同盾科技
www.tongdun.cn

基于HE的MPSI协议

Hang shao

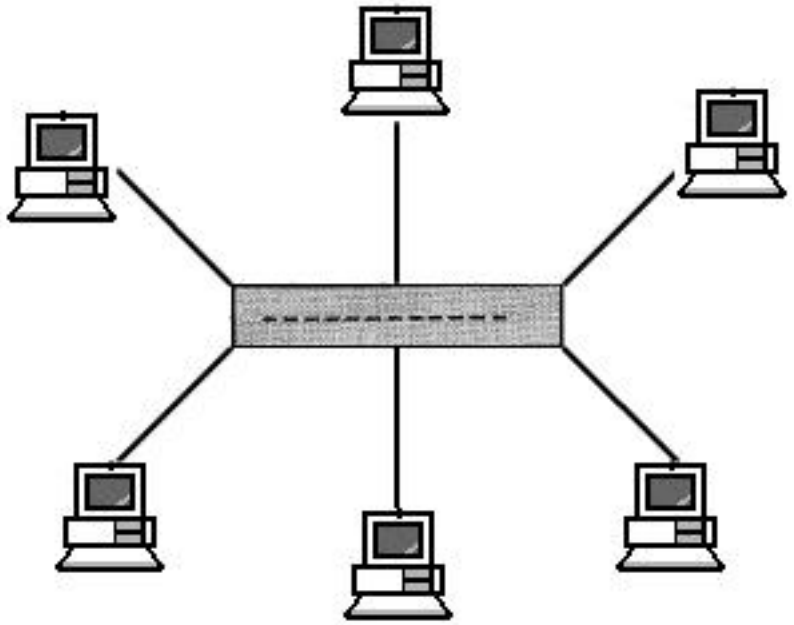
2022.5.31

目录

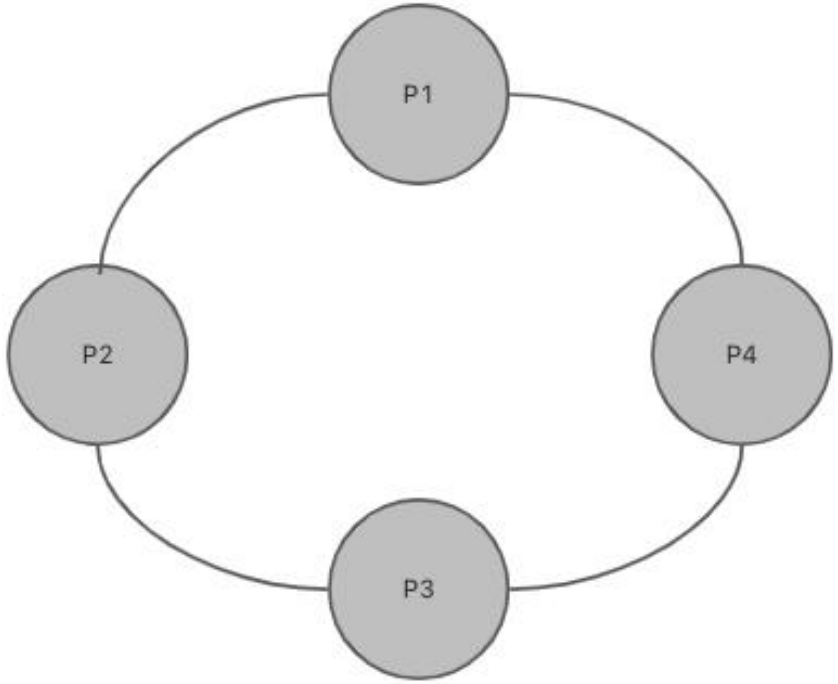
- MPSI（OT）
- 方案1（云环境下集合隐私计算-李顺东(软件学报)）
- 方案2（集合交集问题的安全计算-赵雪玲（密码学报））
- 方案3（ Scalable Multi-Party Private Set-Intersection-2017 ）

MPSI

目前MPSI 协议大多采用2种网络模型：



星型拓扑结构



星型路径网络

MPSI

同盾科技 www.tongdun.cn

Efficient private matching and set intersection—2004，首次提出MPSI

表 3 多方 PSI 复杂性分析

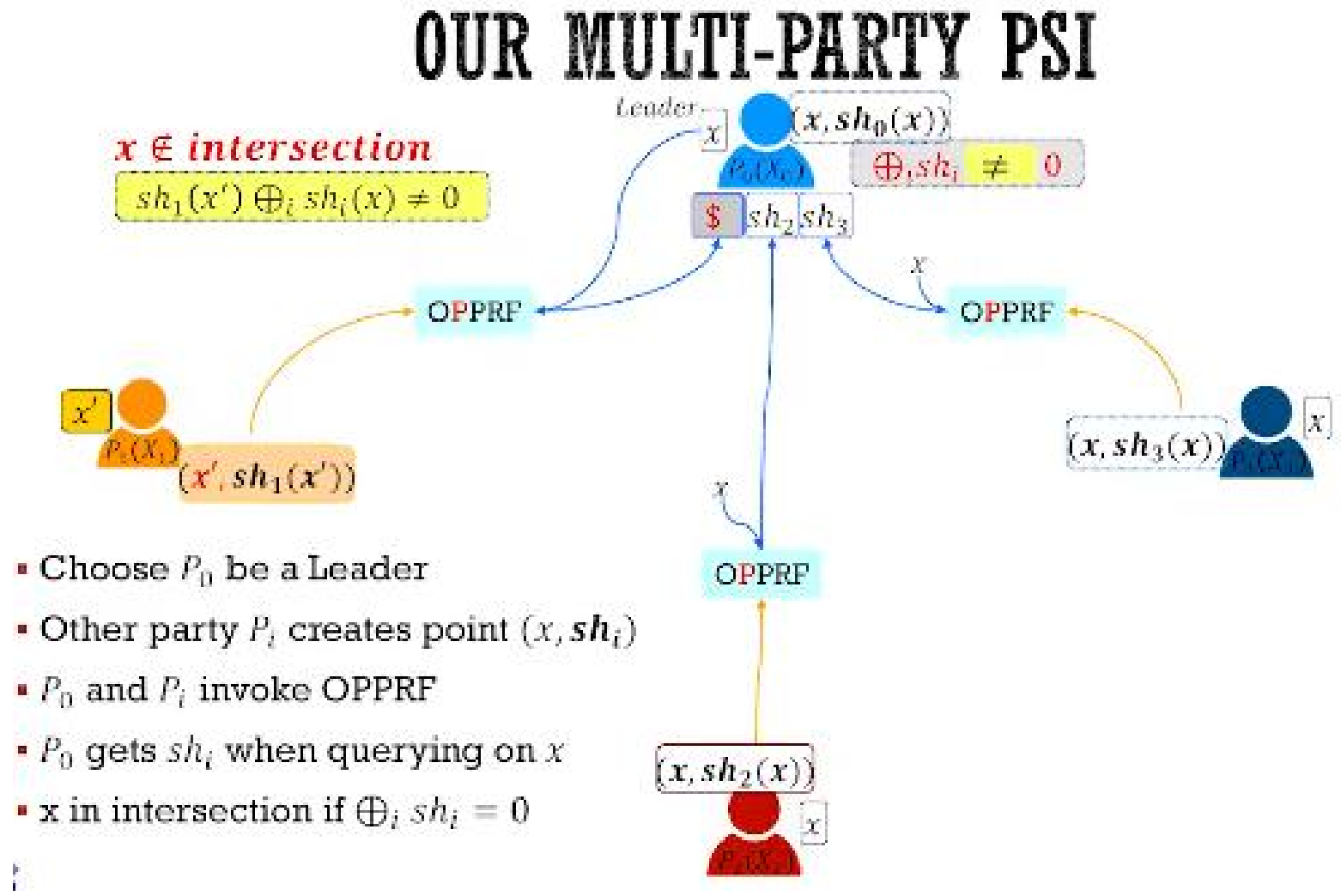
协议	通信量		计算量		安全模型	技术	网络结构
	指定方	其他方	指定方	其他方			
文献[63]	$O(tn\lambda)$	$O(tn\lambda)$	$O(tk)$	$O(tk)$	半诚实	OPRF+OT	星型
文献[63]	$O(tn\lambda)$	$O(n\lambda)$	$O(tk)$	$O(k)$	强半诚实	OPRF+OT	星型
文献[46]	$O(tn\lambda)$	$O(n\lambda)$	$O(tn\log n)$	$O(n)$	恶意	OPE+HE	星型
文献[84]	$O(tn\lambda k)$	$O(tn\lambda k)$	$O(tn\lambda k)$	$O(tn\lambda k)$	半诚实	BF+OT	星型
文献[84]	$O(\log(t)n\lambda k)$	$O(\log(t)n\lambda k)$	$O(tn\lambda k)$	$O(tn\lambda k)$	强半诚实	BF+OT	星型
文献[61]	$O(tnk(\log(nk)+k))$	$O(tnk(\log(nk)+k))$	$O(tnk(\log(nk)+k))$	$O(tnk(\log(nk)+k))$	恶意	BF+OT	星型
文献[74]	$O(tn\lambda)$	$O(n\lambda k)$	$O(tn\lambda)$	$O(n\lambda k)$	半诚实	OPRF+OT	星型-路径

