

Evolutionary Approaches To Minimizing Network Coding Resources

Minkyu Kim, Muriel Médard, Varun Aggarwal, Una-May O'Reilly, Wonsik Kim, Chang Wook Ahn, and Michelle Effros

Abstract—We wish to minimize the resources used for network coding while achieving the desired throughput in a multicast scenario. We employ evolutionary approaches, based on a genetic algorithm, that avoid the computational complexity that makes the problem NP-hard. Our experiments show great improvements over the sub-optimal solutions of prior methods. Our new algorithms improve over our previously proposed algorithm in three ways. First, whereas the previous algorithm can be applied only to acyclic networks, our new method works also with networks with cycles. Second, we enrich the set of components used in the genetic algorithm, which improves the performance. Third, we develop a novel distributed framework. Combining distributed random network coding with our distributed optimization yields a network coding protocol where the resources used for coding are optimized in the setup phase by running our evolutionary algorithm at each node of the network. We demonstrate the effectiveness of our approach by carrying out simulations on a number of different sets of network topologies.

I. INTRODUCTION

It is now well known that network throughput can be significantly increased by employing the novel technique of network coding, where the intermediate nodes are allowed to combine data received from different links [1], [2]. While most network coding solutions employ coding at all possible nodes, it is often possible to achieve the network coding advantage by coding only at a subset of nodes.

Example 1: In the canonical example of network B (Fig. 1(a)) [1], only node z needs to combine its two inputs while all other nodes perform routing only. If we suppose that link (z, w) in network B has capacity 2, which we represent by two parallel unit-capacity links in network B' (Fig. 1(b)), a multicast of rate 2 is possible without network coding. In network C (Fig. 1(c)), where node s wishes to transmit data at rate 2 to the 3 leaf nodes, network coding is required at either node a or node b , but not both. \square

Example 1 leads us to the following question: At which nodes does network coding need to occur to achieve the multicast capacity? If network coding is handled at the application layer, we can minimize the cost of network coding

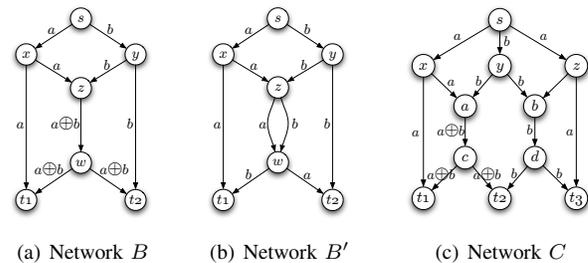


Fig. 1. Sample Networks for Example 1

by identifying the nodes where access up to the application layer is not necessary. If network coding is integrated in the buffer management of a router, it is important to understand where and how many such special routers must be deployed to satisfy the communication demands.

Determining a minimal set of nodes where coding is required is difficult. The problem of deciding whether a given multicast rate is achievable without coding, i.e., whether the minimum number of required coding nodes is zero or not, reduces to a multiple Steiner subgraph problem, which is NP-hard [3]. Hence, the optimization problem to find the minimal number of required coding nodes is NP-hard. Even approximating the minimal number of coding nodes within any multiplicative factor or within an additive factor of $|V|^{1-\epsilon}$ is NP-hard [4].

In [5], we introduce an evolutionary approach for finding a practical multicast protocol that provides the full benefit of network coding with a reduced number of coding nodes. The proposed approach uses a Genetic Algorithm (GA) that operates on a set of candidate solutions which it improves sequentially via mechanisms inspired by biological evolution (e.g., recombination/mutation of genes and survival of the fittest). The algorithm proposed in [5] reduces the number of coding links/nodes relative to prior approaches and applies to a variety of generalized scenarios.

The new algorithm proposed here improves on the one in [5] in three key ways. First, whereas the algorithm in [5] can be applied to only acyclic networks, we devise a modified method that works also with networks with cycles. Second, we introduce a new set of GA components that in our experiments significantly outperforms the one used in [5]. Third, we develop a novel framework where most time consuming computations of the evolutionary algorithm are distributed over the network. This new framework, combined

M. Kim, M. Médard, and W. Kim are with the Laboratory of Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139, USA ({minkyu, medard, wskim14}@mit.edu).

V. Aggarwal and U.-M. O'Reilly are with the Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA ({varun_ag@, unamay@csail.}mit.edu).

C. W. Ahn is with the Department of Information and Communications, Gwangju Institute of Science and Technology, Gwangju 500-712, Korea (cwan@evolution.re.kr).

M. Effros is with the Data Compression Laboratory, California Institute of Technology, Pasadena, CA 91125, USA (effros@caltech.edu).

with the distributed random network coding scheme of [6], can make a distributed network coding protocol where the resources used for coding are optimized in the setup phase as our proposed algorithm running at each node of the network.

The rest of the paper is organized as follows. Section II presents the problem formulation and summarizes related work. Section III describes the improvements of the algorithm using a centralized framework. Section IV extends this approach to a distributed framework. Section V presents experimental results. Section VI concludes with topics for future research.

II. PROBLEM FORMULATION AND RELATED WORK

A. Problem Formulation

We assume that the network is given by a directed multigraph $G = (V, E)$, where each link has a unit capacity. Connections with larger capacities are represented by multiple links. Only integer flows are allowed, hence there is either no flow or a unit rate of flow on each link. We consider the single multicast scenario in which a single source $s \in V$ wishes to transmit data at rate R to a set $T \subset V$ of sink nodes, where $|T| = d$. Rate R is said to be achievable if there exists a transmission scheme that enables all d sinks to receive all of the information sent. We consider only linear coding, where a node's output on an outgoing link is a linear combination of the inputs from its incoming links. Linear coding is sufficient for multicast [2].

