Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Suboptimal network coding subgraph algorithms for 5G minimum-cost multicast networks*

Feng WEI†‡1, Wei-xia ZOU1,2

1MOE Key Lab of Universal Wireless Communications, Beijing University of Posts and Telecommunications, Beijing 100876, China

2State Key Lab of Millimeter Waves, Southeast University, Nanjing 210096, China

†E-mail: weifeng_luck@163.com

Received Jan. 8, 2017; Revision accepted Apr. 20, 2017; Crosschecked May 10, 2018

Abstract: To reduce the transmission cost in 5G multicast networks that have separate control and data planes, we focus on the minimum-power-cost network-coding subgraph problem for the coexistence of two multicasts in wireless networks. We propose two suboptimal algorithms as extensions of the Steiner tree multicast. The critical 1-cut path eliminating (C1CPE) algorithm attempts to find the minimum-cost solution for the coexistence of two multicast trees with the same throughput by reusing the links in the topology, and keeps the solution decodable by a coloring process. For the special case in which the two multicast trees share the same source and destinations, we propose the extended selective closest terminal first (E-SCTF) algorithm out of the C1CPE algorithm. Theoretically the complexity of the E-SCTF algorithm is lower than that of the C1CPE algorithm. Simulation results show that both algorithms have superior performance in terms of power cost and that the advantage is more evident in networks with ultra-densification.

Key words: Network coding subgraph; Minimum power cost; 5G; Separation architecture
https://doi.org/10.1631/FITEE.1700020 **CLC number:** TN92

1 Introduction

With the expansion of mobile games, live casts, and webcasts, multicast connections apply considerable pressure on existing communication networks. Designed as the next generation mobile network, 5G (Chen and Kang, 2018) is expected to be heterogeneous and able to cope with various services. Networking in 5G systems will be different from previous implementations as it adopts an architecture (Ma et al., 2015; Szabó et al., 2015; Akyildiz et al., 2016; Chen et al., 2016) of control plane and data plane separation from the software-defined network (SDN). The separation of the control and transmission architectures with network function virtualization (NFV) and network slicing enables a 5G network to be a comprehensive network for smartwatches, autonomous vehicles, Internet of Things (IoT), and tactile Internet with customized services (Szabó et al., 2015; Akyildiz et al., 2016). A logically centralized network controller in the separation architecture can easily change the routing configuration without affecting the whole network (Kotronis et al., 2012; Akyildiz et al., 2016), thus achieving flexibility in multicast delivery.

In addition, the requirement to increase the aggregate data rate by about 1000× 4G and decrease the roundtrip latency to roughly 1 ms (Andrews et al., 2014) has pushed 5G systems to reorganize their networks through various technologies. The 'big three' 5G technologies (Agyapong et al., 2014; Andrews et al., 2014) are millimeter-wave (mmWave), massive multiple-input multiple-output (MIMO), and ultra-densification. The mmWave frequencies could

‡ Corresponding author

* Project supported by the National Natural Science Foundation of China (No. 61571055), the Fund of SKL of MMW (No. K201815), and the Important National Science & Technology Specific Projects (No. 2017ZX03001028)

ORCID: Feng WEI, http://orcid.org/0000-0002-0484-5946

support 2.6 GHz radio spectrum bands for wireless communications (Rappaport et al., 2011, 2013), which allows the potential for improving the performance in capacity and delay. The smaller wavelengths of mmWave frequencies require new spatial processing techniques (Rappaport et al., 2013; Kulkarni et al., 2016) such as massive MIMO (Rusek et al., 2013; Li et al., 2017) and adaptive beamforming (He et al., 2014; Ali et al., 2017) to match the polarization. The arrow beams created by spatial processing techniques can potentially overcome severe link attenuation (Choi, 2015), and a proper medium access control (MAC) scheme can handle the resulting interference-limited regime and create large capacity links (Shokri-Ghadikolaei and Fischione, 2016). An ultra-dense network (UDN) uses the links and nodes of densification to meet the end-user data rates on the order of 10 Gb/s in localized environments (Baldemair et al., 2015). For access networks in the network architecture, densification in UDN is a challenge for a cellular structure with additional relay nodes added to support multi-hop access (Ma et al., 2015). A 'de-cellular' solution (Chen et al., 2016) or a modification (Su et al., 2016) is needed. With these powerful underlying technologies, looking for a minimum networking strategy to use the technologies effectively is an urgent problem. In this paper we will introduce two networking algorithms as extensions of minimum-cost multicast transmission.

The minimum-cost networking problem is a branch of the optimal resource allocation problem. The fundamental work in Chiang et al. (2007) modeled the network architecture with abstract constraints, which enlarged its extensibility for the resource allocation problem. The basic framework soon evolved into a wireless model (Ribeiro and Giannakis, 2010), the polymatroid flow model (Riemensberger and Utschick, 2014), to outline the wireless broadcast channel, and then a stochastic model (Stai et al., 2015) for a wireless time-varying channel. These models in cooperation with new technologies have made many achievements (Li et al., 2016; Wang et al., 2016), in which network coding is involved to achieve a minimum-cost multicast solution. An approach to solve the minimum-cost multicast problem produced an optimal network coding subgraph. The optimal solutions (Lun et al., 2005; Heindlmaier et al., 2011; Rajawat et al., 2011) are usually derived through a dual approach that decomposes the primary problem and makes its computational complexity and convergence acceptable for small networks.

However, classical multicast is insufficient in 5G networks because the cost depends on the throughput in a wireless link and the coexistence is not within the scope of classical multicast technologies. The collision of different multicast trees with overlapping links will render the destinations unable to receive the desired message in one period, which means there is no network coding solution for multicast tree collisions. In Wang and Shroff (2010), Theorems 1 and 2 derive the equivalent conditions for the existence of a coding solution for two multicasts, which can be a theoretical guidance to find solutions for the collision of two multicast sessions. In a special case where the source and destinations of two multicasts are the same, the scenario describes the throughput increase in one multicast. In Wei and Zou (2017), we created a network coding subgraph for one multicast session with two cut-sets for every destination. A 2-cut-set topology was generated out of a Steiner tree under the target of the least cost increment. Wang et al. (2007) proposed a D-algorithm method to create a coding subgraph for increasing throughput. Jiang et al. (2015) added a coding structure to a multicast routing tree for energy efficiency, increasing the throughput without being specified.