[63]：**Practical Multi—Party Private Set Intersection from Symmetric—Key Techniques—2017**
[46]：**Scalable Multi—Party Private Set—Intersection—2017**
[84]：Efficient scalable multiparty private set—intersection via garbled Bloom filters—2018
[61]：PSImple：Practical Multiparty Maliciously—Secure Private Set Intersection—2021
[74]：Efficient Scalable Multi—Party Private Set Intersection Using Oblivious PRF—2021

[1]：Efficient Linear Multiparty PSI and Extensions to Circuit/Quorum PSI
[2]：**Simple，Fast Malicious Multiparty Private Set Intersection—2021**

MPSI

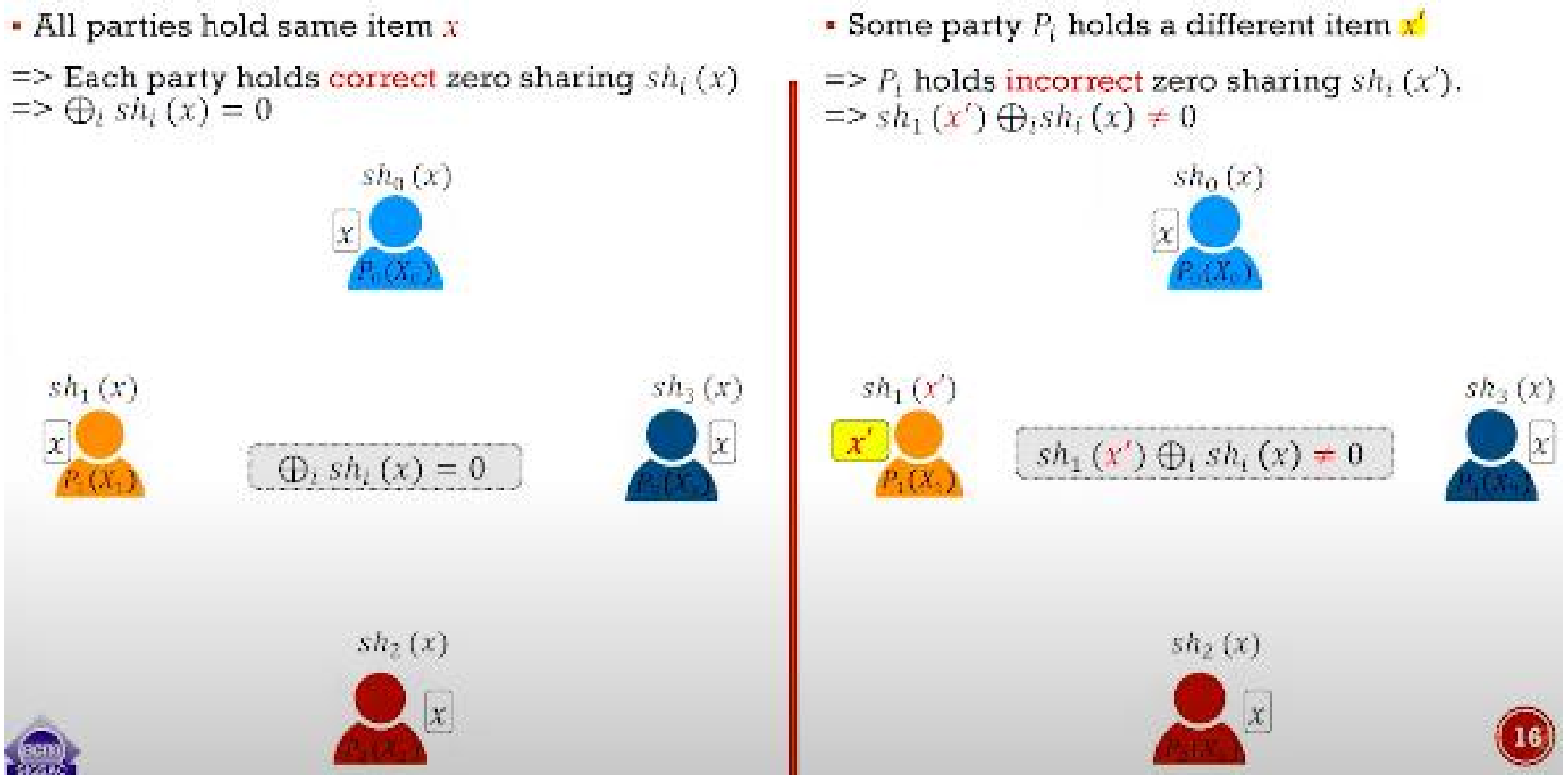
- Zero—Sharing
- OPPRF



[63]：Practical Multi—Party Private Set Intersection from Symmetric—Key Techniques

