A Novel Data Collection Scheme for WSNs

Jie Li and Xiucai Ye
Department of Computer Science
University of Tsukuba
Tsukuba Science City, Japan

Li Xu
Key Laboratory of Network Security and
Cryptology, Fujian Normal University
Fuzhou, China

***Abstract*—In this paper, we present a novel data collection scheme for wireless sensor networks by using separate network coding (SNC). By separately encoding a certain number of data segments in a combined data segment and doing decoding-free data replacement, SNC not only provides efficient storage method for continuous data, but also maintains a high success ratio of data collection. The performance evaluation has been conducted through comprehensive computer simulation. It is shown that SNC outperforms the exiting scheme significantly.**

I. Introduction

Wireless sensor networks (WSNs) are composed of a large number of sensor nodes, which are deployed in an open area without infrastructure. Each sensor node is with limited resources, such as low CPU power, small bandwidth, limited battery and memory storage [1]. As a result, one sensor node can store only a small amount of data collected from its surroundings. To collect the data from sensor nodes, the base stations (BSs) function as intermediate gateways between the sensor network and the application end users.

WSNs have a wide range of applications, including environment monitoring, medical care, smart buildings, industrial and military applications [2]. A very interesting problem that arises in application of WSNs is how to collect data from harsh and extreme environments [3]. In some harsh environments, the data are continuously sensed by the sensor nodes. The communications between the sensor nodes and a BS is expensive and scarce. Thus, data collection by the BS is only performed occasionally. As a result, the sensor nodes have to temporarily store the data and provide the desired data when the BS approaches. Such kind of data collection is known as *continuous data collection* [4]. One typical example is the Great Duck Island habitat monitoring system [5], seabird colonies are extremely sensitive to human interaction. A fast data retrieval is usually desired in each data collection. Our purpose in this paper is to deal with the continuous data collection in WSNs.

Network Coding is an emerging technique that has several interesting applications in practical networking systems. It was first introduced for improving the performance of multicast routing [6]. Different from the original applications in routing, our study extend the idea of network coding for continuous data collection for WSNs with a mobile BS. Similar idea of network coding extent for the applications of data storage and distribution can be found in [7], where the code is created over the connecting of data and storage nodes. The further theoretically study about network coding for data distribution and the practical system for random file distribution are presented in [8] and [9]. Using network coding for ubiquitous data collection was introduced in [10], [11] for wireless sensor networks.

An interesting network coding based data collection scheme was proposed by Dimakis et al. [10] (we call it Dimakis method in the reminder of this paper), which encodes all data segments in each node. The distinct benefit is the improvement of success ratio of data collection. However, the data segments to be collected are static and fixed. This scheme cannot support data collection in which the number of data segments is not predetermined. Wang et al. [4] proposed another interesting network coding base data collection scheme, known as Partial Network Coding (PNC). By encoding only part of the original data segments and removing the older data segments, PNC allows efficient storage for data collection to collect the part of newer data segments. Since the number of original data segments encoded in each combined data segment is a variable respect to time, the combined data segments cannot be always decoded by the BS. Thus, the the success ratio of data collection in PNC is no so high.

In some practical applications, new data has higher value than old ones. Thus, collecting the part of newer data can meet requirements. In this paper, we focus on continuous data collection as [4] to collect the part of newer data segments. To give a further improvement of the existing scheme, we present a novel strategy called Separate Networking Coding (SNC). SNC bases on the the mobile BSs randomly accessing. By separately encoding a certain number of data segments, it does not only provide efficient storage method for continuous data, but also maintains a high success ratio of data collection. We then address a set of practical concerns toward SNC-based continuous data collection in wireless sensor network. The performance evaluation has shown that SNC outperforms the existing scheme significantly.

The remainder of this paper is organized as follows. In Section II, we describe the system model and problem formulation. We present the proposed SNC scheme for continuous data collection in Section III. In Section IV, we evaluate the performance of SNC by simulations. Finally, we conclude the paper in Section V.

978-1-4673-0990-5/12/$31.00 ©2012 IEEE

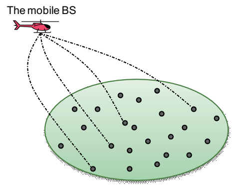

Fig. 1. Data collection by the mobile BS.

