Optimal Index Codes with Near-Extreme Rates

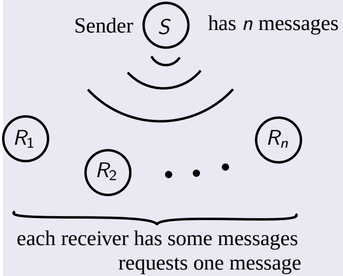
Vitaly Skachek

(joint work with **Son Hoang Dau** and **Yeow Meng Chee**)

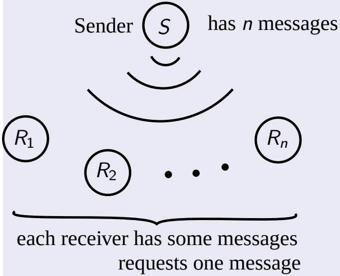
Estonian Theory Days
October 25th, 2013

This work is supported by Research Grant NRF-CRP2-2007-03 (Singapore)

Index Coding with Side Information (ICSI)

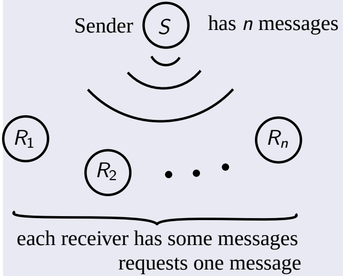


Index Coding with Side Information (ICSI)



Receiver	Demand	Side Info.
R_1	x_1	$\{x_2\}$
R_2	x_2	$\{x_3\}$
R_3	x_3	$\{x_1, x_4\}$
R_4	x_4	$\{x_5\}$
R_5	x_5	$\{x_2, x_4\}$

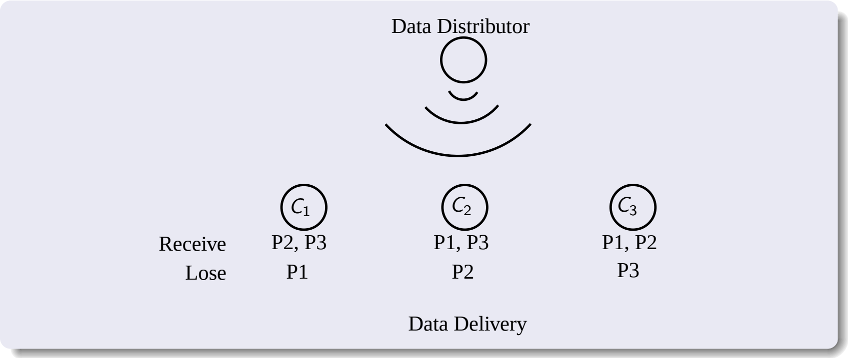
Index Coding with Side Information (ICSI)



Receiver	Demand	Side Info.
R_1	x_1	$\{x_2\}$
R_2	x_2	$\{x_3\}$
R_3	x_3	$\{x_1, x_4\}$
R_4	x_4	$\{x_5\}$
R_5	x_5	$\{x_2, x_4\}$

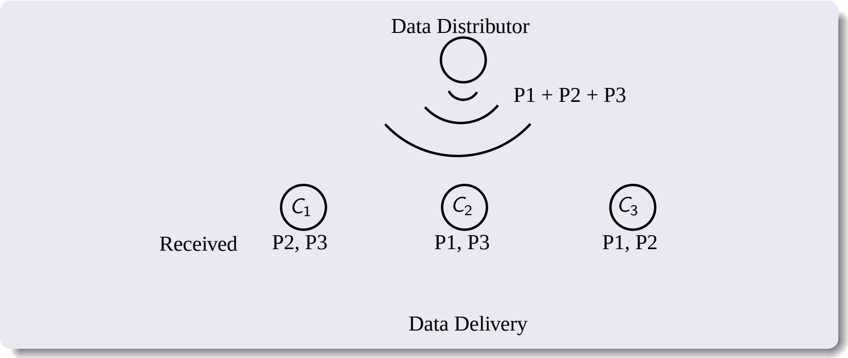
Questions: How can S satisfy all the demands in a minimum number of transmissions?

