Private Information Retrieval with Side Information

Swanand Kadhe, Brenden Garcia, Anoosheh Heidarzadeh, and Alex Sprintson – Texas A&M University
Salim El Rouayheb – Rutgers University

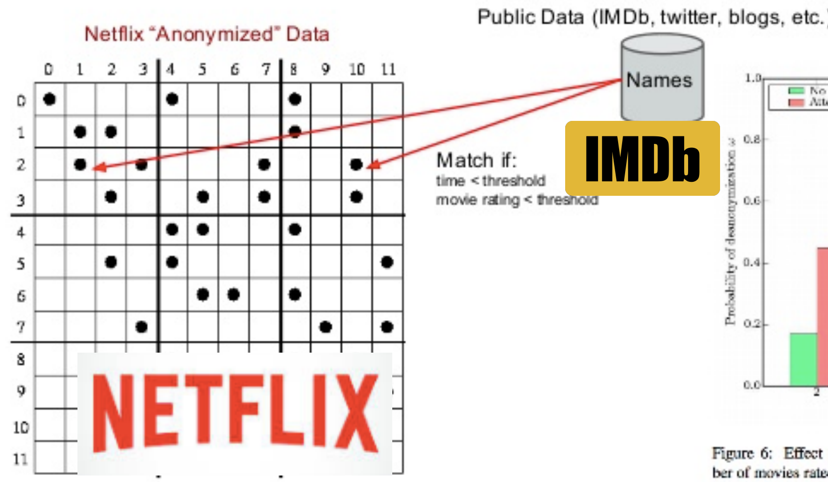
Side Info + Security

Usually

Side “Something” + Security = Bad News

Side-channel attacks

side information and de-anonymization



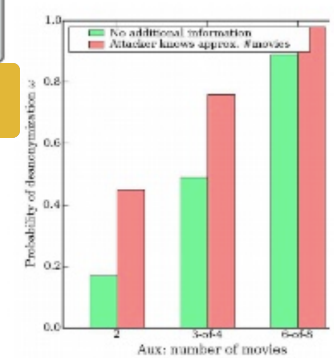


Figure 6: Effect of knowing approximate number of movies rated by victim (±50%.) Adversary knows approximate ratings(±1) and dates (14-day error).

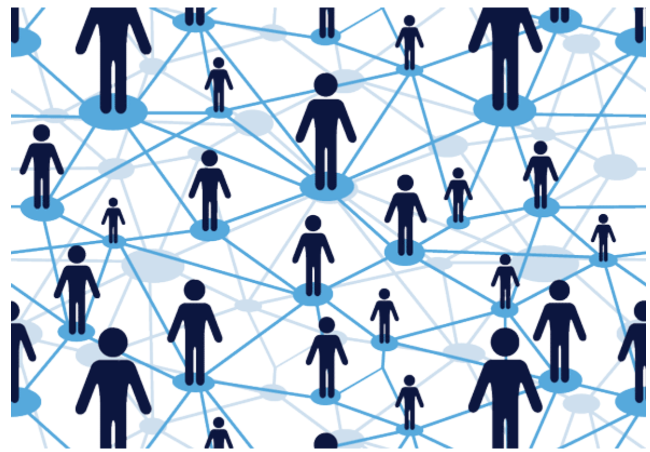
Today

Side Information + PIR = Good News

Why PIR with Side Information?