Given the target rate R , which we assume is achievable if coding is allowed at all nodes, we wish to determine a minimal set of nodes where coding is required in order to achieve this rate. Coding is necessary at a node $v \in V$ if coding is necessary on at least one of node v 's outgoing links. As pointed out also in [4], the number of coding links is a more accurate estimator of the amount of computation incurred by coding. We assume hereafter that our objective is to minimize the number of coding links rather than nodes. Note, however, that as demonstrated in [5], it is straightforward to generalize the proposed algorithm to the case of minimizing the number of coding nodes. Furthermore, [5] shows that, with appropriate changes, the algorithm can be readily applied to more generalized optimization scenarios, e.g., where different links/nodes have different costs for coding.

It is clear that no coding is required at a node with only a single input since it has nothing to combine with. We refer to a node with multiple incoming links as a *merging node*. If the linearly coded output on a particular outgoing link of a particular merging node weights all but one incoming message by zero, then no coding occurs on that link. (Even if the only nonzero coefficient is not identity, there is another coding scheme that replaces the coefficient by identity [4].) Thus, to determine whether coding is necessary on an outgoing link of a merging node, we need to verify whether we can constrain the output on the link to depend on a single input without destroying the achievability of the given rate.

Consider a merging node with $k (\geq 2)$ incoming links and $l (\geq 1)$ outgoing links. For each $i \in \{1, \dots, k\}$ and each $j \in$

$\{1, \dots, l\}$, we set $a_{ij} = 1$ if the input from incoming link i contributes to the linearly coded output on outgoing link j , and $a_{ij} = 0$ otherwise; we call these the *active* and *inactive* states, respectively. Network coding is required over link j only if two or more link states are active. Thus, it is useful to think of $a_j = (a_{ij})_{i \in \{1, \dots, k\}}$ as a *block* of length k (see Fig. 2 for an example).

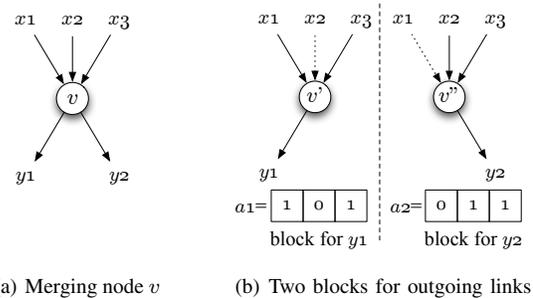


Fig. 2. Node v with 3 incoming and 2 outgoing links has inputs described by vectors $a_1 = (a_{11}, a_{21}, a_{31})$ and $a_2 = (a_{12}, a_{22}, a_{32})$.

As in network C of Example 1, whether node v must code over link y_j varies depending on which other nodes are coding. Thus deciding which nodes should code in general involves a selection out of exponentially many possible choices. We employ a GA-based search method to efficiently address the large and exponentially scaling size of the space.

B. Related Work

Fragouli et al. [7] show that coding is required at no more than $(d-1)$ nodes in acyclic networks with 2 unit-rate sources and d sinks. This result, however, is not easily generalized to more than 2 sources. They also present an algorithm to construct a minimal subtree graph. For target rate R , they first select a subgraph consisting of R link-disjoint paths to each of d sinks and then construct the corresponding labeled line graph in which they sequentially remove the links whose removal does not affect the achievable rate.

Langberg et al. [4] derive an upper bound on the number of required coding nodes for both acyclic and cyclic networks. They give an algorithm to construct a network code that achieves the bounds, where the network is first transformed such that each node has degree at most 3 and each of the links is sequentially examined and removed if the target rate is still achievable without it.

Both of the above approaches remove links sequentially in a greedy fashion, assuming that network coding is done at all nodes with multiple incoming links in the remaining graph. Note that, unless a good order of the link traversal is found, the quality of the solution cannot be much improved as illustrated in [5].

Bhattach et al. [8] give linear programming formulations for the problems of optimizing over various resources used for network coding, based on a model allowing continuous flows. Their optimal formulations, however, involve a number of variables and constraints that grows exponentially with the

number of sinks, which makes it hard to apply the formulations to the case of a large number of sinks, even at the price of sacrificed optimality.

We conclude this section with a brief introduction to GA.

C. A Brief Introduction to GA

GAs [9] operate on a set of candidate solutions, called a *population*. Each solution is typically represented by a bit string, called a *chromosome*. Each chromosome is assigned a *fitness value* that measures how well the chromosome solves the problem at hand, compared with other chromosomes in the population. From the current population, a new population is generated typically using three genetic operators: *selection*, *crossover* and *mutation*. Chromosomes for the new population are selected randomly (with replacement) in such a way that fitter chromosomes are selected with higher probability. For crossover, survived chromosomes are randomly paired, and then two chromosomes in each pair exchange a subset of their bit strings to create two offspring. Chromosomes are then subject to mutation, which refers to random flips of the bits applied individually to each of the new chromosomes. The process of evaluation, selection, crossover and mutation forms one *generation* in the execution of a GA. The above process is iterated with the newly generated population successively replacing the current one. The GA terminates when a certain stopping criterion is reached, e.g., after a predefined number of generations. GAs have been applied to a large number of scientific and engineering problems, including many combinatorial optimization problems in networks (e.g., [10], [11]).

There are several aspects of our problem suggesting that a GA-based method may be a promising candidate: GA has proven to work well if the space to be searched is large, but known not to be perfectly smooth or unimodal, or even if the space is not well understood [9] (which makes traditional optimization methods difficult to apply). Note that the search space of our problem is apparently not smooth or unimodal with respect to the number of coding links and the structure of the space consisting of the feasible binary vectors is not well understood. Since the problem is NP-hard, it is not critical that the calculated solution may not be a global optimum. Note also that, while it is hard to characterize the structure of the search space, once provided with a solution we can verify its feasibility and count the number of coding links therein in polynomial time. Thus, if the use of genetic operations can suitably limit the size of the space to be actually searched, a solution can be obtained fairly efficiently.