MPSI

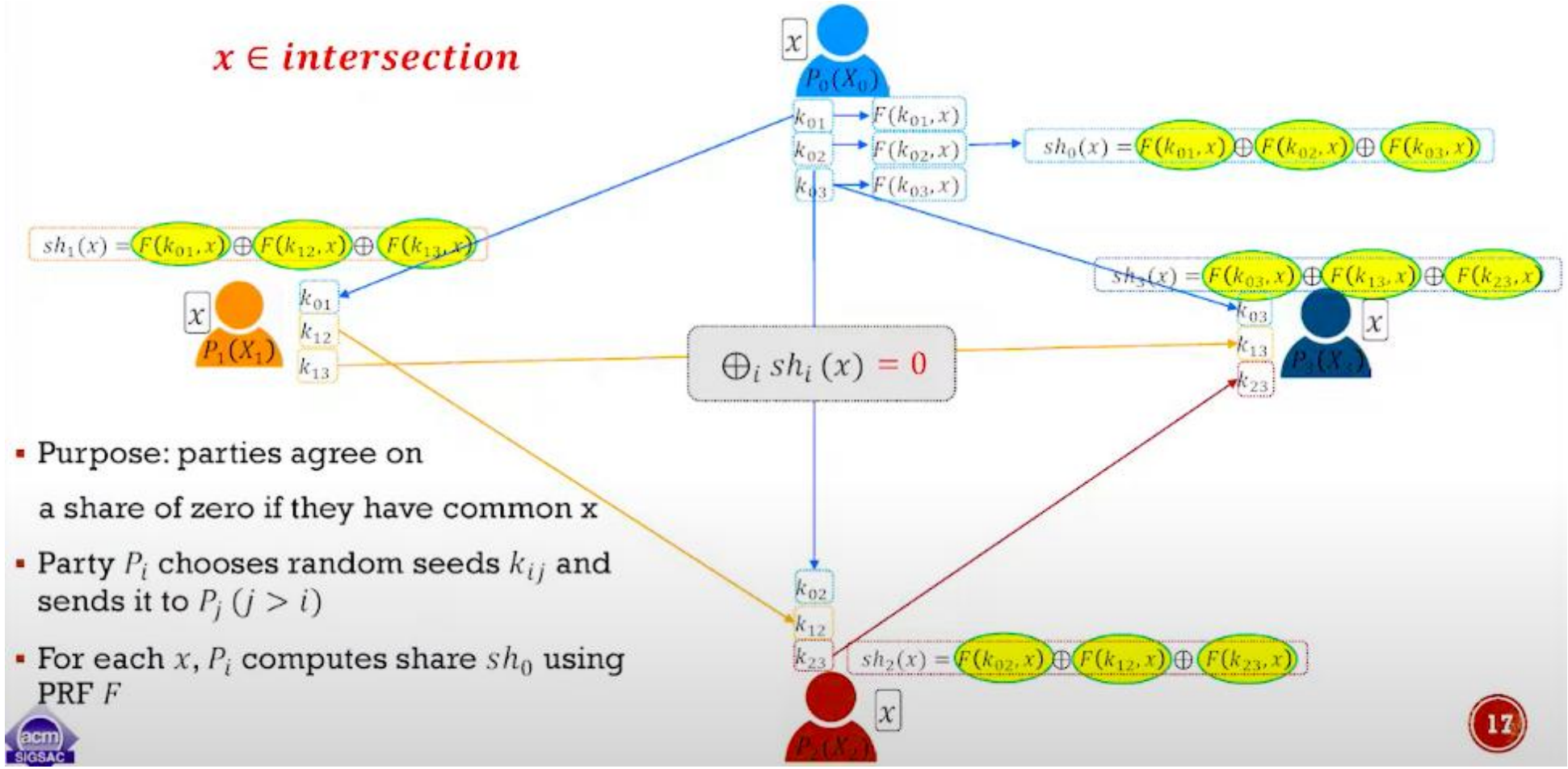
- Zero—Sharing



[63]：Practical Multi—Party Private Set Intersection from Symmetric—Key Techniques

