

NIST WPEC 2024

Circuit-PSI & Applications

Kyoohyung Han¹, **Seongkwang Kim¹**,
Byeonghak Lee¹, Yongha Son²

¹ Samsung SDS

² Sungshin Women's University

Sep. 24. 2024



I. Circuit-PSI

Circuit-PSI

Each parties may not want to reveal the intersection itself

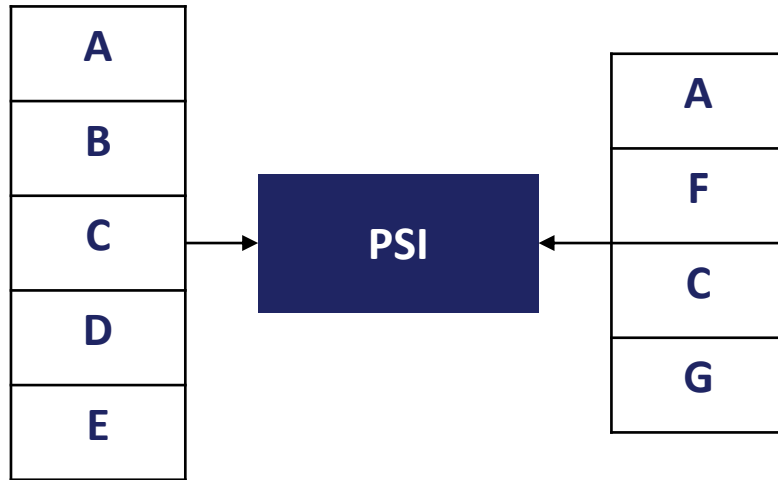
Basic PSI

Circuit-PSI

Circuit-PSI

Each parties may not want to reveal the intersection itself

Basic PSI

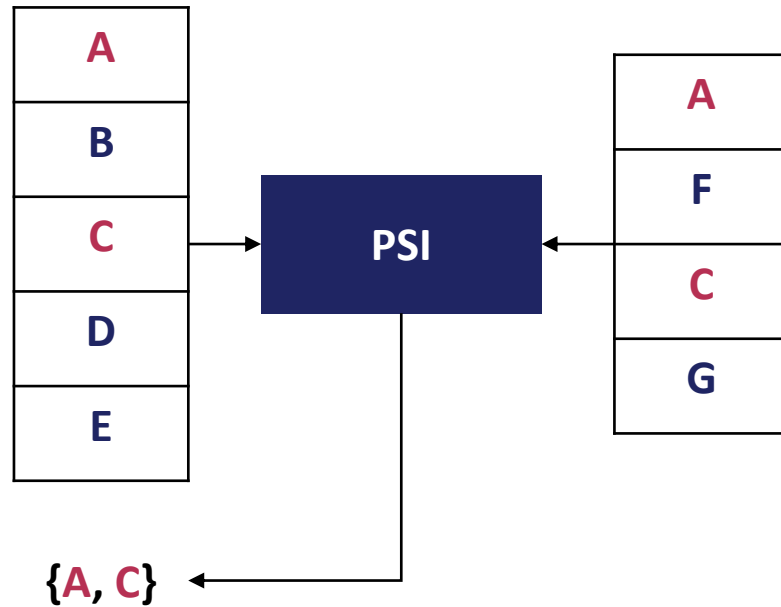


Circuit-PSI

Circuit-PSI

Each parties may not want to reveal the intersection itself

Basic PSI

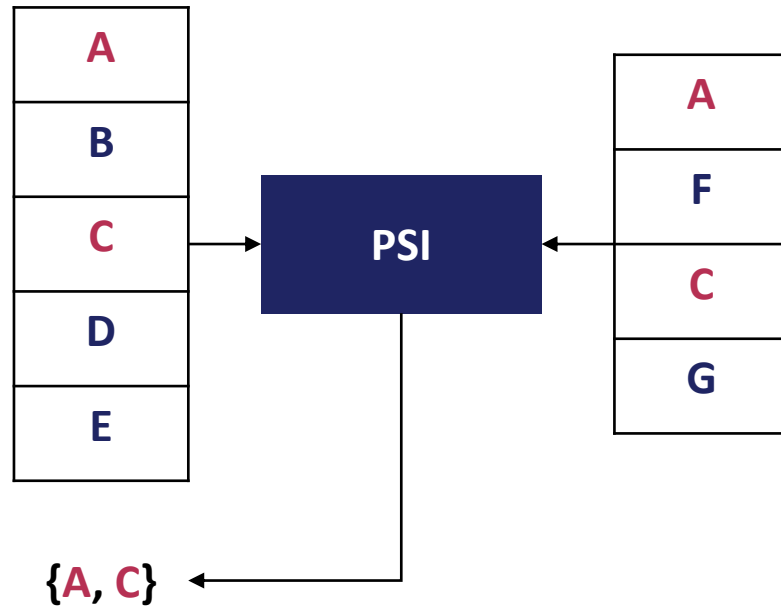


Circuit-PSI

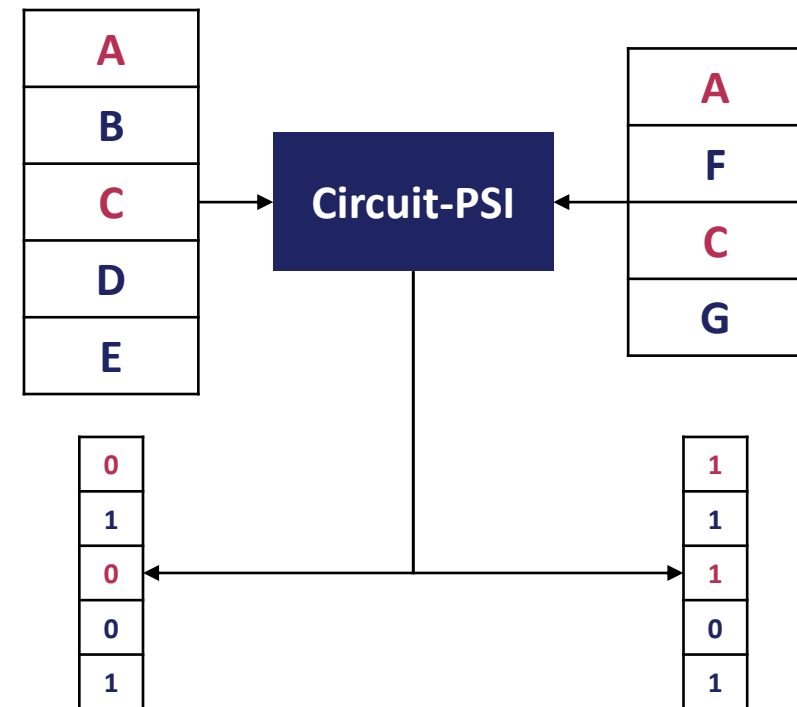
Circuit-PSI

Each parties may not want to reveal the intersection itself

Basic PSI



Circuit-PSI

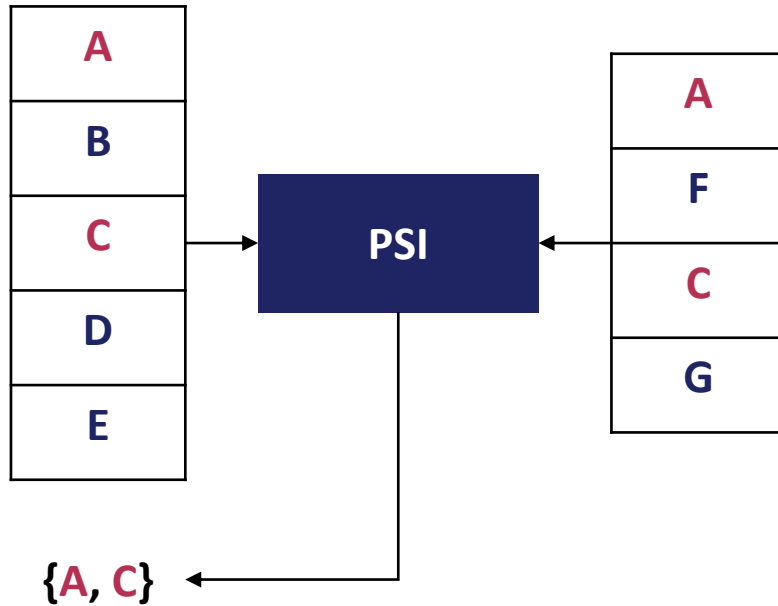


- The XOR sum of two result vector is whether each element is in the other party's set or not

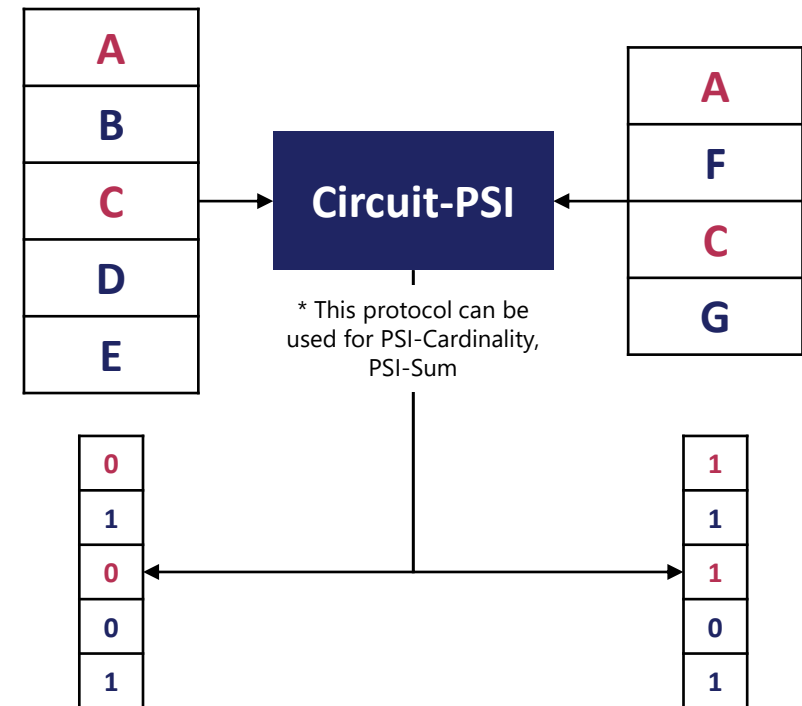
Circuit-PSI

Each parties may not want to reveal the intersection itself

Basic PSI



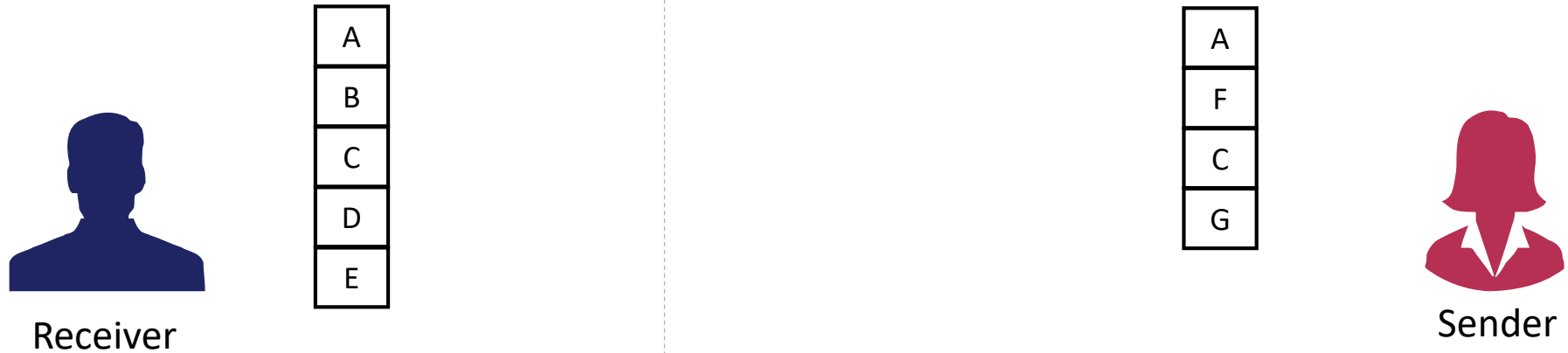
Circuit-PSI



- The XOR sum of two result vector is whether each element is in the other party's set or not

