Efficient Algorithms for Index Coding

Mohammad Asad Rehman Chaudhry

Department of Electrical & Computer Engineering, Texas A & M University, (masadch@tamu.edu)

Joint work with Dr. Alex Sprintson

Background

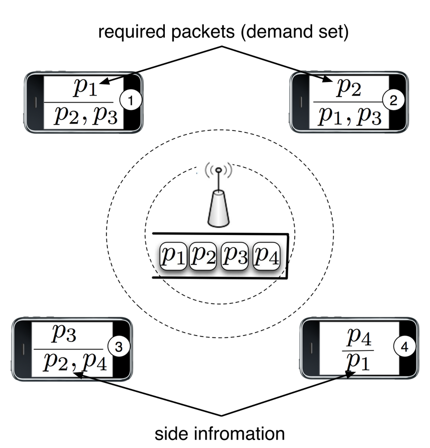
1. There has recently been significant interest in multihop wireless networks.
2. Wireless networks are expected to provide an efficient and reliable service and support a broad range of the emerging applications.
3. Limited wireless spectrum together with interference and fading pose significant challenges for network designers.
4. The novel technique of *Network Coding* has a significant potential to improve the through-put, reliability and of wireless networks
   - By taking advantage of the *broadcast nature* of wireless medium.
5. Design and analysis of efficient network coding schemes for wireless networks have recently attracted a significant interest from the research community.
6. Katti *et al.* presented a practical network coding architecture, referred to as *COPE*, that uses *opportunistic listening* and *opportunistic coding*.

Index Coding

1. A key principle of wireless network coding architectures is *opportunistic listening*.
2. With this approach each network node is snooping on all communications over the wireless medium.
3. The overheard packets are stored for a limited period of time. The key idea is to take advantage of the overheard packets, also referred to as *side information*, to achieve a higher rate of the information exchange.

1. An instance of the Index Coding problem includes:
   (a) A relay node r;
   (b) A set $C = \{c_1, \ldots, c_n\}$ of wireless clients;
   (c) A set $P = \{p_1, p_2, \ldots, p_m\}$ of packets that need to be delivered to clients in C.
2. Each client $c_i \in C$ is associated with two sets:
   (a) *Want(demand)* set $W(c_i) \subseteq P$ - the set of packets required by c_i;
   (b) *Has(side information)* set $H(c_i) \subseteq P$ - the set of packets available at c_i.

Index Coding problem



Minrank problem

	p_1	p_2	p_3	p_4
c_1	1	X	X	0
c_2	X	1	X	0
c_3	0	X	1	X
c_4	X	0	1	0

Fig 1: Index Coding and Min-Rank problems

3. In each *round* of communication the relay can transmit a function of packets in P;
4. *Index Coding problem*: find a scheme that allows each client to decode the packets it requested while minimizing the number of transmissions;
5. Complimentary problem: maximize the number of *saved transmissions*;
6. The Index Coding problem was introduced by Birk and Kol;
7. Both problems are *NP-hard*, the Index Coding problem is hard to approximate.

Our goals

- Present following solutions to Index Coding problem:
  - An exact solution using reduction to *Boolean Satisfiability problem(SAT)*;
  - A heuristic and time efficient solution based on *Graph Coloring*.
- Present following approximate solutions to Complementary Index Coding problem i.e. *transmission saved* based on:
  - An approximate and time efficient solution based on *Color Saving*;
  - An approximate and time efficient solution based on *Cycle Packing*.

Relation to the Min-Rank Problem

- Shown by Bar-Yossef, Birk, Jayram and Kol
- Suppose that the number of clients is equal to the number of packets.
- An instance of Index Coding can be represented by a matrix M.
  - Each row in M corresponds to a client c_i, $i = 1, \cdots, n$, and each column in M corresponds to a packet p_j, $j = 1, \cdots, n$.
  - For a given instance of Index Coding problem Each entry (i,j) in M can take one of these three values: (i) (i,j)=1 if $p_j \in W(c_i)$; (ii) $(i,j) \in 1,0$ if $p_j \in H(c_i)$, i.e., a *don't care* value X as it can take value either 0 or 1; (iii) $(i,j) = 0$ if $p_j \notin \{H(c_i) \cup W(c_i)\}$.
- Problem reduces to finding values of don't care conditions (X) such that rank of matrix is minimized.
- Matrix M for an instance of Index Coding is shown in Fig 1.

