Mobile Access Control Based on Trajectory Analysis and Prediction for Uploading Data

Li Chen, Kun Deng, Jun Yang, Panpan Li, Tao Jiang, and Tao He

***Abstract*—For large-scale mobile data collection, e.g environmental monitoring, mobile sensing nodes upload sensory data which they have collected via WiFi. Within Internet access point's (AP) communication range that they can access Internet by AP to do it. However, unfortunately communication resources of APs are quite scarce to the huge amount of sensory data that have to be uploaded. To address these problems in existing research on access request scheduling, this paper builds a dynamic model for quantifying uncertain space-time features of network and mobile access requests, then predict the future revenues of them. Based on it, reverse affects the optimization of the current strategy. This paper proposes an approximation algorithm to solve the optimization of the current strategy based on dynamic models of mobile access requests. Simulation results verify that SOA consistently outperforms the existing algorithm.**

***Index Terms*—Mobile sensing nodes (MSN), data collection, mobile crowd sensing, time to live (TTL).**

I. Introduction

With the explosive growth of high-capacity mobile communication devices, wireless sensing devices and the rapid development of mobile communication technologies, these mobile communication devices with sensors have become the primary computing environment for large-scale mobile computing and data collection [1]-[3]. For example, in the ocean observation scenario, only if the observation data that meets the requirements can be obtained, and applications such as marine target detection, anomaly and disaster warning, and intelligent management can be made possible.

In mobile wireless sensor network (MWSN) and drive-thru Internet technology, the mobile sensing nodes (MSNs) within access points' (APs') communication range that they can communicate with the AP and access Internet by AP to upload sensory data which they have collected. It's economic and effective to use MWSN for large-scale mobile sensory data collection for crowd-sensing application.

However, the performance of uploading data is terrible in practical applications. It is mainly caused by a huge amounts of sensory data have to upload, the sparse deployment and limited communication capabilities of the AP, and a large number of mobile access requests.

Manuscript received May 5, 2019; revised July 10, 2019. This work is supported in part by Zhejiang Provincial Natural Science Foundation of China (No.LY15F020040,LY16F020026,LY18F020020), the National Natural Science Foundation of China under Grant (No.61572014), the Humanity and Social Science Youth foundation of Ministry of Education of China (No.17YJCZH033) and Scientific Research Fund of Zhejiang Provincial Education Department (Y201636127,Y201840356).

The authors are with the College of Mathematics Physics and Information Engineering, Jiaxing University, Jiaxing 314001, China (e-mail: chenli20040415@163.com).

Motivated by this, we aim at optimal scheduling of scarce communication resources to prove the performance of crowd sensing. It will be determined to a large extent by the current scheduling strategy that related future revenues may be obtained from other APs' scheduling until deadline in networks.

To summarize, the major contributions of this paper lie in the following aspects:

1) We analyze the trajectory information of access requests and build a dynamic model for quantifying uncertain space-time features of network and mobile access requests (Section III).

2) We calculate the future revenues of the access requests based on access requests' trajectory analysis and prediction. And it affects the optimization of the current strategy as shown in definition 1 (Section IV).

3) We propose an approximation algorithm to solve the optimization of the current strategy based on dynamic models of access requests (Section V).

4) We evaluate the performance of the proposed SOA scheduling algorithm. Simulation results verify that SOA consistently outperforms the existing algorithm.

II. Related Work

Data collection application via mobile internet, such as Intelligent Transportation System (ITS), it can use MWSN to collect and monitor data [4], [5] on road traffic environment, and upload data via WiFi. Drive-thru network [6], as a promising wireless technology, provides vehicular users with Internet access service via WiFi access points (e.g. APs). AP's communication coverage area is a disc with the radius of the communication distance. Mobile communication devices with wireless communication devices and various sensors are generally used as mobile sensing nodes (MSN) in MWSN, such as smart phones and vehicles with wireless communication devices (the DSRC devices [7]), participate in communication as on-board units.

The centralized control method [8]-[10] is used to combine terminal control, routing and sensory data flow scheduling, and fine-grained control of the sensory data flow by monitoring the network status. [11], [12] proposed the scheduling scheme based on the priority that the network intermediate node performs "preemptive" scheduling optimization according to the priority of the sensory data flow, thereby alleviating concurrent congestion and improving service quality. [13] Utilized service needs and what the effect is on service quality in SINET, a flow-based

DOI: 10.7763/IJCTE.2019.V11.1248

deadline scheduling scheme for data center networks.