In this paper, we first study a special case where one multicast session doubles its throughput as a supplement to our work in Wei and Zou (2017). We attach an optimization problem to outline the optimal solution, which is solvable with the Gurobi optimizer, and elaborate its relationship with the selective closest terminal first (SCTF) algorithm in Ramanathan (1996). The extended SCTF algorithm (E-SCTF) is a straightforward extension of the SCTF algorithm. With a large enough priority parameter, the steps in SCTF to find a Steiner tree for destinations can be similarly re-expressed as the steps in E-SCTF, implying the algorithm's extensibility.

The second algorithm, a critical 1-cut path eliminating (C1CPE) algorithm, generalizes E-SCTF into the coexistence of two multicasts with the same throughput, aiming to find the one with the least cost from the decodable solution changing from two multicast trees. The C1CPE algorithm is a realization of the theoretical work in Wang and Shroff (2010).

In addition, we attach an optimization problem derived from Heindlmaier et al. (2011) to outline the optimal solution for two multicast sessions with overlapping links.

2 Model

The model is based on an architecture for control and data plane separation. The nodes in the data plane perform multicast sessions and the wireless network, made up of these nodes, is modeled by a directed graph G=(N, A), where N denotes the set of nodes and A denotes the set of links. The multicast session in G is indexed by m, namely, S_m:=(s_m, D_m, a_m), s_m ($s_m \in N$) is the source, D_m ($D_m \subset N$) is the set of destinations, and a_m is the throughput of the multicast.

In this study, we focus on the throughput doubling for one multicast session and the coexistence of two multicast sessions with the same throughput. To simplify the description, we denote the throughput of a link in the original multicast tree as 1, and if the throughput of one link is doubled, we denote the throughput in the link as 2.

Thus, the multicasts with the same throughput can be expressed as (s_1, D_1, 1) and (s_2, D_2, 1). The description for the coexistence of two multicasts is detailed in Section 3.2.1. Furthermore, in a special case where the two multicasts share the same source and destinations, the model describes the variation for one multicast from (s, D, 1) to (s, D, 2), which is detailed in Section 3.1.1.

2.1 Logically centralized controller

We consider a network layer where the underlying link can be flexibly directed to a desired node through multiple antenna technology. In addition, global information, such as the topology and link throughput, can be maintained through a logically centralized controller to obtain the minimum-cost solution as in Fig. 1. The logically centralized controller collects local information from every node to achieve global acknowledge, and then calculates the networking and distributes different instructions for every node. Fig. 1 is one controller for a local area, and the whole separation architecture can be found in Szabó et al. (2015) and Akyildiz et al. (2016).

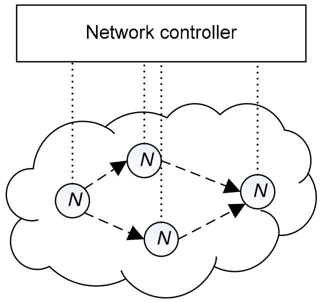

Fig. 1 The control plane and data plane separation

2.2 Power cost analysis

For a link (i, j)∈A with throughput of a, $w_{ij}^{(a)}$ presents the cost that the controller assigns to it, thus constructing the cost matrix $\boldsymbol{W}^{(a)}$. In the following context, $\boldsymbol{W}^{(1)}$ and $\boldsymbol{W}^{(2)}$ present the cost matrices with throughputs of 1 and 2, respectively.

For cost matrices $\boldsymbol{W}^{(1)}$ and $\boldsymbol{W}^{(2)}$, the elements in them represent the cost in the graph. In general, for every $w_{ij}^{(1)} \in \boldsymbol{W}^{(1)}$ and $w_{ij}^{(2)} \in \boldsymbol{W}^{(2)}$, the following inequality is established:

$$w_{ij}^{(1)} \le w_{ij}^{(2)}. \tag{1}$$

In the following analysis, we assume that the controller chooses the transmission power $P^{\rm t}$ of the sender as the weight for the different capacities of the links:

$$w_{ij} = P_{ij}^{\rm t} = L_{ij} P_{\rm r}, \tag{2}$$

where L_{ij} is the loss on link (i, j) and $P_{\rm r}$ is the received power.

Given a destination d, there are two ways to increase the cut-set value from 1 to 2. One is to create a new path; the other is to increase the capacity of each link in the existing path. Creating a new path will lead to the addition of power cost, while doubling the capacity of an existing link would cost more.

Suppose C_0 is the capacity of an existing path:

$$C_0 = \log\left(1 + \frac{P_{\rm r}}{N_0}\right), \tag{3}$$

where $P_{\rm r}$ is the signal power at the receiver and N_0 is the noise power. After doubling the capacity, the power at the receiver becomes $P_{\rm r}'$.

$$2C_0 = \log\left(1+\frac{P_r'}{N_0}\right). \quad (4)$$

Then, we have

$$P_r' = 2P_r + P_r^2/N_0. \quad (5)$$

When the initial signal-to-noise ratio (SNR) $\gamma=P_r/N_0>>2$, we have

$$P_r' \approx \gamma P_r. \quad (6)$$

Eq. (6) means that if doubling the capacity of a link, the receiving power increases by γ times. Considering Eq. (2), the transmission power should also be increased by γ times when the other conditions of the link are unchanged, which means that the cost of the link will be increased by γ times when the link's capacity is doubled. The cost can be re-expressed as follows:

$$w_{ij}^{(1)} = k\gamma d_{ij}^{\alpha}, \quad (7)$$

$$w_{ij}^{(2)} = k\gamma^2 d_{ij}^{\alpha}, \quad (8)$$

$$w_{ij}^{(0)} = 0, \quad (9)$$

where k is a constant, γ is the SNR required by the receiver to build a connection, d_{ij} is the Euclidean distance between nodes i and j, and α is the loss factor for distance.

3 Algorithms

3.1 E-SCTF algorithm for one multicast session with doubled throughput

Here, we present the E-SCTF algorithm to reduce the cost of doubling the throughput in one multicast session. An informal description is given first before a formal specification. The minimum-cost problem supporting the algorithm is described in Section 3.1.2, and the experimental analysis will be given in Section 4.2.

Before the informal description, one note should be mentioned for E-SCTF.

Full message node: a full message node is a node with a cut-set value of 2 to the source. The full message node can acquire all the information from the source, and all of the full message nodes, including the source node, compose a full message set, and we denote this as F in the E-SCTF algorithm.

3.1.1 Informal description

For a given graph $G=(N, \mathcal{A})$, a minimum-cost solution for a multicast session $(s, D, 2)$ needs to be realized. E-SCTF is an extension of Ramanathan (1996) to alter the cut-set value of the topology from 1 to 2.

