Index Coding, Caching & Distributed Storage

Salim El Rouayheb
ECE IIT, Chicago

Mohammad A. Maddah-Ali
Bell Labs, Alcatel-Lucent
Big Data vs. Wireless

Exabytes per month

Cisco forecast

$\infty \equiv 5 \text{ GB (mod at&t)}$

ATT Free Msg: Your data usage on your 4G LTE smartphn is near 5GB this month. Exceeding 5GB during this or future billing cycles will result in reduced data speeds, though you will still be able to email & surf the web. Wi-Fi helps you avoid reduced speeds. Visit www.att.com/datainfo or call 866-344-7584 for more info.
Meanwhile, Storage is Getting Very Cheap

Storage cost per GB (USD)

http://www.mkomo.com/cost-per-gigabyte-update
How Can Storage Help?

- Index coding: [Birk & Kol ’98] + ...
- COPE: [Katti et al. ‘06] + ...
- Caching: [Maddah-Ali & Niesen ’13] + ...
- Femto-caching: [Golrezai et al. ’12] + ...
- ...

Content is cached (stored) on mobile devices during off-peak hours
Index Coding Example

[Birk & Kol ’98]

<table>
<thead>
<tr>
<th>Transmission #</th>
<th>Index code 1</th>
<th>Index code 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$X_1$</td>
<td>$X_1 + X_2$</td>
</tr>
<tr>
<td>2</td>
<td>$X_2$</td>
<td>$X_3$</td>
</tr>
<tr>
<td>3</td>
<td>$X_3$</td>
<td>$X_4$</td>
</tr>
<tr>
<td>4</td>
<td>$X_4$</td>
<td></td>
</tr>
</tbody>
</table>

Can we do better?

- Cached data is NOT a design parameter now
- Even randomly cached data independent of the demands can help
Connections to Other Fields

- Network Coding
- Index Coding
- Interference Management
- Distributed Storage
- Graph Theory

- [R., Sprintson, Georghiades ‘08]
- [Effros, R., Langberg ‘13]
- [Bar-Yossef et al. ‘06]
- [Alon et al. ‘08]
- [Kim et al. ‘13]
- [Dimakis et al. ‘13]
- [Maleki, Cadambe, Jafar ‘12]
- [Shanmugan, Dimakis ‘13]
- [Mazumdar ‘13]
Theorem: [R,Sprintson, Georghiades’08] [Effros,R,Langberg ISIT’13]

Given any network coding problem, one can construct an index coding problem and an integer $L$ such that given any network code, one can efficiently construct a index code of length $L$, and vice versa.
**Index Coding & Graph Coloring**

Side info graph $G_d$

Independence nbr

$$\alpha(G_d) \leq c(G_d) \leq L_{min} \leq \chi_f(\overline{G}) \leq \chi(\overline{G})$$

Shannon capacity [Haemers ‘79]

Fractional Chromatic nbr [Blasiak, Kleinberg, Lubetzky ‘11]

Fractional local chrom. nbr [Shanmugam, Dimakis, Langberg ‘13]
Index Coding & Rank Minimization

- Linear case: \( L_{min}^* = \min_r k(M) \) [Bar-Yossef et al. '06]

- Computing \( L_{min}^* \) is NP hard. [Rouayheb. et al. '07] [Peeters '96]

- Methods for constructing index codes from rank minimization [Esfahanizadeh, F. Lahouti, and B. Hassibi '14] [Huang, R '15]
Index Coding & Rank Minimization

<table>
<thead>
<tr>
<th>Wants: $X_1$</th>
<th>Has: $X_2 X_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_1$</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wants: $X_4$</th>
<th>Has: $X_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_4$</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wants: $X_2$</th>
<th>Has: $X_1 X_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_2$</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wants: $X_3$</th>
<th>Has: $X_2 X_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_3$</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wants: $X_4$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_4$</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$X_3$</th>
<th>$X_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_1$</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$t_2$</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>$t_3$</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$t_4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Matrix $M$

- **Linear case:** $L_{min}^* = \min rk(M)$ [Bar-Yossef et al. '06]
- **Computing** $L_{min}^*$ is NP hard. [Rouayheb. et al. '07] [Peeters '96]
- Methods for constructing index codes from rank minimization [Esfahanizadeh, F. Lahouti, and B. Hassibi '14] [Huang, R '15]
Index Coding & Rank Minimization

- Linear case: $L^*_{min} = \min rk(M)$ [Bar-Yossef et al. '06]
- Computing $L^*_{min}$ is NP hard. [Rouayheb. et al. '07] [Peeters '96]
- Methods for constructing index codes from rank minimization [Esfahanizadeh, F. Lahouti, and B. Hassibi '14] [Huang, R '15]
Performance of different methods

Prob of caching a packet

n=100 users

[Huang, R. '15]
Index Coding on Erdős-Rényi Graphs

\[ \alpha(G) \leq L_{min} \leq \chi(\overline{G}) \]

- When \( n \to \infty \), we have with prob 1

\[ \log n \leq L_{min} \leq \frac{n}{\log n} \]

- Can improve the lower bound [Haviv & Langberg ‘11]

\[ c\sqrt{n} \leq L_{min} \leq \frac{n}{\log n} \]

- Coloring is the best upper bound we know on random graphs. Is it tight? OPEN
Caching Networks

Maddah-Ali, Niesen, 2012

Placement phase: Populate caches (prefetching), Design the caches
  • Demands are not known yet

Delivery phase: Reveal request, deliver content
  • Minimize the rate in the bottleneck

\[ R(M) = K \left( 1 - \frac{M}{N} \right) \frac{1}{1 + KM/N} \]

Gain from entire caches \((KM)\), Even-though caches are isolated
Caching Networks

Gain $\sim O(K)$

number of users

local cache size

uncoded scheme
coded scheme
Caching Networks

Topologies

Scenarios

- Non-Uniform demands
- Online Cache updating
- Delay-Limited Delivery
- Complexity-Limited (Alex Plenary Talk)
- Improving Approximation (Ravi’s Talk)
- Secure Delivery

Research Teams:
Dimakis, Molisch, Caire, Shanmugam, Golrezaei, [2012-15]
Avestimehr. Naderializadeh,[2014]
Sengupta, Tandon and Clancy [2014-15]
Shanmugam, Ji, Tulino, Llorca, Dimakis [2014-2015]
Diggavi, Hachem, Karamchandani, Maddah, Niesen [2013-2015]
Tian [2015]

Tutorials on Cache Networks: An Information Theoretic View, ISIT 2015
QUESTIONS?