Lecture 2, E6713

NETWORK MULTICAST – Throughput

All sources can simultaneously transmit information to all receivers if

1. for each receiver j,
   there are h edge-disjoint paths (S_i, R_j) for $1 \leq i \leq h$, and
2. each node of G can re-encode information.

NETWORK MULTICAST – Codes

- Ahlswede, Cai, Li, Yeung, '00: Random codes, rate $(1-\epsilon)h$.
- Cai, Li, Yeung, '02: Linear codes, rate h.
- Koetter, Medard, '02: Algebraic characterization of linear codes, $|\mathbb{F}| = O(N \cdot |E|)$
- Sanders, Egner, Tolhuizen, and Jaggi, Chou, Jain, '03:
  Polynomial time code construction, $|\mathbb{F}| = O(N)$.
- Ho et al: Randomized (decentralized) codes.
- Fragouli, Soljanin: Polynomial time deterministic, decentralized code construction, $|\mathbb{F}| = O(N)$.
  New alphabet size bound $|\mathbb{F}| = O(\sqrt{N})$.

DECOMPOSITION INTO SUBTREES

- A flow algorithm is run for each receiver to find h edge-disjoint paths (S_i, R_j), $1 \leq i \leq h$.
- Let G'_{ij} be the line graph of (S_i, R_j) and $G' = \bigcup G'_{ij}$.
- Coding points are the nodes of G' with two or more inputs.
- We partition G' into subsets T_i so that
  1. each T_i contains exactly one source node or a coding point, and
  2. every other node belongs to the T_i containing its first ancestral coding point or source node.
- For coding, we need to know how T_i's are connected and
  1. which sources are available at the input of each T_i,
  2. which receiver nodes are in each T_i.

DECOMPOSITION INTO SUBTREES – An Example

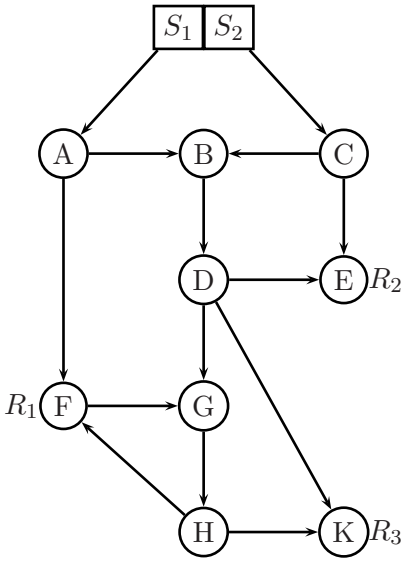
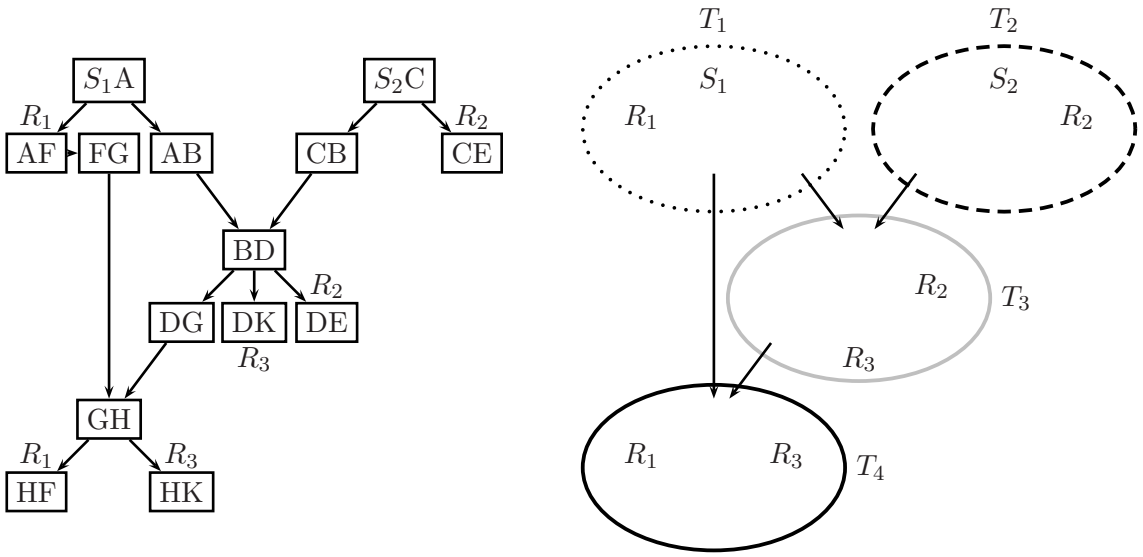


Figure 1: Network with two sources $\{S_1, S_2\}$ and three receivers $\{F, E, K\}$.