At the beginning, initialize the cost matrix $\boldsymbol{W}$ by unit throughput cost matrix $\boldsymbol{W}^{(1)}$ and calculate the Steiner tree with the SCTF algorithm in Ramanathan (1996) with a priority parameter that is large enough, and all the destinations are unmarked. During each iteration, add the source and the nodes with the in-degree equal to 2 to the full message node set. For every unmarked destination d_k, set $w_{ij}^{(2)}$ for the cost of every link (i, j) on the existing connection L_k from the nearest full message node in F to d_k, i.e., increase the cost of (i, j) by γ times. Then set $w_{ij}^{(0)}$ for the cost of the rest of the links in the topology, i.e., set them to zeros. We assume the channel in the graph to be a half-duplex. Recalculate the distance from source s to d_k. Choose the destination whose cost from source s is the least. Mark this destination and attain its new path into the coding subgraph. Repeat the iteration, until all of the destinations are marked. Then we obtain a Steiner-tree-based network coding subgraph with a cut-set value of 2. E-SCTF re-uses the existing 'pipeline' in the topology, thus gaining a low-cost coding subgraph of 2-cut-set.

We illustrate the E-SCTF algorithm with an example. In Fig. 2a, a multicast tree is created for the multicast session $(s, \{d_1, d_2, d_3\}, 1)$ through the SCTF algorithm by the following equivalent steps:

Step 1: Set the cost of the graph by $\boldsymbol{W}^{(1)}$.

Step 2: Find the least cost paths from s to d_1, d_2, and d_3, respectively. Choose the minimum cost in the three paths, suppose it to be the path to d_3, i.e., (s, d_3), and add it to the Steiner tree.

Step 3: Set the cost of sd_3 to $w_{sd_3}^{(0)}$, i.e., zero. Find the least cost paths from s to d_1 and d_2, respectively. Choose the minimum cost in the two paths. Add it to the existing Steiner tree.

Step 4: Set the cost of link (i, j) in the existing Steiner tree to $w_{ij}^{(0)}$, i.e., zero, and find the least cost path for the last destination. Then, we obtain a Steiner tree as given by Fig. 2a.

For the existing multicast tree, we will find an additional path for every destination through the E-SCTF algorithm by the following steps (Wei and Zou, 2017):

Step 1: Find the least cost paths from s to d_1, d_2, and d_3 with a new cost weight.

Take d_1 as an example. Set the cost of the existing path from s, link (s, B), (B, A), and (A, d_1), to $w_{sB}^{(2)}$, $w_{BA}^{(2)}$, and $w_{Ad_1}^{(2)}$, respectively, i.e., increase the cost by γ times. Then, set the cost of the rest of the links in the topology, (s, d_3) and (B, d_2), to $w_{sd_3}^{(0)}$ and $w_{Bd_2}^{(0)}$, respectively. In this situation, the least cost path is found for d_1, supposing it to be the path of (s, d_3), (d_3, B), (B, d_2), and (d_2, d_1). The links with the dashed line in Fig. 2b are the new links added to the topology with the cost of $w_{ij}^{(1)}$.

Step 2: Choose the minimum cost in the three paths, and add it to the existing topology.

Step 3: Find the least cost paths for the rest of the destinations, respectively, and add the one with the minimum cost to the existing topology, until all the destinations acquire their additional path.

E-SCTF adopts the method for the joining order of destinations from SCTF with alternation in the link cost weight. The alternation enables the extensibility of the algorithm.

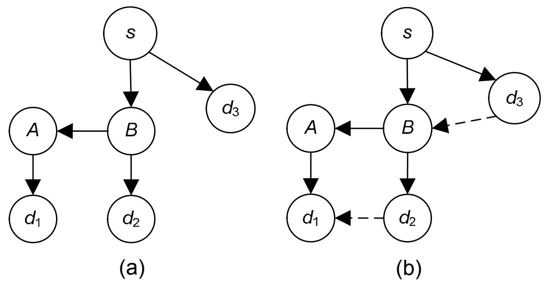

Fig. 2 Steiner tree for three destinations (a) and the coding subgraph for d_1 (b)

If the total number of nodes is N and the number of destinations is M, the complexity of the E-SCTF algorithm is $O(M^2N^2)$. The complexity depends mainly on the shortest path algorithm.

Algorithm 1 gives the E-SCTF algorithm.

3.1.2 Minimum cost problem

In multicast networks, finding the minimum-cost tree formed by routing is a Steiner tree problem. In Steiner trees, the cut-set value between the source and destination is 1. As an alternative to routing, network coding can form a minimum-cost coding subgraph with an arbitrary cut-set value.

Compared to traditional routing, network coding removes the particularity of a single message, as all the information can be acquired as long as a sufficient amount of messages is received. In multicast scenarios, when a network with a 2-cut-set value for every destination is established, the two-dimensional information sent by the source can be received in the same period by each destination with probability 1 (Ho et al., 2006). Using random linear network coding, a link can be shared among paths to different destinations to reduce the power cost of the network.

Algorithm 1 E-SCTF

```
Input: G, s, dest.
Output: topo.
STree=SCTF(G, s, dest)
uMarkD←dest; topo←STree
while ~isempty(uMarkD)
  fullMsgSet←[s, topo(indegree==2)]
  P=[]
  for i=1:length(uMarkD)
    Set all links in topo to zero
    Increase the cost of links along the existing path
    from the nearest full message node to uMarkDest(i)
    by γ times
    P(i)←leastCostPath(s, Dest(i), G)
  end for
  [p0, d0]←min(P)
  topo←topo+p0
  uMarkD←uMarkD–d0
end while
return topo
```

A minimum problem (Lun et al., 2005) that outer bounds our algorithm can be expressed as follows when the cut-set value is specified to 2:

$$\begin{aligned}&\min \sum_{(i,j)\in\mathcal{A}} P_{ij}(z_{ij}) \quad \text{s.t. } z\in\mathbb{N}, \forall(i,j)\in\mathcal{A},\\ &z_{ij}\geq x_{ij}^{(d)}\geq 0,\quad \forall(i,j)\in\mathcal{A}, d\in D,\\ &\sum_{\{j|(i,j)\in\mathcal{A}\}} x_{ij}^{(d)} - \sum_{\{j|(j,i)\in\mathcal{A}\}} x_{ji}^{(d)} = \sigma_i^{(d)},\ \forall d\in D, i\in N,\end{aligned} \tag{10}$$

where

$$\sigma_i^{(d)}=\begin{cases}2, & \text{if } i=s,\\ -2, & \text{if } i=d,\\ 0, & \text{otherwise.}\end{cases}$$

$\mathbb{N}$ is the set of non-negative integers, $\mathcal{A}$ is the set of arcs, and $\{x_{ij}^{(d)}\}$ constructs a virtual flow from s to d. $\{z_{ij}\}$ represents the number for unit throughput and constructs the actual flows in the network, also called a subgraph. $P_{ij}(z_{ij})$ denotes the power cost to transmit actual flows on the arc (i, j). A minimum cost coding subgraph can be obtained after solving the optimization problem above.

