How Network Coding Benefits Converge-Cast in Wireless Sensor Networks

Zhenzhou Tang*†, Hongyu Wang†, Qian Hu*, and Long Hai†
*College of Physics and Electronic Information Engineering, Wenzhou University, Wenzhou, China
†School of Information and Communication Engineering, Dalian University of Technology, Dalian, China
Email: mr.tangzz@gmail.com, whyu@dlut.edu.cn, huqian@wzu.edu.cn, hl850121@163.com

***Abstract*—Network coding is one of the most promising techniques to increase the reliability and reduce the energy consumption for wireless sensor networks (WSNs). However, most of the previous works mainly focus on the network coding for multi-cast or uni-cast in WSNs, in spite of the fact that the converge-cast is the most common communication style in WSNs. In this paper, we investigate, for the first time as far as we know, the feasibility of acquiring network coding benefits in converge-cast, and we present that with the ubiquitous convergent structures self-organized during converge-casting in the network, the reliability benefits can be obtained by applying linear network coding. We theoretically derive the network coding benefits obtained in a general convergent structure, and simulations are conducted to validate our theoretical analysis. The results reveal that the network coding can improve the network reliability considerably.**

I. Introduction

Explosive growth in embedded computing and rapid advances in low power wireless networking technologies are fueling the development of wireless sensor networks (WSNs). And WSNs have been attracting a great attention due to their wide range of potential applications. However, sensor nodes are generally powered by batteries which only provide a limited amount of power, and it is often difficult and costly to recharge or replace the batteries. Moreover, the lossy wireless links which often lead to packet delivery failures further accelerate the energy consumption due to frequent retransmissions, as we all known that data transmission dominates the energy consumption in wireless sensor networks. Therefore, reliability and energy-efficiency are essential for wireless sensor networks.

Network coding is a technique that can increase network throughput and reduce energy consumption by combining data from different input links in an intermediate node [1, 2]. Many applications have incorporated this idea since its initial proposal by Ahlswede [1]. Network coding is particularly well-suited for wireless networks due to the broadcast nature of their communications [3, 4]. It is precisely because of so many advantages of wireless network coding, the applications of network coding in WSNs are intensively studied in recent years. These works mainly involve network coding aided multi-cast and uni-cast, and the specific applications include online reprogramming [5], distributed data storage and acquisition [6], and reliable data transmission [7], etc.

By reviewing the related literature, we have found that the issue of network coding in wireless converge-cast still remains open. The consideration on network coding in converge-cast is stimulated by the special way of how a WSN typically works. Environmental data generated by the sensor nodes scattered in the monitoring regions are gathered towards the sink node hop by hop, as shown in Fig.1. Since the network engages to collect data from the sensor nodes for most of the time, *the most common communication style in WSNs should be converge-cast*. However, to the best of our knowledge, few works have been done on it. To this end, we investigate how the network coding benefits the wireless sensor networks during converge-casting, and how to obtain the network coding benefits as many as possible in this paper.

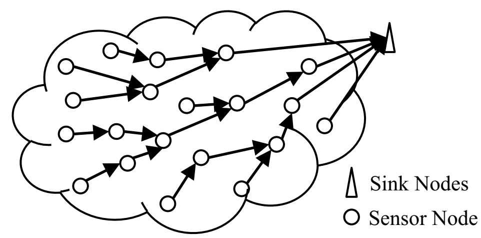


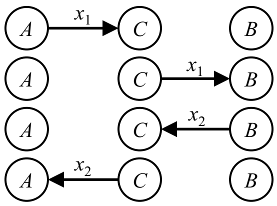
Fig. 1. Converge-cast in wireless sensor networks

The main contributes of our work are summarized as follows:

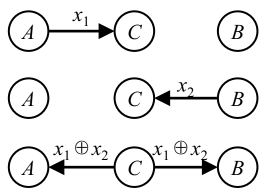
1) In this paper, we investigate, for the first time as far as we know, the feasibility of acquiring network coding benefits in converge-cast. We reveal that with the ubiquitous convergent structures self-organized during converge-casting in the network, the reliability benefits can be obtained by applying linear network coding.
2) We define the typical convergent structures where to perform network coding. And on the basis of that, we also derive the theoretical expressions to calculate the reliability benefits in a general convergent structure.

