ON THE MAC FOR OPTIMAL INFORMATION RETRIEVAL PATTERN IN SENSOR NETWORKS WITH MOBILE ACCESS

Zhiyu Yang, Min Dong, and Lang Tong
School of Electrical and Computer Engineering
Cornell University, Ithaca, NY 14853
{zy26,mdong,ltong}@ece.cornell.edu

and

Brian M. Sadler
Army Research Laboratory
Adelphi, MD 20783
bsadler@arl.army.mil

ABSTRACT

In sensor networks, locations of received data affect the information processing performance. In this paper, we consider how to design multiple access control (MAC) to obtain the optimal information retrieval pattern for signal field reconstruction. Taking both performance and implementation complexity into consideration, besides the optimal centralized scheduler, we propose three other MAC protocols, namely, decentralized scheduling through carrier sensing, Aloha, and decision-directed Aloha. Finally, we provide performance comparison among these protocols.

1. INTRODUCTION

In many applications of sensor network, the sensor network operates in three phases: sensing, information retrieval, and information processing. As a typical example, in physical environmental monitoring, sensors first take measurements of the signal field at a particular time. Then, data are collected from individual sensors. Finally, data from sensors are processed centrally to reconstruct the signal field.

An appropriate network architecture for such applications is SEnsor Networks with Mobile Access (SENMA) [1]. Shown in Fig. 1, SENMA has two types of nodes: low-power low-complexity sensors randomly deployed in large number and a few powerful mobile access points that communicate with sensors. The use of mobile access points enable data collections from specific areas of the network.

We focus on the latter two phases of operation in the SENMA architecture: information retrieval and processing, which are strongly coupled. In SENMA, medium access control is needed to regulate data retrieval from sensors to the access points. A multiple access control (MAC) protocol in SENMA governs how information is retrieved from the sensor field. With a specific MAC, packets collected by mobile access points form a sampled signal field with a specific pattern. This sample pattern directly affects the performance of the signal reconstruction at the last data processing stage. Therefore, it is important to design a MAC protocol that results in the desired retrieval pattern and hence, the optimal signal reconstruction performance. If the mobile access point can schedule transmissions from sensors, it is natural to use a centralized scheduler to poll data from equally spaced sensors locations.

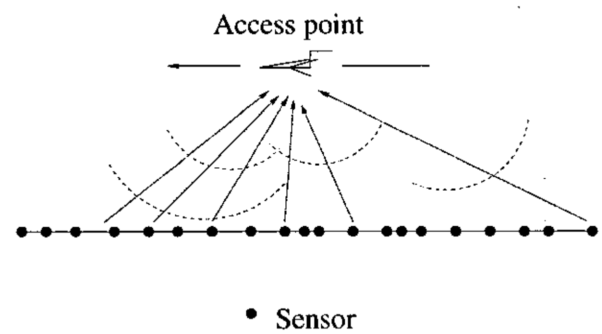

Fig. 1: A 1-D sensor network with a mobile access point

While it may appear obvious that collecting data from optimally chosen locations using centralized scheduler gives better performance, there are several nontrivial practical complications. For sensor networks with finite density, there may not exists a sensor at the desired location. The optimal scheduler must find sensors locations closest to the desired sampling pattern. Such scheduler needs to have the additional information of each sensor locations, which usually is not available at the access point. Furthermore, it comes with nontrivial complications of centralized control and large communications overhead. Decentralized MAC requires much less intervention from the mobile access point and is simple to implement. In order to take advantage of decentralized access but not lose much reconstruction per-

This work was supported in part by the Army Research Office under Grant ARO-DAAB19-00-1-0507, the Multidisciplinary University Research Initiative (MURI) under the Office of Naval Research Contract N00014-00-1-0564, and Army Research Laboratory CTA on Communication and Networks under Grant DAAD19-01-2-0011.

formance, our problem is how to design a (partially) decentralized MAC protocol to achieve the desired retrieval pattern.