The optimization above is an integer programming problem. To keep the linear characteristic of the constraints, we remove the mutually exclusive $x_{ij}^{(d)}$ and $x_{ji}^{(d)}$, which cause the solution to problem (10) to be loose for our algorithm.

The integer z_{ij} represents the number of throughput units, which means the actual throughput on arc (i, j) is $z_{ij}C_0$, where C_0 is the throughput unit. In fact, the cut value in Eq. (10) is 2, and we can infer $z_{ij}\in\{0, 1, 2\}$. For the same link, the cost is different when the throughput is 1 or 2. Then problem (10) can be re-expressed as an integer programming problem:

$$\begin{aligned} &\min \sum_{(i,j)\in\mathcal{A}} w_{ij}^{(1)} I_{z_{ij}=1} + w_{ij}^{(2)} I_{z_{ij}=2} \ \ \text{s.t.}\ z\in\mathbb{N}, \forall (i,j)\in\mathcal{A}, \\ &z_{ij} \ge x_{ij}^{(d)} \ge 0, \quad \forall (i,j)\in\mathcal{A}, d\in D, \\ &\sum_{\{j|(i,j)\in\mathcal{A}\}} x_{ij}^{(d)} - \sum_{\{j|(j,i)\in\mathcal{A}\}} x_{ji}^{(d)} = \sigma_i^{(d)}, \quad \forall d\in D, i\in N, \end{aligned} \tag{11}$$

where $w_{ij}^{(1)}\in W_1$ is the cost weight for link (i, j) when its throughput is 1, and $w_{ij}^{(2)}\in W_2$ is the cost weight for throughput 2. In addition, I_x is the indicator function.

When choosing power as the cost weight, the transmission power on link (i, j) of z_{ij} throughput units is as follows:

$$P_{ij} = \frac{(1+\gamma_0)^{z_{ij}} - 1}{\gamma_0} P_0 = [(1+\gamma_0)^{z_{ij}} - 1] L_{ij} N_0, \tag{12}$$

where

$$\gamma_0 = 2^{C_0} - 1 \tag{13}$$

is the receiving SNR when the throughput is C_0, and

$$P_0 = \gamma_0 L_{ij} N_0. \tag{14}$$

Approximately, $P_{ij}\approx\gamma_0^2 L_{ij}N_0$ when $z_{ij}=2$.

With the discrete value of the object function, i.e., $z_{ij}\in\{0, 1, 2\}$, $P_{ij}(z_{ij})$ can be re-described as an integer quadratic function of z_{ij}:

$$P_{ij} = \left[\left(\frac{\gamma_0^2}{2} - \gamma_0\right) z_{ij}^2 + \left(2\gamma_0 - \frac{\gamma_0^2}{2}\right) z_{ij}\right] L_{ij} N_0. \tag{15}$$

Then problem (11) will be transformed into an integer quadratic programming problem. Problem (16) is the actual optimal problem that outer bounds E-SCTF:

$$\begin{aligned} &\min \sum_{(i,j)\in\mathcal{A}} \left[\left(\frac{\gamma_0^2}{2} - \gamma_0\right) z_{ij}^2 + \left(2\gamma_0 - \frac{\gamma_0^2}{2}\right) z_{ij}\right] L_{ij} N_0 \\ &\text{s.t.} \ \ z_{ij}\in\mathbb{N}, \ \ \forall (i,j)\in\mathcal{A}, \\ &z_{ij} \ge x_{ij}^{(d)} \ge 0, \ \ \forall (i,j)\in\mathcal{A}, d\in D, \\ &\sum_{\{j|(i,j)\in\mathcal{A}\}} x_{ij}^{(d)} - \sum_{\{j|(j,i)\in\mathcal{A}\}} x_{ji}^{(d)} = \sigma_i^{(d)}, \ \ \forall d\in D, i\in N. \end{aligned} \tag{16}$$

With the linear constraints, the transformed problem is solvable with tools like the Gurobi optimizer. Considering the computational complexity, only a few small networks are involved in the simulation for the optimal solution to problem (16).

3.2 C1CPE algorithm for two-multicast session with overlapping links

The coexistence of two multicasts is the generalization of doubling the throughput of one multicast. An algorithm is presented here to reduce the cost of creating decodable topology for a two-multicast situation. An informal description will be made before a formal specification is given. The minimum-cost problem supporting the algorithm is described in Section 3.2.3, and the experimental analysis will be given in Section 4.3.

Before introducing our algorithm, we should clarify some ideas and a definition from Wang and Shroff (2010), which is also important in our algorithm.

Critical 1-edge cut (Wang and Shroff, 2010): For any node u in a directed acyclic/cyclic network, the critical 1-edge cut is the farthest one from u among the 1-edge cuts separating two sources $\{s_1, s_2\}$ and u.

Slave tree: A slave tree is a structure created and shared by one or more destinations of the same source without being shared with the destination(s) of another source.

1-cut path: An un-bifurcated path starting from a critical 1-edge cut ends at the root of the slave tree. A slave tree can connect to only one cut path and a cut path can connect to at most two slave trees in the algorithm. Actually, a 1-cut path is composed of connected overlapping links.

3.2.1 Informal description

For a given graph G=(N, A), two multicast sessions (s_1, D_1, 1) and (s_2, D_2, 1) coexist with overlapping links. To deliver messages to the destinations in one period, we can double the capacity of the overlapping links, or add an assistance path to the topology. The minimum cost is the criterion for our choice.

Adding an assistance path to the topology brings a decodable problem. Theorem 1 in Wang and Shroff (2010) shows the importance of a critical 1-edge cut for a decodable problem of two multicast sessions, and an inference can be made: If the 1-cut path shared between destinations of different sources is doubled or parallel links are obtained, a network coding solution will be available.