MPSI

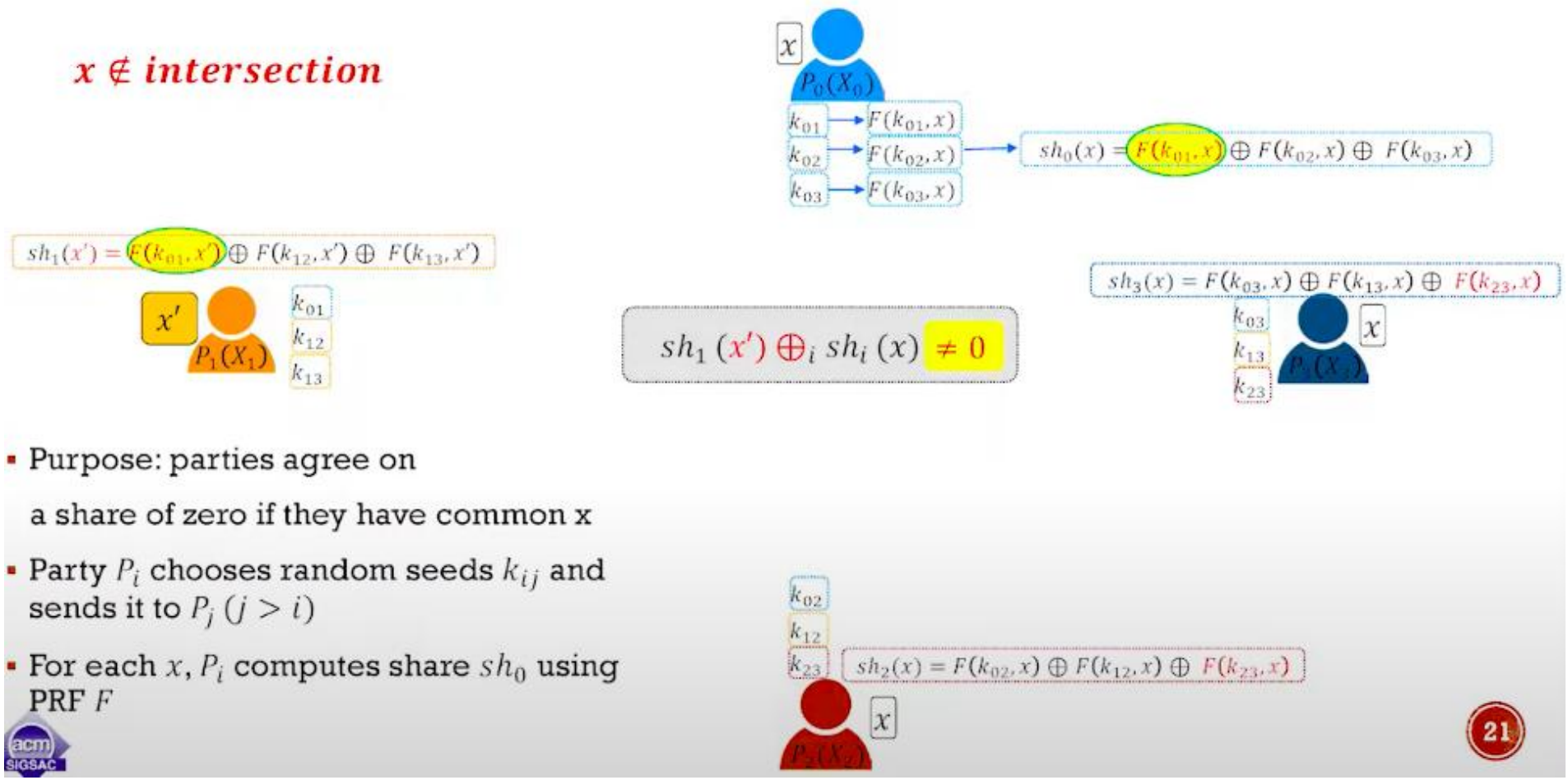
- Zero—Sharing



[63]: Practical Multi—Party Private Set Intersection from Symmetric—Key Techniques

MPSI

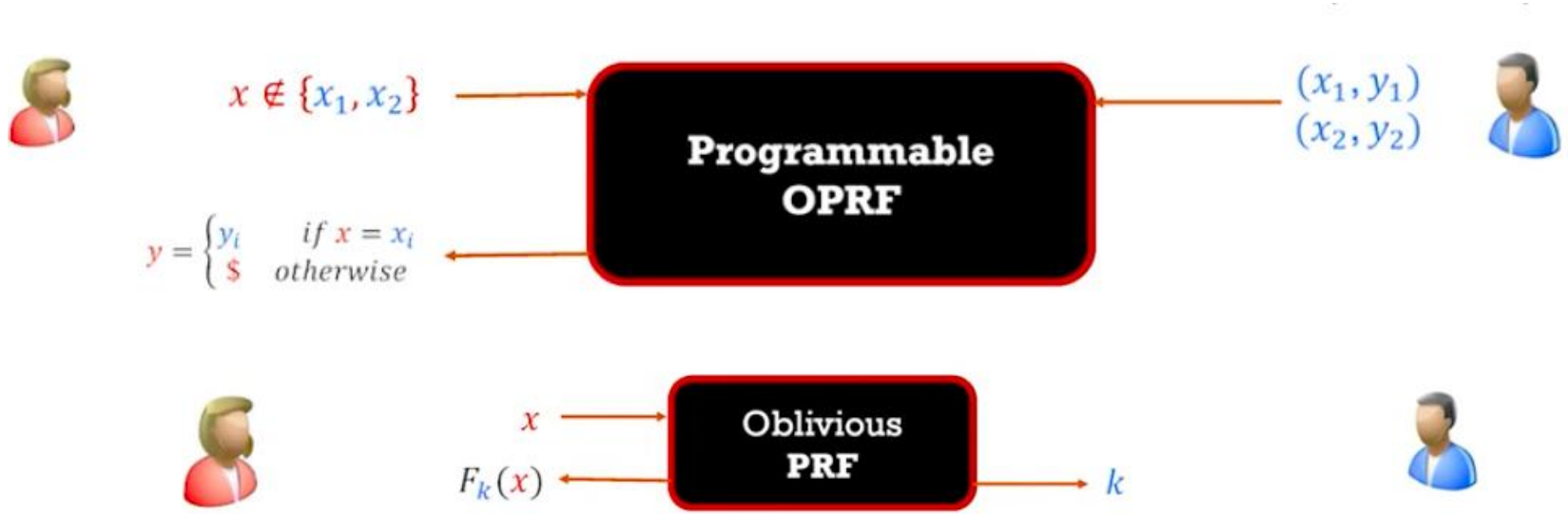
- Zero—Sharing



[63]：Practical Multi—Party Private Set Intersection from Symmetric—Key Techniques

