電子情報通信学会総合大会 2024

秘匿検索のための暗号技術：Private Information Retrievalとその応用について

2024 年 3 月 7 日

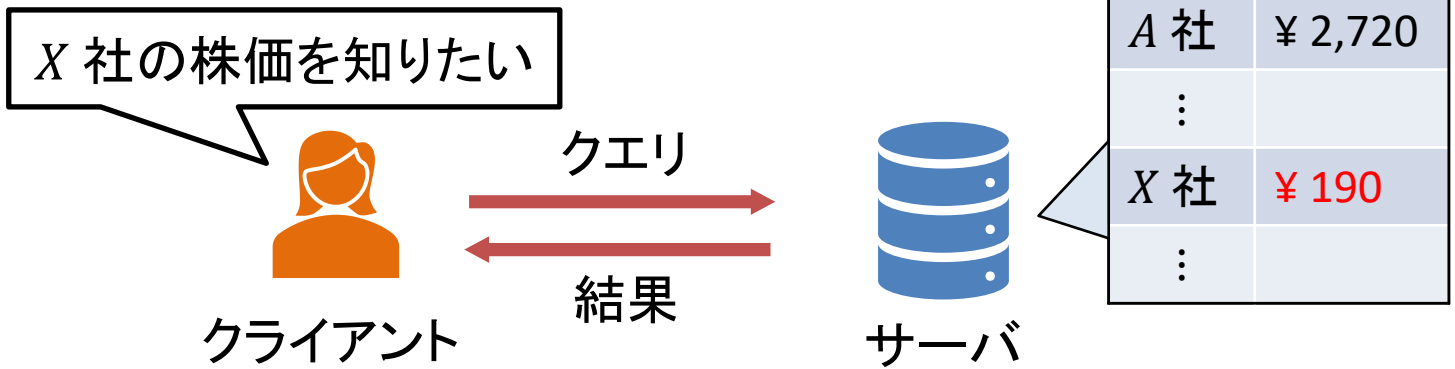
江利口礼央（産業技術総合研究所）

秘匿検索

- 近年データベースビジネス（検索サービス等）が増加
  - 株価データ、化合物データ、ゲノムデータ…

- クエリ内容のプライバシ保護が重要

例



投資戦略が漏洩するとクライアントの損害につながる危険性

秘匿検索

- **自明な解決方法**

  データベース全体をローカルに保存し、検索を行う。

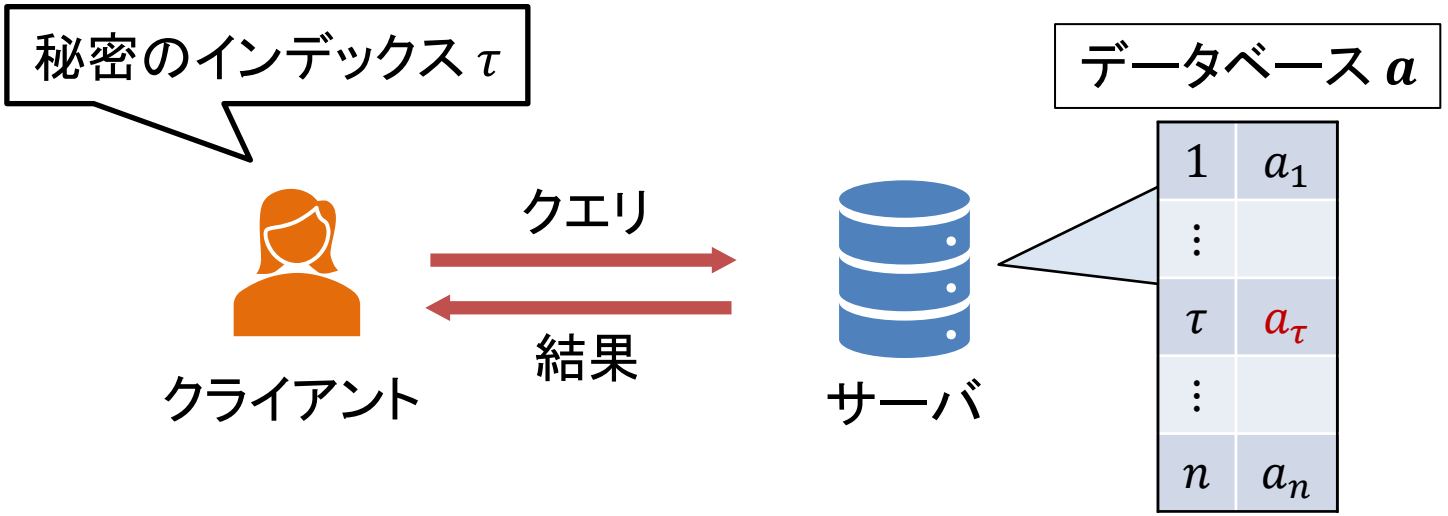
- しかし、データベースのサイズに比例した通信量・計算量が必要

課題

データベースのサイズより小さい通信量で秘匿検索は可能であるか？

Private Information Retrieval (PIR)