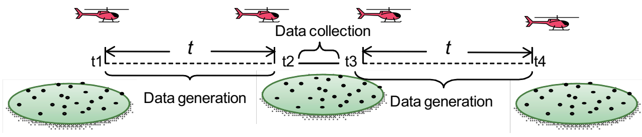

Fig. 2. Continuous data collection. At time t_1 and t_3, data collection is finished. At time t_2 and t_4, data collection is started to perform again.

II. System Description and Problem Formulation

Consider that there are *N* sensor nodes in a WSN with a mobile BS. The sensor nodes are with limited storage space. Each sensor node has *B* buffers, i.e., the buffer size of each sensor node is *B*. Each buffer can store only one data segment. In our scheme, by network coding each buffer stores one encoded data segment which encodes a number of data segments. We denote the *B* buffers as $b_i, i = 1, ..., B$. Consider to collect the information of event (e.g., the temperature goes above $20\,^{\circ}\mathrm{C}$) by using a WSN. The events are generated continuously. Without loss of generality, we consider there is one mobile BS which performs data collection occasionally. For example, a helicopter acts as the mobile BS. The helicopter is with data collection equipment, which is capable of transmitting and receiving radio signals, as illustrated in Fig 1. In each data collection, the mobile BS will randomly contact some sensor nodes to collect data. Let *W* ($\leq N$) denote the number of sensor nodes contacted by the mobile BS in each data collection. As the bad weather may prohibit the mobile BS from performing data collection for a long period of time, the sensor nodes in the WSN have no information about when the BS will perform data collection.

An event is represented by one data segment, denoted by c_j, and generated at a fixed time slot. For convenience, we assume c_q is newer than c_p if $q > p$, i.e., c_q is generated after c_p if $q > p$. The total number of data segments generated in one data generation time interval is $n(t)$, where *t* is a data generation time interval. *t* is a variable and the value depends on how often the mobile BS performs data collection. Note that $n(t)$ may increase as *t* increase. As an example shown in Fig 2, data collection is finished at time t_1. At time t_2, data collection is started to perform again and then stop at time t_3. At time t_4, data collection is started to perform again. The data collection by the mobile BS is performed between time t_2 and t_3. The time interval between t_1 and t_2 and the time interval between t_3 and t_4 are the data generation time interval *t*. During this time interval *t*, the environment data are continuously generated. Consider that the data collection time is much less than the data generation time. The data segments generated during data collection time can be negligible for the sake of convenience.

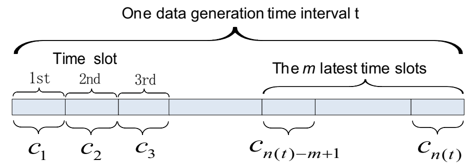

Fig. 3. Data generation in one data generation time interval *t*. c_i is generated in the i^{th} time slot.

In this paper, we consider to collect *m* ($1 < m < n(t)$) data segments during the *m* latest time slots in a time interval *t* by the mobile BS. *m* is the number of latest data segments to be collected in a time interval. *m* latest events are the data segments generated in the *m* latest time slots from $n(t)$, as shown in Fig 3.

We assume that each data segment is recorded by all the sensor nodes whenever generated. More specifically, some data nodes sense the data and transmit to all the storage nodes, then data coding is done in the storage nodes [4]. As our main work focus on the encoding and data replacement in the storage nodes, we omit the data nodes for ease exposition. And the sensor nodes mentioned in this paper function as the storage nodes.

For the network coding, we define a linear function

$$f_i = \sum_{j=1}^{r} \beta_{ij} c_j. \tag{1}$$

It is used to combine an amount of data segments (*r* data segments, $1 \leq r < n(t)$) as a coded data segment f_i. Here, $\vec{\beta} = (\beta_{i1}, \beta_{i2}, \cdots, \beta_{ir})$ is a coefficient vector. Each item β_{ij} is randomly generated from a finite field F_q, $\beta_{ij} \neq 0$, $j = 1, ...r$. Since the coding can be viewed as a combination process, f_i is also referred to as a combined data segment, and c_j as an original data segment. Notice that by network coding one buffer can store one combined data segment which encodes a number of original data segments. We use $C(f_i)$ to denote the number of original data segments combined in f_i. For the sake of convenience, a summary of the notations are given in Table I.