MPSI

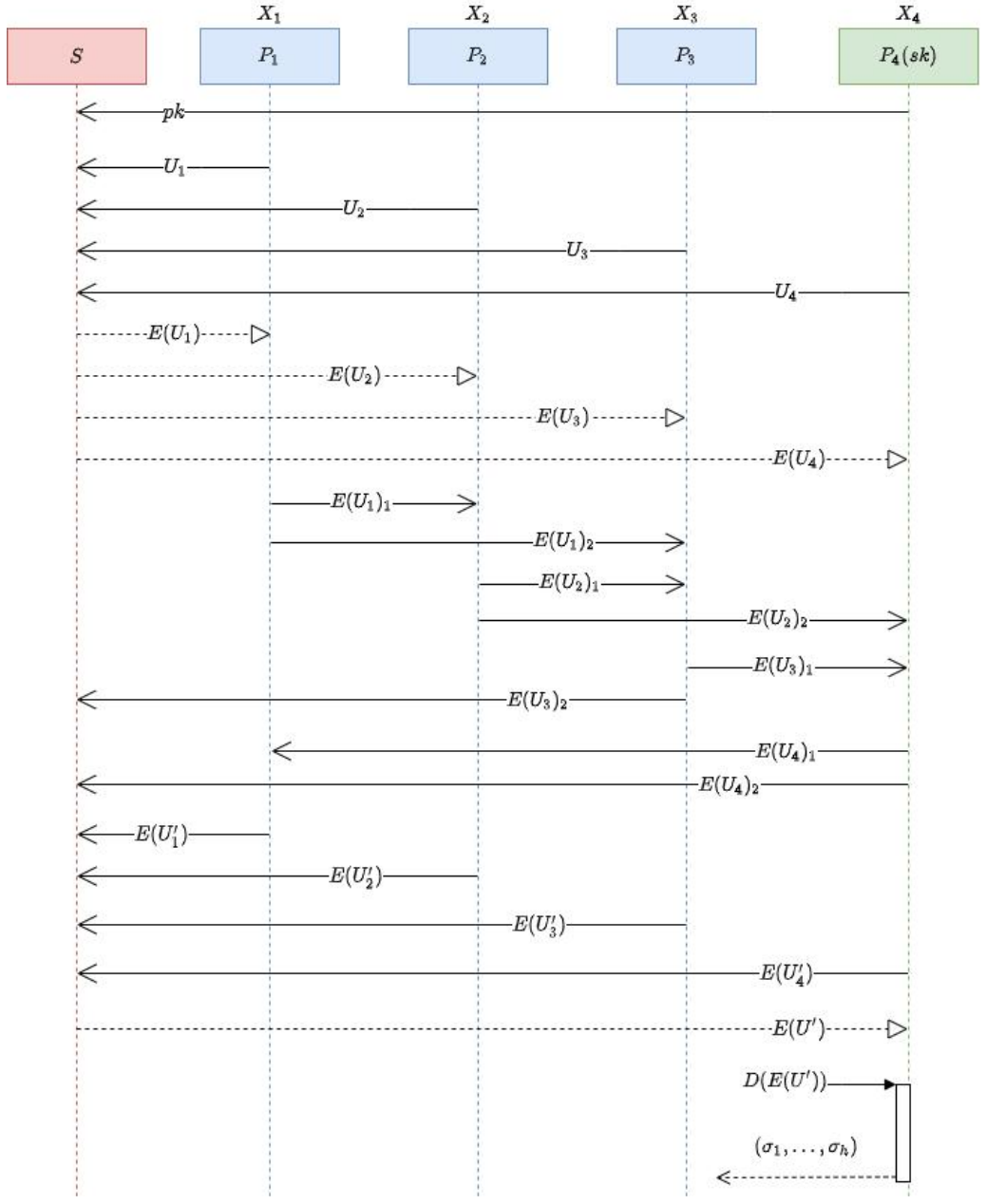
- OPPRF （ Oblivious Programmable PRF ）



[63]：Practical Multi—Party Private Set Intersection from Symmetric—Key Techniques

方案1

- 隐私计算求交/并集
  - 0—r编码
  - 哥德尔编码
  - 密文分割
  - ElGamal加密
  - 半诚实模型



[7]：云环境下集合隐私计算—李顺东(软件学报)

方案1

同盾科技 www.tongdun.cn

- 0—r编码（r—0编码）

集合$X_i \in U = (u_1, \ldots, u_m), (u_1 < \ldots < u_m)$，编码为向量$U_i = (u_{i1}, \ldots, u_{im})$：

```
for i = 0 to m
        if u_i ∈ X_i
        u_ij = r_ij , 其中r_ij ∈ [1, m]
        else
        u_ij = 0
end
```

```
for i = 0 to m
        if u_i ∈ X_i
        u_ij = 0
        else
        u_ij = r_ij , 其中r_ij ∈ [1, m]
end
```

集合	1	2	3	4
U	101	102	103	104
X_1	101	109	103	105
(0—r) U_1	2	0	3	0
(r—0) U_1	0	2	0	3

[7]：云环境下集合隐私计算—李顺东(软件学报)

方案1

- 哥德尔编码

将非负整数序列与自然数建立起一一对应关系

编码：

有穷数列$(a_1,..,a_m)$借助素数序列$(p_1,...,p_m)$，构成编码为一个哥德尔数$[a_1,...,a_m]$

$$[a_1,..,a_m]=\prod_{j=1}^{m}p_j^{a_j}$$

解码：

定理：任意自然数可以唯一分解为多个素数的乘积

已知$(p_1,...,p_m)$，由$[a_1,...,a_m]$求$(a_1,...,a_m)$

例如：$12=2^2*3^1$

[7]：云环境下集合隐私计算-李顺东(软件学报)

方案1

- ElGamal加密

一个传统的公钥加密方案ε包括3种算法:$\mathrm{KeyGen}_\varepsilon$,$\mathrm{Encrypt}_\varepsilon$,$\mathrm{Decrypt}_\varepsilon$.对于$\mathrm{KeyGen}_\varepsilon$算法,给定一个安全参数$\lambda$,输出一个私钥$sk$和对应的公钥$pk$,并给定明文空间$\mathcal{P}$和密文空间$\mathcal{C}$,即:

- $\mathrm{KeyGen}_\varepsilon(\lambda)\leftarrow(sk,pk,\mathcal{P},\mathcal{C})$;
- $\mathrm{Encrypt}_\varepsilon$算法对于给定的公钥$pk$和明文$M\in\mathcal{P}$输出相应的密文$C\in\mathcal{C}$,即:

$$\mathrm{Encrypt}_\varepsilon(pk,M)\leftarrow(C)(M\in\mathcal{P}).$$

- 根据密文C和私钥sk,$\mathrm{Decrypt}_\varepsilon$算法输出相应的明文$M$:

$$\mathrm{Decrypt}_\varepsilon(sk,C)\leftarrow(M)(C\in\mathcal{C}).$$

ElGamal公钥加密算法是具有乘法同态性质的加密算法.对于消息M,其密文可以表示成如下形式:

$$E(M)=(c_1,c_2)=(g^r \bmod p, MH^r \bmod p).$$

其中,r是加密者选择的一个随机数,$h=g^x \bmod p$是私钥x对应的公钥.

对消息M_1,M_2的密文做乘法运算:

$$E(M_1)\times E(M_2)=(g^{r_1},M_1h^{r_1})\times(g^{r_2},M_2h^{r_2})=(g^{r_1+r_2},M_1\times M_2h^{r_1+r_2})=E(M_1\times M_2).$$

即:对M_1,M_2的密文做乘法运算,相当于对M_1,M_2做乘法运算之后再加密.如果$M_1=1$,对M_1的密文与M_2的密文做Evaluate运算后,得到M_2不同的密文.

[7]：云环境下集合隐私计算—李顺东(软件学报)

方案1

同盾科技 www.tongdun.cn

- 密文分割

若P_i直接发送$E(x_i)$给P_n，P_n能直接解密，所以需要混淆密文和参与者的对应关系：

P_i将$E(x_i)$随机分成非零的$k_i(k_i \leq n)$份，即$E(x_i)_1, \ldots, E(x_i)_{k_i}$，且满足$E(x_i) = E(x_i)_1 * \ldots * E(x_i)_{k_i}$，具体来说：

选择k_i个随机数$r_{i1}, \ldots, r_{ik_i}$，满足$r_{i1} * \ldots * r_{ik_i} = 1 \, mod \, p$，从$k_i$个随机数中选择1个随机数与$E(x_i)$相乘构成$E(x_i)_j$，比如$E(x)_1 = r_{i1} * E(x_i)$，$r_{i2}, \ldots, r_{ik_i}$分别表示$E(xz_i)_2, \ldots, E(x_i)_{k_i}$，则：

$$E(x_i)_1 * \ldots * E(x_i)_{k_1} = r_{i1} * E(x_i) * r_{i2} * \ldots * r_{ik_i} = E(x_i)$$

[7]：云环境下集合隐私计算—李顺东(软件学报)