Brief Mechanism of Circuit-PSI

The circuit-PSI protocol consists of Cuckoo hashing, Oblivious PRF



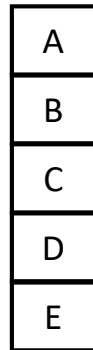
Brief Mechanism of Circuit-PSI #1

The circuit-PSI protocol consists of Cuckoo hashing, Oblivious PRF

Step1. Cuckoo Hashing



Receiver

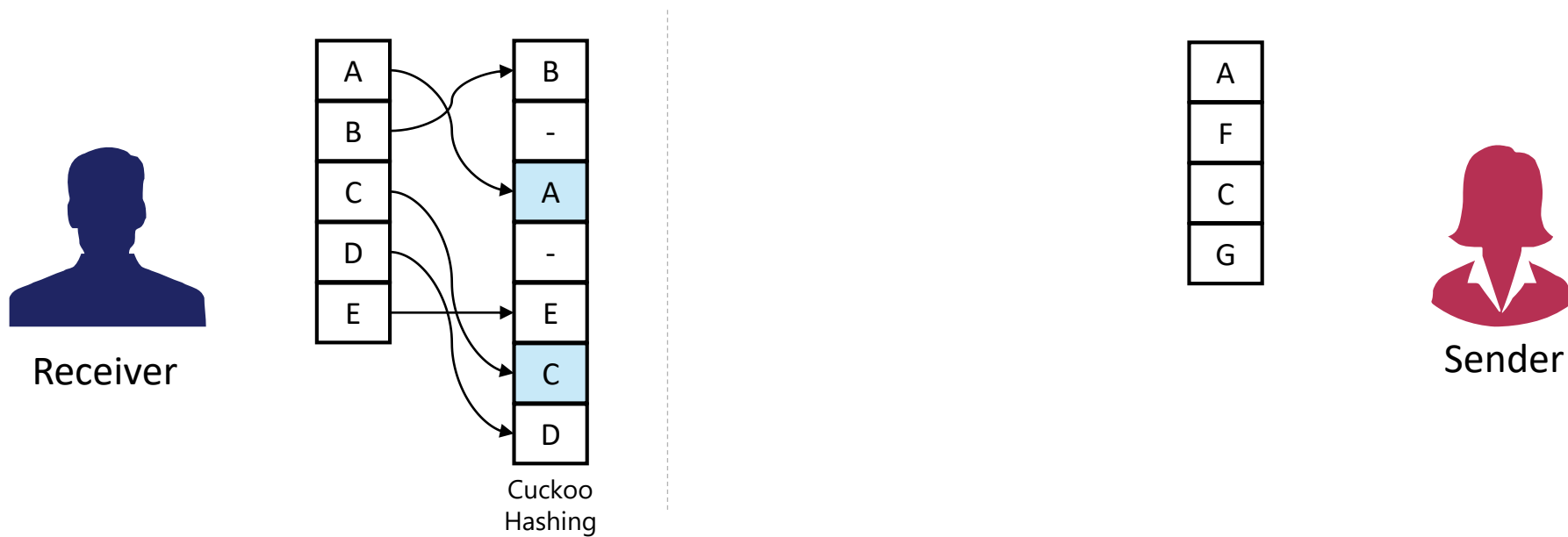


Sender

Brief Mechanism of Circuit-PSI #1

The circuit-PSI protocol consists of Cuckoo hashing, Oblivious PRF

Step1. Cuckoo Hashing



- The receiver compute Cuckoo hash table, and the sender compute Simple hash table

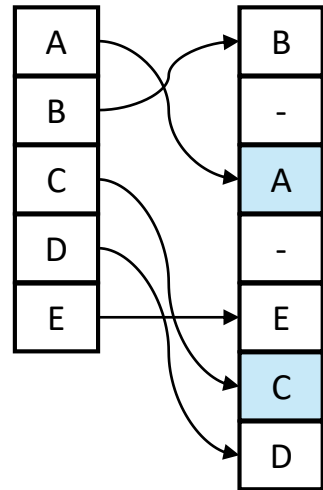
Brief Mechanism of Circuit-PSI #1

The circuit-PSI protocol consists of Cuckoo hashing, Oblivious PRF

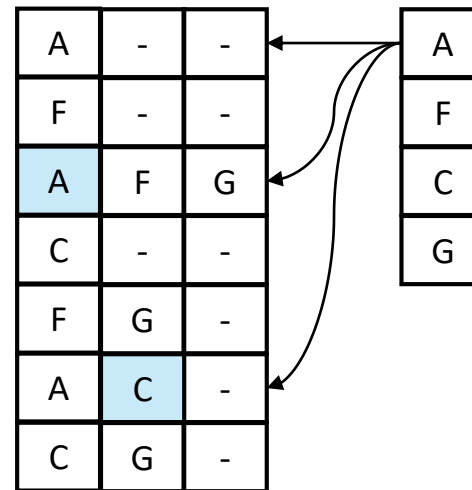
Step1. Cuckoo Hashing



Receiver



Cuckoo Hashing



Simple Hashing



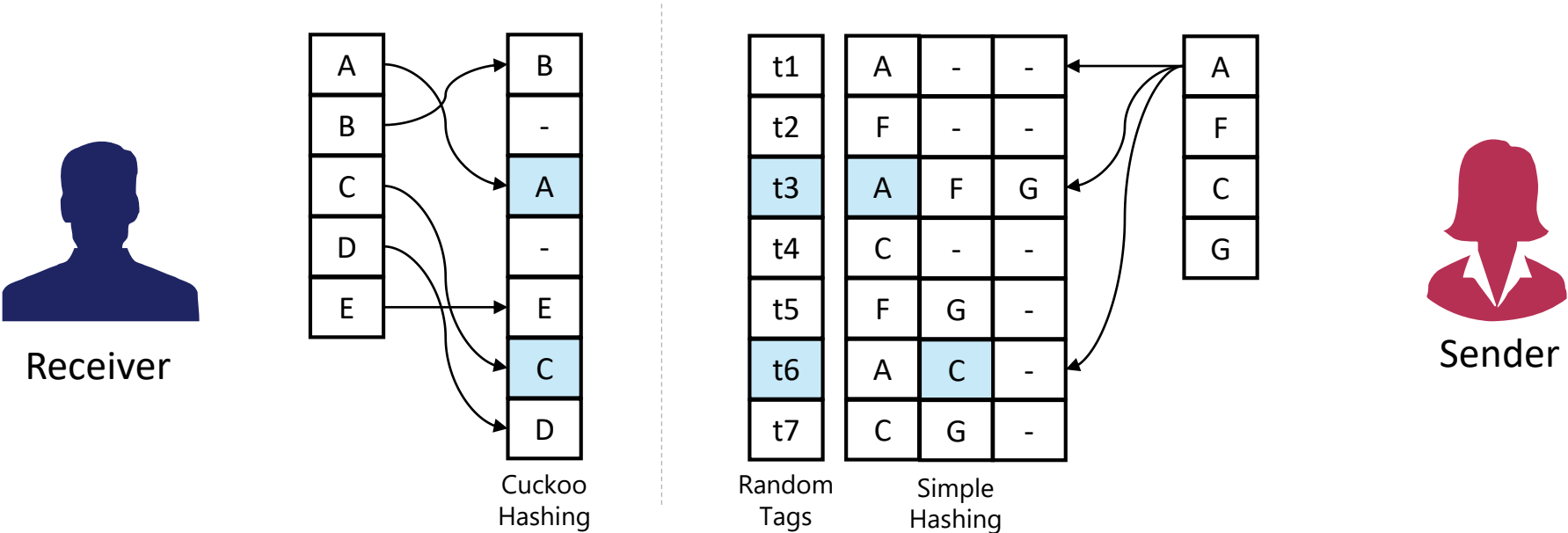
Sender

- The receiver compute Cuckoo hash table, and the sender compute Simple hash table

Brief Mechanism of Circuit-PSI #1

The circuit-PSI protocol consists of Cuckoo hashing, Oblivious PRF

Step 1. Cuckoo Hashing



- The receiver compute Cuckoo hash table, and the sender compute Simple hash table
- The sender picks a random tag for each bin of Cuckoo hash table