The rest of this paper is organized as follows. Section II investigates the feasibility of acquiring network coding benefits in converge-cast. The suitable network coding style and the source of network coding benefits in converge-cast are discussed. In Section III, the linear network coding model is proposed. And also in this section, the amount of reliability benefits is derived, and the numerical and simulation results are presented. Further discussion is presented in Section IV, followed by Section V giving the conclusions.

978-1-4673-1881-5/12/$31.00 ©2012 IEEE



(a) Traditional scheme needs four transmissions



(b) Inter-flow network coding scheme only needs three transmissions

Fig. 2. A simple example of how network coding can improve throughput

II. Feasibility of acquiring network coding benefits in converge-cast

Generally speaking, the wireless network coding can be divided into two categories: inter-flow network coding and intra-flow network coding [8]. In this section, we discuss about the feasibility of acquiring network coding benefits in converge-cast, that is, which kind of benefits can be produced in converge-cast.

A. Inter-flow network coding

In inter-flow network coding, the coding node, in which the data from different flows are combined (coded), broadcasts the newly generated packet to different neighbors in a single transmission. This style of coding is called inter-flow network coding because the coding is done over packets which come from different flows, and differ in their next hop. Consider the example in Fig.2, where node *A* and *B* attempt to exchange a pair of packets. The limited transmission range prevents them from communicating directly and thus they have to turn to node *C* for forwarding. In the traditional store-and-forward scheme, four transmissions are required. While in case of network coding, node *A* and *B* could transmit their respective packets to node *C*, which XORs the two packets and broadcasts the coded packet. Thus, node *A* can recover node *B*'s packet by reXOR-ing with its own packet, and so does node *B*. In this case, the exchange completes only within three transmissions. Hence, the network throughput is improved.

Let $O(C, f)$ denote the set of nodes who can overhear the flow *f* and are known by node *C*. And $D(C, f)$ denotes the set of all downstream nodes of node *C* on flow *f*. Generally, supposing the flow f intersects with f_x at node *C*, *C* is a coding node if and only if both the following conditions are satisfied: $\exists d_1 \in O(C, f) \cap D(C, f_x)$; and $\exists d_2 \in O(C, f_x) \cap D(C, f)$. However, as shown in Fig.1, so long as different flows arrive at the same node in case of converge-cast, their following paths come to the same. Therefore, neither of these two conditions can be satisfied. In other words, inter-flow network coding cannot benefit converge-cast unfortunately.

B. Intra-flow network coding

Intra-flow network coding allows the intermediate nodes to combine the packets heading to the same destination. The encode operation is merely a simple random linear combination of G original packets as $X_k = \sum_{i=1}^{G} g_k^i M_i$, where $M_i(i = 1, 2, ..., G)$ are the original packets, g_k^i are the coefficients randomly selected from a q-order *Galois field* $\mathbb{GF}(q)$, and X_k is the coded packet. The destination node can recover all the original packets if it has received no less than G linear independent encoded packets of M_i [2].

In the presence of intra-flow network coding, each received packet contains some information about all the original packets. No packet is specific for both the sender and the receiver, and no specific packet is indispensable. As a result, unlike the traditional automatic repeat request (ARQ), where the sender needs to know exactly which packets the destination misses so that it can retransmit them, the sender does not need to learn which particular packets the destination misses, and it only needs to get an ACK from the destination once it has received enough packets to decode and recover all the original packets. By this way, Intra-flow network coding is able to dramatically reduce the bandwidth consumed by those feedbacks.