When the mobile BS performs data collection, the sensor nodes communicated with it will upload the stored combined data segments and the related coefficient vectors. We further assume that there are packet acknowledgements. Therefore, no packets are lost. If the mobile BS cannot collect all the desired data, it will perform data collection again and again until succeed. As a result, the WSN will consume more energy and the mobile BS will spend more time. Thus, the success ratio of data collection serves as a major evaluation criterion in our study, and is defined as follows.

Definition 1 (Success Ratio of Data Collection): The success ratio of data collection is the probability that the mobile BS successfully collects all the desired original data segments.

TABLE I
LIST OF NOTATION

Notation	Definition
m	Number of latest data segments to be collected
B	Buffer size of each sensor node
b_i	Buffer of each sensor node with index i
W	Number of sensor nodes contacted by the BS for each access
t	One data generation time interval
$n(t)$	Number of data segments generated in one data generation time interval
c_j	Original data segment with index j
f_i	Combined data segment by sensor node i
β_{ij}	Coding coefficient for c_j in f_i
$C(f_i)$	Number of data segments combined in f_i
q	Size of finite field for coefficients

III. PROPOSED SEPARATE NETWORK CODING BASED DATA COLLECTION

A. Overview of SNC

To collect the *m* latest data segments, we consider that each sensor node is with 2 buffers (i.e., buffer size is 2). We will prove later that buffer size 2 is the minimum buffer size to encode the *m* latest original data segments. The two buffers are denoted by b_1 and b_2. $f^i = (f_{i_1}, f_{i_2})$ are the combined data segments stored in sensor node *i*, f_{i_j} is the combined segment stored in buffer b_j, $j = 1, 2$. The total number of original data segments encoded in the two combined data segments in sensor node *i* is $C(f^i), C(f^i) = C(f_{i_1}) + C(f_{i_2})$.

SNC includes two processes: data encoding and replacement process, and data decoding process. Data encoding and replacement process is done in each sensor node. Each sensor node uses the 2 buffers to encode the sensed new generated data segments and replace the older data segments without decoding the encoded data segments. Data decoding process is done in the mobile BS.

B. Data encoding and replacement

SNC is different from the existing PNC scheme in which each sensor node combines the original data segments in a combined segment and varies the combined number from 1 to *m*. To improve the success ratio, in our scheme each sensor node separately combines *m* original data segments in two combined fashions. In other words, node *i* ($i = 1, ...N$) separately combines *m* original data segments in f_{i_1} and f_{i_2}. The original data segments are first combined in f_{i_1}, as $f_{i_1} = \sum_{j=1}^{r} \beta_{i_1.j} c_j$. *r* is the number of original data segments generated in the environment and sensed by each sensor node. $r \leq m$, $\beta_{i_1.j}$ is the coefficient of c_j in f_{i_1}, randomly generated from a finite field F_q by node *i*, j from 1 to *r*. Since the original data segments are continuously generated as the time interval *t* increases. The new generated data segments are continuously to be encoded in f_{i_1} until $C(f_{i_1})$ increases to *m*. When $C(f_{i_1})$ increases to *m*, the next generated original data segments are then changed to be combined in f_{i_2}, as $f_{i_2} = \sum_{j=m+1}^{r} \beta_{i_2.j} c_j$, $r \leq 2m$, $\beta_{i_2.j}$ is the coefficient of c_j in f_{i_2}. When $C(f_{i_2})$ increases to *m*, the next generated original data segments are combined in a new f_{i_1}, which replace the old f_{i_1} to be stored in buffer b_1. For this new f_{i_1}, when $C(f_{i_1})$ increases to *m*, the next generated original data segments are combined in a new f_{i_2}, which replace the old f_{i_2} to be stored in buffer b_2. This process is going on until the BS arrives to perform data collection.