Brief Mechanism of Circuit-PSI #2

The circuit-PSI protocol consists of Cuckoo hashing, Oblivious PRF

Step2. Oblivious PRF



Receiver

A

t3

A F G

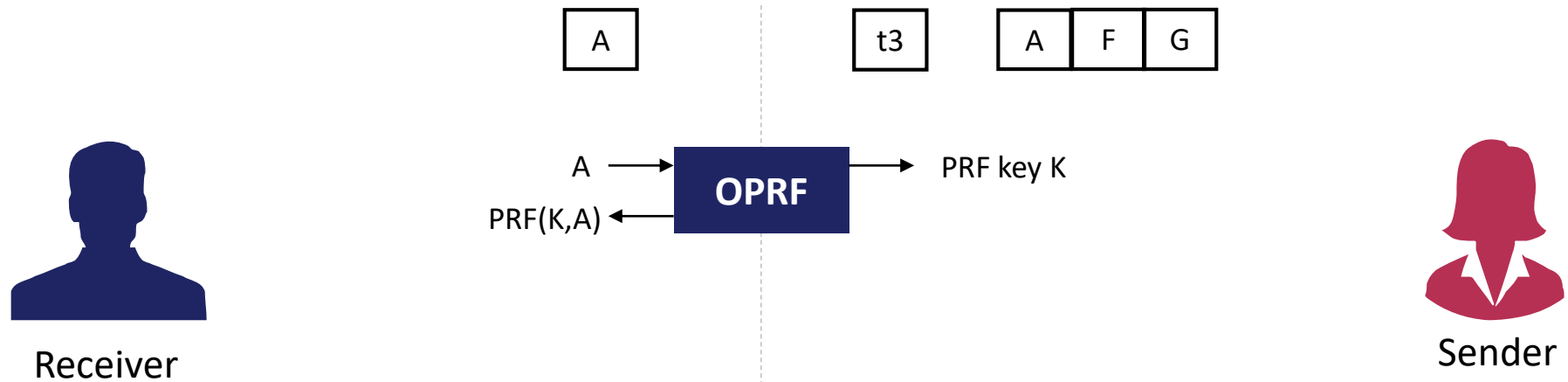


Sender

Brief Mechanism of Circuit-PSI #2

The circuit-PSI protocol consists of Cuckoo hashing, Oblivious PRF

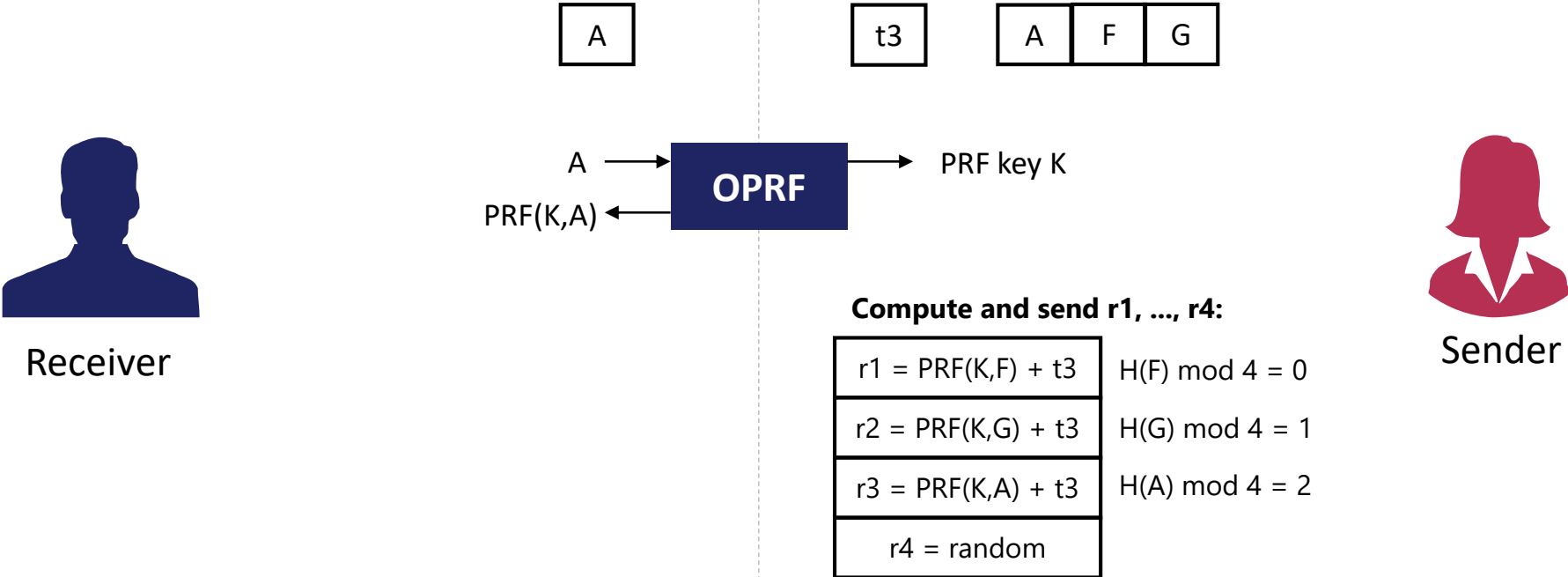
Step2. Oblivious PRF



Brief Mechanism of Circuit-PSI #2

The circuit-PSI protocol consists of Cuckoo hashing, Oblivious PRF

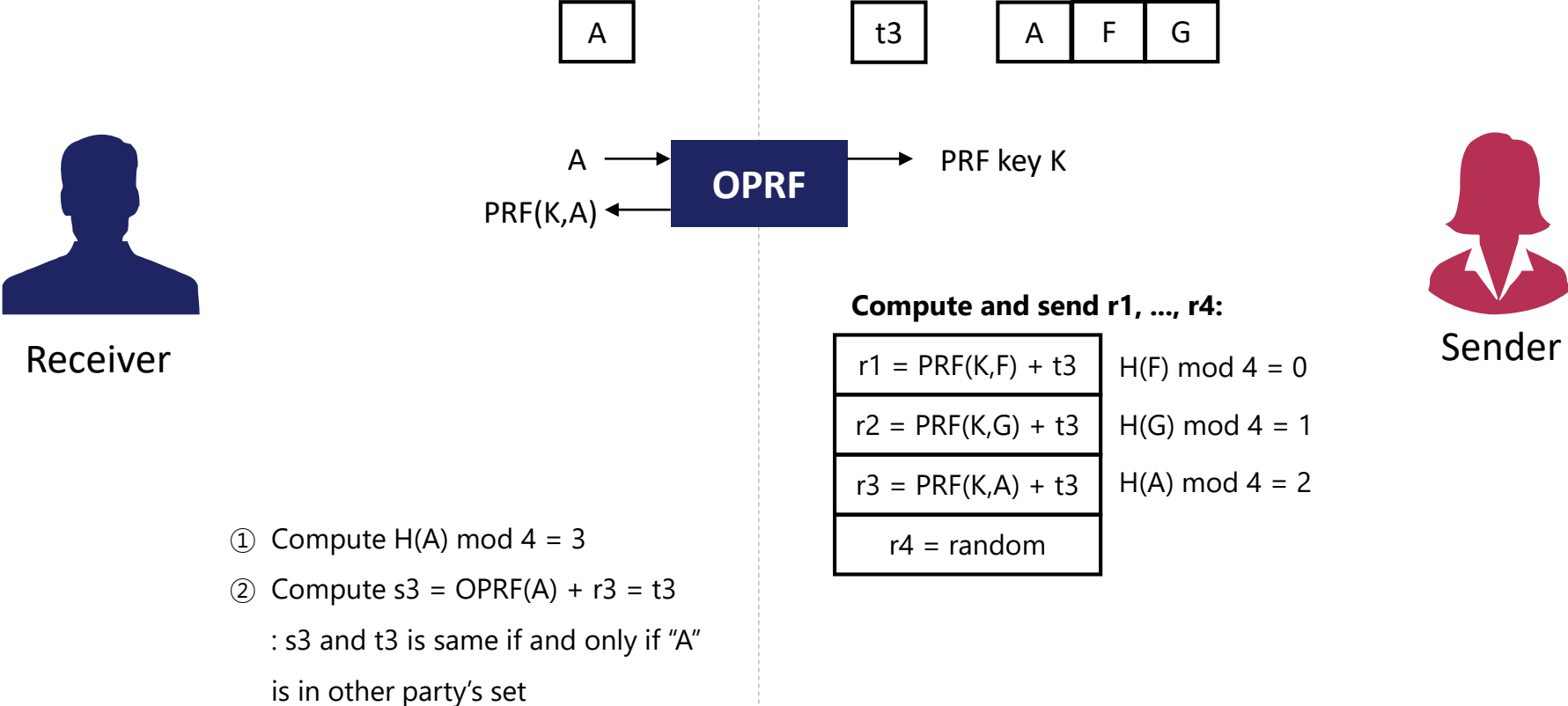
Step2. Oblivious PRF



Brief Mechanism of Circuit-PSI #2

The circuit-PSI protocol consists of Cuckoo hashing, Oblivious PRF

Step2. Oblivious PRF



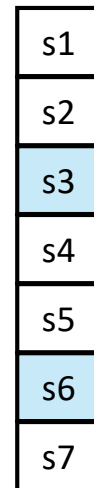
Brief Mechanism of Circuit-PSI #3

The circuit-PSI protocol consists of Cuckoo hashing, Oblivious PRF

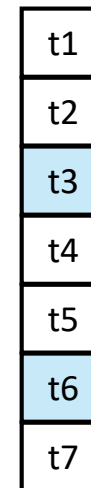
Step3. Private Equality Test



Receiver



Computed
Tags



Random
Tags

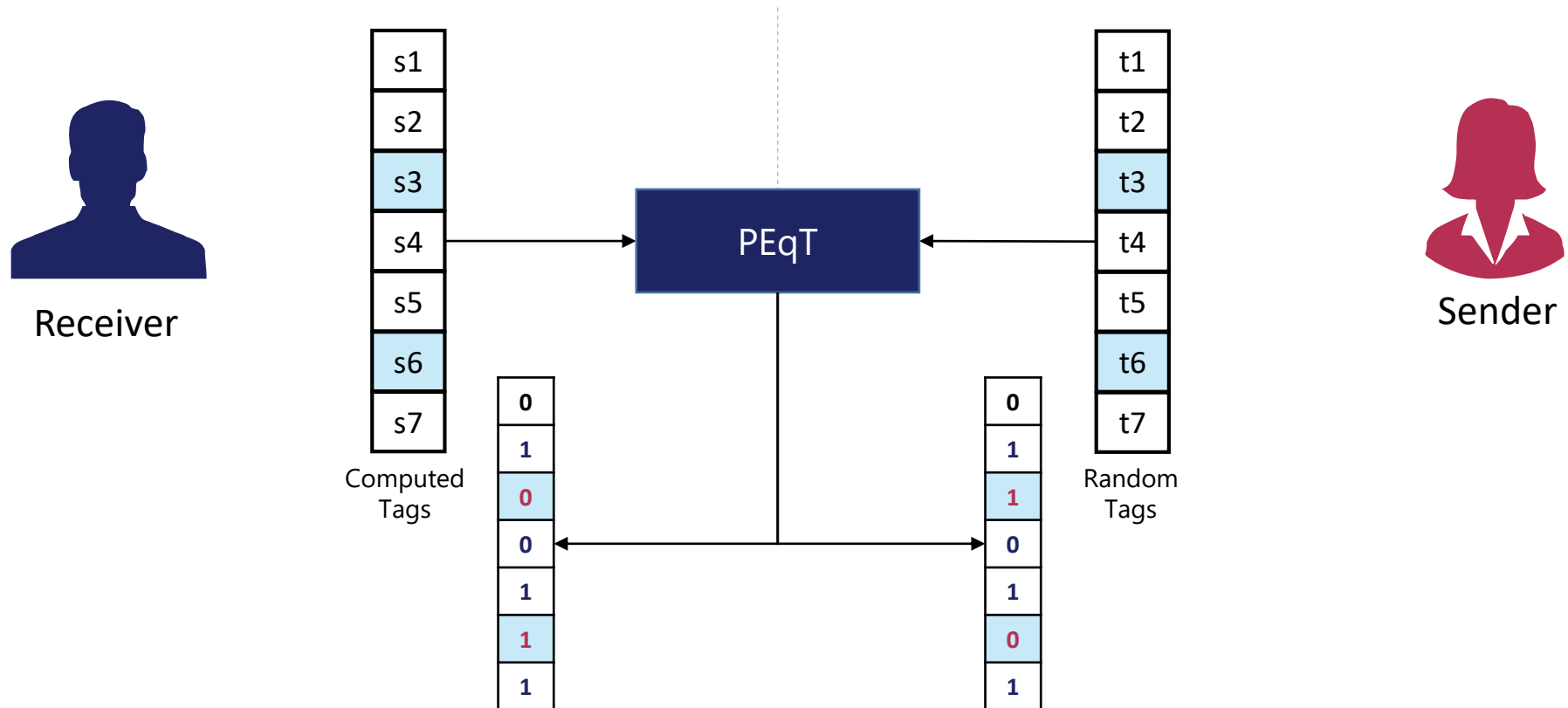


Sender

Brief Mechanism of Circuit-PSI #3

The circuit-PSI protocol consists of Cuckoo hashing, Oblivious PRF

Step3. Private Equality Test



Our Works on Circuit-PSI

HE-based optimizations, and security of building block

Our Works on Circuit-PSI

HE-based optimizations, and security of building block

CPSI with HE (SAC '22)

➤ Reduce communication cost during Step3: equality preserving compression

- Compute below equation using HE scheme

- For $x = \sum x_i B^i, y = \sum y_i B^i,$

$$r + \sum (x_i - y_i)^2 = r \text{ iff } x = y$$

* Ferret-OT is used for OT extension

Network	# of Items	Comm. (MB)	Time (s)
LAN	2^{16}	20.69	2.09
	2^{18}	83.64	5.71
	2^{20}	394.3	20.17

Table 1. Performance of Circuit-PSI

[HMS22] K. Han, D. Moon, and Y. Son. *Improved Circuit-PSI via Equality Preserving Compression*. SAC 2022.