We are delighted to discover that network coding in converge-cast shares certain characteristics with intra-flow schemes. Specifically, in intra-flow network coding, all the coded packets should finally be decoded in the same specific node. And similarly, packets sent by various nodes during converge-casting are tend to be joined up with each other in certain intermediate nodes or the sink node, as shown in Fig.1. That means those converging nodes have the opportunities to recover all the original packets if those packets have been coded elsewhere. Hence, in converge-cast, benefits can be obtained by exploiting the ubiquitous convergent structures self-organized during converge-casting in the network.

C. The source of benefits

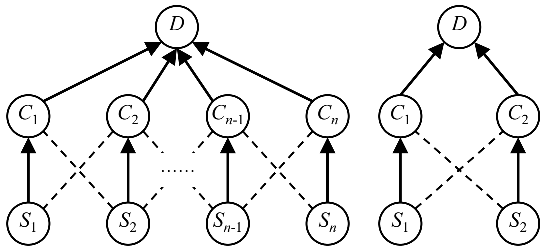
Before discussing about the source of network coding benefits in converge-cast, we first introduce some terminologies which will be used throughout the paper.

Definition 1: The *converge-tree* $\mathbb{T}$ is a directed tree which is a collection of nodes connected to the sink and pathes from those nodes to the sink in the network. The $\mathbb{T}$ is organized by non-network-coding data collection protocols, such as Collection Tree Protocol (CTP) [9]. The root of $\mathbb{T}$ is just the sink node.

Fig.1 shows a converge-tree.

Definition 2: An n-order *converge-structure* $\mathbb{CS}_n$ is a three-layer subtree of the $\mathbb{T}$ composed of n leaves $S_i(i = 1, 2, \ldots, n)$, n interior nodes $C_i(i = 1, 2, \ldots, n)$, and a root D, as shown in Fig.3(a). A $\mathbb{CS}_n$ possesses the following characteristics:

1) There is a one-to-one correspondence between the set of leaves and the set of interior nodes. Assume that S_i is the only child of C_i.
2) $C_i(i = 2, 3, \ldots, n-1)$ can overhear the communications of S_{i-1} and S_{i+1}. While C_1 and C_n can only overhear the communications of S_2 and S_{n-1} respectively.
3) A complete process of linear network coding can be performed in the $\mathbb{CS}_n$. The leaves are source nodes, the interior nodes are encoders and routers, and the root



(a) $\mathbb{CS}_n$ (b) $\mathbb{CS}_2$

Fig. 3. Converge-structures. The nodes and solid lines with arrows form a subtree of converge-tree. The dashed lines represent the overhearing links.

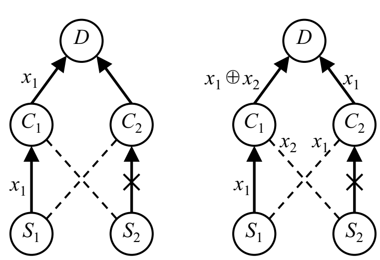
is the decoder. And the $\mathbb{CS}_2$ is the most basic one to perform network coding, as shown in Fig.3(b).

During converge-casting in WSNs, the probability of forming $\mathbb{CS}_n$ is considerably high, especially for low order $\mathbb{CS}_n$. And the computational complexity of performing network coding in $\mathbb{CS}_n$ is low. A encoder only need to combine no more than three original packets.

Next, we explain how the network coding benefits converge-cast. Consider the simple example illustrated in Fig.4(a). It is a $\mathbb{CS}_2$. The packets x_1 and x_2 are to be joined up with each other in D. However, the link from S_2 to C_2 ($\mathcal{L}_{S_2\to C_2}$) fails. Evidently, in case of non-network-coding schemes, at least one retransmission is required to ensure the success of packet delivery. But it is different if the idea of network coding is employed. Since C_1 can overhear x_2, it would combine x_1 and x_2 by means of XOR or linear network coding [2] (in fact, XOR is a special kind of linear network coding scheme with $\mathbb{GF}(2)$). Assume the coded packet is $(x_1 \oplus x_2)$. And likewise, C_2 can overhear x_1 and just only forward it to D for it only get one packet. Finally, D is able to recover x_2 by reXOR the coded packet with x_1, as shown in Fig.4(b). Hence, network coding in converge-cast is able to reduce the probability of retransmission. That is the benefit!