Motivation

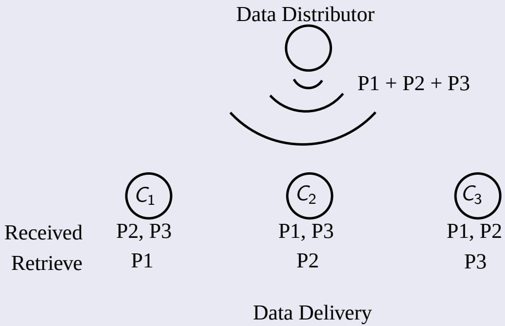


Data Distributor
C_1
C_2
C_3
Receive
Lose
P2, P3
P1
P1, P3
P2
P1, P2
P3
Data Delivery

Motivation


Data Distributor
P1 + P2 + P3
C_1
C_2
C_3
Received
P2, P3
P1, P3
P1, P2
Data Delivery

Motivation

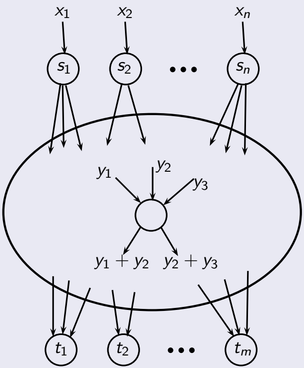


Data Distributor
P1 + P2 + P3
C_1
C_2
C_3
Received
P2, P3
P1, P3
P1, P2
Retrieve
P1
P2
P3
Data Delivery

Motivation

Network Coding

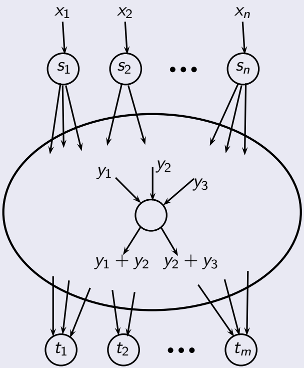
Network Coding (NC): Ahlswede *et al.*, 2000



Motivation

Network Coding

Network Coding (NC): Ahlswede *et al.*, 2000



Index Coding and Network Coding

1. Index coding proposed by Birk and Kol (1998)
2. ICSI is a special case of *non-multicast* network coding
3. ICSI and NC are equivalent (El Rouayheb, Sprintson, Georghiades, 2008; Effros, El Rouayheb, Langberg, 2012)

Literature Overview

- Bar-Yossef *et al.* (2006)
  - Associate each ICSI instance with a digraph
  - Optimal scalar linear transmission rate = minrank of the digraph

Literature Overview

- Bar-Yossef *et al.* (2006)
  - Associate each ICSI instance with a digraph
  - Optimal scalar linear transmission rate = minrank of the digraph
- Peeters (1996): Finding minrank of a graph is NP-hard (deciding whether minrank of a graph is three is NP-hard)

Literature Overview

- Bar-Yossef *et al.* (2006)
  - Associate each ICSI instance with a digraph
  - Optimal scalar linear transmission rate = minrank of the digraph
- Peeters (1996): Finding minrank of a graph is NP-hard (deciding whether minrank of a graph is three is NP-hard)
- Chaudhry and Sprintson (2008): exact and approximate algorithms for minranks

Literature Overview

- Bar-Yossef *et al.* (2006)
  - Associate each ICSI instance with a digraph
  - Optimal scalar linear transmission rate = minrank of the digraph
- Peeters (1996): Finding minrank of a graph is NP-hard (deciding whether minrank of a graph is three is NP-hard)
- Chaudhry and Sprintson (2008): exact and approximate algorithms for minranks
- Bar-Yossef *et al.* (2006); Berliner and Langberg (2011): polynomial time computation of minranks for some families of graphs