Our Works on Circuit-PSI

HE-based optimizations, and security of building block

CPSI with HE (SAC '22)

➤ Reduce communication cost during Step3: equality preserving compression

- Compute below equation using HE scheme
- For $x = \sum x_i B^i, y = \sum y_i B^i,$
$$r + \sum (x_i - y_i)^2 = r \text{ iff } x = y$$

* Ferret-OT is used for OT extension

Network	# of Items	Comm. (MB)	Time (s)
LAN	2^{16}	20.69	2.09
	2^{18}	83.64	5.71
	2^{20}	394.3	20.17

Table 1. Performance of Circuit-PSI

[HMS22] K. Han, D. Moon, and Y. Son. *Improved Circuit-PSI via Equality Preserving Compression*. SAC 2022.

Unbalanced CPSI with HE (AsiaCCS '23)

➤ Convert HE-based PSI to CPSI

- Zero storage required for small set holder
- Recursive HE application for trade-off

[SJ23] Y. Son, and J. Jeong. *PSI with computation or Circuit-PSI for Unbalanced Sets from Homomorphic Encryption*. AsiaCCS 2023.

Our Works on Circuit-PSI

HE-based optimizations, and security of building block

CPSI with HE (SAC '22)

➤ Reduce communication cost during Step3: equality preserving compression

- Compute below equation using HE scheme
- For $x = \sum x_i B^i, y = \sum y_i B^i,$
$$r + \sum (x_i - y_i)^2 = r \text{ iff } x = y$$

* Ferret-OT is used for OT extension

Network	# of Items	Comm. (MB)	Time (s)
LAN	2^{16}	20.69	2.09
	2^{18}	83.64	5.71
	2^{20}	394.3	20.17

Table 1. Performance of Circuit-PSI

[HMS22] K. Han, D. Moon, and Y. Son. *Improved Circuit-PSI via Equality Preserving Compression*. SAC 2022.

Unbalanced CPSI with HE (AsiaCCS '23)

➤ Convert HE-based PSI to CPSI

- Zero storage required for small set holder
- Recursive HE application for trade-off

For Better Building Block (AC '24)

➤ Cryptanalysis on Malicious OPRF

- OPRF for PSI is generally OTe (or VOLE) + OKVS
- We can get at most double OPRF evaluations by overfitting OKVS
- It leads PSI to statistical distance larger than $2^{-\lambda}$

[SJ23] Y. Son, and J. Jeong. *PSI with computation or Circuit-PSI for Unbalanced Sets from Homomorphic Encryption*. AsiaCCS 2023.

[HKLS24] K. Han, S. Kim, B. Lee, and Y. Son. *Revisiting OKVS-based OPRF and PSI: Cryptanalysis and Better Construction*. Asiacrypt 2024, to appear.

II. Applications

1. Genomic Analysis

Safe Sharing of Genomic Data

Pathogen genomic data is generally considered to be non-identifying, but it can be combined with personal information

Scenario

Results

Safe Sharing of Genomic Data

Pathogen genomic data is generally considered to be non-identifying, but it can be combined with personal information

Scenario

Has Tet A	ID
Yes	Abc02
No	Def01
No	Abc01
Yes	Def02

ID	Homeless
Abc01	Yes
Abc02	Yes
Abc03	Yes
Abc04	No

Results

Safe Sharing of Genomic Data

Pathogen genomic data is generally considered to be non-identifying, but it can be combined with personal information

Scenario

Has Tet A	ID
Yes	Abc02
No	Def01
No	Abc01
Yes	Def02

Remain
Has Tet A = Yes

ID
Abc02
Def02

ID	Homeless
Abc01	Yes
Abc02	Yes
Abc03	Yes
Abc04	No

Remain
Homeless = Yes

ID
Abc01
Abc02
Abc03

Results

Safe Sharing of Genomic Data

Pathogen genomic data is generally considered to be non-identifying, but it can be combined with personal information

Scenario

Has Tet A	ID
Yes	Abc02
No	Def01
No	Abc01
Yes	Def02

Remain
Has Tet A = Yes

ID
Abc02
Def02

ID	Homeless
Abc01	Yes
Abc02	Yes
Abc03	Yes
Abc04	No

Remain
Homeless = Yes

ID
Abc01
Abc02
Abc03

PSI-
Cardinality

Tet A \cap Homeless = 1

→ 33% of homeless people
have Shingella with Tet A

Results

Safe Sharing of Genomic Data

Pathogen genomic data is generally considered to be non-identifying, but it can be combined with personal information

Scenario

Has Tet A	ID
Yes	Abc02
No	Def01
No	Abc01
Yes	Def02

Remain
Has Tat A = Yes

ID
Abc02
Def02

ID	Homeless
Abc01	Yes
Abc02	Yes
Abc03	Yes
Abc04	No

Remain
Homeless = Yes

ID
Abc01
Abc02
Abc03

PSI-
Cardinality

Tet A \cap Homeless = 1

→ 33% of homeless people
have Shingella with Tet A

Results

➤ Outputs

- 76.92% of homeless people and 0% of non-homeless people have Shingella bacteria with Tet A

➤ Timing Result

- For 30 rows, it only takes 1 second.
- We estimate that it takes less than a minute for 1M rows

Safe Sharing of Genomic Data from Discovery to Broad Data Sharing
Leveraging the Power of Beacon and Privacy Enhancing Technology (PET),
Galaxy Community Conference 2023 (w/ Soyeon Kim from SFU)

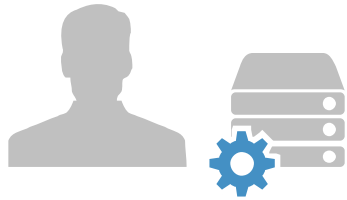
2. Data Aggregation

Data Aggregation System in Korea

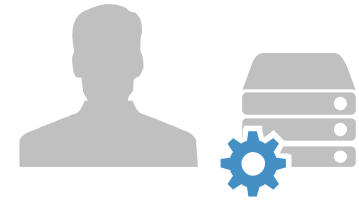
Highly sensitive identifier only goes to trusted 3rd party (government agency)



TTP



Party A



Party B



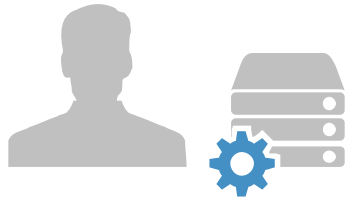
Data aggregation center
(DAC)

Data Aggregation System in Korea

Highly sensitive identifier only goes to trusted 3rd party (government agency)



TTP

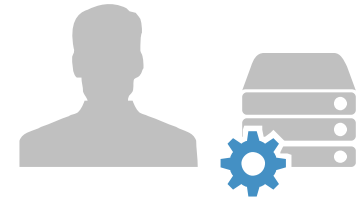


Party A

Data includes:

IDs, indices, features

Index	ID (phone)	ZIP	DOB	...
A01	010-1234-5678	12345	1978-SEP-21	...
A02	010-2345-6789	56789	1996-JUN-05	...



Party B

Data includes:

IDs, indices, features

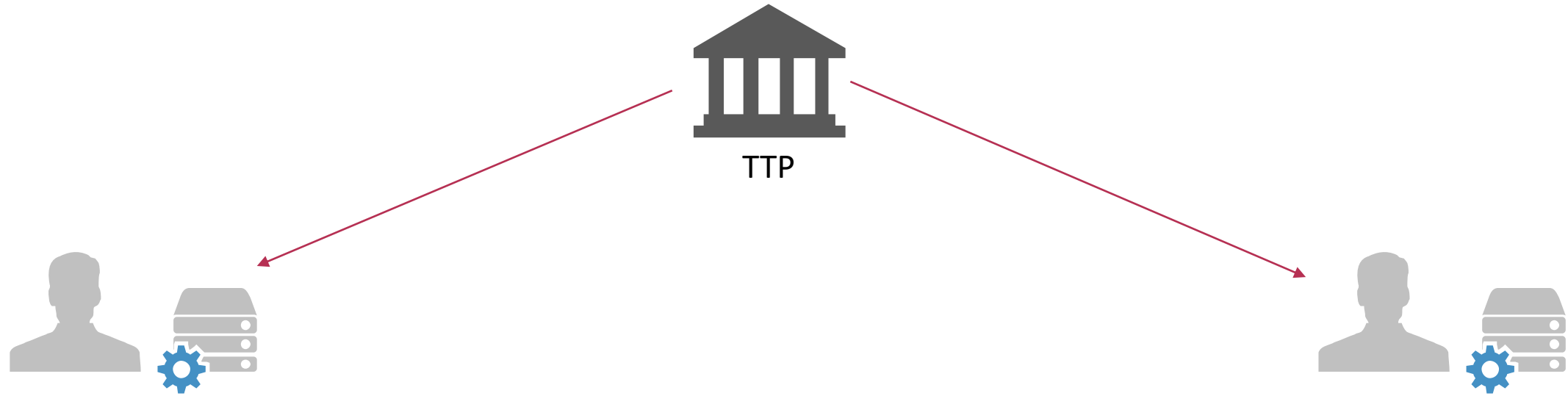


Data aggregation center (DAC)

Data Aggregation System in Korea

Highly sensitive identifier only goes to trusted 3rd party (government agency)

1. TTP sends salt to parties



Party A

Data includes:

IDs, indices, features

Party B

Data includes:

IDs, indices, features

Index	ID (phone)	ZIP	DOB	...
A01	010-1234-5678	12345	1978-SEP-21	...
A02	010-2345-6789	56789	1996-JUN-05	...