Communication resources of APs are quite scarce [14] to the huge amount of sensory data that needs to be uploaded, which has become a bottleneck problem in the collection of large-scale mobile sensory data. This paper found in the investigation that the spatio-temporal distribution of mobile access requests is not balanced. And then do in-depth research the their correlations of access requests, extract space and time dynamic features of access requests and built a dynamic model of mobile access requests.

III. Uncertain Analysis of Network and Access Requests' Trajectory

For crowd sensing data collection, if mobile sensing nodes (MSN) are inside some APs' communication area and their access requests are scheduled by the current APs, then can access the Internet via WiFi to upload sensory data. This paper models and analyzes the uncertain of network and mobile access requests.

A. Dynamic Network Model

It cannot upload sensory data that the access requests aren't scheduled. However, it is fortunate that they still have opportunities to upload data, as long as their trajectories pass through APs' communication area until deadline. Furthermore, this paper further analyses that the dynamic character of the network is predicted that can be abstracted into a time-expanded dynamic network topology model [15].

As shown as in Fig. 1, dynamic network model for mobile crowd sensing is built in this section.

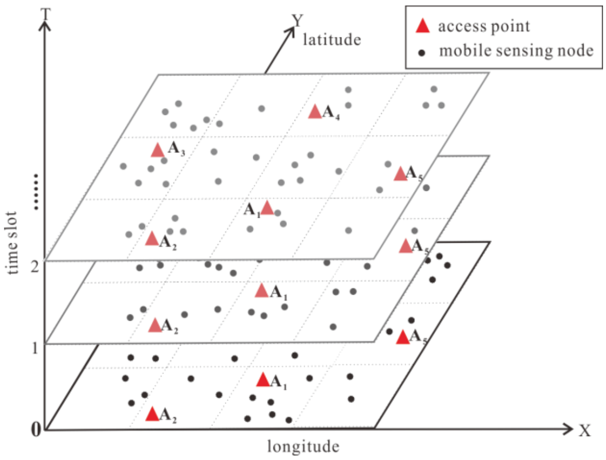


Fig. 1. Slicing network model within deadline in a timeline.

In Fig. 1, triangles represent access points and dots are mobile access requests. The network is divided into different grids with AP communication diameter. There are access points in some grids and not APs in most grids. The MSNs have no chance to upload sensory data when they go through the grids without AP.

B. Dynamic Model of Access Requests

In mobile wireless sensor network, an example of mobile access requests as shown in the Fig. 2 via WiFi for uploading data.

In the Fig. 2, access requests r_1~r_6 in the A_1's communication area at time slot t_0. And r_3 reaches the A_5's in 1 time slot, and r_1 and r_6 reach the A_3's in 2 time slots. We can see that the access requests are dynamically changing in space.

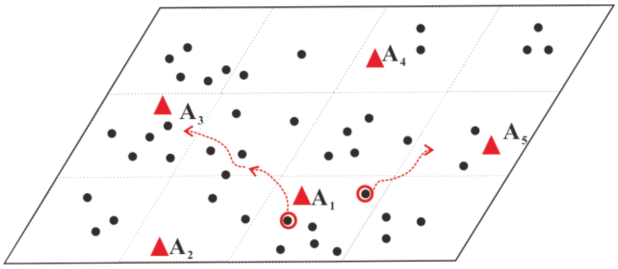


Fig. 2. Example of mobile access requests for uploading data via WiFi.

On the other hand, every mobile sensing node is probably to send an access request to APs at anytime and anyplace. The access requests are temporal dynamic with their emergence and disappearance. 1) The mobile sensing nodes will generate new access requests for uploading data during they continuously collect sensory data. 2) The access requests in network that will extinct in network, shown in the following two cases, a) They are scheduled by APs until deadline. b) They aren't scheduled that will become invalid before their intended expiration date.

This paper extract space and time dynamic features of access requests and as a time-varying graph sequence model. The dynamic model of uploading access requests as shown in Fig. 3.

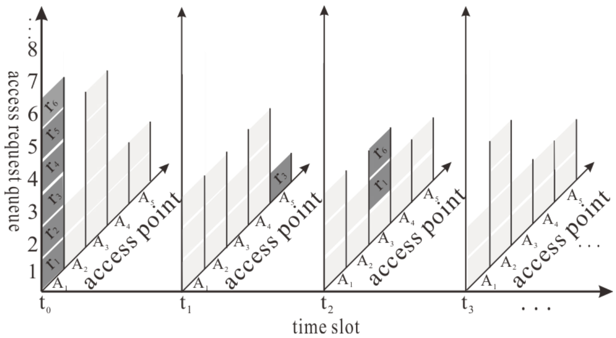


Fig. 3. Dynamic model of access request queue in WMSN.