Other works

- Alon *et al.*(2008)
- Lubetzky, Stav (2009)
- Dau, Skachek, Chee (2011)
- Haviv, Langberg (2012)
- Ong, Lim, Ho (2012)
- Brahma, Fragouli (2012)
- Tehrani, Dimakis (2012)
- Neely, Tehrani, Zhang (2012)
- Shum, Dai, Sung (2012)
- Maleki, Cadambe, Jafar (2012)
- Arbabjolfaei, Bandemer, Kim, Sasoglu (2013)
- Shanmugam, Dimakis, Caire (2013)

Our contributions

- In this work we

Our contributions

- In this work we
  - characterize families of digraphs with some extremely high or low minranks

Our contributions

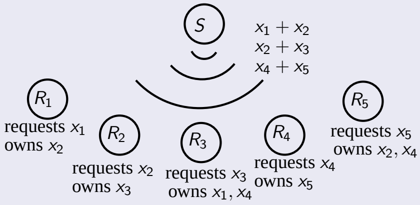
- In this work we
  - characterize families of digraphs with some extremely high or low minranks
  - show that deciding whether minrank of a digraph is two is NP-hard

Our contributions

- In this work we
  - characterize families of digraphs with some extremely high or low minranks
  - show that deciding whether minrank of a digraph is two is NP-hard (trivial for graphs)

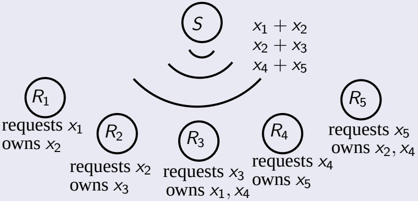
An Example

Consider the following ICSI instance:



An Example

Consider the following ICSI instance:

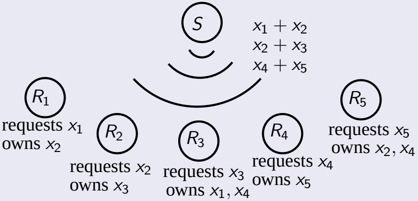


S transmits $x_1 + x_2$, $x_2 + x_3$, and $x_4 + x_5$ (an IC of length 3).

Definitions and Notation

An Example

Consider the following ICSI instance:



S transmits $x_1 + x_2$, $x_2 + x_3$, and $x_4 + x_5$ (an IC of length 3).

1. R_1 decodes: $x_1 = x_2 + (x_1 + x_2)$
2. R_2 decodes: $x_2 = x_3 + (x_2 + x_3)$
3. R_3 decodes: $x_3 = x_1 + (x_1 + x_2) + (x_2 + x_3)$
4. R_4 decodes: $x_4 = x_5 + (x_4 + x_5)$
5. R_5 decodes: $x_5 = x_4 + (x_4 + x_5)$

Describe an ICSI Instance via Side Information Digraphs

Definitions and Notation

Describe an ICSI Instance via Side Information Digraphs

- Vertex set: $\mathcal{V}(\mathcal{D}) = [n] = \{1, 2, \ldots, n\}$ (n messages, n receivers)

Definitions and Notation

Describe an ICSI Instance via Side Information Digraphs

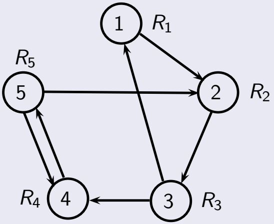
- Vertex set: $\mathcal{V}(\mathcal{D}) = [n] = \{1, 2, \ldots, n\}$ (n messages, n receivers)
- Arc set: $\mathcal{E}(\mathcal{D}) = \big\{(i, j) : \ i \in [n], \ R_i \text{ has } x_j \text{ as side information}\big\}$

Definitions and Notation

Describe an ICSI Instance via Side Information Digraphs

- Vertex set: $\mathcal{V}(\mathcal{D}) = [n] = \{1, 2, \ldots, n\}$ (n messages, n receivers)
- Arc set: $\mathcal{E}(\mathcal{D}) = \big\{(i, j) : \ i \in [n], \ R_i \text{ has } x_j \text{ as side information}\big\}$

Receiver	Demand	Side Info.
R_1	x_1	$\{x_2\}$
R_2	x_2	$\{x_3\}$
R_3	x_3	$\{x_1, x_4\}$
R_4	x_4	$\{x_5\}$
R_5	x_5	$\{x_2, x_4\}$



Definition (Haemer, 1978)

Let $\mathcal{D}$ be a digraph where $\mathcal{V}(\mathcal{D}) = [n]$.