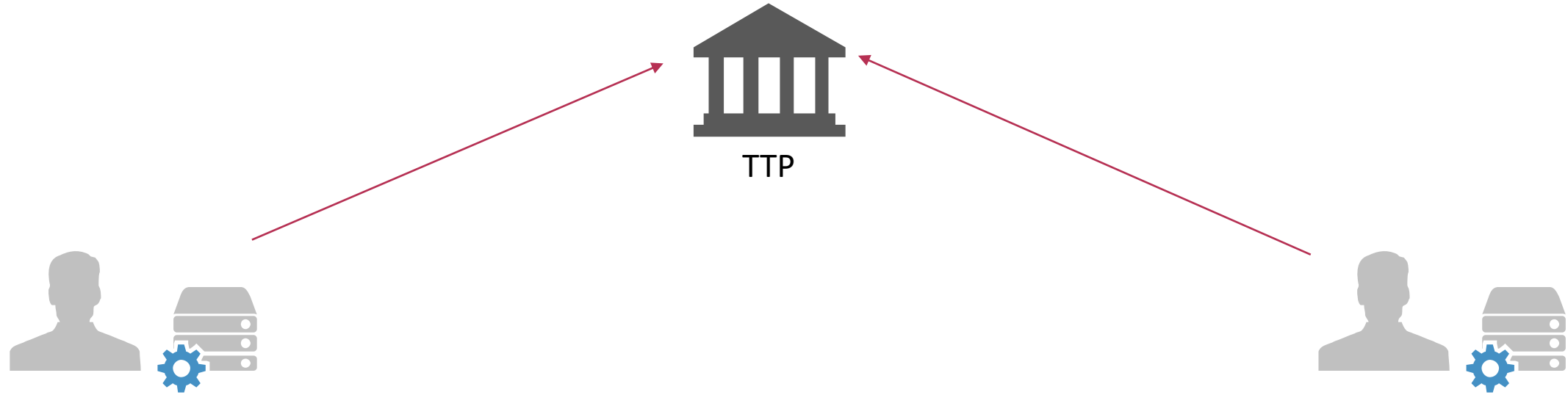


Data aggregation center (DAC)

Data Aggregation System in Korea

Highly sensitive identifier only goes to trusted 3rd party (government agency)

1. TTP sends salt to parties
2. Parties send H(ID), and indices to TTP



Party A
 Data includes:
 IDs, indices, features

Party B
 Data includes:
 IDs, indices, features

Index	ID (phone)	ZIP	DOB	...
A01	010-1234-5678	12345	1978-SEP-21	...
A02	010-2345-6789	56789	1996-JUN-05	...



Data aggregation center (DAC)

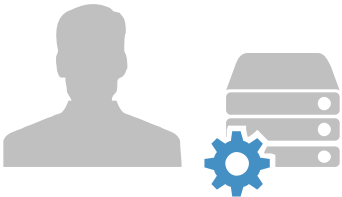
Data Aggregation System in Korea

Highly sensitive identifier only goes to trusted 3rd party (government agency)

1. TTP sends salt to parties
2. Parties send H(ID), and indices to TTP
3. TTP sends index map to DAC



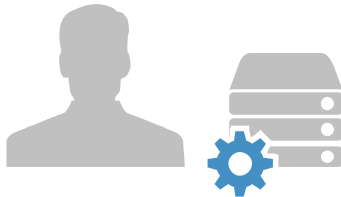
TTP



Party A
 Data includes:
 IDs, indices, features

Index	ID (phone)	ZIP	DOB	...
A01	010-1234-5678	12345	1978-SEP-21	...
A02	010-2345-6789	56789	1996-JUN-05	...

Index A	Index B
A01	B85
A02	B07
...	...



Party B
 Data includes:
 IDs, indices, features



Data aggregation center (DAC)

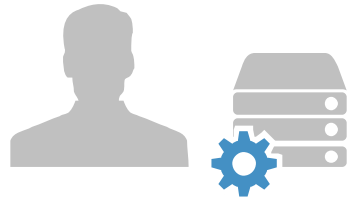
Data Aggregation System in Korea

Highly sensitive identifier only goes to trusted 3rd party (government agency)



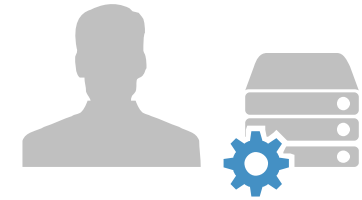
TTP

1. TTP sends salt to parties
2. Parties send H(ID), and indices to TTP
3. TTP sends index map to DAC
4. DAC sends matching ratio to the parties



Party A
Data includes:
IDs, indices, features

Index	ID (phone)	ZIP	DOB	...
A01	010-1234-5678	12345	1978-SEP-21	...
A02	010-2345-6789	56789	1996-JUN-05	...



Party B
Data includes:
IDs, indices, features

Matching ratio = 45%

Matching ratio = 45%



Data aggregation center (DAC)

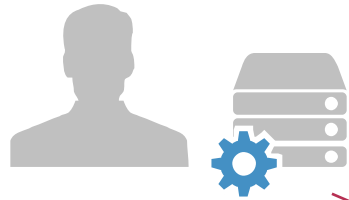
Data Aggregation System in Korea

Highly sensitive identifier only goes to trusted 3rd party (government agency)



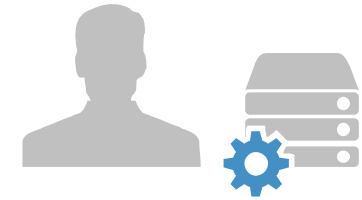
TTP

1. TTP sends salt to parties
2. Parties send H(ID), and indices to TTP
3. TTP sends index map to DAC
4. DAC sends matching ratio to the parties
5. Parties send indices, and features to DAC



Party A
Data includes:
IDs, indices, features

Index	ID (phone)	ZIP	DOB	...
A01	010-1234-5678	12345	1978-SEP-21	...
A02	010-2345-6789	56789	1996-JUN-05	...



Party B
Data includes:
IDs, indices, features



Data aggregation center (DAC)

Circuit-PSI based Data Aggregation System

Samsung SDS is one of the data aggregation center designated by government

Example of Outputs

구분	구분	구분	구분	구분	구분	구분
주민등록번호	이름	주소	이메일	전화번호	회사명	직업명
1 170727-1631706	홍정준	서울특별시 관악구 개포동	xjo@gmail.com	019-743-9150	주식회사 권성희	
3 060821-1241586	김정준	경상북도 문성시 관동81길 (주원서약면)	jaehocoe@dreamwiz.com	064-769-0736	유한회사 유한김	
4 360815-2957489	김지훈	제주특별자치도 제주시 영동대50거리	gimind@hotmail.com	064-783-6075	주식회사 유한김	
5 010120-1704341	최준혁	강원도 수원시 율양구 개포63거리 (영일서동)	jeonghogim@hotmail.com	051-824-4757	주식회사 김김김	
6 380713-2806956	한성남	대구광역시 동대구 석촌로수길 (성호삼우마을)	seoseonjun@hanmail.net	031-411-4426	유한회사 김김희	
7 640827-2059640	최예진	제주특별자치도 구리시 도산대로 (성호역동)	suil@hotmail.com	053-634-1685	주식회사 김김희	
8 720918-2150616	이성민	부산광역시 양정구 봉은사4동 (양정길동)	jeongsun@live.com	032-983-9145	유한회사 김	
9 660804-1847513	장인재	강원도 오산시 백제고분가 (영호성마을)	dohyeon04@nate.com	062-656-1159	(유) 김김김	
10 050225-1943518	이서운	대전광역시 갈사구 태해안8거리 (미경장동)	gimgwangsu@hanmail.net	055-148-1100	(유) 이	
11 320810-1598324	전은영	경기도 여주시 객상70가 (민수마을)	dbag@hanmail.net	051-176-8487	(유) 나	
12 370003-1858733	최민준	대구광역시 수목구 양재천거리 (성희이현리)	myeongaug73@live.com	053-894-3018	유한회사 김김김	
13 650628-2280814	권영욱	서울특별시 서초구 논현2가 (아름아리)	jihyeon@hotmail.com	016-173-6662	주식회사 박희김	
14 420813-1449472	김선영	대구광역시 용산구 석촌로수68거리 (건우백막면)	myeongaug13@live.com	043-410-1908	이서희	
15 040918-2379864	박성준	서울특별시 서초구 송파구 가락거리	hyeonjang@hanmail.net	054-806-9665	오길동	
16 900504-1664274	이연준	경기도 안산시 용유가 (수민김길면)	ggim@gmail.com	051-819-6151	(주) 한택울	
17 840529-1868627	최광수	인천광역시 용산구 석촌로수가 (수민아마울)	jsa3@naver.com	016-309-6563	주식회사 손	
18 340101-1365712	이승민	대구광역시 상동구 개포길	gimseonyeong@hanmail.net	018-078-9754	이길길	
19 740517-1913100	박승경	경상북도 영주시 상당구 석촌로수가 (서운백동동)	gojeongsu@live.com	011-436-2200	주식회사 김김김	
20 430726-2544330	김재현	부산광역시 강북구 봉은사길 (지혜고동)	jeonghunjo@gmail.com	054-868-3861	유한회사 박김이	
21 700804-2145140	김정남	전라남도 가평군 석촌로수12가	ggim@nate.com	070-2461-4163	(유) 서김이	
22 040805-1645345	이민수	강원도 서산시 태해안로 (서명천안마을)	baejeongsu@nate.com	053-917-4775	전업	
23 880220-1243921	이예준	부산광역시 동구 가락길	yunseosim@gmail.com	033-841-3419	김이름	
24 730617-2583714	이준차	경기도 광주시 압구정4로 (영숙백면)	ygim@live.com	064-771-6989	박정	
25 940724-2287928	곽영준	서울특별시 서대문구 반포대57거리	yeongjeoljim@live.com	070-9056-6401	유한회사	
26 290426-1484646	지은서	부산광역시 강북구 서초중앙로	yonseo73@dreamwiz.com	055-611-2323	(주) 김	
27 650011-2861442	자유진	강원도 양양시 봉암사6길 (이영아마을)	tsong@gmail.com	043-161-7487	김승백	
28 050400-1874138	김지혜	경상북도 군포시 서초중앙447길	yeonghwan36@hotmail.com	041-436-8908	회길	
29 780210-2658235	권정호	인천광역시 연동로구 서초대거리	hbag@hotmail.com	041-431-8092	김길	
30 450419-1860155	강도현	인천광역시 강북구 영동대로 (성민마을)	sanghobae@hotmail.com	018-905-1114	이백	