We design MAC protocols for the desired data retrieval pattern. We consider a one-dimensional problem for simplicity which can be extended easily to two-dimensional problem. Taking both performance and implementation com-

[redacted]

pose a decentralized scheduler via carrier sensing which, under the no processing delay assumption, provides minimum performance loss comparing to the optimal scheduler. Then, to simplify the implementation, we describe a MAC scheme which uses Aloha-like random access within a resolution interval centered at the desired retrieval location. Finally, to improve the performance of the previous MAC, we propose a decision-directed Aloha scheme which adaptively chose the desired retrieval location based on the history of retrieved locations. The performance compari-

where $f > 0$, σ are known, $\{W(x) : x \geq 0\}$ is a standard Brownian motion, and $S(x) \sim \mathcal{N}(0, \frac{\sigma^2}{|2f|})$ is the stationary solution of (1). The random field modeled in (1) is essentially a diffusion process which is often used to model many physical phenomenon of interest. Being homogeneous in $\mathcal{A}$, $S(x)$ has the autocorrelation

[redacted]

for $x_0 < x_1$, which is only a function of distance between the two points x_1 and x_0.

2.2. Sensor Network Model

We assume that sensors in $\mathcal{A}$ are deployed randomly and their distribution forms a one-dimensional homogeneous spatial Poisson field with local density ρ sensors/unit area. That is, in an interval of size l, the number of sensors $N(l)$ is a Poisson random variable with probability

2.4. Information Processing and Performance Measure

Assume that, after m slots of data retrieval, there are K packets received from distinct sensors in $\mathcal{A}$. We reconstruct the original signal based on these received data samples. To avoid boundary effect for signal reconstruction, we assume that there is a sensor deployed at each of the two boundaries of $\mathcal{A}$, and we are able to obtain their measurements. The locations of all the received packets[1] are denoted, in ascending order, by $\mathbf{q}_K = \{0, q_1, \cdots, q_K, 1\}$. We estimate $S(x)$ at location x using its two immediate neighbor samples by the MMSE smoothing, *i.e.*, for $q_i < x < q_{i+1}$,

$$\hat{S}(x) = E[S(x)|Y(q_i), Y(q_{i+1})]. \tag{4}$$

Given $\mathbf{q}_K$, we define the maximum field reconstruction distortion by the maximum mean square estimation error in $\mathcal{A}$

$$\mathcal{E}(\mathbf{p}_K) \triangleq \max_{x\in\mathcal{A}} E\{|\hat{S}(x) - S(x)|^2 \mid \mathbf{p}_K\}. \tag{5}$$

The expected maximum distortion of signal reconstruction during the collection time m slots is then given by

$$\bar{\mathcal{E}}(m) \triangleq E\{\mathcal{E}(\mathbf{q}_K)\} \tag{6}$$

where the expectation is taken over the data sample locations $\mathbf{q}_K$ and the received sample size K.

Our objective is to design MAC that results in a smallest signal field reconstruction distortion for a fixed number of retrieval slots.

From [7,8], we have shown that the maximum distortion is determined only by the maximum of distances between any two adjacent data samples

$$\mathcal{E}(\mathbf{p}_K) = \frac{\frac{2f\sigma_Z^2}{\sigma^2} + 1 - e^{-fd_{\max}}}{\frac{2f\sigma_Z^2}{\sigma^2} + 1 + e^{-fd_{\max}}}\frac{\sigma^2}{2f} \triangleq \mathcal{E}(d_{\max}) \tag{7}$$

where $d_{\max} = \max_{0\leq i\leq K} q_{i+1} - q_i$, and $q_0 = 0$, $q_{K+1} = 1$. Therefore, the expected maximum distortion in (6) is only a function of $d_{\max}$. Our objective now is to design MAC so that it results in the minimum $d_{\max}$.

3. MAC FOR OPTIMAL INFORMATION RETRIEVAL PATTERN

3.1. Optimal Centralized Scheduling

Assume that the location information of all sensors is available to the mobile access point. The mobile access point is then able to pre-compute the optimal set of m locations and activate only those sensors. This results in the minimum $d_{\max}$, therefore, the best performance. The performance under this scheduler can be used as a benchmark for performance comparison.

