group moving pattern mining algorithm has four steps: building PST for each object, constructing a similarity graph on PSTs, extracting highly-connected components, and selecting Group Probability Suffix Tree (GPST). After the mining is performed, CH sends the group information to upper layer and gets a group id in return. The group information, group id and GPST are used in the hierarchically prediction-based query and group data aggregation.

Building PSTs for All Objects The movement data set in the CH is a set of moving sequences collected within this cluster. In the step, the CH builds a PST for each object. The Build-PST algorithm is shown in Figure 1.

Build - PST ($P_{\min}, \alpha, \gamma_{\min}, r, L$)

1. Initialization :

Let T consist of a single root node (with an empty label) and let $S \leftarrow \{\sigma/\sigma \in \Sigma \text{ and } P(\sigma) \geq P_{\min}\}$

2. Building the PST skeleton :

While $S \neq \emptyset$, pick any $s \in \Sigma$ and do :

(a) Remove s from S .

(b) If there exists a symbol σ such that $P(\sigma/s) \geq (1 + \alpha) \times \gamma_{\min}$

and $\left\{ \frac{P(\sigma/s)}{P(\sigma/\text{suff}(s))} \geq r \text{ or } \frac{P(\sigma/s)}{P(\sigma/\text{suff}(s))} \leq \frac{1}{r} \right\}$, then add to T the node for s and all the nodes on the

path to s from the deepest node in T that is a suffix of s .

(c) If $|s| < L$ then add the strings $\{\sigma' s / \sigma' \in \Sigma \text{ and } P(\sigma' s) \geq P_{\min}\}$ to S , if any.

3. Smoothing the prediction probabilities :

For each node s in T , let $\gamma_s(\sigma) = (1 - |\Sigma| \times \gamma_{\min}) \times P(\sigma/s) + \gamma_{\min}$

Fig. 1. The Build_PST Algorithm [12]

Constructing a Similarity Graph on PSTs Given n PSTs, we want to group them according to their structural similarities. The similarity score sim_p of two PSTs is defined as follows,

$$sim_p = -\log\left(\sum_{s \in S} \sqrt{\frac{\sum_{\sigma \in \Sigma} (P_1(s) \times P_1(\sigma|s) - P_2(s) \times P_2(\sigma|s))^2}{|\Sigma|}}\right) \quad (1)$$

It is a combination of condition probability and Euclidean distance over the union of node symbol strings of the two PST. The union is identified by S . If sim_p of two PSTs is higher than a given threshold, we consider they are similar. After comparing each pair of the given PSTs, we have pair relationship among all objects that forms a graph. Then the problem is transformed to be a graph connectivity problem.

Extracting highly-connected components Connectivity $k(G)$ of a graph G is defined as the minimum number of edges whose removal results in a disconnected graph. Given a graph with n nodes, we partition the graph into subgraphs such that the subgraphs are highly connected. We partition the graph G if $k(G) < \frac{n}{2}$. The highly connected component is an induced subgraph $G' \subseteq G$ such that G' is highly connected [13]. The nodes in a highly connected subgraph has degree at least $\frac{n'}{2}$ which means that each node in the subgraph has similarity relation with at least half nodes in the subgraph. The HCS cluster algorithm is shown in Figure 2. The clustering problem is then converted to a min-cut problem. A simple min-cut algorithm [14] is utilized in partition of the nodes.

Selecting Group Probabilistic Suffix Tree In last step, the similarity graph is partitioned into highly connected subgraphs. If the size of the subgraphs is higher than a given threshold, the objects corresponding to the subgraph are considered as an efficient group. The threshold is selected such that the efficiency

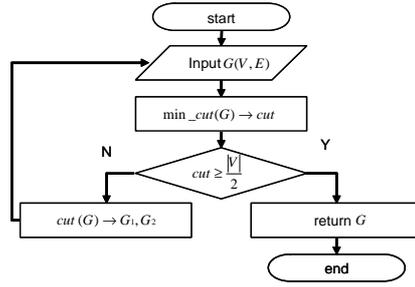


Fig. 2. The HCS Algorithm

gained on group based query algorithm and group data aggregation deserves the mining. After groups are found, a most representative PST named GPST is heuristically selected for each group such that the storage cost is reduced.