Performances

Circuit-PSI based Data Aggregation System

Samsung SDS is one of the data aggregation center designated by government

Example of Outputs

A	B	C	D	E	F	G
1	주민등록번호	이름	주소	이메일	전화번호	회사명
2	170727-1631706	홍정준	서울특별시 관악구 개포동	xjo@gmail.com	019-743-9150	주식회사 권심희
3	060821-1241586	정정준	충청북도 보은시 황호81길 (후원서백면)	jaehocoe@dreamwiz.com	064-769-0736	유한회사 권심희
4	360815-2957489	김지훈	제주특별자치도 원주시 영동대50거리	jeonmin@hotmail.com	064-783-6075	주식회사 윤희갑
5	010120-1704341	최준혁	강원도 수원시 팔달구 개포63거리 (영일서동)	jeonghogim@hotmail.com	051-824-4757	주식회사 김김김
6	380713-2806956	한성남	대구광역시 중로구 석촌로수길 (성호로우마을)	seosequn@hanmail.net	031-411-4426	유한회사 김김희
7	640827-2059640	최예민	제주특별자치도 구리시 도산대로 (성포역동)	suil@hotmail.com	053-634-1685	즈름
8	720918-2150616	이성민	부산광역시 양정구 봉은사4로 (양정길동)	jeongsun@live.com	032-983-9145	유한회사 김
9	660804-1847513	장인재	강원도 오산시 백제고분가 (영호정마을)	dohyeon04@nate.com	062-656-1159	(유) 김관길
10	050225-1943518	이서운	대전광역시 갈사구 태해안8거리 (미경장동)	gimgwangs@hanmail.net	055-148-1100	(유) 이
11	320810-1598324	전은영	경기도 여주시 역삼70가 (민수마을)	dbag@hanmail.net	051-176-8487	(유) 나
12	370003-1858733	최인준	대구광역시 구로구 양재천거리 (영희이현리)	myeongaug73@live.com	053-894-3018	유한회사 김김김
13	650628-2280814	권영욱	제주특별자치시 광진구 가락거리	jihyeon@hotmail.com	016-173-6662	주식회사 박희갑
14	420813-1449472	김선영	대구광역시 용산구 석촌로수68거리 (건우백막면)	myeongaug13@live.com	043-410-1908	이서희
15	040918-2379804	최성훈	제주특별자치시 송파구 가락거리	hyeonjang@hanmail.net	054-806-9665	오길환
16	900504-1664274	이연호	경기도 안산시 화흥가 (수민마을)	ggim@gmail.com	051-819-6151	(주) 한백윤
17	840529-1368627	최광수	인천광역시 용산구 석촌로수가 (수민마을)	jsa53@naver.com	016-309-6563	주식회사 손
18	340101-1365712	이종민	대구광역시 상봉구 개포길	gimseonyeong@hanmail.net	018-078-9754	이길진
19	740517-1913100	박은경	충청북도 청주시 상당구 석촌로수가 (서문백막동)	gojongsu@live.com	011-436-2200	주식회사 김이갑
20	430726-2544330	김재현	부산광역시 강북구 봉은사길 (지혜고동)	jeonhunjoo@gmail.com	054-868-3861	유한회사 박김이
21	700804-2145140	김정남	전라남도 가평군 봉은사12가 (지혜고동)	ggim@nate.com	070-2461-4163	(유) 서갑이
22	040805-1645345	이민우	강원도 서산시 태해안로 (서영전마을)	bagseonja@nate.com	053-917-4775	전안
23	880220-1243921	이예준	부산광역시 중구 가락길	yunseosim@gmail.com	033-841-3419	김이훈
24	730617-2583714	이준자	경기도 광주시 인구정4로 (정숙백면)	vgim@live.com	064-771-6989	박길
25	940724-2267928	곽영준	서울특별시 서대문구 반포대57거리	yeongceolgim@live.com	070-9056-6401	유한회사 권심희
26	290426-1484646	지은서	부산광역시 강북구 서초중앙로	yeongceolgim@live.com	055-611-2323	(주) 김
27	650011-2861442	지유진	강원도 양산시 봉은사6길 (미영마을)	tsong@gmail.com	043-161-7487	김승백
28	050400-1874138	김기재	충청북도 군포시 서초중앙447길	jeonghwan36@hotmail.com	041-436-8908	희갑
29	780210-2658235	권정호	인천광역시 영동로구 서초대거리	hbag@hotmail.com	041-431-8092	김진
30	450419-1860135	강도현	인천광역시 강북구 영동대로 (성민마을)	sanghobae@hotmail.com	018-905-1114	이박

A	B	C	D	E	F	G
16af31a749f8a7	주소	서울특별시 관악구 개포동	회사명	홍정준보수액(만 원)		
baef5e0680203a	충청북도 보은시 황호81길 (후원서백면)	유한회사 권심희	1410			
1c982c965c1a3e	제주특별자치도 원주시 영동대50거리	주식회사 윤희갑	1496			
993645f4e49d2e	대구광역시 중로구 석촌로수길 (성호로우마을)	주식회사 김김김	789			
1490325069af0	대구광역시 중로구 석촌로수길 (성호로우마을)	유한회사 김김희	1262			
72dbb51b0a6a0	제주특별자치도 구리시 도산대로 (성포역동)	즈름	1918			
f714c504b9904	부산광역시 양정구 봉은사4로 (양정길동)	유한회사 김	220			
39635c9d786b0b	강원도 오산시 백제고분가 (영호정마을)	(유) 김관길	144			
11d0f7ddbc86fb	대전광역시 갈사구 태해안8거리 (미경장동)	(유) 이	483			
2ca9657719b3f4	경기도 여주시 역삼70가 (민수마을)	(유) 나	1555			
3a32cc926e567f	대구광역시 구로구 양재천거리 (영희이현리)	유한회사 김김김	1887			
123342033433ce	제주특별자치시 송파구 가락거리	주식회사 박희갑	1926			
aa976399dafc	대구광역시 용산구 석촌로수68거리 (건우백막면)	이서희	1319			
1388bfca6jcd2fb	제주특별자치시 송파구 가락거리	오길환	740			
961cde9a1c088	경기도 안산시 화흥가 (수민마을)	(주) 한백윤	436			
396f5724fcc4e5	대구광역시 용산구 석촌로수가 (수민마을)	주식회사 손	1399			
257304863178b	대구광역시 상봉구 개포길	이길진	559			
1480eb69f49eaa	충청북도 청주시 상당구 석촌로수가 (서문백막동)	주식회사 김이갑	1583			
291a4b13fd2e1	부산광역시 강북구 봉은사길 (지혜고동)	유한회사 박김이	553			
af2e6d8b1b33fb	전라남도 가평군 봉은사12가 (지혜고동)	(유) 서갑이	719			
1bf739af91fcb	강원도 서산시 태해안로 (서영전마을)	전안	1205			
24bf6c5db86370	부산광역시 중구 가락길	김이훈	216			
a57694e4c0134	경기도 광주시 인구정4로 (정숙백면)	박길	297			
33169ba48a173	서울특별시 서대문구 반포대57거리	유한회사 권심희	861			
7746839ae4c5	부산광역시 강북구 서초중앙로	(주) 김	145			
24bf231f6a3c8a	강원도 양산시 봉은사6길 (미영마을)	김승백	1964			
3ac10f099af897	충청북도 군포시 서초중앙447길	희갑	1965			
365e0e9ae9531	인천광역시 영동로구 서초대거리	김진	1088			
2839d042c5b130	인천광역시 강북구 영동대로 (성민마을)	이박	1813			

Unique Identifier

Performances

Matching ratio : 0.43 (43% of your data is in other party's data)

Circuit-PSI based Data Aggregation System

Samsung SDS is one of the data aggregation center designated by government

Example of Outputs

순번	주민등록번호	이름	주소	이메일	전화번호	회사명	직업명
1	170727-1631706	홍정준	서울특별시 관악구 개포길	xjo@gmail.com	019-743-9150	주식회사 권심희	회계사
2	060821-1241586	김정준	충청북도 보은시 황동81길 (주원서백면)	jaehocce@dreamwiz.com	064-769-0736	유한회사 권심희	회계사
3	360815-2957489	김지서	제주특별자치도 영주시 영동대50거리	jeonmin@hotmail.com	064-783-6075	주식회사 윤희갑	회계사
4	010120-1704341	최준현	강원도 수원시 팔달구 개포63거리 (영철서운)	jeonghogim@hotmail.com	051-824-4757	주식회사 김경호	회계사
5	380713-2806956	한성남	대구광역시 중구 서촌로수길 (성호신문마을)	seojeon@hanmail.net	031-411-4426	유한회사 김경희	회계사
6	640827-2059640	최예진	제주특별자치도 구리시 도산대로 (성호역출)	stui@hotmail.com	053-634-1685	주식회사 권심희	회계사
7	720918-2150616	이성민	부산광역시 양정구 봉은사4로 (상철빌딩)	jeongsun@live.com	032-983-9145	유한회사 권	회계사
8	660804-1847513	장인태	경기도 오산시 백제고분가 (영호중앙)	dohyeon04@nate.com	062-656-1159	(유) 김관길	회계사
9	050225-1943518	이서운	대전광역시 강서구 태해안8거리 (미경빌딩)	gimgwangs@hanmail.net	055-148-1100	(유) 이	회계사
10	320810-1598324	전은영	경기도 여주시 역삼70가 (민수이동)	dbag@hanmail.net	051-176-8487	(유) 나	회계사
11	370003-1858733	최민준	대구광역시 구로구 양재천거리 (영희이현리)	myeonggaug73@live.com	053-894-3018	유한회사 김경갑	회계사
12	650628-2280814	권영욱	제주특별자치도 광진구 논현6가 (아름아리)	jiyeon@hotmail.com	016-173-6662	주식회사 박희갑	회계사
13	420813-1449472	김선영	대구광역시 용산구 서촌로수66거리 (건우백백면)	myeonggaug13@live.com	043-410-1908	이서희	회계사
14	040918-2379804	최성훈	제주특별자치도 송파구 가락거리	hyeonjang@hanmail.net	054-806-9665	오길영	회계사
15	900904-1664274	이연호	경기도 안산시 화왕가 (수민이마을)	ggim@gmail.com	051-819-6131	(주) 한백윤	회계사
16	840529-1368627	최광수	인천광역시 용산구 서촌로수가 (수민이마을)	ja53@naver.com	016-309-6563	주식회사 손	회계사
17	340101-1365712	이종민	대구광역시 상동구 개포길	gimseonyeong@hanmail.net	018-078-9754	이길영	회계사
18	740517-1913100	박은경	충청북도 청주시 상당구 서촌로수가 (서촌백백동)	gojeongsu@live.com	011-436-2200	주식회사 김어갑	회계사
19	430726-2544330	김재현	부산광역시 강북구 봉은사길 (지혜고동)	jeonhunjoo@gmail.com	054-860-3861	유한회사 박희어	회계사
20	700804-2145140	김정남	전라남도 가평군 서촌로수12가	ggim@nate.com	070-2461-4163	(유) 서갑어	회계사
21	040805-1649545	이민수	경기도 서산시 태해안로 (서영천마을)	bagsyeongsu@nate.com	053-917-4775	전안	회계사
22	880220-1243921	이예준	부산광역시 중구 가락길	yunseosim@gmail.com	033-841-3419	김어영	회계사
23	730617-2583714	이준자	경기도 광주시 양주정4로 (영호백면)	vgim@live.com	064-771-6989	박영	회계사
24	940724-2267928	권영준	서울특별시 서대문구 방포대57거리	yeongjeolggim@live.com	070-9056-6401	유한회사 권	회계사
25	290426-1484646	지은서	부산광역시 강북구 서초중앙로	seong73@dreamwiz.com	055-611-2323	(주) 권	회계사
26	650011-2861442	지유진	강원도 양산시 봉은사길 (미영아리)	tsong@gmail.com	043-161-7487	김승백	회계사
27	050400-1874130	김기태	경상북도 군포시 서초중앙447길	jeonghwan36@hotmail.com	041-436-8908	유한회사 권	회계사
28	780210-2652335	권정호	인천광역시 영동구 서초대거리	hbag@hotmail.com	041-431-8092	김경	회계사
29	450419-1860135	강도현	인천광역시 강북구 영동대로 (성민이동)	sanghobae@hotmail.com	018-905-1114	이박	회계사