III. CENTRALIZED APPROACH

We first present the centralized version of the algorithm, whose overall structure, based on simple GA [9], is shown in Fig. 3. Sections III-A and III-B present two different methods for mapping the network coding problem to a GA framework (procedure [C1] in Fig. 3) and evaluating the chromosomes ([C3, C8] in Fig. 3). Either of the two methods can be combined with the computational part of the algorithm

(remaining procedures in Fig. 3) which is described in Section III-C.

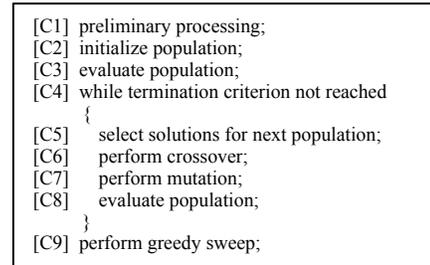


Fig. 3. Flow of Centralized Algorithm

A. Algebraic Method

We first describe the algebraic method by which a choice of coding links is mapped to a GA problem and a given candidate solution (chromosome) is evaluated. Owing to space limitations, here we present only the main concepts; the reader is referred to [5] for details. This algebraic method will also be used later in the distributed version of the algorithm. This method applies only to acyclic networks; for cyclic networks, it can be very inefficient as discussed in Section III-B.

Given an acyclic graph $G = (V, E)$, we first construct the corresponding labeled line graph $G' = (V', E')$ [12], where each node in V' represents a link in E and each link $(v', w') \in E'$ implies that the links $e, f \in E$ corresponding to nodes $v', w' \in V'$, respectively, are connected in G via some node $u \in V$ such that $u = \text{head}(e) = \text{tail}(f)$. To construct a network code, we assign a coefficient to each link in G' as in [12]. Note that there is a one-to-one correspondence between the binary variables a_{ij} introduced in Section II-A and the coefficients assigned to the incoming links to a node with *multiple* incoming links in G' . Thus, for each binary variable a_{ij} we can consider the associated coefficient. If there are m such coefficients in G' , a chromosome is represented by a vector consisting of m binary variables; if we denote by d_{in}^v and d_{out}^v the in-degree and the out-degree of node $v \in V$, m is given by $m = \sum_{v \in \mathcal{V}} d_{in}^v d_{out}^v$, where $\mathcal{V} \subset V$ is the set of all merging nodes in G .

To evaluate a given chromosome, we first verify its feasibility. If $a_{ij} = 0$ in the chromosome, then input x_i is inactive with respect to output y_j and we set associated coefficient to zero. If $a_{ij} = 1$, then we let the associated coefficient be an indeterminate nonzero value. To determine whether the target rate R is achievable, we rely on random linear coding; i.e., to each of the remaining coefficients we assign a random element from a finite field and check whether the system matrix is nonsingular. Note that this feasibility test entails a bounded error, which is shown in [5] not to be critical since the error is one-sided, i.e., a feasible chromosome may mistakenly be declared infeasible but not vice versa, and we can lower the bound on the error probability as much as we desire at an additional cost of computation.

We then define the fitness value F of chromosome \underline{z} as

$$F(\underline{z}) = \begin{cases} \text{number of coding links,} & \text{if } \underline{z} \text{ is feasible,} \\ \infty, & \text{if } \underline{z} \text{ is infeasible,} \end{cases} \quad (1)$$

where the number of coding links can be easily calculated by counting the number of blocks in the chromosome with at least two 1's. It is not hard to verify that the computational complexity required to evaluate a single chromosome is $O(d \cdot (|E|^{2.376} + R^3))$.

B. Graph Decomposition Method

Note that the above algebraic method deals explicitly with the scalar coefficients that appear in the system matrix assuming that the network operates with zero delay (and thus the network is cycle-free). In the presence of cycles, delay must be taken into account, hence the system matrix becomes a matrix over the polynomial ring with coefficients that are *rational functions* in the delay variable D [12]. In this case, the matrix computation involves calculating the coefficient for each power of the delay variable D , which in general renders the feasibility test prohibitively inefficient.

In this subsection, we show that, with an appropriate graph decomposition, the evaluation can be done by calculating the max-flows between the source and the sinks. Note that the minimum of those max-flows equals the maximum achievable multicast rate regardless of the existence of cycles in the network [1]. Unlike the algebraic one, this modified method operates on the actual graph G rather than the labeled line graph.

In the first stage of the algorithm ([C1] in Fig. 3), we decompose each merging node that is not a sink as follows. (For a sink node with nonzero out-degree, introduce a virtual sink connected via R links and decompose the original sink.) Consider a merging node v with $d_{in} (\geq 2)$ incoming links and d_{out} outgoing links (see Fig. 4). We introduce d_{in} nodes $u_1, \dots, u_{d_{in}}$, which we call *incoming auxiliary nodes*, and redirect the i -th ($1 \leq i \leq d_{in}$) incoming link of node v to node u_i . Similarly, we create d_{out} nodes $w_1, \dots, w_{d_{out}}$, which we call *outgoing auxiliary nodes*, and let the j -th ($1 \leq j \leq d_{out}$) outgoing link of node v be the only outgoing link of node w_j . We then insert a link (u_i, w_j) between each pair of nodes u_i and w_j ($1 \leq i \leq d_{in}, 1 \leq j \leq d_{out}$).

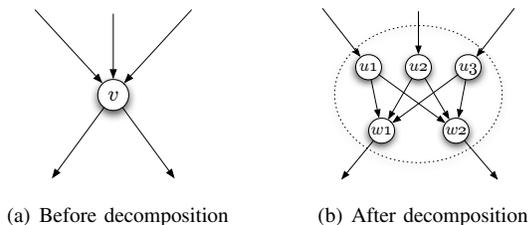


Fig. 4. Decomposition of a node with $d_{in} = 3$ and $d_{out} = 2$.