III. Quantitative analysis of network coding benefits in converge-cast

In this section, we study how many benefits can be obtained by applying network coding to converge-cast in WSNs. We first introduce the typical network coding model, and then we derive the theoretical benefits introduced by network coding.



(a) (b)

Fig. 4. The reliability benefits introduced by employing network coding in converge-structure.

A. The network coding model

A $\mathbb{CS}_n$ can be represented by a matrix defined as:

Definition 3: A n-order *converge-matrix* $\mathbf{M}_n$ is mapped from $\mathbb{CS}_n$. The element of $\mathbf{M}_n$, denoted as $c_{i,j}(i,j = 1,2,...,n)$, represents the coefficient of the original packet sent by S_j when performing linear network coding in C_i. The coefficients are chosen over all elements of a Galois Field $\mathbb{GF}(q)$. $\mathbf{M}_n(i,j) = 0$ if the corresponding $\mathcal{L}_{S_j\to C_i}$ fails. And if all the links work, the $\mathbf{M}_n$, which is called as the perfect $\mathbf{M}_n$, should be:

$$\mathbf{M}_n = \begin{pmatrix} c_{1,1} & c_{1,2} & 0 & \cdots & 0 & 0 & 0 \\ c_{2,1} & c_{2,2} & c_{2,3} & \cdots & 0 & 0 & 0 \\ 0 & c_{3,2} & c_{3,3} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & c_{n-2,n-2} & c_{n-2,n-1} & 0 \\ 0 & 0 & 0 & \cdots & c_{n-1,n-2} & c_{n-1,n-1} & c_{n-1,n} \\ 0 & 0 & 0 & \cdots & 0 & c_{n,n-1} & c_{n,n} \end{pmatrix}$$

The rank of $\mathbf{M}_n$ is denoted as $\mathbb{R}(\mathbf{M}_n)$.

Hence, let $\mathbf{X}_n = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$ be the vector of original packets. The vector of coded packets, which is denoted as $\mathbf{Y}_n = (y_1, y_2, \ldots, y_n)^{\mathrm{T}}$, can be calculated as $\mathbf{Y}_n = \mathbf{M}_n \cdot \mathbf{X}_n$. And D is able to recover all the original packets with $\mathbf{X}_n = \mathbf{M}_n^{-1} \cdot \mathbf{Y}_n$. Obviously, this equation has a unique solution if and only if $\mathbb{R}(\mathbf{M}_n) = n$.

Taking into consideration the limited computing capacity of a sensor node, it is reasonable to assign predetermined values to these coefficients. Since a coded packet includes no more than three original ones, for example, $\mathbb{GF}(2^3)$ is adopted. Hence, the coefficients should be chosen from $(0, 1, \alpha, \alpha^3)$ where α is the primitive element of $\mathbb{GF}(2^3)$. And the binary and decimal notations are $(000, 001, 010, 011)$ and $(0, 1, 2, 3)$ respectively. It is simple to prove that the perfect $\mathbb{R}(\mathbf{M}_n) = n$.

$$\mathbf{M}_n = \begin{pmatrix} \alpha & \alpha^3 & 0 & \cdots & 0 & 0 & 0 \\ 1 & \alpha & \alpha^3 & \cdots & 0 & 0 & 0 \\ 0 & 1 & \alpha & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \alpha & \alpha^3 & 0 \\ 0 & 0 & 0 & \cdots & 1 & \alpha & \alpha^3 \\ 0 & 0 & 0 & \cdots & 0 & 1 & \alpha \end{pmatrix} \overset{\text{decimal notation}}{=\!=\!=\!=} \begin{pmatrix} 2 & 3 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 2 & 3 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & 3 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 2 & 3 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 2 \end{pmatrix}$$