$$N_1: \{ f_{1_1} = \sum_{j=1}^{10} \beta_{1_1.j} c_j, f_{1_2} = \sum_{j=1}^{10} \beta_{1_2.j} c_j \}$$
$$N_2: \{ f_{2_1} = \sum_{j=1}^{10} \beta_{2_1.j} c_j, f_{2_2} = \sum_{j=1}^{10} \beta_{2_2.j} c_j \}$$
$$N_3: \{ f_{3_1} = \sum_{j=1}^{10} \beta_{3_1.j} c_j, f_{3_2} = \sum_{j=1}^{10} \beta_{3_2.j} c_j \}$$
$$N_4: \{ f_{4_1} = \sum_{j=1}^{10} \beta_{4_1.j} c_j, f_{4_2} = \sum_{j=1}^{10} \beta_{4_2.j} c_j \}$$

(a) Dimakis method

$$N_1: \{ f_{1_1} = [c_{10}], f_{1_2} = [c_9, c_{10}] \}$$
$$N_2: \{ f_{2_1} = [c_9, c_{10}], f_{2_2} = [c_{10}] \}$$
$$N_3: \{ f_{3_1} = [c_8, c_9, c_{10}], f_{3_2} = [c_9, c_{10}] \}$$
$$N_4: \{ f_{4_1} = [c_9, c_{10}], f_{4_2} = [c_{10}] \}$$

(b) PNC

$$N_1: \{ f_{1_1} = [c_7, c_8, c_9], f_{1_2} = [c_{10}] \}$$
$$N_2: \{ f_{2_1} = [c_7, c_8, c_9], f_{2_2} = [c_{10}] \}$$
$$N_3: \{ f_{3_1} = [c_7, c_8, c_9], f_{3_2} = [c_{10}] \}$$
$$N_4: \{ f_{4_1} = [c_7, c_8, c_9], f_{4_2} = [c_{10}] \}$$

(c) SNC

Fig. 4. Data distribution in 4 sensor nodes (N_1 through N_4), each node is with 2 buffers, *m*=3. (a) Dimakis method, (b) PNC, (c)SNC.

An illustration example is shown in Fig 4. It includes the corresponding Dimakis method and PNC scheme. Assume that the original data segments generated are c_1, c_2, $\cdots$, c_{10}. In this example, for simply we let *m*=3, that is, the mobile BS wants to collect the 3 latest data segments c_8, c_9 and c_{10}. In the application, *m* could be much larger than 3. That is the reason why each sensor node with limited storage space need network coding to store these amount of data segments. For the sake of convenience, here we define

$$[c_1, ..., c_r] = \sum_{j=1}^{r} \beta_{i_k.j} c_j \tag{2}$$

to denote the combined data segments in node N_i, omiting the the coefficients $\beta_{i_k.j}$ ($j = 1, ..., r, k = 1, 2.$). Dimakis method in Fig 4 (a), combining all the data segments without replacement, it is not suitable for continuous data collection as too much data segments encoded in the combined data segments will lead undecodable in the decoding. PNC is suitable for continuous data collection as with data replacement, but it is not always successful. As in Fig 4 (b), if the BS contacts N_1, N_2 and N_4, the combined data segments in these three nodes do not encode c_8, it can not collect c_8. SNC allows data replacement, and the combined data segments stored in each sensor node encode all the desired data segments, as shown in Fig 4 (c).

Every time an original data segment c_j is encoded with the combined data segments f^i. The data encoding algorithm (Algorithm 1) is locally executed at each sensor node.

Algorithm 1 Data Encoding(f^i, c_j)

Let $u = \lceil j/m \rceil$
if u is odd **then**
    Randomly generate $\beta_{i_1.j}$ from F_q
    if $C(f_{i_1}) \leq m$ **then**
        $f_{i_1} = f_{i_1} + \beta_{i_1.j} c_j$
    else
        $f_{i_1} = \beta_{i_1.j} c_j$
    end if
else
    Randomly generate $\beta_{i_2.j}$ from F_q
    if $C(f_{i_2}) \leq m$ **then**
        $f_{i_2} = f_{i_2} + \beta_{i_2.j} c_j$
    else
        $f_{i_2} = \beta_{i_2.j} c_j$
    end if
end if

In Algorithm 1, $\lceil j/m \rceil$ is the upper integer bound of *j/m* (e.g., $\lceil 1/3 \rceil = 1$ and $\lceil 4/3 \rceil = 2$). The encoding process is with the data replacement to encode only the amount of latest original data segments in the combined data segments. To obtain the theorem in the next subsection, we first obtain the following 3 lemmas. The proofs of Lemma 1 and Lemma 2 are omitted due to the space limit.