Before the informal description, we will briefly illustrate the C1CPE algorithm with an example:

Two multicast sessions, (s_1, {d_{11}, d_{12}}, 1) and (s_2, {d_{21}, d_{22}, d_{23}}, 1), share link AB in the multicast trees. The dotted circle nodes have acquired message from S_1 and the bold circle nodes from S_2. The destinations d_{12}, d_{21}, and d_{22} cannot obtain their desired message. In Fig. 3a, link AB is a critical path for d_{12}, d_{21}, and d_{22}. For those destinations, there are two solutions to obtain their desired message. One is to find the path from the nodes with the desired message and abandon the original path. As illustrated in Fig. 3b, a path from d_{23} to d_{21} and d_{22} provides the desired message for them. The other solution is to acquire another message, which is not its desired one and is different from the acquired message, to acquire the desired message by decoding. As illustrated in Fig. 3c, a path from d_{11} to C gives node C an additional message, and then the message desired by d_{21} and d_{22} can be decoded from the acquired messages. Fig. 3 shows that nodes B, C, d_{21}, and d_{22} compose a tree structure. Once node B or C acquires the desired message, d_{21} and d_{22} will receive it subsequently. We call the tree structure the slave tree for d_{21} and d_{22}.

The method in C1CPE is to eliminate the 1-cut path tangled between the destinations of different sources with minimum cost. The algorithm contains two main processes, backward searching and forward painting, to determine the quantity of the 1-edge cut path. After the topology Topo of two Steiner trees is created with the SCTF algorithm (Ramanathan, 1996), backward searching divides the destinations into D_i^s and D_i^p, and then finds slave trees SlvT and the corresponding 1-cut path Ct for destinations in D_i^s. Forward painting follows. Forward painting specifies the color for every node in the topology. Destinations in D_i^s will not acquire the desired color.

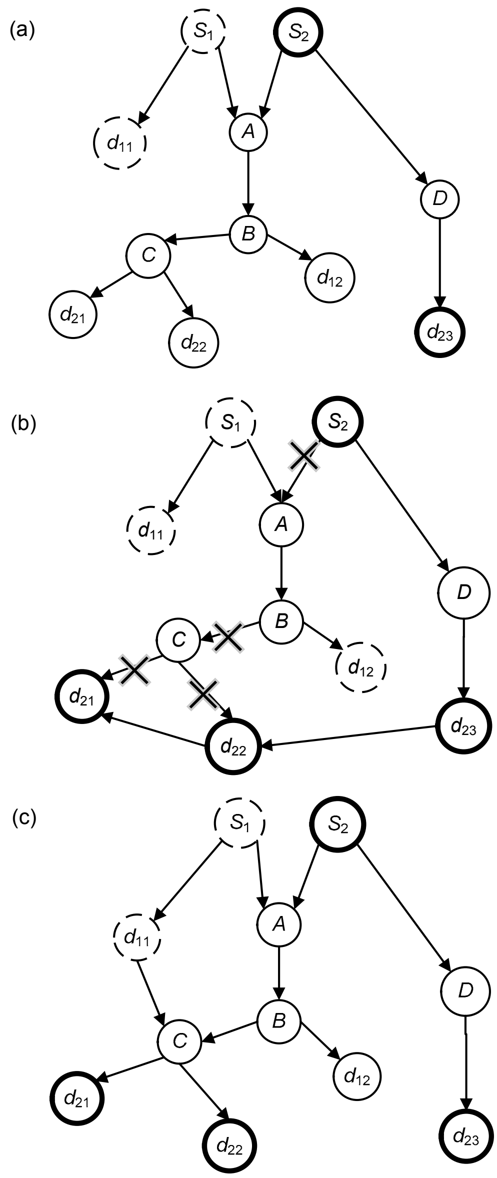

Fig. 3 An illustration of the C1CPE algorithm: two-multicast session with overlapping links (a), routing solution for d_{21} and d_{22} obtaining desired message (b), and decoding solution for d_{21} and d_{22} obtaining the desired message (c)

The dotted circle nodes have acquired message from S_1 and the concentric circle nodes have acquired message from S_2

In the algorithm, a color represents a message flow. The flow sourcing from s_1 has color 1, and s_2 has color 2. The other flow is the mixture of color 1 and color 2 with random linear network coding. Any two different colors can be decoded as color $\{1, 2\}$, which is the desired color for the two destination groups, respectively.

After forward painting, the topology acquires its color for every node and a slave tree with the same desired color needs to find the lowest cost way to obtain an additional color. The additional color can be acquired from the node set smcl with the slave tree's desired color, or the node set othcl with another different color. When the additional color is the slave tree's desired color, delete the useless part of the original path. The solution should be the minimum-cost increment for the whole slave tree. Since there might be many slave trees in the topology, the algorithm chooses the minimum-cost solution at each iteration. Once the optimal strategy optStra is acquired, add it to the existing topology Topo. The process is repeated until D_i^{s} is empty.

In summary, the algorithm selects the minimum-cost path of the minimum path(s) of the slave trees to release some destinations from the 1-edge cut. The algorithm requires calculating the shortest path at each iteration with a Floyd algorithm. If the total number of nodes is N, and the number of slave trees is M, the complexity of the algorithm is $O(M^2N^3)$. The complexity counts mainly on the shortest path algorithm.

Algorithm 2 gives the C1CPE algorithm.

3.2.2 Minimum cost problem

The minimum-cost problem for two multicasts is derived from Heindlmaier et al. (2011). Here, the optimal problem is restricted by two actual multicasts and one virtual multicast. The virtual multicast session index is 3, and the destination set for the virtual multicast is D^3. The optimal problem is as follows:

$$\min \sum_{(i,j)\in\mathcal{A}} \left[\left(\frac{\gamma_0^2}{2}-\gamma_0\right) z_{ij}^2 + \left(2\gamma_0 - \frac{\gamma_0^2}{2}\right) z_{ij}\right] L_{ij} N_0$$

$$\text{s.t.} \quad z_{ij} \in \mathbb{N},$$

$$z_{ij} \geq \sum_{m=1,2,3} y_{ij}^{(m)} \geq 0, \quad \forall i \in N,$$

$$y_{ij}^{(m)} \geq x_{ij}^{(m,d)}, \ \forall i \in N, d \in D^m, \ m = 1, 2, 3,$$