Definitions and Notation

Definition (Haemer, 1978)

Let $\mathcal{D}$ be a digraph where $\mathcal{V}(\mathcal{D}) = [n]$.

1. A matrix $\mathbf{M} = (m_{i,j}) \in \mathbb{F}_q^{n \times n}$ is said to *fit* $\mathcal{D}$ if

Definitions and Notation

Definition (Haemer, 1978)

Let $\mathcal{D}$ be a digraph where $\mathcal{V}(\mathcal{D}) = [n]$.

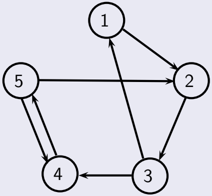
1. A matrix $\mathbf{M} = (m_{i,j}) \in \mathbb{F}_q^{n \times n}$ is said to *fit* $\mathcal{D}$ if

$$\begin{cases} m_{i,j} \neq 0, & i = j, \\ m_{i,j} = 0, & i \neq j, \ (i,j) \notin \mathcal{E}(\mathcal{D}). \end{cases}$$

Definitions and Notation

Definition (Haemer, 1978)

Let $\mathcal{D}$ be a digraph where $\mathcal{V}(\mathcal{D}) = [n]$.



A digraph of minrank 3

1. A matrix $\mathbf{M} = (m_{i,j}) \in \mathbb{F}_q^{n\times n}$ is said to *fit* $\mathcal{D}$ if

$$\begin{cases} m_{i,j} \neq 0, & i = j, \\ m_{i,j} = 0, & i \neq j,\ (i,j) \notin \mathcal{E}(\mathcal{D}). \end{cases}$$

2. The *minrank* of $\mathcal{D}$ over $\mathbb{F}_q$ is defined to be

$$\mathrm{minrk_q}(\mathcal{D}) \triangleq \min \left\{ \mathrm{rank}(\mathbf{M}) :\ \mathbf{M} \in \mathbb{F}_q^{n\times n} \text{ and } \mathbf{M} \text{ fits } \mathcal{D} \right\}.$$

$$\mathbf{A}+\mathbf{I} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix} \implies \mathbf{M} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & \underline{\mathit{0}} & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & \underline{\mathit{0}} & 0 & 1 & 1 \end{pmatrix}$$

Theorem (Bar-Yossef *et al.*, 2006)

$\mathrm{minrk_q}(\mathcal{D})$ = *length of the shortest (scalar linear) index code*

Graphs and Digraphs of Extreme MinRanks

Theorem (Bar-Yossef *et al.*, 2006)

$\mathsf{minrk_q}(\mathcal{D})$ = *length of the shortest (scalar linear) index code*

Theorem

For any digraph $\mathcal{D}$ we have

$$\alpha(\mathcal{D}) \leq \mathsf{minrk_q}(\mathcal{D}) \leq \mathsf{cc}(D).$$

Graphs and Digraphs of Near-Extreme MinRanks

Summary

Min-Rank	Graph $\mathcal{G}$	Digraph $\mathcal{D}$
1	$\mathcal{G}$ is complete (trivial)	$\mathcal{D}$ is complete (trivial)
2	$\mathcal{G}$ is not complete and $\overline{\mathcal{G}}$ is 2-colorable (Peeters, '96)	$\mathcal{D}$ is not complete and $\overline{\mathcal{D}}$ is fairly 3-colorable*
$n-2$	$\mathcal{G}$ (connected) has a maximum matching of size two and does not contain F as a subgraph*	unknown
$n-1$	$\mathcal{G}$ (connected) is a star graph	unknown
n	$\mathcal{G}$ has no edges (trivial)	$\mathcal{D}$ has no circuits (from Bar-Yossef *et al.*, '06)