Lemma 1: After data replacement, for the combined data segments f^i stored in each sensor node *i*, the number of data segments $C(f^i)$ is with an upper bound *2m* and a lower bound *m+1*.

Lemma 2: At any time, the two sets of data segments encoded in the two combined data segments in a sensor node contain the *m* latest data segments.

Lemma 3: By SNC, the minimum buffer size of each sensor node for storing the combined data segments which encode the *m* latest data segments at any time is 2.

Proof. From Lemma 2, each sensor node with 2 buffers can encode the *m* latest data segments at any time. We then show that each sensor node with 1 buffers cannot do it at any time. If each sensor node is only with one buffer to do encoding and data replacement, the probability of encoding the *m* latest data segments is only $1/m$, since the number of new data segments encoded in a combined segment is from 1 to *m*. □

C. Data decoding

The decoding process is performed by the mobile BS. The mobile BS randomly contacts *W* sensor nodes to collect $2W$ combined data segments $\mathbf{f} = [f^1, f^2, ..., f^W]$, where $f^i = (f_{i_1}, f_{i_2})$ is the combined data segments collected from sensor node *i*, $i = 1, ..., W$. The mobile BS also collects the related coefficient vectors $(\overrightarrow{\beta_{i_1}}, \overrightarrow{\beta_{i_2}})$ about the combined data segments (f_{i_1}, f_{i_2}), $i = 1, ..., W$. The decoding process of $\mathbf{f}$ includes two stages, decoding about $\mathbf{f}_1$ and $\mathbf{f}_2$, where $\mathbf{f}_1 = [f_{1_1}, f_{2_1}, ..., f_{W_1}]$ and $\mathbf{f}_2 = [f_{1_2}, f_{2_2}, ..., f_{W_2}]$ are the combined data segments collected from buffer 1 and buffer 2 of *W* sensor nodes.

TABLE II
SYSTEM PARAMETERS AND SETTINGS

System Parameters	Settings
Length $\times$ Width	300 m $\times$300 m
Number of sensor nodes	1000
Transmit range between sensors	20 m
Transmit range between sensors and BS	150 m $\sim$ 250 m
Data generate time interval	30 sec
Energy consumption for sending a message	20nAh
Size of finite field for coefficients *q*	2^8
Confidence interval	95%

From the data encoding and replacement process, data replacement happens when the number of original data segments encoded in the combined data segments increases to *m*. As a result, the number of replaced older original data segments is an integer multiple of *m*. Consider that when the mobile BS performs data collection, the replaced older data segments are $c_1, c_2, ..., c_{km}$, and the original data segments encoded in the two combined data segments which is stored in each sensor node are $c_{km+1}, c_{km+2}, ..., c_{n(t)}$, k is an integer and $(n(t) - m - 1)/m \leq k \leq (n(t) - 2m)/m$.

Without loss of generality, we assume $f_{i_1} = \sum_{j=km+1}^{(k+1)m} \beta_{i_1.j} c_j$ and $f_{i_2} = \sum_{j=(k+1)m+1}^{n(t)} \beta_{i_2.j} c_j$. We first focus on the decoding process about $\mathbf{f}_1$. The corresponding coefficient vectors are $[\overrightarrow{\beta_{1_1}}, \overrightarrow{\beta_{2_1}}, ..., \overrightarrow{\beta_{W_1}}]$, where $\overrightarrow{\beta_{i_1}} = [\beta_{i_1.km+1}, ..., \beta_{i_1.(k+1)m}]$, $i = 1, ..., W$. The coefficient vectors form a $W \times m$ coefficient matrix $\mathbf{A}$.