Solutions to Index Coding

Optimal Solution Based on SAT

- *SAT:* a boolean decision problem:
  - Input: A Boolean expression written using AND, OR, NOT (Conjunctive normal form);
  - Output: For Input expression, is there some assignment of TRUE and FALSE values to the variables that will make the entire expression true;
  - SAT can be efficiently solved by available SAT solvers such as *Chaff* or *Minisat*.
- *Reduction to SAT:*
  - Index Coding problem is transformed into a boolean expression by substituting the summation and multiplication over $GF(2)$ by the AND ($\wedge$) and XOR ($\oplus$) operations;
  - g_j^t: Encoding coefficient for j^{th} transmission and t^{th} packet;
  - q_i^j: Decoding coefficient for i^{th} client and j^{th} transmission;
  - For each packet in the demand set of client c_i : $\oplus_{j=1}^{k}\left(g_j^t \wedge q_i^j\right) \equiv 1 \ \forall c_i \in C$;
  - For each packet that is neither in demand set nor in the side infromation set of client c_i : $\oplus_{j=1}^{k}\left(g_j^t \wedge q_i^j\right) \equiv 0 \ \forall c_i \in C$;
  - Boolean expression for client c_1 shown in Fig 1 : $(g_1^1 \wedge q_3^1) \oplus (g_2^1 \wedge q_3^2) \wedge \overline{(g_1^1 \wedge q_3^4) \oplus (g_2^1 \wedge q_3^4)}$.

Heuristic Solution Based on Graph Coloring

- *Time Efficient* as compared to optimal solution.
- In graph coloring, we need to assign a color to each vertex such that for any edge the vertices are assigned different colors.
- *Reduction to Graph Coloring:*
  - For each client c_i there is a corresponding vertex v_i in the graph;
  - Each two vertices v_i and v_j are connected by an edge if one of the following holds:
    * Clients c_i and c_j have identical demand sets;
    * Demand set of client c_i is subset of side information set of client c_j and side information set of client c_j is subset of the side information set of client c_i.
  - Find cliques in resultant graph;
  - All clients that correspond to nodes in a clique can be satisfied by one transmission, which includes a linear combination of all packets in their demands sets;
  - Minimize the number of transmissions by finding the least number of cliques that cover whole graph also known as *clique partition* problem;
  - The clique partitioning problem for a graph $G(V,E)$ is equivalent to the graph coloring problem of the complimentary graph $\bar{G}(V,E)$, the complimentary graph $\bar{G}(V,E)$ contains all edges in $V \times V \setminus E$.

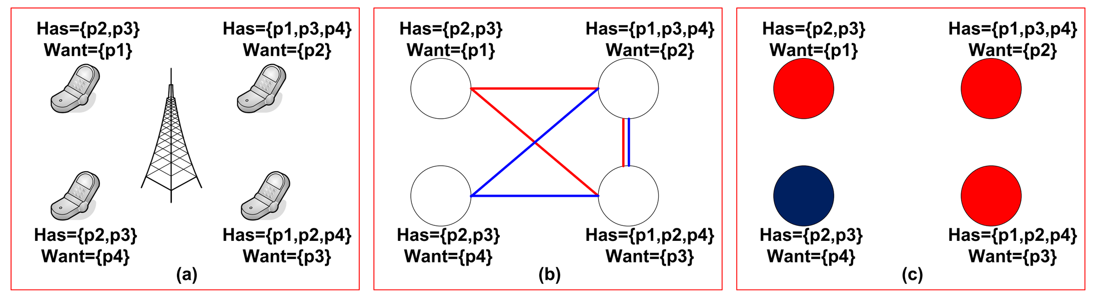


Fig 2: (a)An instance of Index Coding (b) Corresponding instance of Clique Partition (c) Corresponding instance of Graph Coloring on complement graph

Solutions to Complimentary Problem to Index Coding

Approximate Solution Based on Color Saving

- Time Efficient.
- *Color Saving:*
  - Color Saving maximizes the *number of unused colors*;
  - "Number of unused colors" is the difference between the number of vertices in the graph and the number of colors used for the coloring;
  - *Maximizing* the number of unused colors is equivalent to *minimizing* number of used colors resulting in minimizing required number of transmissions.
- *A Simple Color Saving Based Algorithm:*
  - Construct an undirected graph $G(V,E)$ as described in Section "Graph Coloring Based Solution";
  - While there exists a clique of size 3 in $G(V,E)$ do
    * Find a clique $\{v_i, v_j, v_k\}$ of size 3 in $G(V,E)$;
    * Create a packet that satisfies all clients in $\{v_i, v_j, v_k\}$;
    * $V := V \setminus \{v_i, v_j, v_k\}$;
  - Compute a maximum matching of $G(V,E)$;
  - For each pair $\{v_i, v_j\}$ in the matching create a packet that satisfies all clients in $\{v_i, v_j\}$;
  - Create a new packet for each one of the remaining vertices of V.
- Using *Semi-local Optimization* a solution with approximation ratio of $\frac{5}{6}$ to the optimal solution of Color Saving can be achieved.