B. Quantitative analysis

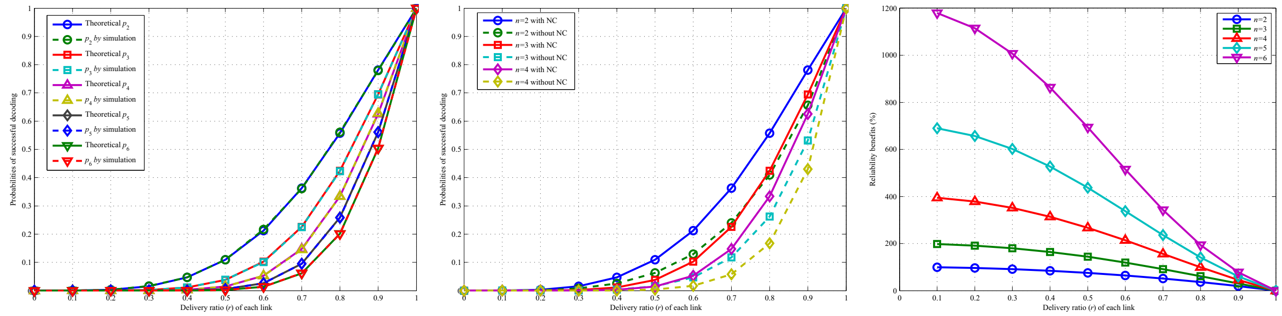
Theorem 1: For a feasible converge-cast in a $\mathbb{CS}_n$ with lossy links and a network coding in which the coefficients are determined by the corresponding $\mathbf{M}_n$, assuming that all the links share the same delivery ratio r, the probability p_n that D can recover all the original packets can be calculated as:

$$p_n = \begin{cases} r^4(2 - r^2) & \text{if } n = 2 \\ p_{n-1}(r + r^2 - r^3)r^n & \\ \quad + (p_{n-2} - p_{n-3}r^2)r^{2+n}(1 - r) & \text{if } n \geqslant 3 \end{cases}$$

Particularly, $p_0 \triangleq 1$ and $p_1 \triangleq r^2$.

Proof: First, p_2 can be calculated easily with the exhaust algorithm.

$$p_2 = \left[2r^2(1-r)^2 + \binom{3}{4} r^3(1-r) + r^4\right] r^2 = r^4(2 - r^2)$$



(a) (b) (c)

Fig. 5. Numerical and simulation results. (a). Probabilities of successful decoding for $\mathbb{CS}_n(n=2,...,6)$ under different delivery ratios; (b). Comparisons between the probabilities of successful decoding with and without network coding; (c). Reliability benefits of employing network coding in $\mathbb{CS}_n(n=2,...,6)$.

Then we calculate $p_n(n \geqslant 3)$. D can recover all the n original packets if and only if $\mathbb{R}(\mathbf{M}_n)=n$ and all the links from the encoders to the decoder are successful. In other words, $p_n = p[\mathbb{R}(\mathbf{M}_n)]r^n$ where $p[\mathbb{R}(\mathbf{M}_n)]$ is the probability of $\mathbb{R}(\mathbf{M}_n)=n$. According to the total probability formula, we have

$$p[\mathbb{R}(\mathbf{M}_n)] = (p_n^1 + p_n^2)$$
$$p_n^1 = p[\mathbb{R}(\mathbf{M}_{n-1})] \cdot p[\mathbb{R}(\mathbf{M}_n) \mid \mathbb{R}(\mathbf{M}_{n-1})]$$
$$p_n^2 = p[\overline{\mathbb{R}(\mathbf{M}_{n-1})}] \cdot p[\mathbb{R}(\mathbf{M}_n) \mid \overline{\mathbb{R}(\mathbf{M}_{n-1})}]$$

where $p[\overline{\mathbb{R}(\mathbf{M}_n)}] = p[\mathbb{R}(\mathbf{M}_n) < n]$.