3.2 The Energy Efficient Object Tracking Sensor Network

The energy efficient OTSN includes a hierarchically prediction-based query algorithm and a group data aggregation algorithm. The prediction-based query algorithm flexibly determines to query for an individual object or a group of objects by one query. If on-tracking objects belong to a group, a group query can be applied such that not only flooding-based network traffic is eliminated but long-distance network traffic is reduced. In addition, the total data amount for the updates is further compressed under acceptable location precision by the group data aggregation algorithm. In the best situation, only one query is required in querying for a group of objects, and only one update is required in each updating interval for a group of objects.

Hierarchically Prediction-based Query Algorithm After the groups and their best fit GPSTs are produced. The information is sent to the CHs such that a group query and group update can be achieved efficiently. For the query-based OTSN, while receiving a query, the sink first predicts the most possible cluster that the object is currently located by using GPST and then sends the query to the CH. While the CH receives the query, it performs another prediction to get the most possible sensor that can detect the object. After receiving the query, the sensor invites its neighbors within GDAR to participate in tracking the object.

Group Data Aggregation Algorithm For the update-based OTSN, while an object is detected by a sensor, a group data aggregation process is initiated. The sensor performs as the master sensor that invites its neighbors within GDAR to collaborate in tracking objects and handles the local data collection for a period. After that, the master sensor reports to the CH about the detected objects and

the id of the sensor that detects most objects. Finally, the CH further compresses total data amount by using group id and filters redundant data according to the specified precision.

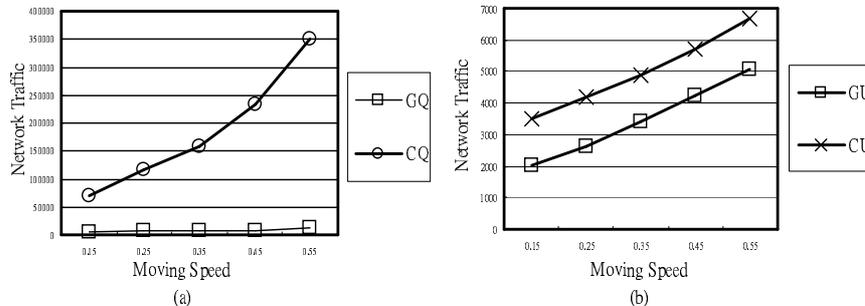


Fig. 3. Network traffic comparison.

4 Experiments

We implemented an event driven simulator in C++ with SIM [15] to evaluate the performance of our design. In the simulation, we use a Location-Dependent Parameterization of a Random Direction Mobility Model [16] to simulate the roaming behavior of a group leader. The other members are followers that are uniformly distributed around the leader within a specified Group Dispersion Range (GDR). The metrics used in evaluating our design are the amount of transmitted data in bytes per sensor (network traffic) and the average distance between observed location and real location of all objects (location distance).

4.1 Experiment 1

In the first experiment, we compare the performance between conventional query-based OTSN (CQ) and our group prediction query-based OTSN (GQ) as well as conventional update-based OTSN and our group aggregation update-based OTSN (GU). The sink has to persistently query for an object in query-base OTSN, and the sensors surrounding target objects persistently update the locations of objects in update-base OTSN. Figure 3(a) shows that the network traffic of CQ is about 20 times of GQ at the same location distance (~ 1.0). While the moving speed is higher, the rate is higher. Figure 3(b) shows that the network cost in CU about 1.48 time of GU at the same location distance (~ 1.0). While object moving speed is low, CU can achieve high location precision at a comparable network cost. However, if the object moving speed is higher, in order to

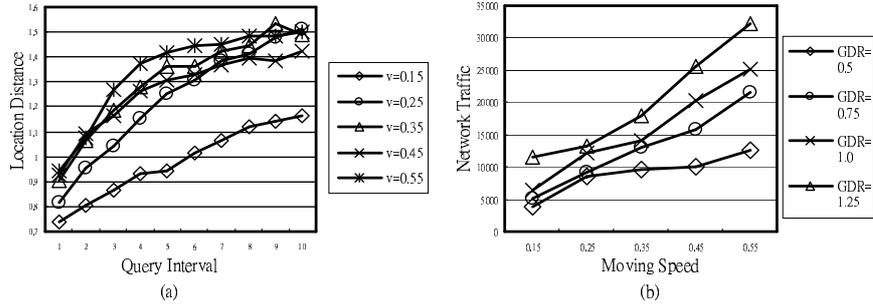


Fig. 4. (a)The impact of object moving speed on the location precision. (b)The impact of GDR on the network traffic.

achieve the same location precision, the update frequency increases and incurs much more traffic than GU.