Approximate Solution Based on Cycle Packing

- *Cycle Packing:*
  - Packing maximum number of edge disjoint cycles in a graph.
- The basic idea of the solution is to pack maximum number of edge disjoint cycles in a graph, $H(V,E)$, where $H(V,E)$ is constructed as:
  - For each client c_i there is a corresponding vertex v_i in the graph;
  - There is an arc from vertex v_i to vertex v_j if two if one of the following holds:
    * Clients c_i and c_j have identical $Wants$ sets;
    * $Wants$ set of one client c_i is subset of Has set of client c_j.
- *Feed-Back Vertex Set:*
  - Least number of vertices whose removal shall make the graph acyclic.
- *Feed-Back Arc Set:*
  - Least number of arcs whose removal shall make the graph acyclic.
- Feed-Back Arc Set and Cycle Packing are dual problems.
- *Approximation Ratio: The number of transmissions saved by the cycle packing is $\Omega(\frac{1}{\sqrt{n \cdot \log n \cdot \log \log n}})$ of the number of transmissions saved by the optimal solution to Index Coding problem, where n is the number of clients.*

Simulation Results

- Optimal SAT based solution shows the best performance in terms of coding gain followed by Graph Coloring and Color Saving.
- Time taken by Color Saving is much less as compared to SAT based solution and Graph Coloring.
- Sparsest Set Clustering shows least coding gain performance and best time performance.

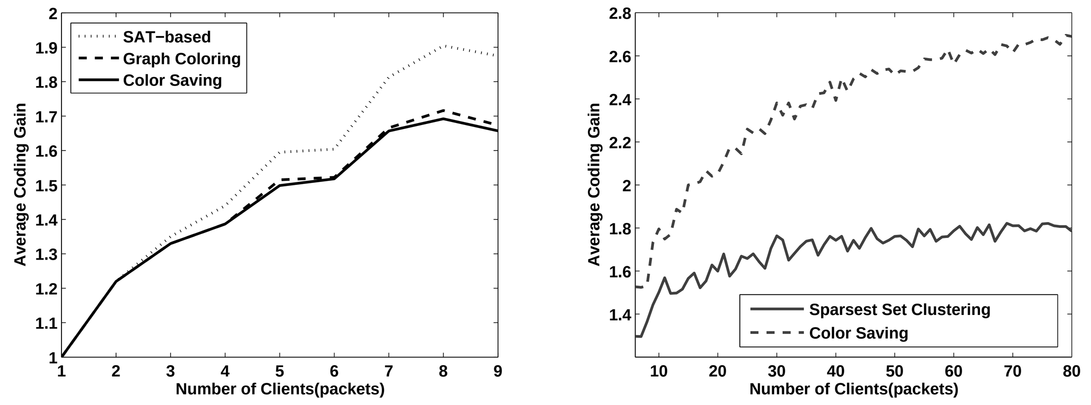


Fig: (a)Comparison of different technique used (b)Comparison of Sparsest Set Clustering and Color Saving

No. of Clients	SAT-based	Graph Coloring	Color Saving
2	0.63	0.03	0.00506
4	0.7341	0.06942	0.01524
5	0.9391	0.07914	0.01492
6	0.9622	0.1043	0.0454
8	11.93	0.2637	0.09848
9	82.97	0.7654	0.05246

Table: Time comparison (CPU(sec)) of different techniques used

No. of Clients	Sparsest Set Clustering	Color Saving
40	3.28	0.514
80	6.764	5.829
100	9.375	15.719
160	15.656	143.188

Table: Time comparison (CPU(sec)) of Sparsest Set Clustering and Color Saving

Refernces

Y. Birk and T. Kol. Coding on Demand by an Informed Source (iscod) for Efficient Broadcast of Different Supplemental Data to Caching Clients. *IEEE/ACM Trans. Netw.*, 14(SI):2825–2830, 2006.
Z. Bar-Yossef, Y. Birk, T. S. Jayram, and T. Kol. Index Coding with Side Information. *FOCS, 197– 206, 2006.*

M.A.R Chaudhry, and A. Sprintson. Efficient Algorithms for Index Coding. *IEEE INFOCOM Workshops, 1 – 4, 2008.*

S. Katti, H. Rahul, W. Hu, D. Katabi, M. Medard, and J. Crowcroft. XORs in the Air: Practical Wireless Network Coding. *ACM SIGCOMM, 36(4): 243 – 254, 2006.*

M. Langberg, and A. Sprintson. On the Hardness of Approximating the Network Coding Capacity. *IEEE ISIT, 6 – 11, 2008.*

S. Y. El Rouayheb, M. A. R. Chaudhry, and A. Sprintson. On the Minimum Number of Transmissions in Single-Hop Wireless Coding Networks. *IEEE ITW, 120 – 125, 2007.*

http://minisat.se/

R. Duh, and Martin Frer. Approximation of k-Set Cover by Semi-Local Optimization. *ACM STOC, 256–264, 1997.*

M. Krivelevich, Z. Nutov, M. R. Salavatipour,J. V. Yuster, and R. Yuster. Approximation Algorithms and Hardness Results for Cycle Packing Problems. newblock *ACM Transactions on Algorithms, 3(4): 1549 – 6325, 2007.*