With each of the newly introduced links between the auxiliary nodes, we associate a binary variable in a chromosome. Since each of those new links corresponds to a pair of

connected incoming and outgoing links in the original graph, we can see a one-to-one correspondence between the binary variables here and those introduced in the algebraic method.

To verify the feasibility of a given chromosome \underline{z} , we first delete all the links associated with 0 in \underline{z} and then compute the max-flow between the source and each of the d sinks. If an outgoing auxiliary node has only one incoming link, we can replace the incoming link, together with its outgoing link, by a single link without changing any of the max-flows. If an auxiliary outgoing node has no incoming link, we simply delete the node. If all d max-flows are at least R , rate R is achievable with network coding only at the outgoing auxiliary nodes with two or more incoming links. The number of remaining outgoing auxiliary nodes equals the number of coding links in the original graph. Hence, by counting the number of such outgoing auxiliary nodes, we can calculate the fitness value F of chromosome \underline{z} , defined as in (1).

The max-flow based evaluation of a single chromosome requires $O(d \cdot |E'|^2 \sqrt{|V'|})$ time [13], where $|E'|$ and $|V'|$ are the number of links and nodes, respectively, in the decomposed graph.

Note that, unlike the algebraic method, this feasibility test incurs no error. Since it works both with and without cycles, the graph decomposition method may be preferable to the algebraic method when centralized operation is feasible. Central operation requires that the topology of the whole network be known to a central computing node and may be slow. This approach may be appropriate, for example, in the planning stage of a network. However, the algebraic method plays a crucial role in the distributed version of the algorithm, as shown in Section IV.

C. Computational Part of GA

In describing the computational part of the GA, we first discuss two novel elements designed specifically for the problem and then describe other typical GA components. The discussion applies for both the algebraic and graph decomposition methods.

1) *Block-Wise Representation and Operators*: A block of length k is defined to be a subset of a chromosome consisting of k binary variables that indicate the link states for the transmission onto a particular outgoing link fed by k incoming links, i.e., the k components of vector $a_j = (a_{ij})_{i \in \{1, \dots, k\}}$ introduced in Section II-A. We may allow for a length- k block to take all possible 2^k strings as in [5], which we refer to as *bit-wise representation*¹. Note, however, that once a block has at least two 1's, replacing all the remaining 0's with 1 has no effect on whether coding is done and that substituting 0 with 1, as opposed to substituting 1 with 0, does not hurt the feasibility. Therefore, for a feasible chromosome, any block with two or more 1's can be treated the same as the block with all 1's.

¹In the GA community, the method for representing a candidate solution as a bit string is called *genotype encoding*; we avoid the use of this term to minimize confusion with the term encoding in the context of network coding.

The above observation leads to *block-wise representation*, where each block of length k is allowed to take one of the following $(k+2)$ strings: [000...0] (1 string for *no* transmission state), [100...0], [010...0], [001...0], ..., [000...1] (k strings for *uncoded* transmission state), [111...1] (1 string for *coded* transmission state). If we let w be the total number of blocks (i.e., $w = \sum_{v \in \mathcal{V}} d_{out}^v$) and k_i denote the length of the i -th block ($i = 1, \dots, w$), the search space size is reduced to $\prod_{i=1}^w (k_i + 2)$, from $2^{\sum_{i=1}^w k_i}$ in the case of bit-wise representation.

To preserve the structure of block-wise representation, we need a set of new genetic operators. In [5], uniform crossover [9] and binary mutation [9] are used, which for comparison we refer to as *bit-wise* operators². Let us now define *block-wise* operators as follows. For block-wise uniform crossover, we let each pair of chromosomes subject to crossover exchange each full block, rather than each individual bit, independently. For block-wise mutation, we let each block of length k subject to mutation take another string chosen uniformly at random out of the other $(k + 1)$ allowed strings.

It is interesting to note that the benefit of the smaller search space size in fact comes at the price of losing the information on the blocks with partially active link states that may serve as intermediate steps toward an uncoded transmission state. Also, whereas the *average* number of bits flipped by block-wise mutation of a length- k block using mutation rate α is $\frac{4k^2}{(k+1)(k+2)}\alpha$, which is smaller than that by the bit-wise mutation ($k\alpha$), the probability that 2 or more bits are flipped is often much larger for block-wise mutation; this may negatively affect the GA's ability to improve the solution through fine random changes. Hence, the overall effect of block-wise representation and operators on the algorithm's performance is not easy to predict theoretically. Section V includes an experimental evaluation of this question.

2) *Greedy Sweep*: We introduce another novel operator, referred to as *greedy sweep*, where we inspect the best chromosome obtained at the end of the iteration and switch each of the remaining 1's to 0 if it can be done without violating feasibility ([C8] in Fig. 3). This procedure can only improve the solution, and sometimes the improvement is substantial. Moreover, if we denote by \underline{z} the chromosome after the greedy sweep, then \underline{z} gives an upper bound on the number of coding links the same as in [4].

Lemma 1: The number of coding links associated with \underline{z} is upper bounded by $R^3 d^2$ for an acyclic network and $(2B + 1)R^3 d^2$ for a cyclic network, where B is the minimum number of links that must be removed from the network in order to eliminate cycles.

Proof: Let us consider the graph decomposition method. Here, switching 1 to 0 in a chromosome implies that we delete the associated link in the decomposed graph. In the decomposed graph associated with \underline{z} , there is no link between auxiliary

²For uniform crossover, a pair of chromosomes exchanges each bit independently with a given probability, and for binary mutation, each bit of a chromosome is flipped independently with a given probability.

links that can be removed without violating the achievability. One can easily verify that there is a one-to-one correspondence between the links between auxiliary nodes in our decomposed graph (see Fig. 4) and the set of the *paths within* the gadget Γ_v (see Fig. 2 in [4]). Now we can replace all non-merging nodes with a degree larger than 3 by the gadgets and greedily remove links in the same way as above, which, however, is irrelevant of the number of coding links. Therefore, from \underline{z} we can construct a simple instance, as defined in [4], and it gives the desired upper bounds on the number of coding links (Lemma 14 in [4]). \square

Lemma 1 provides our algorithm with a guarantee on its performance which is at least no worse than that of the algorithm in [4].