IV. Problem Formulation

Given a mobile wireless sensor network with APs, MSNs and their trajectory information. The MSNs are inside AP's communication area and waiting for being scheduled by the current AP (A_c) and their access request set is represented as R_c.

The access requests in AP's communication area that their total revenue include two parts: the current revenue and possible future one. The future revenues is dependent on the current task scheduling scheme in the t_c time slot, which is S_c the set of access requests being scheduled currently by A_c until deadline.

In this paper, our goal is to maximize the total revenue of S_c the set of access requests being scheduled currently by A_c until deadline. It is to a Stochastic Optimization Problem of Uploading Access Control (SOPUA for short) that it is formally described as shown in definition 1.

Definition 1 SOPUA problem: It is defined as an optimization collaboration problem of accessing request set, and its total revenue is maximum until deadline. Its objective function is as follows:

$$Max \quad G(X,\mathrm{P}) = \sum_{i=1}^{n} [x_i \cdot v_i + (1-x_i)\varphi \cdot \Pi \cdot \mathrm{P}_i \cdot v_i] \quad (1)$$

$$\text{subject to}: \sum_{i=1}^{n} x_i \cdot c_i \leq \mathrm{C},\ x_i \in \{0,1\} \quad (2)$$

In the formula (2), the current communication resource of the current AP (A_c) is C, and c_i is consumed by the i^{th} access request when it upload sensory data.

v_i is the revenue of the i^{th} access request, $v_i > 0$. If $x_i = 1$ then the i^{th} access request will be scheduled. The access requests can earn current revenues $(x_i \cdot v_i > 0)$ if and only if they will be scheduled. If $x_i = 0$ then it won't be done, then its current revenue is 0. Fortunately, its future revenue $(1-x_i)\varphi \cdot \Pi \cdot \mathrm{P}_i \cdot v_i > 0$. $\mathrm{P}_i = (p_1, p_2, ..., p_T)$ is a probability vector, p_t is a probability of the i^{th} access request is scheduled in the t^{th} time slot before deadline. Its value depends on the probability of going through APs and the spatio-temporal distribution of APs' communications load until deadline. Π is an attenuation coefficient vector of the prediction that its value decays with time.

Their future revenues are related to the randomness of being scheduled by AP, and it depends on the probability of going through APs and the spatio-temporal distribution of APs' communications load until deadline.

V. SOA Algorithm

To solve SOPUA problem, this paper proposes stochastic optimization access control algorithm (SOA) based on access requests' trajectory analysis and prediction. Use SOA algorithm to find the decision vector X=[x_1,x_2,…,x_n] to maximize the objective function value of the formula 2.

SOA is a collaboration scheme based on load balancing for timeline with trajectory information that the specific description of the algorithm is as follows:

1)To predict the trajectory data of access requests in Sc based on the GPS and the current velocity and acceleration data of the access request, and so on;

2)To preprocess the trajectory data of the access requests. This paper obtains based on the analysis results which APs the request will pass through. That is a sequence of APs sorted by time slot until its TTL;

3) To calculate the AP load condition and access request data in the network, and then calculate the probability that each access request may be scheduled before TTL. Thereby the possible future revenue of this request is calculated;

4)Access requests in Sc are scheduled by current AP based on comprehensive analysis results.

VI. Performance Evaluation

The proposed work is based on load-balanced collaboration in a timeline to maximize the total revenue of current AP until deadline. In this section, we evaluate the performance of the proposed SOA scheduling scheme as shown in Fig. 4~6.

A. Simulation Experiment

In this section, we compare SOA with the classical first-in first-out algorithm (called FIFO for short) and random algorithm. This paper model a drive-thru network in detail using LINGO simulator and process data based on MATLAB.

All of the access requests are received by APs that they are stored in the scheduling queue. A client request sequence is exponentially distributed. This paper supposes that the length of access request queue follows exponential distributions based on an analogy. Therefore, the exponential distribution is used in this paper to perform simulation experiments and analysis of experimental results. Access requests arrival follows a Poisson process with arrival rate λ in the random-generated network.

This paper constructs a model as an inverted index table that it is an inverted index time series of access requests, as shown as Table I.

TABLE I: An Inverted Index Time Series of Access Requests

Requests	Time shot 1	Time shot 2	...	Time shot T
r_1	$x_{0,1}$	$x_{1,1}$	...	$x_{T,1}$
r_2	$x_{0,2}$	$x_{1,2}$	...	$x_{T,2}$
...	...	...	...	$x_{0,1}$
r_n	$x_{0,n}$	$x_{1,n}$	...	$x_{T,n}$

This paper built an inverted index time series based on access requests. As shown as in Table I, indicates the access reqest r_i will not go through any AP's communication area at the t^{th} time slot represents AP's label that r_i goes through.