$$x_{ij}^{(m,d)} \geq 0, \ \forall i, j \in N, d \in D^m, \ m = 1, 2, 3,$$

$$\sum_{(i,j)\in\mathcal{A}} x_{ij}^{(m,d)} - \sum_{(j,i)\in\mathcal{A}} x_{ji}^{(m,d)} = \begin{cases} 1 + \mu_i^{(m,d)} - \lambda_i^{(m)}, & i = s^m, \\ \mu_i^{(m,d)} - \lambda_i^{(m)}, & \text{else}, \end{cases}$$

$$\forall i \in N \setminus \{d\}, d \in D^m, m = 1, 2,$$

$$\mu_i^{md} = 0, \forall i \notin D^3, d \in D^m, m = 1, 2, \tag{17}$$

where $\mathbb{N}$ is the set of non-negative integers, $\mathcal{A}$ is the set of arcs, and $\{x_{ij}^{(m,d)}\}$ constructs a virtual flow from s^m to d. $\{z_{ij}\}$ represents the number for unit throughput and constructs the actual flows in the network, also called a subgraph. As in Heindlmaier et al. (2011), at each node i, $\lambda_i^{(m)}$ denotes the amount of flow extracted from the original connection m and injected into session 3, and $\mu_i^{(m,d)}$ is the amount of traffic re-injected at node i into session m for destination d from session 3. For the virtual multicast, the following is added (Heindlmaier et al., 2011):

Algorithm 2 C1CPE

Input: G, s_1, dest_1, s_2, dest_2
Output: Topo

```
Obtain the Steiner tree with source s_i and destination
   set D_i, i=1, 2; ST_i=SCTF(s_i, D_i), i=1, 2
Topo←ST_1∪ST_2
Find D_i^p and D_i^s in Topo
while ~isempty(D_1^s∪D_2^s)
  for d∈D_i^s
    Find slave tree set {SlvT}_i and 1-cut path set {Ct}_i
  end for
  for d∈D_i^p
     Paint every node on a pure path with the
        corresponding color
  end for
  Paint other nodes with as many colors as possible
  for slv∈{SlvT}_i
    smcl←getDesiredColorNode()
    othcl←getOtherColorNode()
    smStra← arg min_path Cost(Path(smcl, slvt))
    otlStra← arg min_path Cost(Path(othcl, slvt))
  end for
  optStra← argmin(Cost([smStra_1, smStra_2, othStr_1,
     othStr_2]))
  Add optStra to Topo
  Find D_i^p and D_i^s in Topo
end while
return Topo
```

$$\sum_{(i,j)\in\mathcal{A}} x_{ij}^{(3,d)} - \sum_{(j,i)\in\mathcal{A}} x_{ji}^{(3,d)} = \sum_{m=1,2} \lambda_i^{(m)},$$
$$\forall i \in N \setminus \{d\}, d \in D^3, \quad (18)$$
$$\sum_{i\in D^3} \mu_i^{(m,d)} = \sum_{i\in\mathcal{N}} \lambda_i^{(m)}, \forall d \in D^m, m = 1,2.$$

As can be seen in Eqs. (17) and (18), the virtual multicast destinations are responsible for re-injecting the decoded information into the actual session. However, the optimal method to pre-specify the virtual destinations is not included in the optimization above, which makes the optimization not a complete outer bound for C1CPE and the simulation results for problem (17) is not included in this paper.

4 Simulation

We examine the behavior of two algorithms with simulations for many randomly generated graphs. In the following, we will introduce simulation parameters in the random graph, and then present the results of our simulation.

4.1 Random graph parameter

Our random graph is similar to that in Jiang et al. (2015) and Ramanathan (1996). Vertices are randomly placed in a square with a uniform distribution for the x and y coordinates. The costs for the unit data rate and double unit data rate between nodes i and j are the costs in Eqs. (7) and (8).

The cost of the whole resulting topology is the key parameter that we are concerned with, which will be simulated for both of the algorithms. The optimal solution for E-SCTF is simulated with the aid of Gurobi.

For simulation results of the E-SCTF algorithm, the power cost of the Steiner tree with unit throughput in the first simulation condition of each figure is adopted as the reference value for relative power cost. The vertical axis in Figs. 4–6 shows the multiple of the reference value.

For simulation results of the C1CPE algorithm, we adopt the power cost of a link with unit throughput and distance of 1 m as the reference value for relative power cost. The vertical axis in Figs. 7 and 8 shows the multiple of the reference value.

4.2 Simulation results for E-SCTF

The relative power cost is the sum of the cost of all the links with a normalized constant. The number of destinations and the number of total nodes affect the observations. Fig. 4 shows the relative power cost varying with the destination number when the total node number is 100. The relative power cost increases with the destination number, as the additional cost should be paid when more destinations are involved. The cost increases rapidly for the first few destinations and then flattens out as the ultra-densification reduces the cost increment for new destinations. E-SCTF saves about 1/3 of the power cost of NCBM, which is the extension of the D algorithm in a 2-cut-set situation.

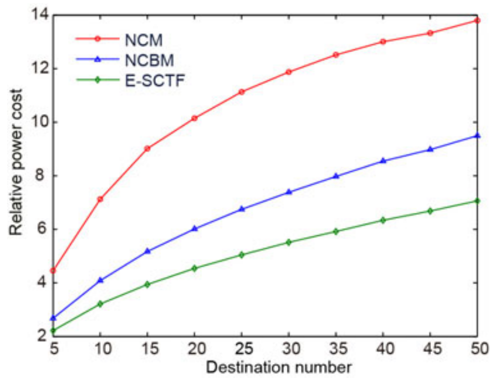

Fig. 4 Relative power cost for different destination numbers

Fig. 5 is the relative power cost against the node numbers when the destination number is fixed at 10. The relative power cost decreases with the node number as the increasing number of nodes improves the opportunity to obtain a lower cost multicast session. After the total number goes beyond 35, the gaps between the algorithms seem to be constant.

Compared with other algorithms, network coding based multicast (NCM) (Jiang et al., 2015) costs the most (Figs. 4 and 5). Its core idea is to add the coding structure based on a Steiner tree to maximize the energy efficiency, and the fixed throughput situation here degrades its performance. Besides, there is a main issue not considered in NCM: network coding can be valid only between different information paths. When performing network coding between two paths with common links, the cost of the common links will increase, and this is the main reason for the increase of cost for the NCM algorithm.