3) *Typical GA Components*: When initializing the population ([C2] in Fig. 3), we randomly generate each block and insert an all-one vector, whose effect as a feasible starting point is crucial as discussed in [5]. The iteration is terminated if the generation number reaches the predefined limit ([C4] in Fig. 3). For selection ([C5] in Fig. 3), we employ *tournament selection* [9], where we repeat a tournament between a predefined number of randomly selected chromosomes out of which the best one is selected (with replacement) for the next generation.

IV. DISTRIBUTED APPROACH

Noting that the main advantage of network coding based multicast is that an efficient capacity-achieving code can be constructed in a distributed fashion [6], [14], a motivation for decentralization of the algorithm becomes apparent. That is, given that the actual multicast can proceed in a decentralized manner, an algorithm used for resource optimization should be more desirable if it does not require centralized computation.

Moreover, as will be discussed below, such decentralization enables the most time consuming task, fitness evaluation, to be distributed over the network such that the computational complexity required at each node depends only on local parameters. The size of the population often serves as an important factor for the ability of a GA to find a good solution [15]. Though it is not an easy task to predict the accurate population size required for a specific problem, it is always desirable to devise an evaluation method with a low complexity, which allows for a flexibility in adopting a large-sized population when needed.

In this section, we present a novel distributed framework for our evolutionary algorithm, in which the feasibility test is done locally at each sink while the intermediate nodes actually construct random linear codes. With a limited amount of feedback information from the sinks and the merging nodes, fitness evaluation can be done with a substantially lower complexity. Furthermore, the population can be managed in a distributed manner such that each merging node locally manages a subset of the population that determines the local operations at that node (see Fig. 5). Also, it will be shown that, with some amount of coordination, all genetic operations can be done locally at individual merging nodes.

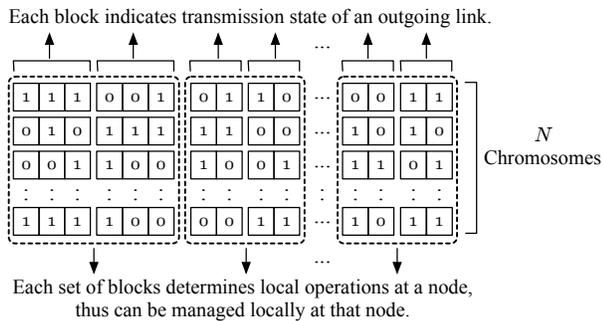


Fig. 5. Structure of Population

In addition to computational efficiency, this distributed approach has an important benefit that the coding resource optimization can be done on the fly while a network is operational, allowing for the following network coding protocol: As the source node sends an “optimize” signal, all the nodes participating in the multicast go into the optimization mode, and as the distributed evolutionary algorithm proceeds, the links/nodes where coding is not required are identified. When the source node sends a “transmit” signal, the network starts to multicast data based on the best network code found, in which coding is done only at the required links/nodes.

Since the distributed version of the algorithm is based on the algebraic method described in Section III-A, we need to take different approaches depending on whether the network has cycles or not. In the following, we begin to describe the details of the distributed approach assuming that the network is acyclic, and later in this section we extend the approach to cyclic networks, highlighting the changes to be made.

A. Assumptions and Preliminaries

We assume that each link can transmit one packet of a fixed size per time unit in the given direction. Each link is also assumed to be able to send some amount of feedback data, typically much smaller than the packet size, in the *reverse* direction. Also, we assume that each interior node operates in a burst-oriented mode; i.e., for the forward (backward) evaluation phase, each node starts updating its output only after an updated input has been received from all incoming (outgoing) links.

The overall structure of our distributed algorithm is shown in Fig. 6 with the locations of each procedure specified. We now proceed to describe the detailed procedures of the algorithm in the order of their occurrences.

B. Details of Algorithm

1) **Preliminary Processing [D1]**: The source initiates the algorithm by transmitting the “optimize” signal containing the following predetermined parameters: target multicast rate R , population size N , the size q of the finite field to be used, crossover probability, and mutation rate. Each participating node that has received the signal passes the signal to its downstream nodes.

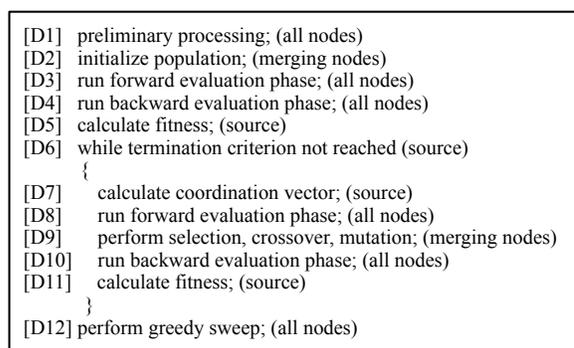


Fig. 6. Flow of Distributed Algorithm

2) **Population Initialization [D2]**: Let us consider a merging node with $d_{in} (\geq 2)$ incoming links. For *each* of its d_{out} outgoing links, the node has to manage a binary vector of length d_{in} , which we refer to as a *coding vector*, to indicate the link states for a single chromosome. Hence, for the population of size N , the node must have $N \cdot d_{out}$ coding vectors to determine the operations at that node. To initialize this subset of the population, the merging node randomly generates $N \cdot d_{in} \cdot d_{out}$ binary numbers and set all the components to 1 for the coding vectors that correspond to the first of the N chromosomes.