方案1

同盾科技 www.tongdun.cn

- 求交集

输入：$P_1,...,P_n$方的数据$X_1,...,X_n \in U = (u_1,...,u_m)$

输出：$X_1 \cap ... \cap X_n$

1、P_n保留sk，公布$pk, params$和素数集合$P = (p_1,...,p_m)$

2、每个P_i计算：

（1）利用r-0编码将其数据X_i编码为U_i；再用哥德尔编码将其编码为一个自然数x_i

（2）用pk加密x_i：$E(x_i)$

（3）将$E(x_i)$随机分为k_i份发送给所有参与者中的k_i个（可以包含自己）

（4）将收到的所有密文相乘得到$E(X_i')$，并发送给P_n

3、P_n计算：

（1）将收到的密文相乘得到$E(x)$

（2）用sk解密，得到x

（3）将x展开得到向量U'

（4）根据U中为0的分量得到并集，并公布

[7] 云环境下集合隐私计算 李顺东(软件学报)

方案1

同盾科技
www.tongdun.cn

- 总结

1、用户需要知道全集

2、通信和计算量：

（E和D是加解密算法）

计算开销	通信开销
nE+D	n(k+1)

[7]：云环境下集合隐私计算—李顺东(软件学报)

方案2

同盾科技 www.tongdun.cn

- 保密判定交/并集的势与给定阈值的大小关系
- 保密判定交/并集与给定元素的包含关系
- 保密判定交/并集与给定集合的包含关系

  - 0—1编码
  - Left EIGamal门限加密
  - 密文重加密

[8]：集合交集问题的安全计算—赵雪玲（密码学报）

方案2

同盾科技 www.tongdun.cn

- 0—1编码

对于全集$Q = (q_1, \ldots, q_l)$，集合$X = (x_1, \ldots, x_m) \in Q$，将$X$编码为一个数组$M = (m_1, \ldots, m_l)$：

$$m_j = \begin{cases} 1, q_j \in X \\ 0, q_j \notin X \end{cases}$$

例如：

Q=(1,2,3,4,5)
P1：X1=(1,2,3)
P2：X2=(2,3,4)
P3：X3=(3,4,5)
Alice：t=2
对于P1：M1=（1，1，1，0，0）
对于P2：M2=（0，1，1，0，0）
对于P3：M3=（0，0，1，0，0）
L=0+0+1+0+0=1
L-t<0，即输出为0

[8]：集合交集问题的安全计算—赵雪玲（密码学报）

方案2

- Left ElGamal门限加密

密钥生成: 给定安全参数 k, 密钥生成算法生成一个 k 比特的大素数 p, 以及 Z_p^* 的一个生成元 g. n 个参与者 $P_1, P_2, \cdots, P_n$ 商定一个小素数 ρ (lifted ElGamal 门限加密算法解密结果为 ρ^M, 此时, 需进一步计算离散对数, 为简化计算, 选择小素数 ρ. 同时, 本文解密结果均为小数据, 因此可以通过离散对数表获取离散对数计算结果). 选取 $k_i \in Z_p^*$ 作为 P_i $(i = 1, 2, \cdots, n)$ 的私钥, 计算 $K_i = g^{k_i} \bmod p$, 并联合

私钥k_i有n份，解密时需要n份密钥才能正确解密，公钥为$\mathrm{h} = \prod_{i=1}^{n} K_i \bmod p$。

加密：明文$M \in Z_p^*$，随机数r ，$C = (C_1, C_2) = (g^r \bmod p, \rho^M \mathrm{h}^r \bmod p)$

联合解密：$t_i = C_1^{k_i} \bmod p$，$t = \prod_{i=1}^{n} C_i^{k_i} \bmod p = C_1^{\sum_{i=1}^{n} k_i} \bmod p$，$\rho^M = \frac{C_2}{t} \bmod p$

该门限加密方案有“加法同态性”，即$E(m_1).E(m_2) = E(m_1 + m_2)$

[8]：集合交集问题的安全计算—赵雪玲（密码学报）

方案2

同盾科技 www.tongdun.cn

- 密文重随机数

密文重加密，密文更换，对应明文不变。

对密文消息 M 重随机化的具体步骤为: 设明文消息 M 加密后的密文为

$$E(M) = (C_1, C_2) = (g^r \bmod p,\ \rho^M h^r \bmod p).$$

参与者选择随机数 r', 并计算

$$\begin{aligned}(C_1 g^{r'} \bmod p, C_2 h^{r'} \bmod p) &= (g^r g^{r'} \bmod p,\ \rho^M h^r h^{r'} \bmod p)\\ &= (g^{r*} \bmod p,\ \rho^M h^{r*} \bmod p)\\ &= (C_1', C_2') = E(M).\end{aligned}$$

经过随机化后, 相当于对明文 M 重新加密, 和选择的密文不再相同.

[8]：集合交集问题的安全计算–赵雪玲（密码学报）

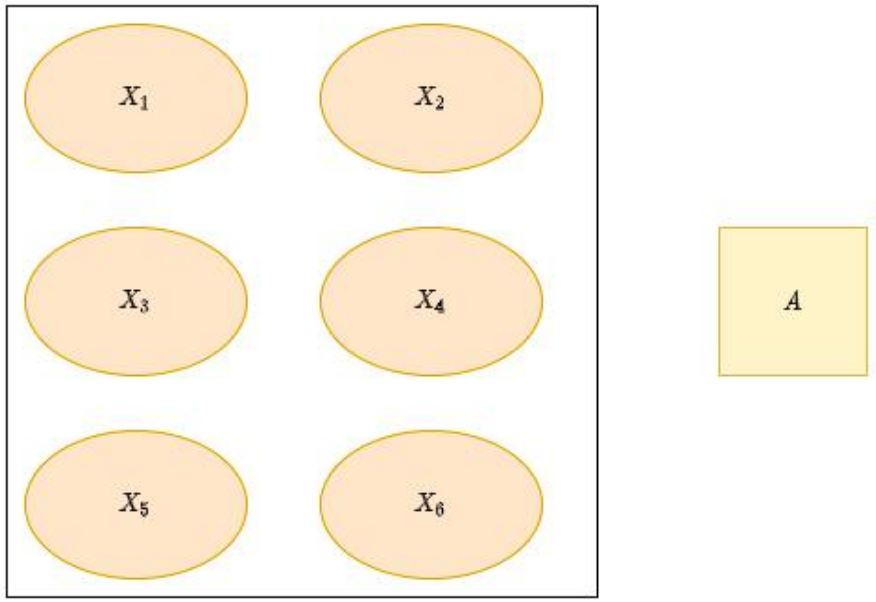
方案2

- 判定交集与集合的包含关系

问题描述: 假设 n 个参与者 $P_i(i=1,2,\cdots,n)$, 分别拥有集合 $X_i=\{x_{i1},x_{i2},\cdots,x_{il_i}\}\subseteq Q$, Alice 拥有集合 $A=\{a_1,a_2,\cdots,a_s\}\subseteq Q$. 其中 $Q=\{q_1,q_2,\cdots,q_l\}$ 且 $q_1<q_2<\cdots<q_l$. n 个参与者和 Alice 想要保密判断集合 A 是否包含于 n 个集合的交集, 且不泄露参与者和集合交集的任何信息.