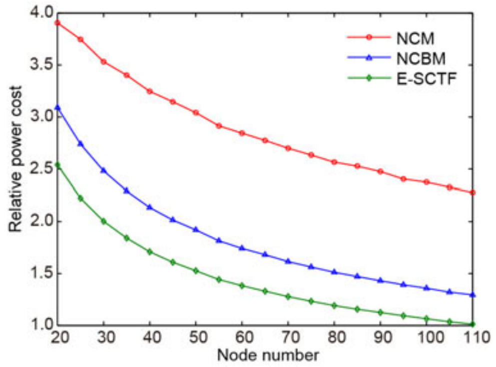

Fig. 5 Relative power cost for different node numbers with 10 destinations

Fig. 6 is the relative power cost when the destination number is fixed at 3. It involves additional simulation results for problem (16). As the optimal solution loosens the mutual exclusiveness of links (i, j) and (j, i), the cost of the optimal solution is further reduced. Considering the computational complexity of the optimal solution, we reduce the networks' scale and simulation time, but the results still show that the curves in Fig. 6 keep their tendency as in Fig. 5 and that the E-SCTF algorithm is the closest one to the optimal solution.

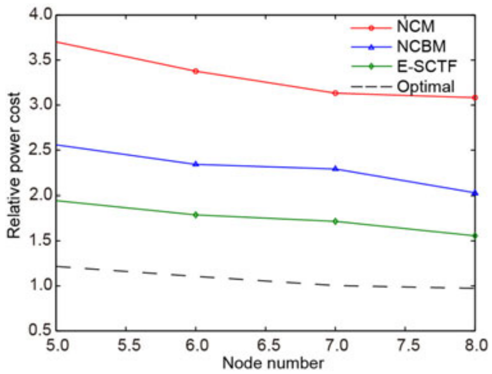

Fig. 6 Relative power cost for different node numbers

4.3 Simulation results for C1CPE

The total number in Fig. 7 is 100. The cost for two multicast trees increases because new destinations will cost more and the cost increment of doubling the overlapping links' capacity and the cost increment of C1CPE increase with the destination number of each session, as the increase in the destination number will raise the risk of the overlapping of links. The slopes of the strategies for doubling the overlapping links and two original trees are almost the same, while the cost increment of C1CPE is much slower. C1CPE's cost increment is under 1/5 of that of doubling the overlapping links' capacity.

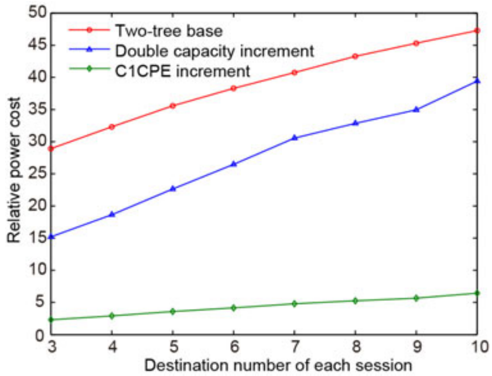

Fig. 7 Relative power cost for different destination numbers of each session

The power saving in C1CPE is also impressive (Fig. 8), which is the relative power cost against the total number of nodes when the destination of each session is six. The cost for two multicast trees decreases because new nodes in the graph may bring a new path of lower cost. The decline in the cost increment of doubling the overlapping links' capacity and the cost increment of C1CPE mean that increase in the node number reduces the chance of overlapping for two multicast sessions.

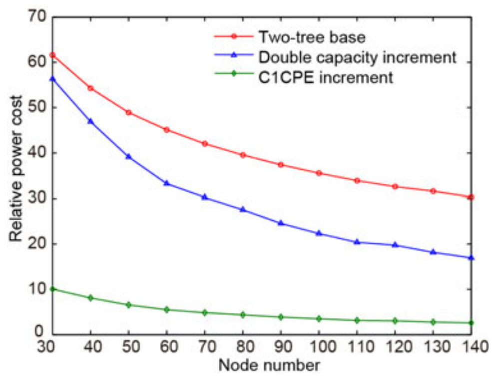

Fig. 8 Relative power cost for different node numbers

5 Conclusions

The routing configuration modification for a multicast session may be necessary in the coexistence of two multicasts. The throughput variation is a special coexistence case as the new throughput can be seen as another session. Our previous work is enriched in the E-SCTF algorithm and the algorithm searches for the lowest incremental cost solution changing from a 1-cut-set Steiner tree to a 2-cut-set coding subgraph when the multicast doubles its throughput. The algorithm shows the extensibility of the SCTF algorithm, which builds a Steiner tree for one multicast session. Then, we solve the generalized routing problem for the coexistence of two multicasts with the same throughput in the C1CPE algorithm. The algorithm searches for the lowest incremental cost solution changing from an un-decodable topology to a decodable topology. As suboptimal algorithms, both of them are suitable for large-scale networks with lower computational complexity.

The algorithms are opposite to the logically centralized controller in a 5G separation architecture with the minimum-cost target effectively reducing the waste of the resource. Given the boom in live broadcast and webcasts, the algorithms will be useful in 5G multicast networks.

References

Agyapong PK, Iwamura M, Staehle D, et al., 2014. Design considerations for a 5G network architecture. *IEEE Commun Mag*, 52(11):65-75. https://doi.org/10.1109/MCOM.2014.6957145

Akyildiz IF, Nie S, Lin SC, et al., 2016. 5G roadmap: 10 key enabling technologies. *Comput Netw*, 106:17-48. https://doi.org/10.1016/j.comnet.2016.06.010

Ali E, Ismail M, Nordin R, et al., 2017. Beamforming techniques for massive MIMO systems in 5G: overview, classification, and trends for future research. *Front Inform Technol Electron Eng*, 18(6):753-772. https://doi.org/10.1631/FITEE.1601817

Andrews JG, Buzzi S, Choi W, et al., 2014. What will 5G be? *IEEE J Sel Areas Commun*, 32(6):1065-1082. https://doi.org/10.1109/JSAC.2014.2328098

Baldemair R, Irnich T, Balachandran K, et al., 2015. Ultra-dense networks in millimeter-wave frequencies. *IEEE Commun Mag*, 53(1):202-208. https://doi.org/10.1109/MCOM.2015.7010535

Chen SZ, Kang SL, 2018. A tutorial on 5G and the progress in China. *Front Inform Technol Electron Eng*, 19(3):309-321. https://doi.org/10.1631/FITEE.1800070

Chen SZ, Qin F, Hu B, et al., 2016. User-centric ultra-dense networks for 5G: challenges, methodologies, and directions. *IEEE Wirel Commun*, 23(2):78-85. https://doi.org/10.1109/MWC.2016.7462488