Because side information is everywhere nowadays

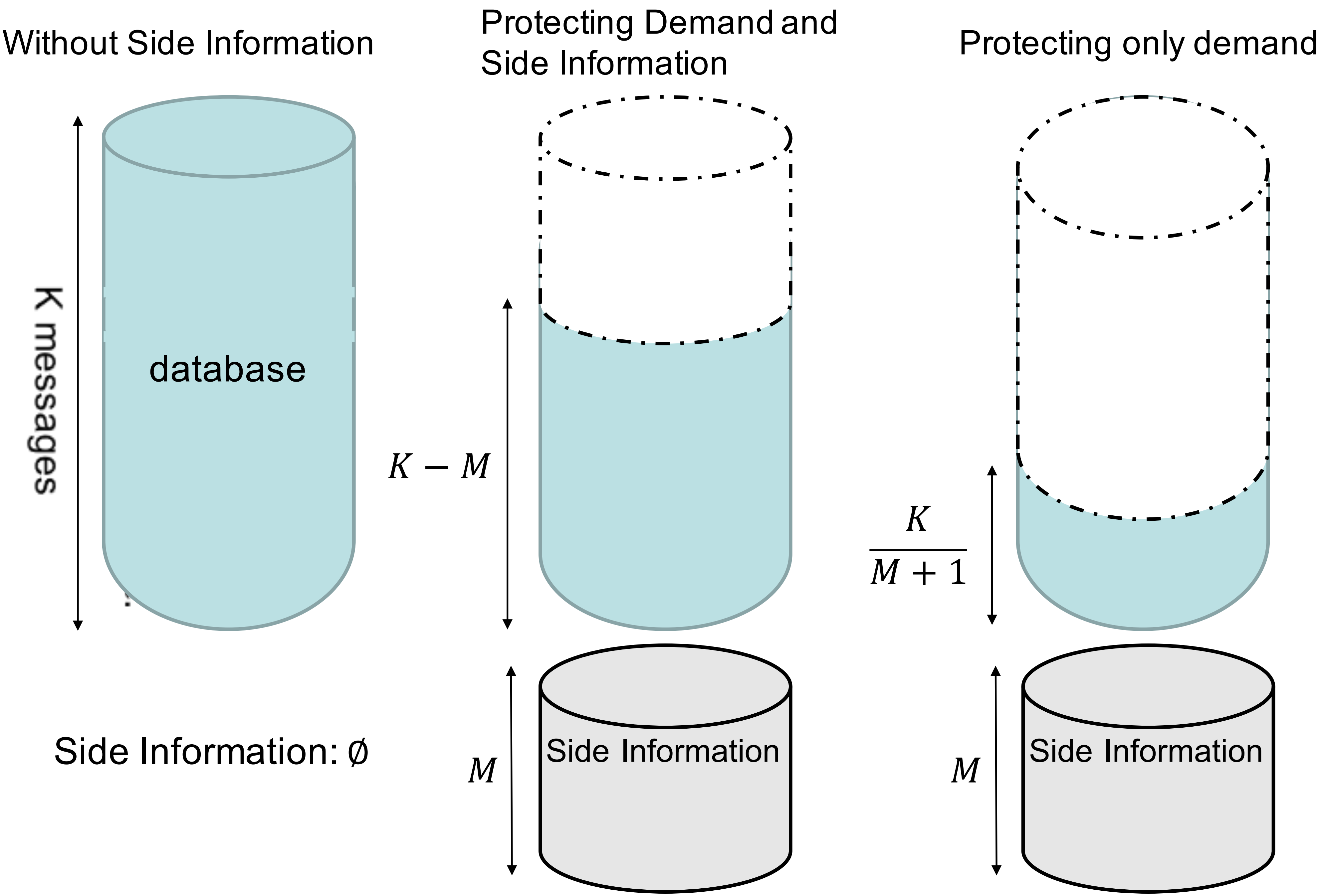
Overheaing in a
wireless channel



Communication
with peers

Previous PIR
Sessions*

TAKEAWAY MESSAGE



Related Work – PIR with Side Information

Tandon '17:

N servers, K messages of L bits, SL bits in side information

Side Information:

- Function of replicated messages
- Known to the server

Wei-Banawan-Ulukus '17

N servers, K messages, rK bits in side information

Side Information:

- Unknown to server
- Random fraction r of bits from each message

Piecewise linear inner and outer bounds are found for low and high caching ratios (r)

Chen-Wang-Jafar '17:

Model is similar to ours. Results discussed later

Holy Grail: Single Server : Example

X_1, X_2, X_3, X_4

Without any side information how can the user download X_1 privately?

User downloads everything from the server.

X_1 X_2 X_3 X_4

$W = 1$ *demand index*

Holy Grail: Single Server : Example

X_1, X_2, X_3, X_4

Without any side information how can the user download X_1 privately?

User downloads everything from the server.

With One Packet of Side Information how can the user download X_1 privately?

$$\frac{W = 1}{S = \{2\}.} \quad \begin{matrix} \textit{demand index} \\ \textit{side info index} \end{matrix}$$

Holy Grail: Single Server : Example

X_1, X_2, X_3, X_4

Without any side information how can the user download X_1 privately?