(a) (b)

Figure 2: Line graph for the network with 2 sources and 3 receivers with coding points BD and GH and its subtree graph.

For networks with two sources, we will use the points on the projective space of dimension 1, *i.e.,* the projective line $\mathbb{PG}(1, q)$:

$$[0\,1],\ [1\,0],\ \text{and}\ [1\,\alpha^i]\ \text{for}\ 0 \leq i \leq q-2, \tag{1}$$

where α is a primitive element of $\mathbb{F}_q$. Any two different points on the projective line $\mathbb{PG}(1, q)$ form a basis for $\mathbb{F}_q^2$, and thus any point on the line is in their span. Consequently, any code which assigns different points of $\mathbb{PG}(1, q)$ to different subtrees is a valid network multicast code. Note that this type of coding is decentralized since we only need to use information available locally at a subtree, without taking into account how the subtrees are connected or how the receiver nodes are distributed over the graph.

Example 1 *Two codes for the network in Fig. 1*

1. *A Decentralized Code: Since there are two source and two coding trees, we take the first four points from the list (1):*

   $$c(T_1) = [1\,0],\ c(T_2) = [0\,1],\ c(T_3) = [1\,1],\ c(T_4) = [1\,\alpha].$$

   For this code, we need the field with three elements.

2. *The Smallest Alphabet Code: If we take into account the information on the receivers in trees, we see that T_2 and T_4 can be assigned the same coding vector since they do not share any receivers:*

   $$c(T_1) = [1\,0],\ c(T_2) = [0\,1],\ c(T_3) = [1\,1],\ c(T_4) = [0\,1].$$

   For this code, the field with two elements is sufficient.

Scalability: Suppose that a new node J with the receiver R_4 is introduced in our example network in Fig. 1, as shown in Fig. 3. We are interested in finding out if the two codes described in the Example 1

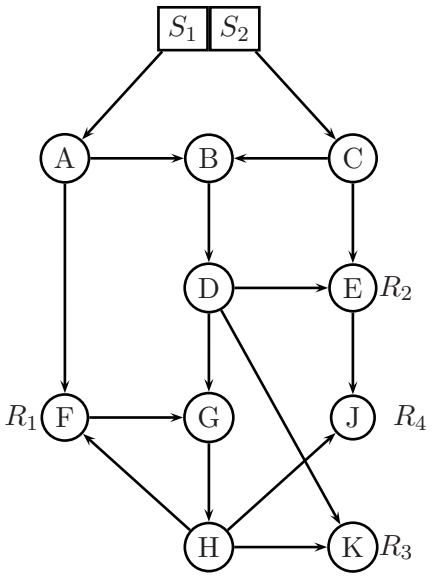


Figure 3: Network with two sources $\{S_1, S_2\}$ and four receivers F, E, K, and J.

are still valid. *If the network changes, its decentralized code will not have to be changed if the same subtree decomposition is still possible.* Thus the first code in Example 1 is also valid for the network in Fig. 3. To see if the other code in Example 1 is also valid, we have to see which two subtrees in Fig. 2 will contain the receiver R_4. We see that R_4 will have to be in the subtree T_4 and in either subtree T_2 or T_3 (*i.e.,* observe one of the two corresponding coding vectors). If R_4 is placed in T_2, the code is not valid since T_2 and T_4 have identical labels. If R_4 is placed in T_3, the code remains valid.

RANDOMIZED NETWORK CODING – by Ho etal.

- The components of local coding vectors are chosen uniformly at random from $\mathbb{F}_q$.
- The components of global coding vectors are passed along with data.
- Let η be the number of edges with associated randomized coefficients.
- The probability that all N receivers can decode all h sources is at least $(1 - N/q)^\eta$.
- Advantages: decentralized, scalable.
- Disadvantages: very large alphabets may be required, h-symbol overhead.

A CODE CONSTRUCTION ALGORITHM – by Jaggi etal.

- A flow algorithm is run for each receiver to find h edge-disjoint paths (S_i, R_j), $1 \leq i \leq h$.
- At each subsequent stage of the algorithm,
  the global coding vectors for certain edges belonging to these paths are determined.
- These are the edges with all upstream coding vectors already determined.
- At stage k, matrix $\mathbf{C}_j^{(k)}$ is computed for receiver j, $1 \leq j \leq N$. Initially $\mathbf{C}_j^{(0)} = I$.
- The i-th raw of $\mathbf{C}_j^{(k)}$ represents the global coding vector of the edge of (S_i, R_j) processed at stage k.
- Note that the i-th raw of $\mathbf{C}_j^{(k)}$ will be identical to say l-th raw of some $\mathbf{C}_m^{(k)}$.
- The coding vectors determined at stage k ensure that matrices $\mathbf{C}_j^{(k)}$, $1 \leq j \leq N$, are regular.
- Algorithms for finding such vectors, field size, scalability?