Chiang M, Low SH, Calderbank AR, et al., 2007. Layering as optimization decomposition: a mathematical theory of network architectures. *Proc IEEE*, 95(1):255-312. https://doi.org/10.1109/JPROC.2006.887322

Choi J, 2015. Iterative methods for physical-layer multicast beamforming. *IEEE Trans Wirel Commun*, 14(9):5185-5196. https://doi.org/10.1109/TWC.2015.2434374

He SW, Huang YM, Wang HM, et al., 2014. Leakage-aware energy-efficient beamforming for heterogeneous multicell multiuser systems. *IEEE J Sel Areas Commun*, 32(6): 1268-1281. https://doi.org/10.1109/JSAC.2014.2328142

Heindlmaier M, Lun DS, Traskov D, et al., 2011. Wireless inter-session network coding—an approach using virtual multicasts. IEEE Int Conf on Communications, p.1-5. https://doi.org/10.1109/icc.2011.5963472

Ho T, Medard M, Koetter R, et al., 2006. A random linear network coding approach to multicast. *IEEE Trans Inform Theory*, 52(10):4413-4430. https://doi.org/10.1109/TIT.2006.881746

Jiang DD, Xu ZZ, Li WP, et al., 2015. Network coding-based energy-efficient multicast routing algorithm for multi-hop wireless networks. *J Syst Softw*, 104:152-165. https://doi.org/10.1016/j.jss.2015.03.006

Kotronis V, Dimitropoulos X, Ager B, 2012. Outsourcing the routing control logic: better Internet routing based on SDN principles. Proc 11th ACM Workshop on Hot Topics in Networks, p.55-60. https://doi.org/10.1145/2390231.2390241

Kulkarni MN, Ghosh A, Andrews JG, 2016. A comparison of MIMO techniques in downlink millimeter wave cellular networks with hybrid beamforming. *IEEE Trans Commun*, 64(5):1952-1967. https://doi.org/10.1109/TCOMM.2016.2542825

Li JZ, Ai B, He RS, et al., 2017. Indoor massive multiple-input multiple-output channel characterization and performance evaluation. *Front Inform Technol Electron Eng*, 18(6):773-787. https://doi.org/10.1631/FITEE.1700021

Li SY, Sun W, Hua CC, 2016. Optimal resource allocation for heterogeneous traffic in multipath networks. *Int J Commun Syst*, 29(1):84-98. https://doi.org/10.1002/dac.2800

Lun DS, Ratnakar N, Koetter R, et al., 2005. Achieving minimum-cost multicast: a decentralized approach based on network coding. Proc IEEE 24th Annual Joint Conf of IEEE Computer and Communications Societies, p.1607-1617. https://doi.org/10.1109/INFCOM.2005.1498443

Ma Z, Zhang ZQ, Ding ZG, et al., 2015. Key techniques for 5G wireless communications: network architecture, physical layer, and MAC layer perspectives. *Sci China Inform Sci*, 58(4):1-20. https://doi.org/10.1007/s11432-015-5293-y

Rajawat K, Gatsis N, Giannakis GB, 2011. Cross-layer designs in coded wireless fading networks with multicast. *IEEE/ACM Trans Netw*, 19(5):1276-1289. https://doi.org/10.1109/TNET.2011.2109010

Ramanathan S, 1996. Multicast tree generation in networks with asymmetric links. *IEEE/ACM Trans Netw*, 4(4):558-568. https://doi.org/10.1109/90.532865

Rappaport TS, Murdock JN, Gutierrez F, 2011. State of the art in 60-GHz integrated circuits and systems for wireless communications. *Proc IEEE*, 99(8):1390-1436. https://doi.org/10.1109/JPROC.2011.2143650

Rappaport TS, Sun S, Mayzus R, et al., 2013. Millimeter wave mobile communications for 5G cellular: it will work! *IEEE Access*, 1:335-349. https://doi.org/10.1109/ACCESS.2013.2260813

Ribeiro A, Giannakis G B, 2010. Separation principles in wireless networking. *IEEE Trans Inform Theory*, 56(9): 4488-4505. https://doi.org/10.1109/TIT.2010.2053897

Riemensberger M, Utschick W, 2014. A polymatroid flow model for network coded multicast in wireless networks. *IEEE Trans Inform Theory*, 60(1):443-460. https://doi.org/10.1109/TIT.2013.2287498

Rusek F, Persson D, Lau BK, et al., 2013. Scaling up MIMO: opportunities and challenges with very large arrays. *IEEE Signal Process Mag*, 30(1):40-60. https://doi.org/10.1109/MSP.2011.2178495

Shokri-Ghadikolaei H, Fischione C, 2016. The transitional behavior of interference in millimeter wave networks and its impact on medium access control. *IEEE Trans Commun*, 64(2):723-740. https://doi.org/10.1109/TCOMM.2015.2509073

Stai E, Loulakis M, Papavassiliou S, 2015. Cross-layer design of wireless multihop networks over stochastic channels with time-varying statistics. *IEEE Trans Wirel Commun*, 14(12):6967-6980. https://doi.org/10.1109/TWC.2015.2462845

Su LY, Yang CY, Chih-Lin I, 2016. Energy and spectral efficient frequency reuse of ultra dense networks. *IEEE Trans Wirel Commun*, 15(8):5384-5398. https://doi.org/10.1109/TWC.2016.2557790

Szabó D, Németh F, Sonkoly B, et al., 2015. Towards the 5G revolution: a software defined network architecture exploiting network coding as a service. Proc ACM Conf on Special Interest Group on Data Communication, p.105-106.

Wang C, Guo ST, Yang YY, et al., 2016. An optimization framework for mobile data collection in energy-harvesting wireless sensor networks. *IEEE Trans Mobile Comput*, 15(12):2969-2986. https://doi.org/10.1109/TMC.2016.2533390

Wang CC, Shroff NB, 2010. Pairwise intersession network coding on directed networks. *IEEE Trans Inform Theory*, 56(8):3879-3900. https://doi.org/10.1109/TIT.2010.2050932

Wang J, Li Y, Wang XM, 2007. Network coding based multicast in Internet. Int Conf on Parallel Processing Workshops, p.44. https://doi.org/10.1109/ICPPW.2007.58

Wei F, Zou WX, 2017. Steiner-tree-based 2-cut-set network coding subgraph algorithm in wireless multicast network. Proc Int Conf in Communications, Signal Processing, and Systems, p.373-381. https://doi.org/10.1007/978-981-10-3229-5_40