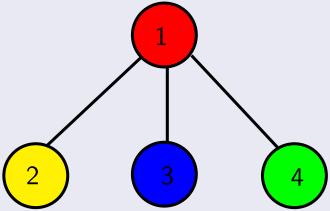
Graphs and Digraphs of Near-Extreme MinRanks

Summary

Min-Rank	Graph $\mathcal{G}$	Digraph $\mathcal{D}$
1	$\mathcal{G}$ is complete (trivial)	$\mathcal{D}$ is complete (trivial)
2	$\mathcal{G}$ is not complete and $\overline{\mathcal{G}}$ is 2-colorable (Peeters, '96)	$\mathcal{D}$ is not complete and $\overline{\mathcal{D}}$ is fairly 3-colorable*
$n-2$	$\mathcal{G}$ (connected) has a maximum matching of size two and does not contain F as a subgraph*	unknown
$n-1$	$\mathcal{G}$ (connected) is a star graph	unknown
n	$\mathcal{G}$ has no edges (trivial)	$\mathcal{D}$ has no circuits (from Bar-Yossef *et al.*, '06)

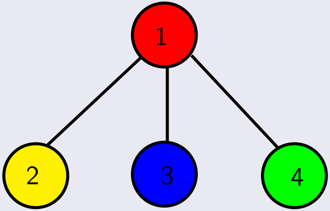
Digraphs of minranks two

$\mathrm{minrk}_2(\mathcal{D}) = 2$ iff $\mathcal{D}$ is not complete and $\overline{\mathcal{D}}$ is fairly 3-colorable