User downloads everything from the server.

$X_1 + X_2 + X_3 + X_4$

$X_1 + 2X_2 + 3X_3 + 4X_4$

$X_1 + 4X_2 + 4X_3 + X_4$

With One Packet of Side Information how can the user download X_1 privately?

Send $4 - 1$ random linear combinations

$W = 1$ *demand index*

$S = \{2\}.$ *side info index*

Can we do better?

Problem Model

N non-colluding servers

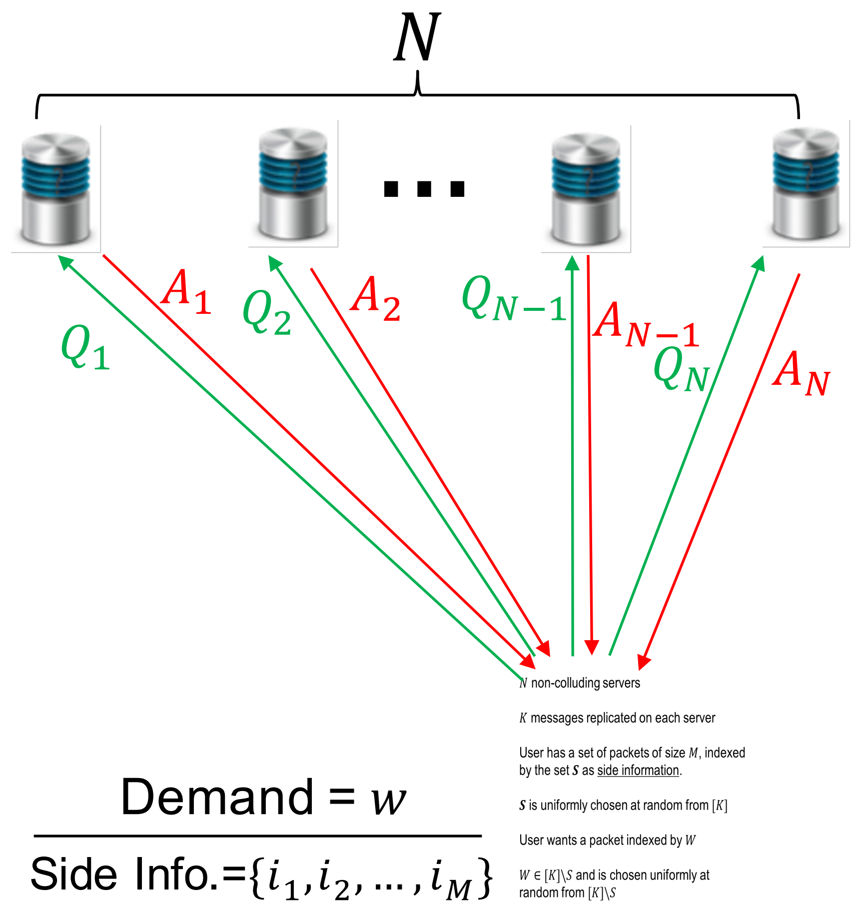
K messages replicated on each server

User has a set of packets of size M, indexed by the set $\boldsymbol{S}$ as side information.

$\boldsymbol{S}$ is uniformly chosen at random from $[K]$

User wants a packet indexed by W

$W \in [K] \backslash S$ and is chosen uniformly at random from $[K] \backslash S$