Let $\mathbf{x}$, $x_1 \leq x_2 \leq \ldots$, be the sensor locations. The optimal $d_{\max}$ is

$$d_{\max} = \min_{i_1<i_2<\cdots<i_m} \max(x_{i_1}, x_{i_2} - x_{i_1}, \cdots, x_{i_m} - x_{i_m-1}, 1 - x_{i_m}).$$

The optimization problem can be solved using a brute force search. To reduce the computation complexity, we propose the following efficient algorithm. It first find an initial set of locations and the corresponding $d_{\max}$. Based on this, it finds another set of locations resulting in a smaller $d_{\max}$. Iteratively, $d_{\max}$ converges to its minimum value.

Algorithm: The search scheme consists of three steps.

- Step 1: Location initialization. A set of sensor locations is chosen as the initial set. The maximum distance of the chosen set is assigned as $d_{\max}^{(0)}$.
- Step 2: For $d_{\max}^{(i)}$, within interval $(0, d_{\max}^{(i)})$, the access point finds the sensor location closest to $d_{\max}^{(i)}$ and set it as q_1. Continue this procedure, for $(q_j, \min\{q_j + d_{\max}^{(j)}, 1\})$, find the sensor location closest to the right boundary of the interval and set it as q_{j+1}, for $j \leq m-1$. The access point then find the maximum distance of $0, q_1, \cdots, q_m, 1$ and denote it as $d_{\max}^{(i+1)}$. If there does not exist a sensor in the interval, the search stops and $d_{\max}^{(i)}$ obtained previously is the minimum $d_{\max}$.
- Step 3: If $d_{\max}^{(i+1)} < d_{\max}^{(i)}$, let $i = i+1$ and repeat step 2. Otherwise, the search stops, and the $d_{\max}^{(i+1)}$ is the minimum $d_{\max}$.

3.2. Decentralized Scheduling through Carrier Sensing

In practice, the sensors locations information is not available at the mobile access point. Each senor only knows its own locations. In this case, in order to retrieve data with the desired pattern and in a decentralized fashion, we propose decentralized scheduling through carrier sensing. Since the propagation delay is relatively small as compared to the slot length, we assume perfect carrier sensing with no propagation delay. Specifically, the access point schedule sensor transmission through carrier sensing, where the distance of sensors from the desired locations is used in the backoff scheme. The sensor backoff time is a function of the

[1] For convenience, K only denotes the number of packets not from the two boundary sensors of $\mathcal{A}$.

distance to the desired location. Similar idea of using carrier sensing for decentralized transmission is first proposed in [9–11], where the channel state information is used in the backoff function of carrier sensing for opportunistic transmission.

Algorithm: For the desired location q^*_m, the mobile access point first broadcast a beacon containing the first location q^*_1 information. A sensor computes its distance to q^*_1: $d_i = |x_i - q^*_1|$. Based on the backoff function $g(d)$ which maps the distance to a backoff time, it then chooses a backoff time τ and listens to the channel. A sensor transmits its packet after its backoff delay only if there is no other sensors transmit during its backoff period. In order to let only the sensor closest to the desired location transmit, the function $g(d)$ should be designed to be a strictly increasing function. An example of $g(d)$ is given in Fig. 2.

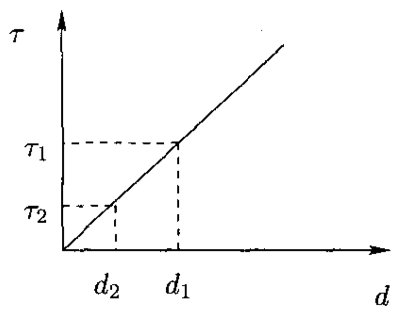

Fig. 2: Backoff function $g(d)$

3.3. Aloha

In the previous algorithm, the effect of propagation delay is ignored and perfect carrier sensing is assumed. However, these assumptions may not hold in practice and the performance is affected accordingly. In this subsection, we consider a simpler algorithm that does not depend on these factors.