Unique Identifier	Join Key	주소	이름	이메일	전화번호	회사명	직업명
16af31a749f8a7	16af31a749f8a7	서울특별시 관악구 개포길	홍정준	xjo@gmail.com	019-743-9150	주식회사 권심희	회계사
bac5e0680203a	bac5e0680203a	충청북도 보은시 황동81길 (주원서백면)	김정준	jaehocce@dreamwiz.com	064-769-0736	유한회사 권심희	회계사
c982c965c1a3e	c982c965c1a3e	제주특별자치도 영주시 영동대50거리	김지서	jeonmin@hotmail.com	064-783-6075	주식회사 윤희갑	회계사
9936d5f4e4942e	9936d5f4e4942e	강원도 수원시 팔달구 개포63거리 (영철서운)	최준현	jeonghogim@hotmail.com	051-824-4757	주식회사 김경호	회계사
1490325069af0	1490325069af0	대구광역시 중구 서촌로수길 (성호신문마을)	한성남	seojeon@hanmail.net	031-411-4426	유한회사 김경희	회계사
72dbb51b0a6a0	72dbb51b0a6a0	제주특별자치도 구리시 도산대로 (성호역출)	최예진	stui@hotmail.com	053-634-1685	주식회사 권심희	회계사
f714c504b9904	f714c504b9904	부산광역시 양정구 봉은사4로 (상철빌딩)	이성민	jeongsun@live.com	032-983-9145	유한회사 권	회계사
39035c9d780b0b	39035c9d780b0b	경기도 오산시 백제고분가 (영호중앙)	장인태	dohyeon04@nate.com	062-656-1159	(유) 김관길	회계사
11d0f7ddb86fb	11d0f7ddb86fb	대전광역시 강서구 태해안8거리 (미경빌딩)	이서운	gimgwangs@hanmail.net	055-148-1100	(유) 이	회계사
2ca9657719b3f4	2ca9657719b3f4	경기도 여주시 역삼70가 (민수이동)	전은영	dbag@hanmail.net	051-176-8487	(유) 나	회계사
3a32cc926e567f	3a32cc926e567f	대구광역시 구로구 양재천거리 (영희이현리)	최민준	myeonggaug73@live.com	053-894-3018	유한회사 김경갑	회계사
123342033433ce	123342033433ce	제주특별자치도 광진구 논현6가 (아름아리)	권영욱	jiyeon@hotmail.com	016-173-6662	주식회사 박희갑	회계사
aa976399d4fe	aa976399d4fe	대구광역시 용산구 서촌로수66거리 (건우백백면)	김선영	myeonggaug13@live.com	043-410-1908	이서희	회계사
1388bfca6dcd2fb	1388bfca6dcd2fb	제주특별자치도 송파구 가락거리	최성훈	hyeonjang@hanmail.net	054-806-9665	오길영	회계사
961cde9a1c088	961cde9a1c088	경기도 안산시 화왕가 (수민이마을)	이연호	ggim@gmail.com	051-819-6131	(주) 한백윤	회계사
398f5c7247cc4e5	398f5c7247cc4e5	인천광역시 용산구 서촌로수가 (수민이마을)	최광수	ja53@naver.com	016-309-6563	주식회사 손	회계사
2747304863178b	2747304863178b	대구광역시 상동구 개포길	이종민	gimseonyeong@hanmail.net	018-078-9754	이길영	회계사
1480eb69f49eaa	1480eb69f49eaa	충청북도 청주시 상당구 서촌로수가 (서촌백백동)	박은경	gojeongsu@live.com	011-436-2200	주식회사 김어갑	회계사
291a4b13f0d2e1	291a4b13f0d2e1	부산광역시 강북구 봉은사길 (지혜고동)	김재현	jeonhunjoo@gmail.com	054-860-3861	유한회사 박희어	회계사
af2e60db1b31fb	af2e60db1b31fb	전라남도 가평군 서촌로수12가	김정남	ggim@nate.com	070-2461-4163	(유) 서갑어	회계사
1bf7739af91fc	1bf7739af91fc	경기도 서산시 태해안로 (서영천마을)	이민수	bagsyeongsu@nate.com	053-917-4775	전안	회계사
24bf6c5db86370	24bf6c5db86370	부산광역시 중구 가락길	이예준	yunseosim@gmail.com	033-841-3419	김어영	회계사
a57694e4c0134	a57694e4c0134	경기도 광주시 양주정4로 (영호백면)	이준자	vgim@live.com	064-771-6989	박영	회계사
335169ba48a173	335169ba48a173	서울특별시 서대문구 방포대57거리	권영준	yeongjeolggim@live.com	070-9056-6401	유한회사 권	회계사
7746839a6e4c5	7746839a6e4c5	부산광역시 강북구 서초중앙로	지은서	seong73@dreamwiz.com	055-611-2323	(주) 권	회계사
24bf231fa3c8a	24bf231fa3c8a	강원도 양산시 봉은사길 (미영아리)	지유진	tsong@gmail.com	043-161-7487	김승백	회계사
3ac10f09a8967	3ac10f09a8967	경상북도 군포시 서초중앙447길	김기태	jeonghwan36@hotmail.com	041-436-8908	유한회사 권	회계사
366e0eae9531	366e0eae9531	인천광역시 영동구 서초대거리	권정호	hbag@hotmail.com	041-431-8092	김경	회계사
2839d042c5b130	2839d042c5b130	인천광역시 강북구 영동대로 (성민이동)	강도현	sanghobae@hotmail.com	018-905-1114	이박	회계사

Matching ratio : 0.43 (43% of your data is in other party's data)

Performances

How it works

- For the unique ID generation, we used OPRF instead of salted hash
- Matching ratio is computed using circuit-PSI

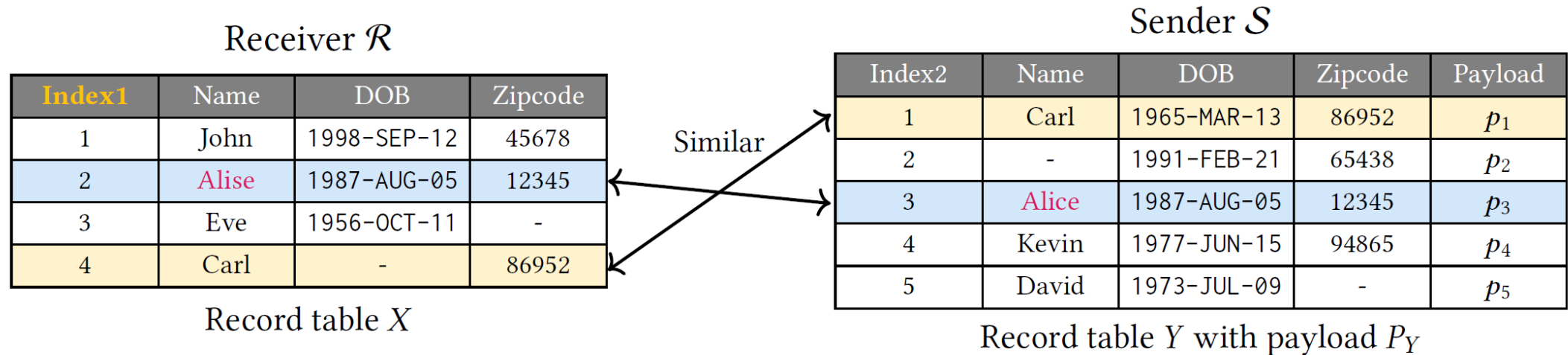
Timing Result (LAN)

	Dataset Size	Running Time
Join Key Generation & Ratio Computation	1m	1m 28s
	10m	11m 38s
	20m	20m 40s
Multiple ID Col Support	-	Checked

3. Privacy-Preserving Record Linkage

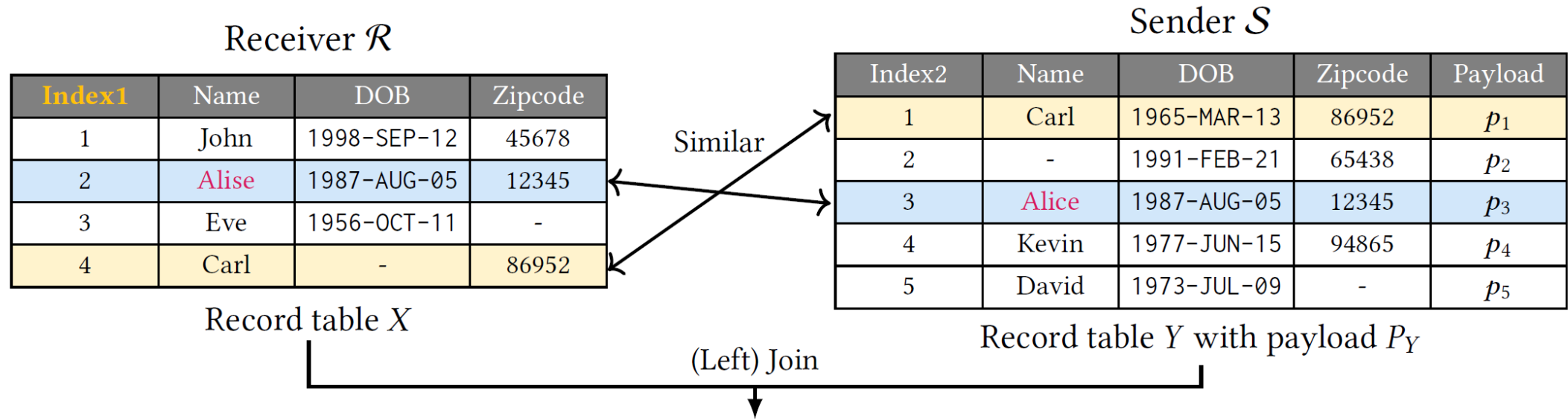
Privacy-Preserving Record Linkage (PETS '25)

What is target functionality?