PROBLEM MODEL

N non-colluding servers

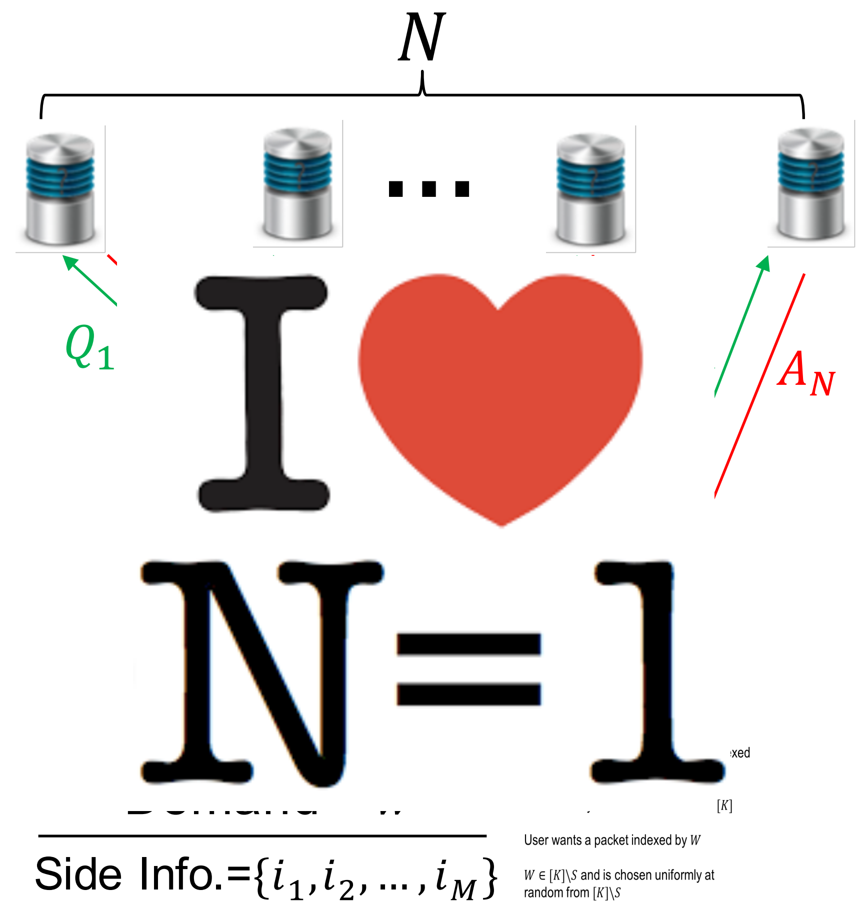
K messages replicated on each server

User has a set of packets of size M, indexed by the set $\boldsymbol{S}$ as <u>side information</u>.

$\boldsymbol{S}$ is uniformly chosen at random from $[K]$

User wants a packet indexed by W

$W \in [K]\backslash S$ and is chosen uniformly at random from $[K]\backslash S$



RESULTS

For K messages, and M packets as side information

Theorem 1 (Single Server): The following are the capacities when enforcing the respective privacy:

$$W\text{-privacy: } C_W = \left\lceil \frac{K}{M+1} \right\rceil^{-1}$$

$$(W,S)\text{-privacy: } C_{W,S} = (K-M)^{-1}$$

Theorem 2 (Multi-Server): The capacity is lower bounded, when enforcing W-privacy, by:

Conjecture

$$C_W \overset{=}{\geq} \left(1 + \frac{1}{N} + \frac{1}{N^2} + \cdots + \frac{1}{N^{\left(\frac{K}{M+1}-1\right)}}\right)^{-1}$$

(W,S) Privacy:

Or how to protect both your demand & side information

Single Server: (W,S)-Privacy General Achievability

$X_1, X_2, \ldots, X_K$

M-K random linear combinations over a high field
Or
MDS coded packets

In general the user can query the parity symbols of a $(2K+M,K)$ systematic MDS code from the server to decode.

Converse later.

Demand = w

Side Info.=$\{i_1, i_2, \ldots, i_M\}$

W Privacy:

Or how to protect your demand

Single Server: W-Privacy Achievability Example

$X_1, X_2, \ldots, X_9$

$\frac{W = 5}{S = \{1,8\}}$

Partition and Code Scheme:

Say K = 9 and M = 2.

1.) User Partitions $\{1,2, \ldots, 9\}$ into 3 equal sized sets

a.) One set is {1,5,8}

b.) The others are randomly chosen from remaining elements to be {2,6,9} and {3,4,7}.

2.) Client sends sets of partition to server in random order

Partition: $\{1,5,8\}\ \{2,6,9\}\{3,4,7\}$

Single Server: W-Privacy Achievability Example

$X_1, X_2, \ldots, X_9$

$\{1,5,8\}$

$$\frac{W = 5}{S = \{1,8\}}$$

<u>Partition and Code Scheme:</u>

Say K = 9 and M = 2.

1.) User Partitions $\{1,2,\ldots,9\}$ into 3 equal sized sets

a.) One set is {1,5,8}

b.) The others are randomly chosen from remaining elements to be {2,6,9} and {3,4,7}.