3) **Forward Evaluation Phase [D3, D8]**: For the feasibility test of a chromosome, each node transmits a vector consisting of R components, which we refer to as a *pilot vector*. Each of its components is from the finite field \mathbb{F}_q and the i -th component represents the coefficient used to encode the i -th source data. We assume that a set of N pilot vectors is transmitted together by a single packet.

The source initiates the forward evaluation phase by sending out on each of its outgoing links a set of N random pilot vectors. Each non-merging node simply forwards all the pilot vectors received from its incoming link to all its outgoing links.

Each merging node transmits on each of its outgoing links a random linear combination of the received pilot vectors, computed based on the node’s coding vectors as follows. Let us consider a particular outgoing link and denote the associated d_{in} coding vectors by $v_1, v_2, \dots, v_{d_{in}}$. For the i -th ($1 \leq i \leq N$) output pilot vector u_i , we denote the i -th input pilot vectors received from the incoming links by $w_1, w_2, \dots, w_{d_{in}}$. Define the set J of indices as

$$J = \{1 \leq j \leq d_{in} \mid \text{the } i\text{-th component of } v_j \text{ is } 1\}. \quad (2)$$

Then,

$$u_i = \sum_{j \in J} w_j \cdot \text{rand}(\mathbb{F}_q), \quad (3)$$

where $\text{rand}(\mathbb{F}_q)$ denotes a random element from \mathbb{F}_q . If the set J is empty, u_i is assumed to be zero.

4) **Backward Evaluation Phase [D4, D10]**: To calculate a chromosome’s fitness value, two kinds of information need to be gathered: 1) whether each sink can decode data of rate R

and 2) how many links are used for coding at each merging node.

Each sink can determine whether data of rate R is decodable for each of the N chromosomes by computing the rank of the collection of received pilot vectors. It is worth to point out that this is the same algebraic evaluation method described in Section III-A, with the difference that, rather than computing the system matrix with randomized elements centrally, we now actually construct random linear codes over the network in a decentralized fashion. Hence, this feasibility test also bears the same, but uncritical, possibility of errors as in the centralized case. Regarding the number of coding links, each merging node can simply count the number links where coding is required by inspecting its coding vectors used in the forward evaluation phase.

For the feedback of this information, each node transmits a vector consisting of N components, which is referred to as a *fitness vector*. Each of the components must be at least $\lceil \log(|E| + 2) \rceil$ bits long since for each chromosome the number of coding links can range from zero to $|E|$ and an additional symbol (infinity) is needed to signify infeasibility. The backward evaluation phase proceeds as follows:

- After the feasibility tests of the N chromosomes are done, each sink generates a fitness vector whose i -th ($1 \leq i \leq N$) component is zero if the i -th chromosome is feasible at the sink, and infinity otherwise. Each sink then initiates the backward evaluation phase by transmitting its fitness vector to all of its parents.
- Each interior node calculates its own fitness vector whose i -th ($1 \leq i \leq N$) component is the number of coding links at the node for the i -th chromosome plus the sum of all the i -th components of the received fitness vectors. Each node then transmits the calculated fitness vector to *only one* of its parents, and an all-zero fitness vector (for just signaling) to the other parent nodes.

Note that, since the network is assumed to be acyclic, each coding link of a chromosome contributes exactly once to the corresponding component of the source node's fitness vector, and thus the above update procedure provides the source with the correct total number of coding links.

5) **Fitness Calculation [D5, D11]**: The source calculates the fitness values of N chromosomes simply by performing component-wise summation of the received fitness vectors. Note that if an infinity were generated by *any* of the sinks, it should dominate the summations all the way up to the source, and thus the source can calculate the correct fitness value for the infeasible chromosome.

6) **Termination Criterion [D6]**: The source can determine when to terminate the optimization by counting the number of generations iterated thus far.

7) **Coordination Vector Calculation [D7]**: Since the population is divided into subsets that are managed at the merging nodes, genetic operations also need to be done locally at the merging nodes. However, some amount of coordination is required for consistent genetic operations throughout all the merging nodes; more specifically, for 1) consistent selection

of chromosomes, 2) consistent pairing of chromosomes for crossover, and 3) consistent decision on whether each pair is subject to crossover. This information is carried by a *coordination vector*, calculated at the source, consisting of the indices of selected chromosomes that are randomly paired and 1-bit data for each pair indicating whether the pair needs to be crossed over. The coordination vector is *transmitted together with* the pilot vectors in the next forward evaluation phase.

8) **Genetic Operations [D8]**: Based on the received coordination vector, each merging node can locally perform genetic operations and renew its portion of the population as follows:

- For selection, each node only retains the coding vectors that correspond to the indices of selected chromosomes.
- For block-wise crossover, each node independently determines whether each block is crossed over. Since no block is shared by multiple merging nodes, this can be done independently at each merging node.
- For block-wise mutation, each node independently determines whether each block is mutated without any coordination with other nodes.

9) **Greedy Sweep [D12]**: Greedy sweep requires an additional protocol where the source is notified of the merging nodes with at least one coding link in the best solution obtained at the end of the iteration. Then, for each of such merging nodes, the source sends out a packet to test if uncoded transmission is possible on the link(s) where currently coding is required. Since this additional protocol requires more extensive coordination between nodes, we may leave this procedure optional, whose detailed description is omitted owing to space limitations. Note, however, that in our experiments the solutions obtained with the *block-wise* representation/operations are already good enough so that further improvement by greedy sweep has never been observed. Nevertheless, greedy sweep may be useful as a safeguard that prevents the algorithm's poor performance due to misadjusted parameters, e.g., too small population size.