In case $\mathbf{M}_{n-1}$ is a full-rank matrix, $\mathbf{M}_n$ is also full rank if either of the following conditions is satisfied.

1) $\mathcal{L}_{S_n \to C_n}$ works. Or
2) $\mathcal{L}_{S_n \to C_n}$ fails, but both $\mathcal{L}_{S_{n-1} \to C_n}$ and $\mathcal{L}_{S_n \to C_{n-1}}$ are successful. Moreover, $\mathbf{M}_{n-2}$ must be a full-rank matrix.

Hence, we have

$$\begin{aligned} p_n^1 &= p[\mathbb{R}(\mathbf{M}_{n-1})] \cdot p[\mathbb{R}(\mathbf{M}_n) \mid \mathbb{R}(\mathbf{M}_{n-1})] \\ &= p[\mathbb{R}(\mathbf{M}_{n-1})]r + p[\mathbb{R}(\mathbf{M}_{n-1})\mathbb{R}(\mathbf{M}_{n-2})]r^2(1-r) \end{aligned} \tag{1}$$

And because

$$\begin{aligned} p[\mathbb{R}(\mathbf{M}_{n-1})\mathbb{R}(\mathbf{M}_{n-2})] &= p[\mathbb{R}(\mathbf{M}_{n-1})(\Omega - \overline{\mathbb{R}(\mathbf{M}_{n-2})})] \\ &= p[\mathbb{R}(\mathbf{M}_{n-1})] - p[\mathbb{R}(\mathbf{M}_{n-1})\overline{\mathbb{R}(\mathbf{M}_{n-2})}] \end{aligned}$$

therefore

$$\begin{aligned} p_n^1 = p[\mathbb{R}(\mathbf{M}_{n-1})](r+r^2-r^3) \\ - p[\mathbb{R}(\mathbf{M}_{n-1}) \cdot \overline{\mathbb{R}(\mathbf{M}_{n-2})}]r^2(1-r) \end{aligned} \tag{2}$$

Moreover, $\mathbf{M}_n$ must be a singular matrix if $\mathbb{R}(\mathbf{M}_{n-1}) < n-2$. Assuming that $\mathbb{R}(\mathbf{M}_{n-1}) < n-1$, $\mathbf{M}_n$ is full rank if both of the following conditions are satisfied.

1) $\mathbf{M}_{n-2}$ is a full-rank matrix. And
2) Both $\mathcal{L}_{S_{n-1} \to C_n}$ and $\mathcal{L}_{S_n \to C_{n-1}}$ are successful.

Therefore, we have

$$\begin{aligned} p_n^2 &= p[\overline{\mathbb{R}(\mathbf{M}_{n-1})}] \cdot p[\mathbb{R}(\mathbf{M}_n) \mid \overline{\mathbb{R}(\mathbf{M}_{n-1})}] \\ &= p[\mathbb{R}(\mathbf{M}_n)\overline{\mathbb{R}(\mathbf{M}_{n-1})}] = p[\overline{\mathbb{R}(\mathbf{M}_{n-1})}\mathbb{R}(\mathbf{M}_{n-2})]r^2 \end{aligned} \tag{3}$$

By utilizing (3), it can be inferred that

$$p[\mathbb{R}(\mathbf{M}_{n-1})\overline{\mathbb{R}(\mathbf{M}_{n-2})}] = p[\overline{\mathbb{R}(\mathbf{M}_{n-2})}\mathbb{R}(\mathbf{M}_{n-3})]r^2 \tag{4}$$

Substitute (2), (3) and (4) into (1), we have

$$\begin{aligned} p_n = {} & p[\mathbb{R}(\mathbf{M}_{n-1})](r+r^2-r^3) + p[\overline{\mathbb{R}(\mathbf{M}_{n-1})}\mathbb{R}(\mathbf{M}_{n-2})]r^2 \\ & - p[\overline{\mathbb{R}(\mathbf{M}_{n-2})}\mathbb{R}(\mathbf{M}_{n-3})]r^4(1-r) \end{aligned} \tag{5}$$