2.) Client sends sets of partition to server in random order

Partition: $\{2,6,9\}\{3,4,7\}$

Single Server: W-Privacy Achievability Example

$X_1, X_2, \ldots, X_9$

$\{2,6,9\}$ $\{1,5,8\}$ $\{3,4,7\}$

$$\frac{W = 5}{S = \{1,8\}}$$

Partition:

<u>Partition and Code Scheme:</u>

Say K = 9 and M = 2.

1.) User Partitions $\{1,2,\ldots,9\}$ into 3 equal sized sets

a.) One set is {1,5,8}

b.) The others are randomly chosen from remaining elements to be {2,6,9} and {3,4,7}.

2.) Client sends sets of partition to server in random order

Single Server: W-Privacy Achievability Example

$X_1, X_2, \ldots, X_K$

$\{2,6,9\}$ $\{1,5,8\}$ $\{3,4,7\}$

$$\frac{W = 5}{S = \{1,8\}}$$

<u>Partition and Code Scheme:</u>

Say K = 9 and M = 2.

3.) Client sends user the 3 packets corresponding to the bitwise XOR of the partition.

Single Server: W-Privacy Achievability Example

$X_1, X_2, \ldots, X_K$

$X_2 + X_6 + X_9 \quad X_1 + X_5 + X_8 \quad X_3 + X_4 + X_7$

<ins>Partition and Code Scheme:</ins>

Say K = 9 and M = 2.

3.) Client sends user the 3 packets corresponding to the bitwise XOR of the partition.

$$\frac{W = 5}{S = \{1,8\}}$$

Single Server: W-Privacy Achievability Example

$X_1, X_2, \ldots, X_K$

Partition and Code Scheme:

Say K = 9 and M = 2.

3.) Client sends user the 3 packets corresponding to the bitwise XOR of the partition.

$X_2 + X_6 + X_9$ $X_1 + X_5 + X_8$ $X_3 + X_4 + X_7$

$$\frac{W = 5}{S = \{1,8\}}$$

Single Server: W-Privacy Achievability Example

$X_1, X_2, \ldots, X_K$

<u>Partition and Code Scheme:</u>

Say K = 9 and M = 2.

3.) Client sends user the 3 packets corresponding to the bitwise XOR of the partition.

$\frac{W = 5}{S = \{1,8\}}$

$X_2 + X_6 + X_9 \quad \boxed{X_1 + X_5 + X_8} \quad X_3 + X_4 + X_7$

User Decodes with this packet

Single Server: W-Privacy Achievability Scheme

$X_1, X_2, \ldots, X_K$

Demand $W = w$

Side Info.$(S)=\{i_1, i_2, \ldots, i_M\}$

<u>Partition and Code Scheme:</u>

Assume $(M+1)|K$.

1.) User Partitions $\{1,2,\ldots,K\}$ into $\frac{K}{M+1}$ equal sized sets:

a.)One set is $\{w\} \cup S$

b.) The others are randomly chosen from remaining elements

2.) User sends sets of partition to server in random order

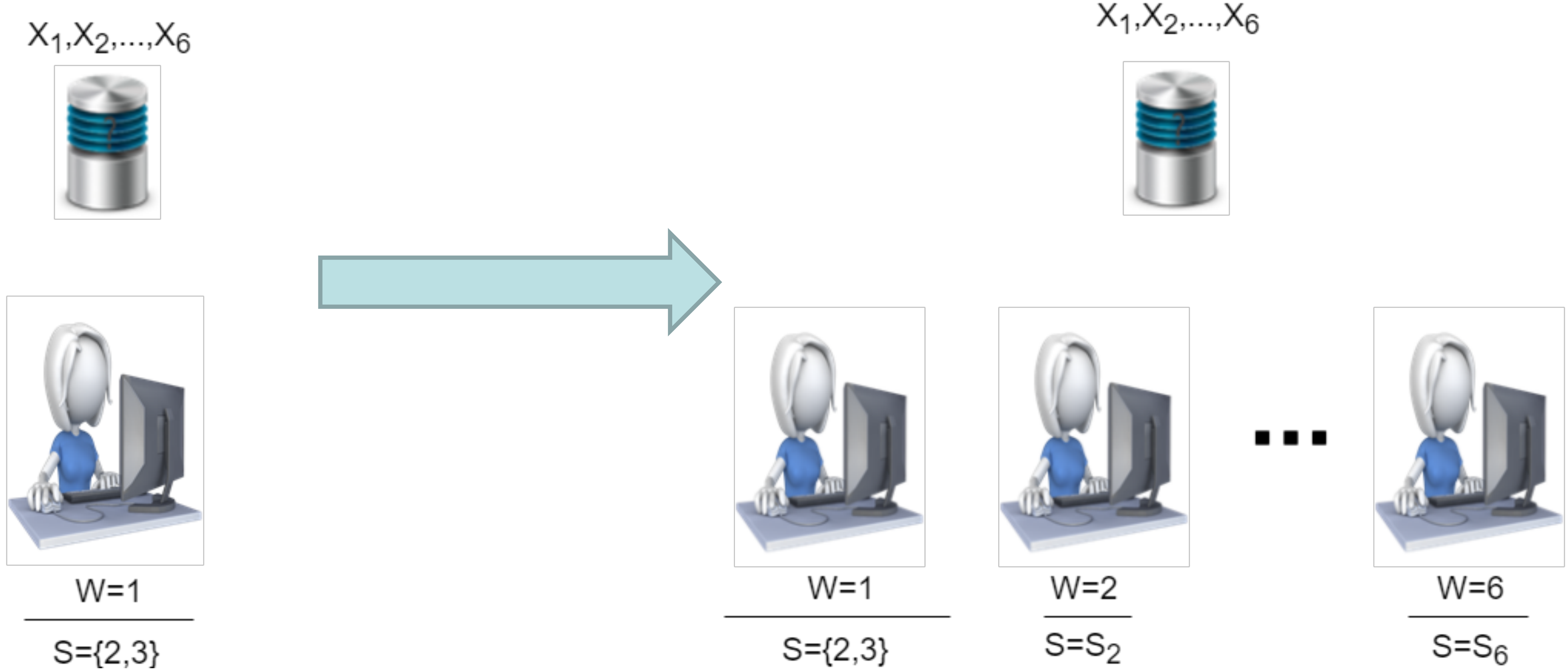
3.) The server sends back the bitwise XOR of the packets indexed by each set in the partition.

Single Server: W-Privacy Converse

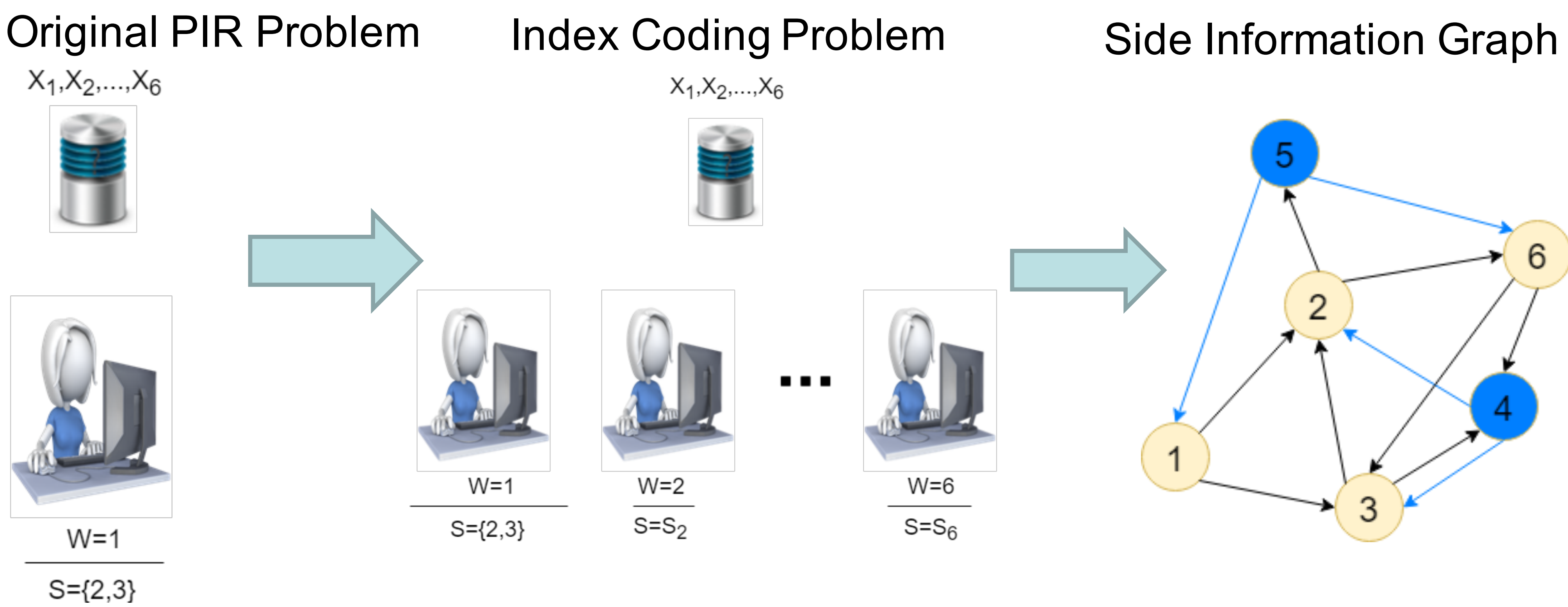
Proposition (Necessary Condition for W-Privacy):
A solution to the PIR problem for W-Privacy must have the property that $\forall i \in [K], \exists S_i \subseteq [K] \setminus \{i\}, |S_i| = M$, such that X_i can be decoded with the response from the server and S_i.