C. Complexity

For evaluation of a single chromosome, each merging node v computes random linear combinations of inputs in the forward evaluation phase, which requires $O(d_{in}^v \cdot d_{out}^v \cdot R)$, and each non-merging node w simply forwards the received data, which requires $O(d_{out}^w)$. Feasibility test at each sink t is done by calculating the rank of a $d_{in}^t \times R$ matrix, where we assume $d_{in}^t \geq R$, hence it requires $O(d_{in}^t{}^2 R)$. In the backward evaluation phase, update of a fitness vector takes $O(d_{in}^v + d_{out}^v)$. Therefore, the computational complexity required for evaluation of a single chromosome is $O(\sum_{v \in V} d_{in}^v d_{out}^v R + \sum_{w \in V \setminus V} d_{out}^w + \sum_{t \in T} d_{in}^t{}^2 R)$, which can be substantially less than that for the centralized version of the algorithm.

D. Networks with Cycles

Cycles can be dealt with in two different ways as in other network coding problems. First, we can select a subgraph that does not contain a directed cycle, based on which we proceed

to code construction and decoding in essentially the same manner as in the acyclic case [16], [17]. Alternatively, we may directly apply coding over cycles by combining information from possibly different time periods at intermediate nodes and deploying memory at the receivers for decoding [12], [18], [19], where the network code can be considered essentially a convolutional code.

The former of the above two scenarios allows for simple coding and decoding, but it may necessitate coding at the links/nodes where coding is not necessary if some link connections were not removed in the earlier stage [5]. On the other hand, the latter scenario may allow us to explore the full-fledged tradeoff between coding and capacity, but both specifying and decoding the code are more complex than in the former case.

Here we focus on the first scenario and describe how our distributed algorithm can be incorporated in the whole framework of such network coding schemes. Note that, if the original coding scheme is designed to operate on an acyclic subgraph selected beforehand, which seems more practicable, there is no reason to employ more complex network codes based on the original cyclic graph to minimize coding resources. However, we expect that a similar approach can be readily applied to the convolutional network coding scenario with an appropriate cycle-avoiding mechanism for the transmission of the control messages such as the feedback information.

To set up an acyclic set of connections on a given network, we use the distributed algorithm in [17], where a binary variable is assigned to each pair (l, l') of incident links indicating that the connection from link l to link l' is allowed or not. The value of each binary variable is determined such that the transmission along a directed cycle is prohibited. It is interesting to note that those binary variables used for subgraph selection are assigned to the link coefficients the same way as in our algorithm. Hence, our algorithm can be incorporated into the whole framework as follows:

- i) Use the algorithm in [17] to select the set of link coefficients to be used for transmission.
- ii) Each node then exchanges the binary variables assigned to its links with its neighbors so that each node can identify the allowed connections.
- iii) We then apply the our distributed algorithm ignoring the link coefficients that correspond to the disallowed connections.

Alternatively, if minimizing the link cost is our primary concern, we may use the algorithms that find the minimum cost subgraph in a decentralized fashion, such as the one in [14]. The resulting subgraph does not contain a directed cycle if the link costs are all positive. Hence, a two-stage method is possible where the minimum cost subgraph selection is followed by our distributed algorithm. This two-stage method may perform very well in practice, as will be demonstrated in the next section.

V. EXPERIMENTAL RESULTS

The parameters used for the experiments are as follows: Population size is 150 and the iteration terminates after 1000 generations. Tournament size (for selection) and mutation rate are 100 and 0.012, respectively, for the block-wise case, and 10 and 0.006 for the bit-wise case. Crossover rate is fixed at 0.8.

A. Effects of Block-Wise Representation and Operators

We evaluate the performance of our algorithm with the block-wise and the bit-wise representations and operators, using the centralized version of the algorithm with the graph decomposition method. The experiments are based on the two topologies generated by the algorithm in [20] with the following parameters: (50 nodes, 87 links, 10 sinks, rate 5) and (75 nodes, 156 links, 15 sinks, rate 7).

For comparison, we also perform experiments with two existing greedy approaches by Fragouli et al. [7] ("Minimal 1") and Langberg et al. [4] ("Minimal 2"). For both approaches, link removal is done in a random order. For Minimal 1, the subgraph is selected also in a greedy fashion by sequentially removing links. Table I shows the best and the average values, as well as standard variation, obtained in 30 trials.

	(50,87,10,5)			(75,156,15,7)		
	Best	Avg.	Std.	Best	Avg.	Std.
Block-wise	2	2.40	0.62	3	3.63	0.61
Bit-wise	2	3.33	1.03	5	6.43	1.30
Minimal 1	3	4.90	1.37	6	9.50	2.16
Minimal 2	3	4.33	1.37	4	7.90	1.71

TABLE I
NUMBER OF REQUIRED CODING LINKS

Table I shows that the block-wise representation and operators clearly outperform the bit-wise counterpart in all aspects. We can also observe that the performance of our algorithm, with either of the two representation and operators, is at least as good and often better than that of both Minimal 1 and Minimal 2, except only in the best value of the bit-wise case for the larger network. More in-depth comparisons between these algorithms can be found in our subsequent paper [21], where our algorithm with the block-wise representation and operators is found to exhibit a far greater performance advantage over the other three cases.

B. Performance of Distributed Algorithm

Since the distributed algorithm shares the same computational part of GA with the centralized one, the two algorithms show the same performance in terms of solution quality. However, as shown in Section IV-C, the computational complexity required by the distributed algorithm depends on local topological parameters, and this can often lead to a significant gain in terms of the running time.

To compare the running times of the two approaches, we generate a set of highly connected topologies such that there exists a link between each pair of nodes i and j ($i < j$), where the source is node 1 and the sinks are the last 10 nodes. This test is pessimistic in the sense that the distributed

algorithm is simulated on a single machine while each node's function is performed by a separate thread, thus it cannot benefit from the multi-processing gain whereas it only suffers from additional computational burdens for managing a number of threads. Table II shows that, nevertheless, the distributed algorithm exhibits an advantage in running time as the size of the network grows.