TABLE II: Experimental Parameters

Parameter	Description
R	the number of access requests
C	total channel resources of AP
F_P	the probability density function of the number of access requests
λ	access requests' arrival rate
φ	revenue impact factor

In this part, the performance of the access request scheduling SOA is analyzed by simulation experiments, and the effectiveness of the method is verified. The parameters involved in the experiment are shown in Table II.

B. Performance Analysis

The proposed work in this paper is a stochastic optimization load-balanced collaboration scheme based on AP wireless communication resource given and delay tolerance constraints.

The prediction result of the time series of access requests is shown in Fig. 4. The blue line in the figure is the trend of the number of access requests, and the red line is the predicted load situation in the future. The error analysis of the time sequence prediction is shown in Fig. 5. The modified picture consists of two parts. The upper part is the output. The trend graph of the value, below the error before the output value domain target value. The dash line indicates the error condition, "*" is the data target value, and "+" is the analysis output value.

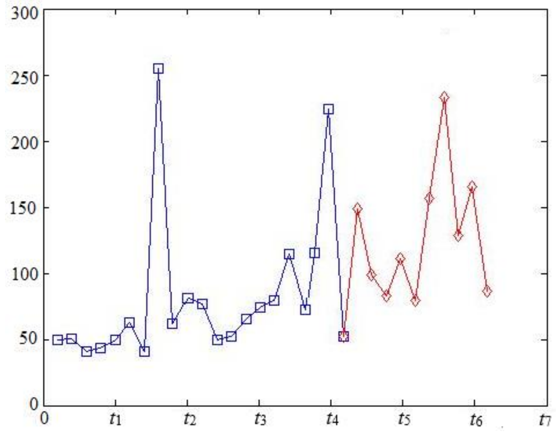


Fig. 4. Prediction of AP's load in timeline in the drive-thru networks.

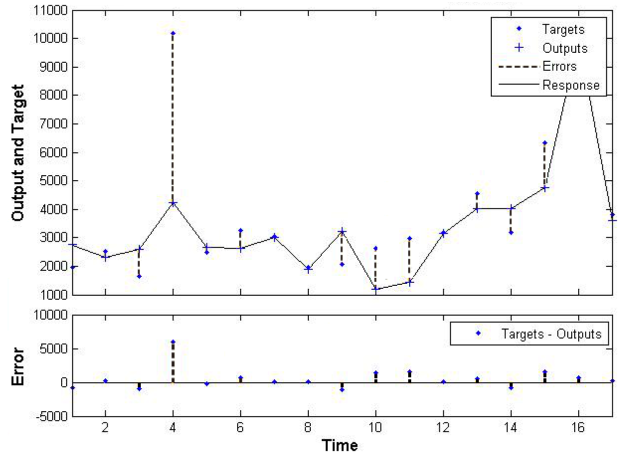


Fig. 5. Access request time series forecast trend and original trend graph.

The prediction result of the time series of access requests is shown in Fig. 4. The blue line in the figure is the trend of the number of access requests, and the red line is the predicted load situation in the future. The error analysis of the time sequence prediction is shown in Fig. 5. The modified picture consists of two parts. The upper part is the output. The trend graph of the value, below the error before the output value domain target value. The yellow line indicates the error condition, "*" is the data target value, and "+" is the analysis output value.

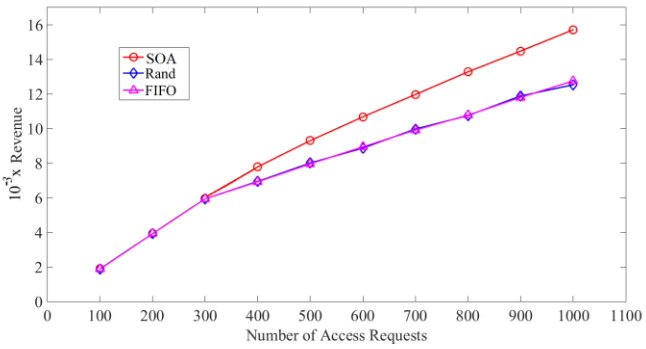


Fig. 6. Revenue vs the number of access requests.

As shown in Fig. 6, if the AP has a light load, the SOA algorithm has a similar total revenue to the traditional method, and their future revenues are all 0. However, if the AP is overloaded, the gain of SOA is infinitely close to the theoretical maximum revenue threshold with the number of mobile access requests increasing.