The decoding process about $\mathbf{f}_1$ is basically to solve a set of linear equations $\mathbf{f}_1 = \mathbf{A}\mathbf{c}_1$, with the *m* original data segments $\mathbf{c}_1 = [c_{km+1}, c_{km+2}, ..., c_{(k+1)m}]$ to be the variables. We apply Gaussian Elimination [12]. If $|\mathrm{A}| \neq 0$, $\mathbf{c}_1 = \mathbf{A}^{-1}\mathbf{f}_1$, otherwise the rank of $\mathbf{A}$ is less than *m* and the set of equations is insoluble.

The decoding about $\mathbf{f}_2$ is similar with $\mathbf{f}_1$. For the linear equations of $\mathbf{f}_1$ and $\mathbf{f}_2$, there are at most *m* variables. By randomly contact *m* sensor nodes (W=m), the BS could decode the linear equations of $\mathbf{f}_1$ and $\mathbf{f}_2$. The probability of decode these linear equations are very close to 100% for a large enough field size q, as mention in [4].

Theorem 1: The success ratio of data collection by SNC with buffer size *B=2* to collect the *m* latest data segments is 100% (neglecting linear dependency of the coefficients).

The proof of Theorem 1 is obviously from Lemma 2 and the decoding process.

IV. PERFORMANCE EVALUATION

In this section, we evaluate the performance of proposed scheme by simulation. we deploy 1000 sensor nodes randomly into a field of 300m $\times$ 300m. The solution of linear equations in network coding are using the Gaussian Elimination [13]. The coefficient field is $q = 2^8$, which can be efficiently implemented in a 8-bit or more advanced microprocessor [11]. Table II demonstrates part of the important parameters and settings in the simulation.

We first compare the success ratio of SNC with PNC. Fig 5 shows that the success ratio as a function of the *m* latest

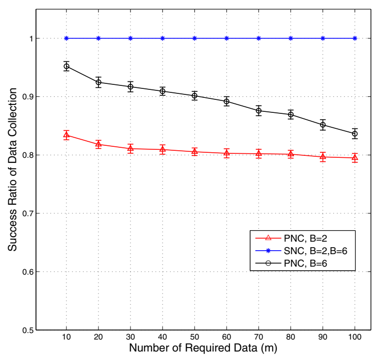

Fig. 5. Success ratio vs. number of data m.

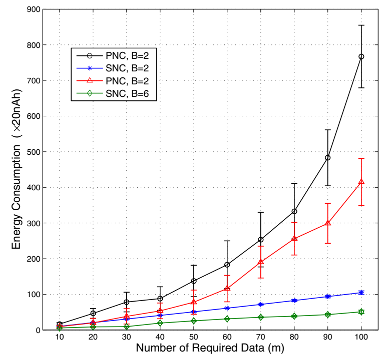

Fig. 6. Energy consumption vs. number of data m.

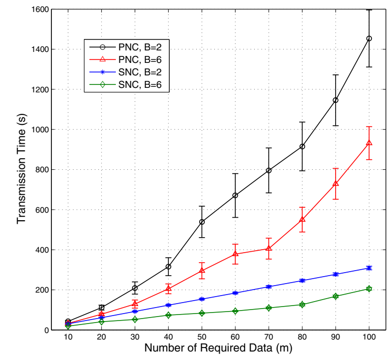

Fig. 7. Transmission time vs. number of data m.

data segments (required data segments). It is clear that SNC performs better than PNC for different m. The success ratio of SNC maintains at 100% when B is set to be 2 or 6. This fact is also proved in Theorem 1. The success ratios of PNC with B=2 and B=6 decrease as m increases.

We compare energy consumption of SNC with PNC in one time interval of data collection. We use the energy model by Mainwaring et al [5]. The energy consumption for transmitting one packet is 20 nAh (10.9 Ampere hour). As shown in Fig 6, the energy consumption of SNC is much less than PNC for different m no matter when B is set to be 2 or 6. It can be seen that the energy consumptions increases as the number of required data (m) increases, but the slope for SNC is much smaller, the energy consumption with SNC is nearly proportional to the number of required data m.