Number of nodes	15	20	25	30	35	40
Centralized	0.3	1.5	4.3	13.5	29.5	65.6
Distributed	1.8	2.7	4.4	6.3	10.8	15.4

TABLE II
ELAPSED TIME PER GENERATION (SECONDS)

C. Effectiveness of Two-Stage Method

We introduced in Section IV-D a two-stage method where we first select a minimum-cost subgraph assuming network coding is done everywhere and then apply our evolutionary algorithm to the resulting subgraph. Though not optimal, this two-stage method can be very useful when optimization over both link cost and coding cost is required; the minimum link cost is guaranteed, and the resulting acyclic subgraph can often be substantially smaller than the original network.

We test the two-stage method on ISP 1755 and 3967 topologies from the Rocketfuel project [22], using the algorithm in [14] to obtain a minimum cost subgraph. With 10 randomly selected sinks and target rates 2, 3, and 4 on both topologies, each of 30 runs always ends up with *zero* coding links. These results may suggest that, while *assuming* network coding enables to calculate a minimum-cost subgraph, there may be very few links/nodes where network coding is actually required in the end.

VI. CONCLUSIONS AND FUTURE WORK

We have presented evolutionary approaches to minimizing the resources used for network coding in a single multicast scenario. The proposed algorithms have been shown to have advantages over our previously proposed algorithm [5], as well as other existing greedy algorithms, in terms of the applicability to general topologies, the solution quality, and the practicability in a distributed environment.

For future research, we may further improve the distributed algorithm, by a smarter management of population and packet transmissions, such that it converges faster to a better solution and works asynchronously, providing robustness against delay, failure, or topological changes in the network. Also, we could observe a tradeoff between coding and link usage in the sense that in some networks [5], reducing link usage first by subgraph selection may increase coding in the remaining subgraph, which is not the case in our experiments in Section V-C. Hence, whether there exists such a tradeoff can be considered a topological property of a network and thus the effectiveness of the two-stage method discussed in Section IV-D may depend on the network topology. We may further investigate this tradeoff using evolutionary algorithms designed for multi-objective optimization.

ACKNOWLEDGMENT

The first author would like to thank Fang Zhao for her help on the experiments.

REFERENCES

- [1] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, 2000.
- [2] S.-Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Trans. Inf. Theory*, vol. 49, no. 2, pp. 371–381, 2003.
- [3] M. B. Richey and R. G. Parker, "On multiple steiner subgraph problems," *Networks*, vol. 16, no. 4, pp. 423–438, 1986.
- [4] M. Langberg, A. Sprintson, and J. Bruck, "The encoding complexity of network coding," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2386–2397, 2006.
- [5] M. Kim, C. W. Ahn, M. Médard, and M. Effros, "On minimizing network coding resources: An evolutionary approach," in *Proc. NetCod*, 2006.
- [6] T. Ho, R. Koetter, M. Médard, D. R. Karger, and M. Effros, "The benefits of coding over routing in a randomized setting," in *Proc. IEEE ISIT*, 2003.
- [7] C. Fragouli and E. Soljanin, "Information flow decomposition for network coding," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 829–848, 2006.
- [8] K. Bhattad, N. Ratnakar, R. Koetter, and K. R. Narayanan, "Minimal network coding for multicast," in *Proc. IEEE ISIT*, 2005, pp. 1730–1734.
- [9] M. Mitchell, *An Introduction to Genetic Algorithms*. MIT Press, 1996.
- [10] R. Elbaum and M. Sidi, "Topological design of local-area networks using genetic algorithms," *IEEE/ACM Trans. Netw.*, vol. 4, no. 5, pp. 766–778, 1996.
- [11] B. Dengiza, F. Altıparmak, and A. E. Smith, "Efficient optimization of all-terminal reliable networks, using an evolutionary approach," *IEEE Trans. Rel.*, vol. 46, no. 1, pp. 18–26, 1997.
- [12] R. Koetter and M. Médard, "An algebraic approach to network coding," *IEEE/ACM Trans. Netw.*, vol. 11, no. 5, pp. 782–795, 2003.
- [13] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network Flows: Theory, Algorithms, and Applications*. Prentice Hall, 1993.
- [14] D. S. Lun, M. Médard, T. Ho, and R. Koetter, "Network coding with a cost criterion," MIT-LIDS, Tech. Rep. P-2584, 2004.
- [15] G. Harik, E. Cantú-Paz, D. E. Goldberg, and B. L. Miller, "The gambler's ruin problem, genetic algorithms, and the sizing of populations," *Evolutionary Computation*, vol. 7, no. 3, pp. 231–253, 1999.
- [16] S. Jaggi, P. Sanders, P. A. Chou, M. Effros, S. Egner, K. Jain, and L. Tolhuizen, "Polynomial time algorithms for multicast network code construction," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 1973–1982, 2005.
- [17] T. Ho, B. Leong, R. Koetter, and M. Médard, "Distributed asynchronous algorithms for multicast network coding," in *Proc. NetCod*, 2005.
- [18] T. Ho, M. Médard, J. Shi, M. Effros, and D. R. Karger, "On randomized network coding," in *Proc. Annual Allerton Conference on Communication, Control, and Computing*, 2003.
- [19] E. Erez and M. Feder, "Efficient network codes for cyclic networks," in *Proc. ISIT*, 2005.
- [20] G. Melançon and F. Philippe, "Generating connected acyclic digraphs uniformly at random," *Inf. Process. Lett.*, vol. 90, no. 4, pp. 209–213, 2004.
- [21] M. Kim, V. Aggarwal, U.-M. O'Reilly, M. Médard, and W. Kim, "Genetic representations for evolutionary optimization of network coding resources," in *Applications of Evolutionary Computing*, 2007.
- [22] N. Spring, R. Mahajan, and D. Wetherall, "Measuring ISP topologies with rocketfuel," in *Proc. ACM/SIGCOMM '02*, 2002, pp. 133–145.