Algorithm: At slot $t = 1, 2, \ldots, m$, the mobile access point activates a resolution interval $(t/(m+1)-\epsilon/2, t/(m+1)+\epsilon/2)$ through a beacon signal. The sensors within the activated range transmit their packets independently with probability P. (Fig. 3)

The expected throughput of each polling is

$$\sum_{i=1}^{\infty} \frac{(\epsilon\rho)^i e^{-\epsilon\rho}}{i!} iP(1-P)^{i-1} = \epsilon\rho P e^{-\epsilon\rho P},$$

which is maximized when $\epsilon\rho P = 1$. Therefore, $P^* = 1/\epsilon\rho$. To achieve as even sample locations as possible, the interval length ϵ should be as small as possible, so that if a packet goes through, it will be as close to the center of the resolution interval as desired. For fixed ρ, reducing ϵ results in increased P^*. Since $P^* \leq 1$, we have $\epsilon^* = 1/\rho$ and $P^* = 1$.

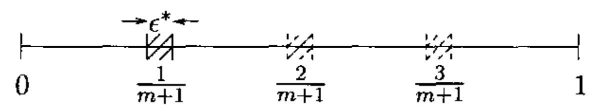

Fig. 3: Aloha scheme: The mobile access point determines m evenly spaced locations and activates an interval of length ϵ^* centered at one of this location at a time. The sensors within the activated range transmit with probability P^*.

3.4. Decision-Directed Aloha

In the previous algorithm, the locations of the activated intervals are predetermined before the operation. To obtain a better retrieval result, the access point may utilize the history of the collision channel output to optimize the polling locations.

Algorithm: The basic polling strategy is similar to the previous algorithm. The mobile access point first activates an interval of length $\epsilon^* = 1/\rho$ at a time. The sensors within the range transmit with probability $P^* = 1$. The difference is that the locations of the polling intervals depend on the previous polling results, which is described as follows.

After obtaining a new packet, the access point checks all the previous received data and find the two adjacent sample locations that have the maximum distance. The access point then locates the next polling interval in the middle of these two samples locations. If an empty slot occurs, the access point then activates the interval of ϵ^* adjacent (either left or right) to the previous polling interval until a success or collision occurs. If a collision occurs, the access point resolves the collision by splitting the previous polling interval until a packet is successfully transmitted (similar as the splitting algorithms). If a packet goes through, the access point starts to look for $d_{\max}$ again among the current received samples. The algorithm keeps running until it has polled m times.

3.5. A Lower Bound

We now provide a lower bound on the expected $d_{\max}$ of different protocols. If the number of packets the mobile access point receives is K, we have

$$E[d_{max}] \geq E\Big[\frac{1}{K+1}\Big] \geq \frac{1}{E[K]+1}$$

where the last inequality is due to the convexity of $1/x$. For the Aloha and decision-directed Aloha, $E[K] = m/e$.

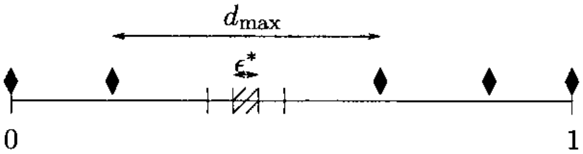

Fig. 4: Decision-directed Aloha scheme: The mobile access point activates an interval of length ϵ^* at a time. The sensors within the activated range transmit with probability $P^* = 1$. The solid diamonds indicates the received packets. The algorithm tries to break the maximum distance by placing the next polling interval at the center of the two received data samples locations whose distance is $d_{\max}$.

Therefore,

$$E[d_{max}] \geq \frac{1}{m/e + 1} \tag{8}$$

for the two Aloha schemes.

4. SIMULATIONS

In this section, we compare the performance the MAC protocol proposed in the last section through simulations. Sensors are randomly deployed according to Poisson distribution with density ρ. For convenience, we name these MAC protocols as following:

π_1: the optimal centralized scheduler
π_2: decentralized scheduling through carrier sensing
π_3: Aloha
π_4: decision-directed Aloha

We use the $d_{\max}$ found using π_2 as the initial maximum distance for the iteration algorithm in π_1. The search stops after 1-2 iterations typically. In the comparison, we use $E[d_{\max}]$ as the performance metric.