Because

$$p[\overline{\mathbb{R}(\mathbf{M}_{n-1})}\mathbb{R}(\mathbf{M}_{n-2})] = p[\mathbb{R}(\mathbf{M}_{n-2})] - p[\mathbb{R}(\mathbf{M}_{n-1})\mathbb{R}(\mathbf{M}_{n-2})] \tag{6}$$
$$p[\overline{\mathbb{R}(\mathbf{M}_{n-2})}\mathbb{R}(\mathbf{M}_{n-3})] = p[\mathbb{R}(\mathbf{M}_{n-3})] - p[\mathbb{R}(\mathbf{M}_{n-2})\mathbb{R}(\mathbf{M}_{n-3})] \tag{7}$$

And according to (1), we have

$$\begin{aligned} & p[\mathbb{R}(\mathbf{M}_{n-1})\mathbb{R}(\mathbf{M}_{n-2})] \\ & = p[\mathbb{R}(\mathbf{M}_{n-2})]r + p[\mathbb{R}(\mathbf{M}_{n-2})\mathbb{R}(\mathbf{M}_{n-3})]r^2(1-r) \end{aligned} \tag{8}$$

Substitute (6), (7) and (8) into (5), and simplify it, we have

$$\begin{aligned} p[\mathbb{R}(\mathbf{M}_n)] = {} & p[\mathbb{R}(\mathbf{M}_{n-1})](r+r^2-r^3) \\ & + p[\mathbb{R}(\mathbf{M}_{n-2})]r^2(1-r) - p[\mathbb{R}(\mathbf{M}_{n-3})]r^4(1-r) \end{aligned}$$

That is:

$$p_n = p_{n-1}(r+r^2-r^3)r^n + (p_{n-2} - p_{n-3}r^2)r^{2+n}(1-r)$$

The theorem is proved. ■

The probability that all the packets are successfully received by D without network coding, denoted as p'_n, can be calculated as $p'_n = r^{2n}$. And the reliability benefit B introduced by network coding in $\mathbb{CS}_n$ is $B = \frac{p_n - p'_n}{p'_n}$.

C. Numerical and Simulation Results

Simulations have been conducted to validate Theorem 1. Network coding benefits of $\mathbb{CS}_n(n=2,3,4,5,6)$ are verified under various link delivery ratios. Altogether 10^5 experiments are done to count the number of successful decoding in D. The comparisons between the theoretical and simulation results are shown in Fig.5(a), where we can find that the theoretical results are just the same as those of simulations.

Fig.5(b) presents the comparisons between the probabilities of successful recovery of data in $\mathbb{CS}_n(n=2,3,4)$ with and without network coding. Apparently, network coding improves the reliabilities dramatically. It is also shown in Fig.5(c), where the reliability benefits introduced by employing network coding in $\mathbb{CS}_n(n=2,3,4,5,6)$ are illustrated, where p'_n is the probability that all the packets are successfully received

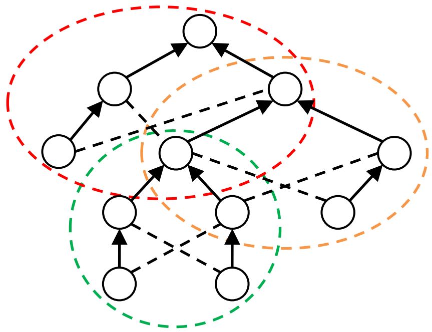


Fig. 6. The overlapped $\mathbb{CS}_n$. Each ellipse contains a $\mathbb{CS}_n$.

by D without network coding. We can see that the benefits are surprisingly high, particularly in case of poor link quality.