VII. Conclusion

To improve collecting data performance and quality of service, this paper presents an AP' load-balanced collaboration scheme (SOA) based on probabilistic trajectory prediction in a timeline for uploading mobile sensing data. Firstly, analyzed and predict trajectory information of mobile access requests; Secondly, slicing the network within deadline as a means to increase scheduling revenue, and built models for quantifying uncertain dynamic features of network and mobile access requests; Thirdly, calculated their possible future revenues based on the spatio-temporal correlation; Lastly, optimizes scheduling strategy based on trajectory information.

References

[1] E, Fasolo, M. Rossi, J. Widmer *et al.*, "In-network aggregation techniques for wireless sensor networks: A survey," *Wireless Communications*, vol. 14, no. 2, pp. 70-87, 2007.

[2] S. Ji and Z. Cai, "Distributed data collection in large-scale asynchronous wireless sensor networks under the generalized physical interference model," *IEEE/ACM Transactions on Networking*, 2013.

[3] K. W. Fan, S. Liu, and P. Sinha, "Structure-free data aggregation in sensor networks," *IEEE Transactions on Mobile Computing*, vol. 6, no. 8, pp. 929-942, 2007.

[4] S. Ji and Z. Cai, "Distributed data collection and its capacity in asynchronous wireless sensor networks," in *Proc. IEEE INFOCOM*, 2012, vol. 131, no. 5, pp. 2113-2121.

[5] S. Ji, R. Beyah, and Z. Cai, "Snapshot/continuous data collection capacity for large-scale probabilistic wireless sensor networks," in *Proc. IEEE INFOCOM*, 2012.

[6] O. Jörg and D. Kutscher, "A disconnection-tolerant transport for drive-thru internet environments," in *Proc. IEEE Joint Conference of the IEEE Computer & Communications Societies*, 2005.

[7] A. A. Carter, "The status of vehicle-to-vehicle communications as a means of improving crash prevention performance," Tech. Rep. 05-0264, 2005.

[8] A. Munir, G. Baig, S. M. Irteza *et al.*, "Friends, not foes: synthesizing existing transport strategies for data center networks," *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 4, pp. 491-502, 2014.

[9] Y. Zhao, K. Chen, W. Bai *et al.*, "RAPIER: Integrating routing and scheduling for coflow-aware data center networks," in *Proc. IEEE Computer Communications*, 2015.

[10] L. Liu, J. Li, and J. Wu, "TAPS: Task-aware preemptive flow scheduling," in *Proc. IEEE International Workshop on Local & Metropolitan Area Networks*, 2015.

[11] D. Zats, T. Das, P. Mohan *et al.*, "DeTail: Reducing the flow completion time tail in datacenter networks," *ACM SIGCOMM Computer Communication Review*, vol. 42, no. 4, pp. 139-150, 2012.

[12] M. Alizadeh, S. Yang, S. Katti *et al.*, "Deconstructing datacenter packet transport," in *Proc. the 11th ACM Workshop on Hot Topics in Networks - HotNets-XI*, 2012, pp. 133-138.

[13] Y. Xu, H. Luo, and F. Ren, "Is minimizing flow completion time the optimal way in meeting flow's deadline in datacenter networks?" *China Communications*, vol. 13, pp. 6-15, 2016.

[14] Z. Zheng, P. Sinha, and S. Kumar, "Sparse WiFi deployment for vehicular internet access with bounded interconnection gap," *IEEE/ACM Transactions on Networking*, vol. 20, no. 3, pp. 956-969, 2012.

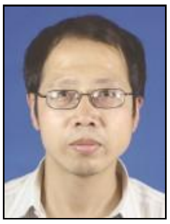
[15] G. Yan and S. Olariu, "A probabilistic analysis of link duration in vehicular ad hoc networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 4, pp. 1227-1236, 2011.

Li Chen was born in 1975. She received the PhD and is a lecturer at present. She is a member of China Computer Federation. Her main research interests include machine learning, and mobile communications technology.

Deng Kun was born in 1980. PhD, lecturer. His main research interests include data mining and complex networks analysis.

Yang Jun was born in 1978. He holds the PhD and is a lecturer now. He is a member of China Computer Federation. His current research interests include image and video processing, pattern recognition.

Li Panpan was born in 1983. He is a lecturer with the PhD. He is a member of China Computer Federation. His research interests include parallel processing, information security, and privacy preserving.



Jiang Tao was born in 1973. He is an associate professor with a PhD. His research interests include spatial-temporal database query processing, and uncertain data management.

He Tao was born in 1999. He is an undergraduate and student member of China Computer Federation. His research interests include data mining and information security technology.