Fig. 5 and Fig. 6 plot $d_{\max}$ under various m for sensor density $\rho = 40$ and 200, respectively. As expected, as m increases, the number of data samples received at the mobile access point increases and $d_{\max}$ decreases. We see that there is little performance loss by using π_2. Notice that, when m is larger than ρ, under π_1 and π_2 on average data from all sensors can be retrieved. Therefore, the performance gap under these protocols is diminished. The performance under π_3 is worse than other schemes even when m greater than ρ. This is due to missing data samples from those scheduled interval because of either collision or void of sensors. Unlike π_3, the location and length of each polling interval of π_4 are decision-based. When m is large, it has enough slot to search for intervals where sensors exist and resolve collision, therefore, avoids the problem in π_3. From

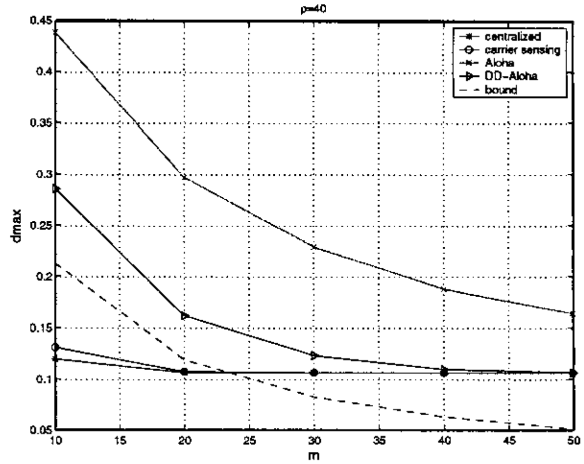

Fig. 5: $d_{\max}$ vs. m. $\rho = 40$.

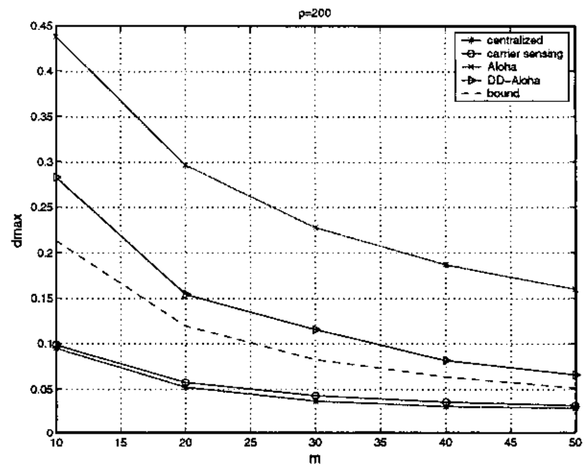

Fig. 6: $d_{\max}$ vs. m. $\rho = 200$.

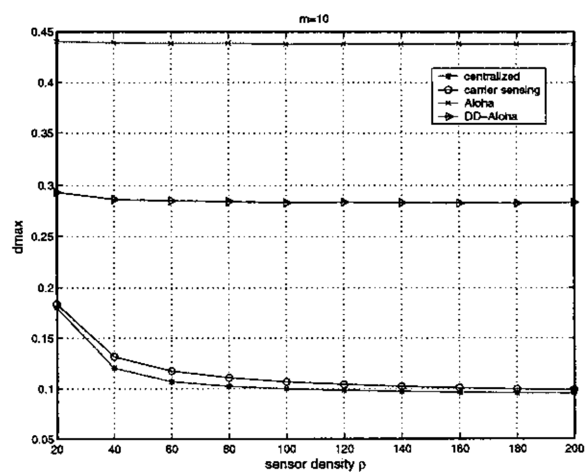

Fig. 7: $d_{\max}$ vs. ρ. $m = 10$.

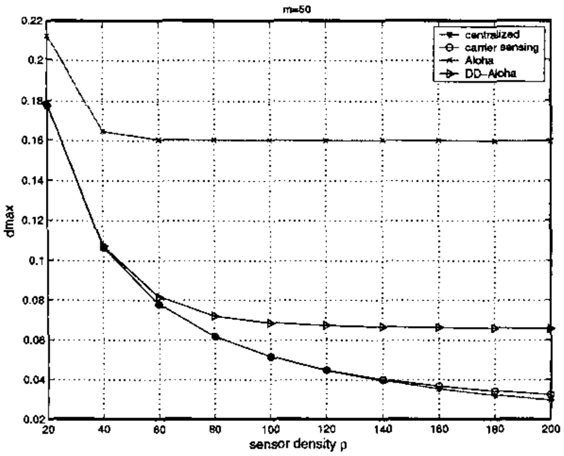

Fig. 8: $d_{\max}$ vs. ρ. $m = 50$.