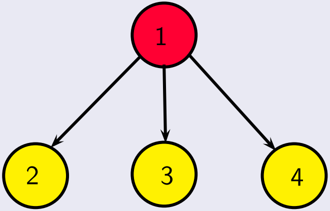


A 4-coloring of a graph

Graphs and Digraphs of Extreme MinRanks

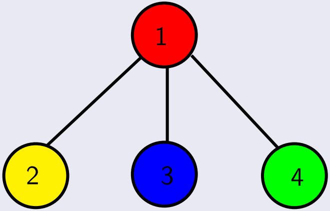


A 4-coloring of a graph

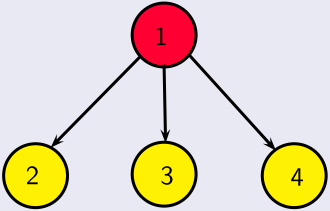


A fair 2-coloring of a digraph

A fair coloring: out-neighbors of each vertex have the same color



A 4-coloring of a graph



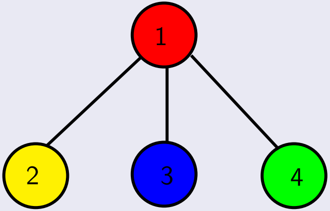
A fair 2-coloring of a digraph

A fair coloring: out-neighbors of each vertex have the same color

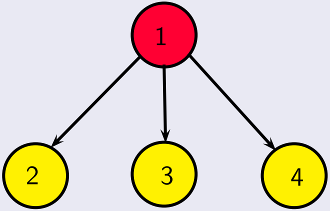
Theorem

The fair k-coloring problem is NP-complete for $k \geq 3$

Graphs and Digraphs of Extreme MinRanks



A 4-coloring of a graph



A fair 2-coloring of a digraph

A fair coloring: out-neighbors of each vertex have the same color

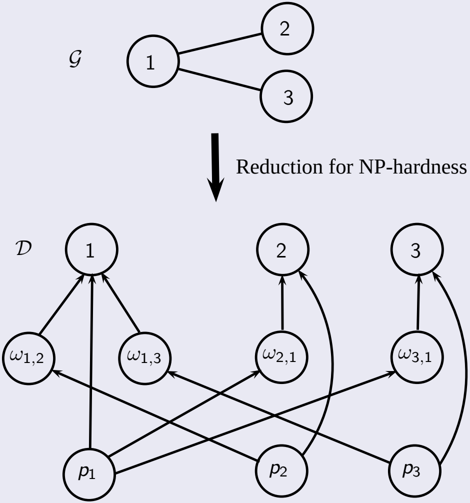
Theorem

The fair k-coloring problem is NP-complete for $k \geq 3$

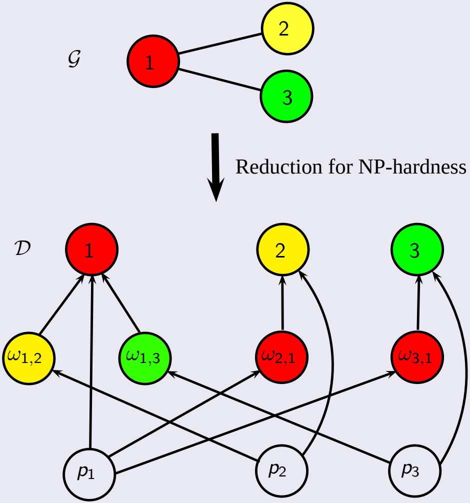
Corollary

Deciding whether $\mathrm{minrk}_2(\mathcal{D}) = 2$ *is an NP-complete problem*

Reduction from k-Coloring to Fair k-Coloring



Reduction from k-Coloring to Fair k-Coloring



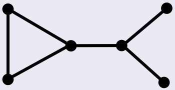
Graphs and Digraphs of Near-Extreme MinRanks

Summary

Min-Rank	Graph $\mathcal{G}$	Digraph $\mathcal{D}$
1	$\mathcal{G}$ is complete (trivial)	$\mathcal{D}$ is complete (trivial)
2	$\mathcal{G}$ is not complete and $\overline{\mathcal{G}}$ is 2-colorable (Peeters, '96)	$\mathcal{D}$ is not complete and $\overline{\mathcal{D}}$ is fairly 3-colorable*
$n-2$	$\mathcal{G}$ (connected) has a maximum matching of size two and does not contain F as a subgraph*	unknown
$n-1$	$\mathcal{G}$ (connected) is a star graph	unknown
n	$\mathcal{G}$ has no edges (trivial)	$\mathcal{D}$ has no circuits (from Bar-Yossef *et al.*, '06)

Theorem

Suppose $\mathcal{G}$ is a connected graph of order $n \geq 6$. Then $\mathrm{minrk_q}(\mathcal{G}) = n - 2$ *iff $\mathcal{G}$ has a maximum matching of size two and does not contain a subgraph isomorphic to the graph depicted below*



The forbidden subgraph F

Graphs and Digraphs of Near-Extreme MinRanks

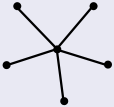
Summary

Min-Rank	Graph $\mathcal{G}$	Digraph $\mathcal{D}$
1	$\mathcal{G}$ is complete (trivial)	$\mathcal{D}$ is complete (trivial)
2	$\mathcal{G}$ is not complete and $\overline{\mathcal{G}}$ is 2-colorable (Peeters, '96)	$\mathcal{D}$ is not complete and $\overline{\mathcal{D}}$ is fairly 3-colorable*
$n-2$	$\mathcal{G}$ (connected) has a maximum matching of size two and does not contain F as a subgraph*	unknown
$n-1$	$\mathcal{G}$ (connected) is a star graph	unknown
n	$\mathcal{G}$ has no edges (trivial)	$\mathcal{D}$ has no circuits (from Bar-Yossef *et al.*, '06)

Digraphs of Extreme MinRanks

Theorem

Let $\mathcal{G}$ be a connected graph of order $n \geq 2$. Then $\mathrm{minrk_q}(\mathcal{G}) = n - 1$ if and only if $\mathcal{G}$ is a star graph.



Open Problems

Determine the hardness of the problem of deciding whether $\mathrm{minrk_q}(\mathcal{D}) = 2$ for $q > 2$?

Characterize families of graphs of order n with minrank $n - k$, for a constant $k > 2$