输入: $P_i(i=1,2,\cdots,n)$ 输入集合 $X_i=\{x_{i1},x_{i2},\cdots,x_{il_i}\}\subseteq Q$. Alice 输入 $A=\{a_1,a_2,\cdots,a_s\}\subseteq Q$.

输出: $F(A,X_1,X_2,\cdots,X_n)$



[8]：集合交集问题的安全计算—赵雪玲（密码学报）

方案2

同盾科技 www.tongdun.cn

- 总结：

1、**用户需要知道全集**

2、输入范围有限

3、通信和计算量

计算开销	通信开销	参与者
$2n(l+1)+5$	$3n+1$	$n+1$

[8]：集合交集问题的安全计算–赵雪玲（密码学报）

方案3

- 使用加法同态性的门限密码算法。
- 使用两方PSI设计多方PSI。
- 给出**半诚实模型的协议**和恶意模型的协议。

- 拥有加法同态性的门限密码算法
- 2PC协议（两方PSI）— Efficient private matching and set intersection—2004
- 改进的2PC协议
- 半诚实模型的协议

[46]：Scalable Multi-Party Private Set-Intersection-2017

方案3

- 拥有加法同态性的门限密码算法

  ElGamal变体和Paillier

- Left ElGamal门限加密

密钥生成: 给定安全参数 k, 密钥生成算法生成一个 k 比特的大素数 p, 以及 Z_p^* 的一个生成元 g. n 个参与者 $P_1, P_2, \cdots, P_n$ 商定一个小素数 ρ (lifted ElGamal 门限加密算法解密结果为 ρ^M, 此时, 需进一步计算离散对数, 为简化计算, 选择小素数 ρ. 同时, 本文解密结果均为小数据, 因此可以通过离散对数表获取离散对数计算结果). 选取 $k_i \in Z_p^*$ 作为 P_i $(i = 1, 2, \cdots, n)$ 的私钥, 计算 $K_i = g^{k_i} \bmod p$, 并联合

私钥k_i有n份，解密时需要n份密钥才能正确解密，公钥为$h = \prod_{i=1}^{n} K_i mod\, p$。

加密：明文$M \in Z_p^*$，随机数r，$C = (C_1, C_2) = (g^r mod p, \rho^M h^r mod\, p)$

联合解密：$t_i = C_1^{k_i} mod\, p$，$t = \prod_{i=1}^{n} C_i^{k_i} mod\, p = C_1^{\sum_{i=1}^{n} k_i} mod\, p$，$\rho^M = \frac{C_2}{t} mod\, p$

该门限加密方案有“加法同态性”，即$E(m_1).E(m_2) = E(m_1 + m_2)$

[46]：Scalable Multi-Party Private Set-Intersection-2017

方案3

- 2PC协议（两方PSI）

（1）P_2计算出多项式$Q(x) = (x - x2_1)...(x - x2_{m_2})$，并将系数加密$C_q$发送给$P_1$

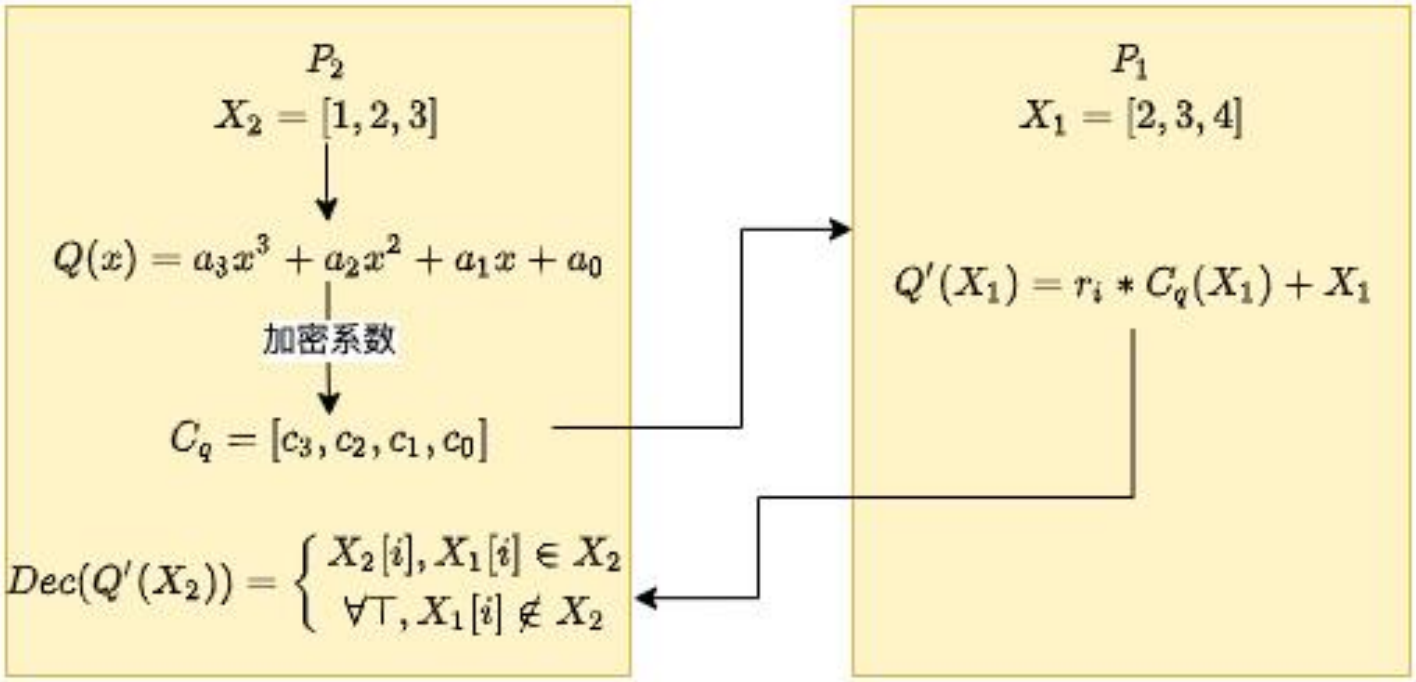
（2）P_1收到C_q后，对于每个元素 $x1_j \in X_1$，同态的计算$Q'(x1_j)$，并发送给P_2