Fig. 5, we see that, when m is large, the performance under π_4 is as good as the optimal case. Finally, we also plot the lower bound in (8) for the two Aloha schemes.

Fig. 7 and Fig. 8 plot $d_{\max}$ under various ρ for retrieval slots $m = 10$ and 50, respectively. As expected, as ρ increases, sensor field is denser, and the received data locations are closer to the desired locations. Therefore, $d_{\max}$ decreases to the minimum value. Again, we see the performance under π_2 closely follows the optimal one. As ρ increases, we see the performance gap between the two Aloha schemes and π_1 increases. The performance loss under π_3 is mainly due to its throughput which limits the number of received samples. We see that there is a significant performance improvement of π_4 over π_3 by adaptively optimizing the retrieval pattern based on the retrieval history.

5. CONCLUSION

To obtain the signal field using sensor networks, the locations of retrieved data affect the signal field reconstruction performance. In this paper, we design MAC protocols to obtain the desired data retrieval pattern. Taking both performance and implementation complexity into consideration, besides the optimal centralized scheduler, we propose three other MAC protocols. Our simulations show that using the decentralized scheduling through carrier sensing results in little performance loss comparing to that of the optimal scheduler. For the two Aloha schemes, by exploring the history of retrieved data locations, the decision-directed Aloha provides a significant performance gain over the simple Aloha scheme.

6. REFERENCES

[1] L. Tong, Q. Zhao, and S. Adireddy, "Sensor Networks with Mobile Agents," in *Proc. 2003 Military Communications Intl Symp.*, (Boston, MA), Oct. 2003.

[2] A. Woo and D. Culler, "A transmission control scheme for medium access in sensor networks," in *Proc. ACM/IEEE Conference on Mobile Communication and Networking*, (Rome, Italy), pp. 221–235, July 2001.

[3] W. Ye, J. Heidemann, and D. Estrin, "An Energy-Efficient MAC Protocols for Wireless Sensor Networks," in *Proc. of IEEE INFOCOM*, (New York, NY), pp. 1567–1576, June 2002.

[4] K. Sohrabi, J. Gao, V. Ailawadhi, and G. Pottie, "Protocols for self-organization of a wireless sensor network," *IEEE Personal Communication*, pp. 16–27, October 2000.

[5] R. Iyer and L. Kleinrock, "QoS Control for Sensor Networks," in *Proc. of 2003 ICC*, May 2003.

[6] P. Venkitasubramaniam, S. Adireddy, and L.Tong, "Sensor Networks with Mobile Agents: Optimal Random Access and Coding," *IEEE Journal on Sel. Areas in Comm.*, vol. 22, pp. 1058–1068, Aug. 2004.

[7] M. Dong, L. Tong, and B. Sadler, "Impact of MAC design on signal field reconstruction in dense sensor networks." submitted to *IEEE Trans. on Signal Processing*, November 2003.

[8] M. Dong, L. Tong, and B. Sadler, "Information retrieval and processing in sensor networks: deterministic scheduling vs. random access," in *ISIT 2004*, (Chicago, IL), June-July 2004.

[9] Q. Zhao and L. Tong, "QoS Specific Medium Access Control for Wireless Sensor Network with Fading," in *Proc. of Eighth International Workshop on Signal Processing for Space Communications*, (Catania, Italy), Sept. 2003.

[10] Q. Zhao and L. Tong, "Distributed opportunistic information retrieval in sensor networks: CSI-based carrier sensing." submitted to *IEEE Transactions on Wireless Communications*, Feb. 2004.

[11] Q. Zhao and L. Tong, "Distributed opportunistic transmission for wireless sensor networks," in *Proc. of IEEE ICASSP 2004*, (Montreal, Canada), May 2004.