IV. Discussion

A. Prerequisites of obtaining network coding benefits

It can be found that: Firstly, without overhearing, none of $C_i(i = 1, 2, \ldots, n)$ can receive any packet sent by source node other than its own child, which makes network coding impossible. The existence of overhearing can be interpreted as the introduction of redundant links, which is certainly capable of improving transmission reliability. Secondly, if all the links were completely reliable, no retransmission would be needed, and the network coding would not introduce any benefit at all. Therefore, both the feasibility of overhearing and the unreliability of links are the prerequisites to obtain the benefits of network coding in $\mathbb{CS}_n$.

B. The need of network coding

There may be a puzzle about the necessity of network coding in $\mathbb{CS}_n$. For example, as shown in Fig.4(b), instead of sending the coded packet $(x_1 \oplus x_2)$, C_1 only need to forward x_2 to D, which can also achieve the aim. Moreover, there would be overhead on combining x_1 and x_2. But the question is how C_1 could learn about which packet to forward. In this case, information exchange between C_1 and C_2 is indispensable for them to coordinate the forwarding. And what's more, in a $\mathbb{CS}_n$, a $C_i(i = 1, 2, \ldots, n)$ has to negotiate with up to five adjacent interior nodes about the right packet to forward, which introduces much more overhead and latency than network coding.

C. How to explore more coding opportunities

As mentioned above, there may be plenty of $\mathbb{CS}_n$ formed during converge-casting. Some of them could be overlapped to explore more network coding opportunities, as shown in Fig.6. A node may act as a source node in a $\mathbb{CS}_n$, and meanwhile, it is the decoder or one of the encoder in another $\mathbb{CS}_n$. The overlap can increase the network coding opportunities, and thus more network coding benefits can be achieved.

V. Conclusion and Future Work

In this paper, we have demonstrated that linear network coding can introduce considerably high reliability benefits during converge-cast in WSNs. We have presented the typical converge-structures where to perform network coding, and have theoretically analyzed the amount of benefits for converge-structures with various scales.

Potential areas of future work include distributed converge-structures planning and the design of a complete network based converge-cast scheme.

Acknowledgment

This work is supported by the Zhejiang Provincial Natural Science Foundation of China (LQ12F02009), the Non-profit Technology Application Research Projects of Zhejiang Province (2011C310290029) and the Scientific Research Fund of Zhejiang Provincial Education Department (Y201121034).

References

[1] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, 2000.

[2] H. Tracey, M. Muriel, S. Jun, E. Michelle, and R. K. David, "On randomized network coding," in *The 41st Annual Allerton Conference on Communication, Control, and Computing*, vol. 41, 2003, pp. 11–20.

[3] S. Chachulski, M. Jennings, S. Katti, and D. Katabi, "Trading structure for randomness in wireless opportunistic routing," in *the 2007 conference on Applications, technologies, architectures, and protocols for computer communications*, 2007, pp. 169–180.

[4] S. Katti, H. Rahul, W. Hu, D. Katabi, M. M, and J. Crowcroft, "Xors in the air: practical wireless network coding," in *The 2006 conference on Applications, technologies, architectures, and protocols for computer communications*, 2006, pp. 243–254.

[5] I. H. Hou, T. Yu-En, T. F. Abdelzaher, and I. Gupta, "Adapcode: Adaptive network coding for code updates in wireless sensor networks," in *The IEEE 27th Conference on Computer Communications*, 2008, pp. 1517–1525.

[6] 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, 2008.

[7] Y. Yang, C. Zhong, Y. Sun, and J. Yang, "Network coding based reliable disjoint and braided multipath routing for sensor networks," *Journal of Network and Computer Applications*, vol. 33, no. 4, pp. 422–432, 2010.

[8] C. Fragouli, D. Katabi, A. Markopoulou, M. Medard, and H. Rahul, "Wireless network coding: Opportunities & challenges," in *The 2007 IEEE Military Communications Conference*, 2007, pp. 1–8.

[9] O. Gnawali, R. Fonseca, K. Jamieson, D. Moss, and P. Levis, "Collection tree protocol," in *the 7th ACM Conference on Embedded Networked Sensor Systems*, 2009, pp. 1–14.