As a fast data retrieval in the application is usually desired, we then compare SNC with PNC about the time cost to successfully collect the data in one time interval of data collection. For simplicity, we neglect other affecting factors which is much less than the transmission time. As shown in Fig 7, the mobile BS spends much less time with SNC no matter when B is set to be 2 or 6. The time cost with SNC is nearly proportional to the number of required data m. For PNC, the cost time increase rapidly as m increases when B is set to be 2 and 6.

V. Conclusion

In this paper, we present a novel data collection scheme for WSNs by using separate networking coding (SNC). SNC not only provides efficient storage method for continuous data to collect the m latest data segments, but also maintains a high success ratio of data collection. We introduce the process of data encoding and replacement, and the process of data decoding. We also show that the success ratio of data collection by SNC with buffer size 2 is 100% (neglecting linear dependency of the coefficients). Furthermore, we evaluate the performance of SNC by simulations. It is shown that SNC outperforms the exiting scheme significantly.

Acknowledgments

This work has been partially supported by Grand-in-Aid for Scientific Research from Japan Society for Promotion of Science (JSPS), Research Collaboration Grant from NII of Japan, and National Natural Science Foundation of China with grant NO.61072080.

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “Wireless sensor networks: A survey,” *Comm. ACM*, vol. 38, no. 4, pp. 393–422, 2002.
[2] D. Culler, D. Estrin, and M. Srivastava, “Overview of sensor networks,” *IEEE Comput.*, vol. 37, no. 8 (Special Issue on Sensor Networks), pp. 41–49, 2004.
[3] A. Cerpa, J. Elson, D. Estrin, L. Girod, M. Hamilton, and J. Zhao, “Habitat monitoring: Application driver for wireless communications technology,” in *Proceeding of SIGCOMM LA ’01 Workshop on Data communication in Latin America and the Caribbean*. San Jose, Costa Rica: ACM, 2001, pp. 20–41.
[4] D. Wang, Q. Zhang, and J. Liu, “Partial network coding: Concept, performance, and application for continuous data collection in sensor networks,” *ACM Transactions on Sensor Networks*, vol. 4, no. 3, pp. 1–22, 2008.
[5] A. Mainwaring, J. Polaster, R. Szewczyk, D. Culler, and J. Anderson, “Wireless sensor networks for habitat monitoring,” in *Proceeding of the 1st ACM international workshop on Wireless sensor networks and applications*. Atlanta, GA: ACM, 2002, pp. 88–97.
[6] R. Ahlswede, N. Cai, S. LI, and R. Yeung, “Network information flow,” *IEEE Trans. on Information Theory*, vol. 46, no. 4, pp. 1204–1216, 2000.
[7] A. G. Dimakis, P. B. Godfrey, Y. Wu, M. J. Wainwright, and K. Ramchandran, “Network coding for distributed storage systems,” *IEEE Trans. Info. Theory*, vol. 56, no. 9, pp. 4539–4551, 2010.
[8] S. Acedanski, S. Deb, M. Medard, and R. Koetter, “How good is random linear coding based distributed networked storage,” in *Proceedings of First Workshop Network Coding, Theory, and Applications (NetCod05)*, Riva del Garda, Italy, 2005.
[9] C. GKANTSIDIS and P. RODRIGUEZ, “Network coding for large scale content distribution,” in *Proceedings of INFOCOM’2005*. Miami, FL. DEB: IEEE, 2005, pp. 2235–2245.
[10] A. Dimakis, V. Prabhakarna, and K. Ramchandran, “Ubiquitous access to distributed data in large-scale sensor networks through decentralized erasure codes,” in *Proceeding of of the International Conference on Information Processing in Sensor Networks (IPSN05)*. ACM, 2005, pp. 111–117.
[11] J. Widmer and J. Andboudec, “Network coding for efficient communication in extreme networks,” in *Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking*. Philadelphia, PA: ACM, 2005, pp. 284–291.
[12] S. Lindsey and C. Raghavendra, “Pegasis: Power-efficient gathering in sensor information systems,” in *IEEE In Aerospace Conference Proceedings*. IEEE, 2002, pp. 1125–1130.
[13] J. Gentle, *Numerical Linear Algebra for Applications in Statistics*. Springer, 1998.