Original PIR Problem

Index Coding Problem



Single Server: W-Privacy Converse



Length of Index Coding Solution lower bounded by size of Maximum Acyclic Induced Subgraph (MAIS) of side information graph.

Lemma :

Let G be an M-regular simple directed graph with K nodes. Then

$$|MAIS(G)| \geq \left\lceil \frac{K}{M+1} \right\rceil$$

Multi Server Results

To establish capacity bound:

$$C_W \geq \left(1 + \frac{1}{N} + \frac{1}{N^2} + \cdots + \frac{1}{N^{\left(\frac{K}{M+1}-1\right)}}\right)^{-1}$$

User partitions packets and sends partition to server; server forms super-messages.

The user and servers then run the classical PIR scheme

[Sun-Jafar '16] on the super messages, where the user wants the packet corresponding to the set $\{w\} \cup S$ in the partition.

Partition:
$\{\{1,5,8\},\{2,6,9\},\{3,4,7\}\}$

Super Messages:

$$\hat{X}_1 = X_1 + X_5 + X_8$$
$$\hat{X}_2 = X_2 + X_6 + X_9$$
$$\hat{X}_3 = X_3 + X_4 + X_7$$

Server 1	**Server 2**
$\hat{X}_{1,1}, \hat{X}_{2,1}, \hat{X}_{3,1}$	$\hat{X}_{1,2}, \hat{X}_{2,2}, \hat{X}_{3,2}$
$\hat{X}_{1,3} + \hat{X}_{2,2}$	$\hat{X}_{1,5} + \hat{X}_{2,1}$
$\hat{X}_{1,4} + \hat{X}_{3,2}$	$\hat{X}_{1,6} + \hat{X}_{3,1}$
$\hat{X}_{2,3} + \hat{X}_{3,3}$	$\hat{X}_{2,4} + \hat{X}_{3,4}$
$\hat{X}_{1,7} + \hat{X}_{2,4} + \hat{X}_{3,4}$	$\hat{X}_{1,8} + \hat{X}_{2,3} + \hat{X}_{3,3}$

Table: Packets sent to user from servers

Recently the capacity for (W, S)-privacy in this scenario was found to be:

$$\left(1 + \frac{1}{N} + \cdots + \frac{1}{N^{K-M-1}}\right)^{-1}$$ [Chen-Wang-Jafar '17]

Summary

- Introduced PIR with Side information
- Gains even in the case of a single server
- Two types of privacy: W privacy (demand) and (W,S) privacy (demand and side info)
- Achievability: Partition and Code
- Combinatorial converse proof based on index codes and directed graph
- Open problems
- PIR capacity: Multi-server W-privacy. We presented an achievability scheme
- Many possible formulations
  - Coded vs. uncoded side inform
  - How much the servers know about the side info
  - ….