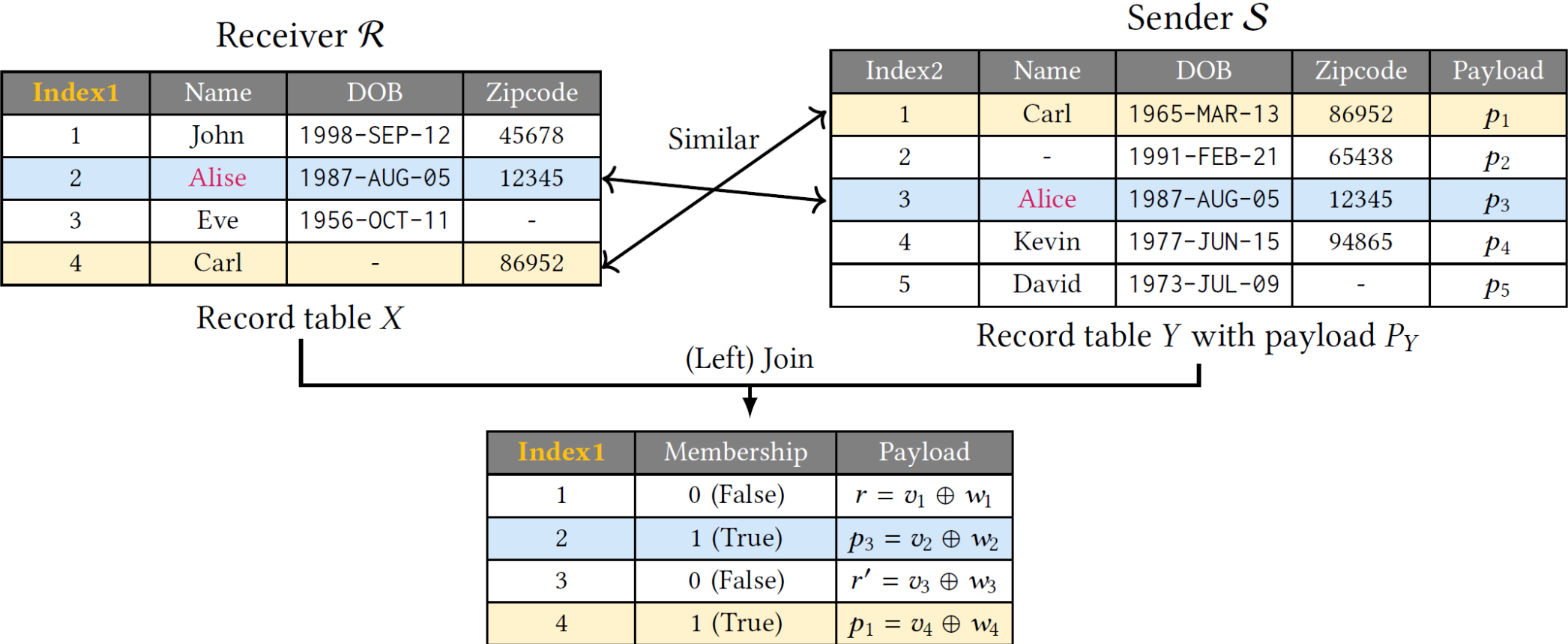
Privacy-Preserving Record Linkage (PETS '25)

What is target functionality?



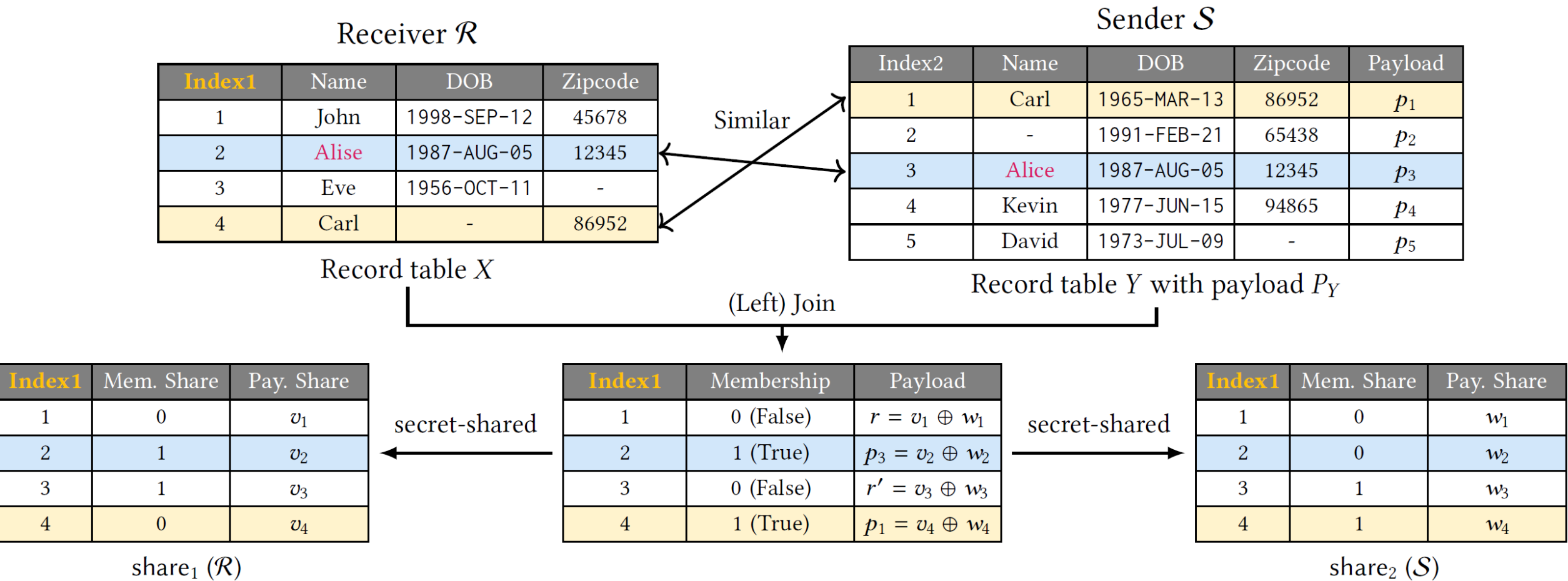
Privacy-Preserving Record Linkage (PETS '25)

What is target functionality?



Privacy-Preserving Record Linkage (PETS '25)

What is target functionality?



Privacy-Preserving Record Linkage (PETS '25)

from CPSI and some MPC techniques (simplified)

1. Encode features (e.g., concatenation, LSH)

Index1	Name	DOB	Zipcode
1	John	1998-SEP-12	45678
2	Alise	1987-AUG-05	12345
3	Eve	1956-OCT-11	-
4	Carl	-	86952

Record table X with quasi-identifiers

Encode

Index1	feature1	feature2	feature3	feature4
1	d330ca	719cf0	0e0348	79457c
2	01191e	bf9a47	0c1452	c41b44
3	e60467	b91fc3	0e79d1	005dae
4	f6fabd	a8fdd6	a80394	737963

Encoded table \hat{X} with features

Privacy-Preserving Record Linkage (PETS '25)

from CPSI and some MPC techniques (simplified)

1. Encode features (e.g., concatenation, LSH)
2. Invoke CPSI protocols feature-wisely

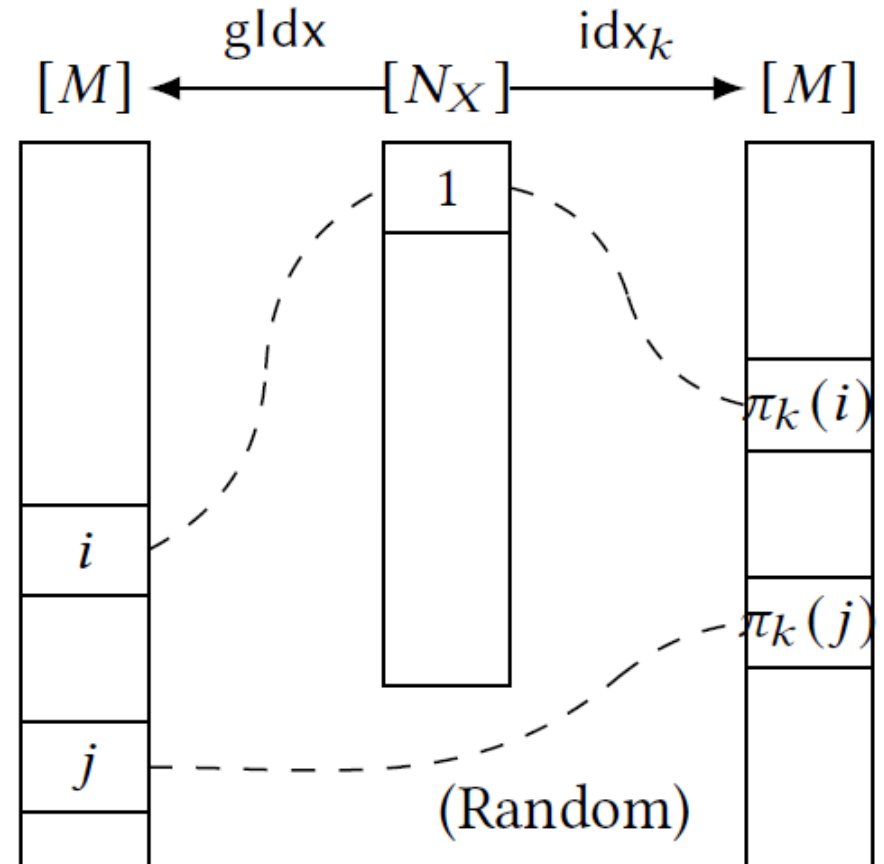
Index1	feature1	feature2	feature3	feature4
1	d330ca	719cf0	0e0348	79457c
2	01191e	bf9a47	0c1452	c41b44
3	e60467	b91fc3	0e79d1	005dae
4	f6fabd	a8fdd6	a80394	737963

Encoded table \hat{X} with features

Privacy-Preserving Record Linkage (PETS '25)

from CPSI and some MPC techniques (simplified)

1. Encode features (e.g., concatenation, LSH)
2. Invoke CPSI protocols feature-wisely
3. Align the shares using Permute-and-Share protocol



Privacy-Preserving Record Linkage (PETS '25)

from CPSI and some MPC techniques (simplified)

1. Encode features (e.g., concatenation, LSH)
2. Invoke CPSI protocols feature-wisely
3. Align the shares using Permute-and-Share protocol
4. Compute matching criterion using generic MPC (e.g., GMW)

$$\sum_i \text{Eq}(\text{feature}_i) \geq t$$

Index1	feature1	feature2	feature3	feature4
1	d330ca	719cf0	0e0348	79457c
2	01191e	bf9a47	0c1452	c41b44
3	e60467	b91fc3	0e79d1	005dae
4	f6fabd	a8fdd6	a80394	737963

Encoded table \hat{X} with features

Privacy-Preserving Record Linkage (PETS '25)

from CPSI and some MPC techniques (simplified)

1. Encode features (e.g., concatenation, LSH)
2. Invoke CPSI protocols feature-wisely
3. Align the shares using Permute-and-Share protocol
4. Compute matching criterion using generic MPC (e.g., GMW)

$$\sum_i \text{Eq}(\text{feature}_i) \geq t$$

Dataset	Data Size	#Features	Comm. (MB)	Setup (s)	Online (s)	F1-score
European Census	24K	6	76.2	4.47	1.13	0.948
NCVR	1M	3	1615	169	23.7	0.976

References

- Data table in p.37-39 is generated by Faker (owner: Francois Zaninotto, url: <https://faker.readthedocs.io/en/master/>) at Oct. 1. 2021
- [HMS22] K. Han, D. Moon, and Y. Son. Improved Circuit-PSI via Equality Preserving Compression. SAC 2022.
- [SJ23] Y. Son, and J. Jeong. PSI with computation or Circuit-PSI for Unbalanced Sets from Homomorphic Encryption. AsiaCCS 2023.
- [HKLS24] K. Han, S. Kim, B. Lee, and Y. Son. Revisiting OKVS-based OPRF and PSI: Cryptanalysis and Better Construction. Asiacrypt 2024, to appear.
- [HKS24] K. Han, S. Kim, and Y. Son. Private Computation on Common Fuzzy Records. PETS 2025, to appear.

Thank you

Q&A

SAMSUNG SDS