（3）P_2收到$Q'(x1_j)$后，解密，判断解密结果是否在X_2中，如果在则就在交集中，否则不在。

输入：

P_1： $X_1 = (x1_1, ..., x1_{m_1})$

P_2： $X_2 = (x2_1, ..., x2_{m_2})$



[46]：Scalable Multi-Party Private Set-Intersection-2017

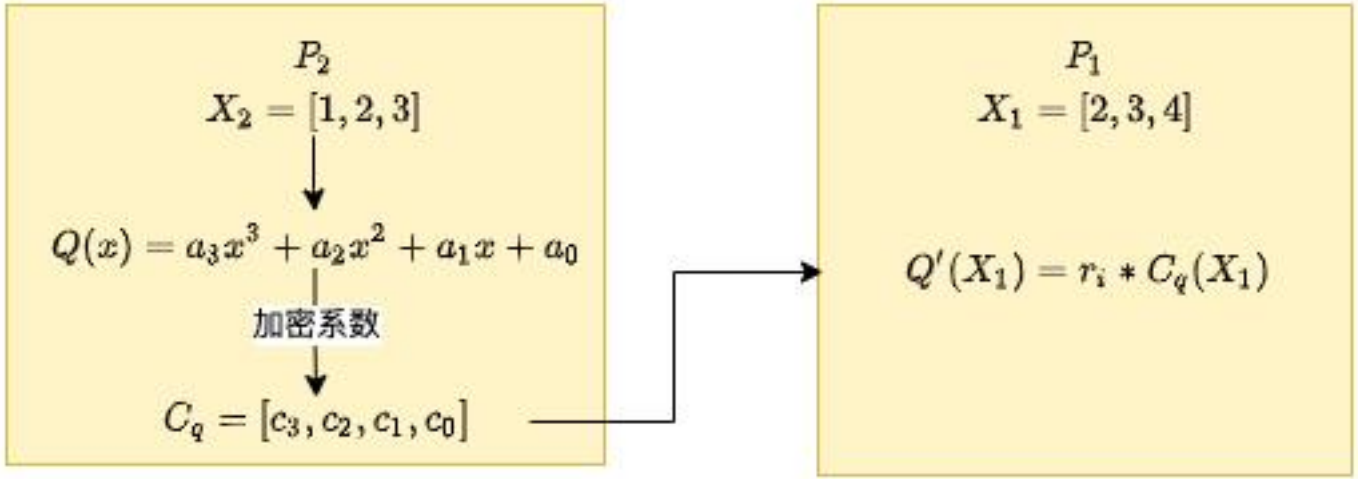
方案3

- 改进的2PC协议

（1）P_2计算出多项式$Q(x) = (x - x2_1)...(x - x2_{m_2})$，并将系数加密$C_q$发送给$P_1$

（2）P_1收到C_q后，对于每个元素 $x1_j \in X_1$，

同态的计算$Q'(x1_j)$，并保存



[46]：Scalable Multi-Party Private Set-Intersection-2017

方案3

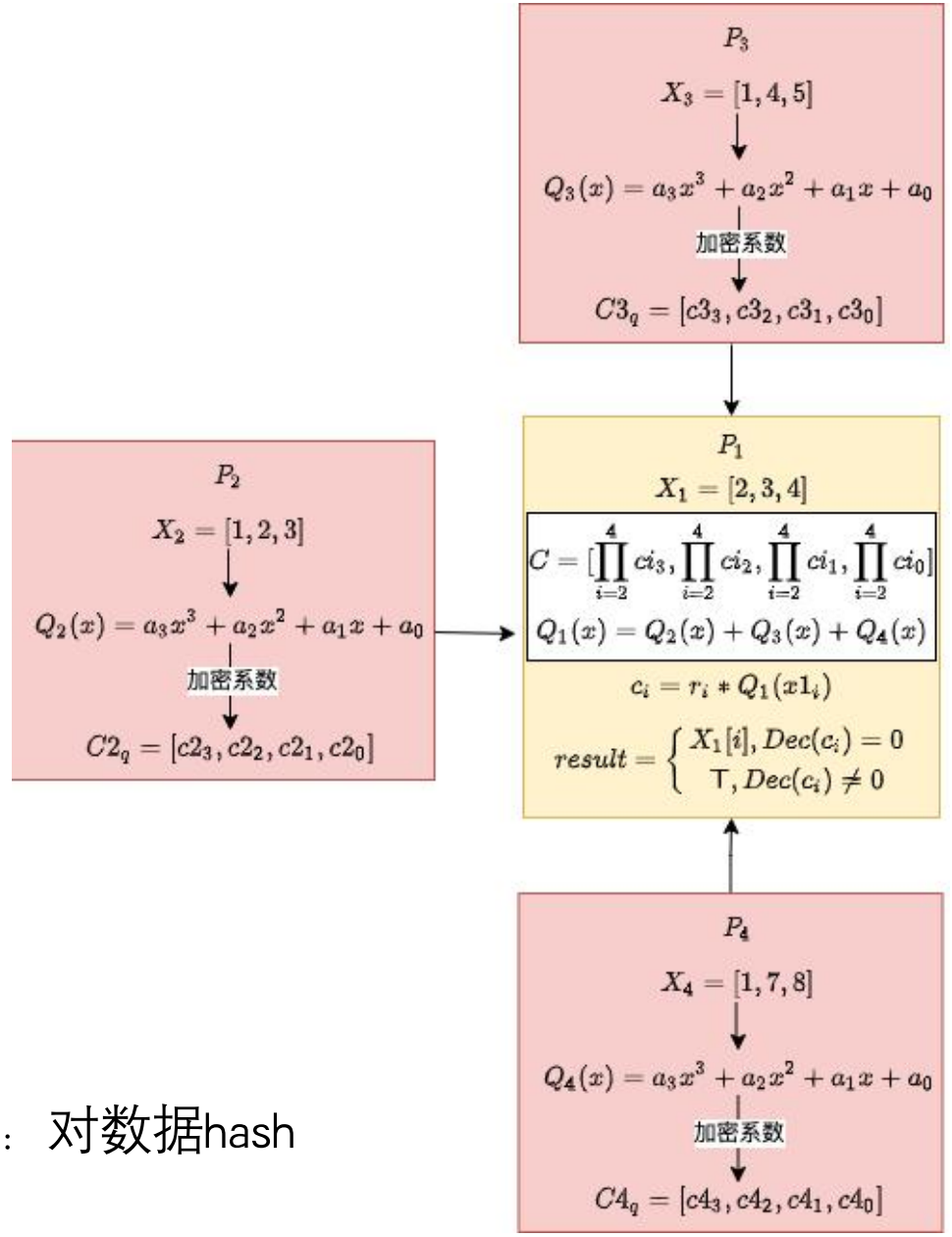
- 半诚实模型的协议

协议分为三步：

（1）参与方协商出一个公钥PK，每个参与者有一个私钥SK_i

（2）P_1和其它参与者逐一执行交互（改进的2PC协议），获得加密的多项式

（3）P_1执行求交，并**联合**所有参与方解密计算结果，根据计算结果是否为0判断数据是否属于交集！



通信量：$O(m_i * n)$

计算量：n * $O(m_i)$指数

为了减少计算和通信，可以数据进行预处理：对数据hash

[46]：Scalable Multi-Party Private Set-Intersection-2017

Thanks