4.2 Experiment 2

In the experiment, we study the impact of object moving speed and query interval on the location precision in GQ. The results in Figure 4(a) show that higher query frequency leads to a better location precision. Besides, at the same query interval, if the object moving speed is higher, the location precision is lower. To achieve higher location precision, higher query frequency is required.

4.3 Experiment 3

Figure 4(b) shows the impact of GDR. Higher GDR means that objects are more scattered such that a larger GDAR is required. GDAR is an important parameter that influences both the location information collecting range and object query snooping range. If GDR is higher, GDAR is set to be larger, and therefore more local in-network traffic is incurred.

5 Conclusions

In the paper, a group moving pattern mining algorithm, a prediction-based routing algorithm, and a group data aggregation algorithm are proposed. A simulation is conducted in light of the performance of the algorithms. Our contribution has two folds: first, exploring group relationship among object moving patterns. Second, based on the group moving pattern, an efficient prediction-based query algorithm and an efficient data aggregation algorithm for OTSN are provided. According to our experiments, the explored group relationship and group moving pattern are adopted to predict the group location such that the amount of query network traffic is significant reduced. The amount of update network traffic that

is incurred by reporting the location of monitored objects is also greatly reduced especially while the group density of monitored objects are high. In addition, the adaptive data aggregation range improves the prediction hit rate.

Acknowledgements

The work was supported in part by the National Science Council of Taiwan, R.O.C., under Contracts NSC93-2752-E-002-006-PAE.

References

1. Akyildiz, I.F., W. Su, Y.S., Cayirci, E.: Wireless sensor networks: A survey. *Computer Networks* **38**(4) (2002) 393–422
2. Culler, D., Estrin, D., Srivastava, M.: Overview of sensor networks. *IEEE Computer*, Special Issue in Sensor Networks (2004)
3. Al-Karaki, J.N., Kamal, A.E.: Routing techniques in wireless sensor networks: a survey. *IEEE Wireless Commun.* **11**(6) (2004) 6–28
4. Zhang, W., Cao, G.: Dctc: Dynamic convoy tree-based collaboration for target tracking in sensor networks. *IEEE Trans. on Wireless Commun.* **3**(5) (2004)
5. Lin, C.Y., Peng, W.C., Tseng, Y.C.: Efficient in-network moving object tracking in wireless sensor networks. *IEEE Trans. on Mobile Comput.* **5**(8) (2006) 1044–1056
6. Xu, Y., Winter, J., Lee, W.C.: Prediction-based strategies for energy saving in object tracking sensor networks. *IEEE Int. Conf. on Mobile Data Manag.* (2004) 346–357
7. Yang, L., Feng, C., Rozenblit, J.W., Qiao, H.: Adaptive tracking in distributed wireless sensor networks. *13th Annual IEEE Int. Symp. and Workshop on Engineering of Computer Based Systems* (2006) 103–111
8. Tseng, V.S., Lin, K.W.: Mining temporal moving patterns in object tracking sensor networks. *Int. Workshop on Ubiquitous Data Manag.* (2005)
9. Tseng, Y.C., Kuo, S.P., Lee, H.W., Huang, C.F.: Location tracking in a wireless sensor network by mobile agents and its data fusion strategies. *Int. Workshop on Information Processing in Sensor Networks* (2003)
10. Peng, W.C., Ko, Y.Z., Lee, W.C.: On mining moving patterns for object tracking sensor networks. *7th Int. Conf. on Mobile Data Manag.* (2006)
11. Ma, S., Tang, S., Yang, D., Wang, T., Han, J.: Combining clustering with moving sequential pattern mining: A novel and efficient technique. *8th PAKDD* (2004) 419–423
12. Ron, D., Singer, Y., Tishby, N.: Learning probabilistic automata with variable memory length. *7th annual Conf. on Computational learning theory* (1994)
13. Hartuv, E., Shamir, R.: A clustering algorithm based on graph connectivity. *Information Processing Letters* **76**(4-6) (2000) 175–181
14. Stoer, M., Wagner, F.: A simple min-cut algorithm. *Journal of the ACM* **44**(4) (1997) 585–591
15. Bolier, D.: *Sim : a c++ library for discrete event simulation* (1995)
16. Gloss, B., Scharf, M., Neubauer, D.: Location-dependent parameterization of a random direction mobility model. *IEEE 63rd Conf. on Veh. Technol.* **3** (2006) 1068–1072