- 最も基本的なインデックスクエリを実現する秘匿検索技術 [CGKS98]



- より高度な検索機能を実現するものもある。
  - キーワード検索 [BI05]、レンジクエリ [BGI16] 、任意のクエリ [DHRW16]

不可能性

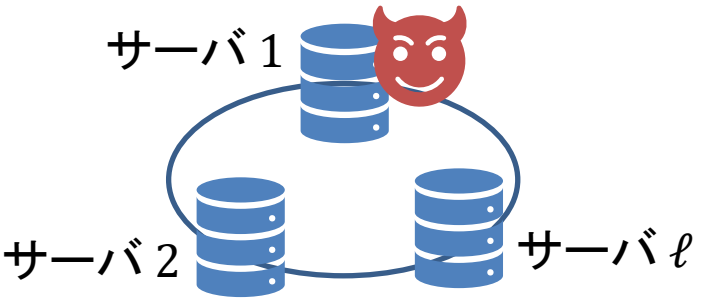
[CGKS98]

PIR 方式が
　サーバが**一台しか**存在せず かつ **情報理論的**安全性を満たす ならば
通信量は n（データベースサイズ）より大きい。

本発表

ℓ-Server PIR

同じ DB を持つ複数台のサーバを仮定



Computational PIR

計算量的仮定をおく

- 数論的仮定
  [KO97, CMS99, Cha04, Lip05]
- 耐量子仮定
  [GR05, CK20, CHK22, ZLTS23]

Multi-Server PIR

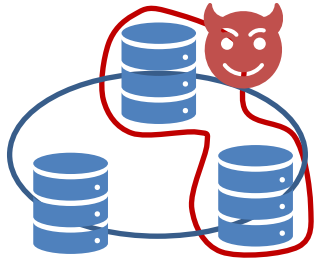
これまでの研究動向

1-Private ℓ-server PIR

多項式通信量 $n^{\Omega(1)}$

$O(n^{1/\ell})$ [CGKS98]

t-Private ℓ-server PIR



これまでの研究動向

1-Private ℓ-server PIR

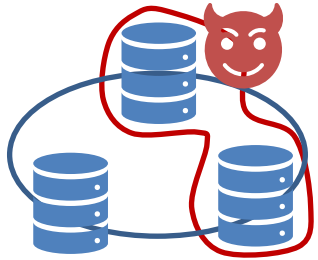
多項式通信量 $n^{\Omega(1)}$

$O(n^{1/\ell})$ [CGKS98]

$O(n^{1/(2\ell-1)})$ [Amb97]

$n^{o\left(\frac{\log\log\ell}{\ell\log\ell}\right)}$ [BIKR02]

t-Private ℓ-server PIR



これまでの研究動向

1-Private ℓ-server PIR

多項式通信量 $n^{\Omega(1)}$

$O(n^{1/\ell})$ [CGKS98]

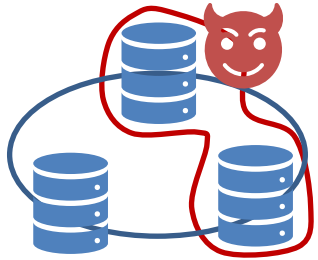
$O(n^{1/(2\ell-1)})$ [Amb97]

$n^{o\left(\frac{\log\log\ell}{\ell\log\ell}\right)}$ [BIKR02]

劣多項式通信量 $n^{o(1)}$

$\ell = 3$ [Yek08, Efr12, IS10]

t-Private ℓ-server PIR



これまでの研究動向

1-Private ℓ-server PIR

多項式通信量 $n^{\Omega(1)}$

$O(n^{1/\ell})$ [CGKS98]

$O(n^{1/(2\ell-1)})$ [Amb97]

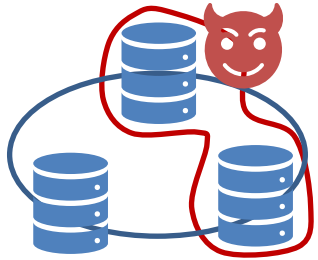
$n^{O\left(\frac{\log\log\ell}{\ell\log\ell}\right)}$ [BIKR02]

劣多項式通信量 $n^{o(1)}$

$\ell = 3$ [Yek08, Efr12, IS10]

$\ell = 2$ [DG16]

t-Private ℓ-server PIR



これまでの研究動向

1-Private ℓ-server PIR

多項式通信量 $n^{\Omega(1)}$

$O(n^{1/\ell})$ [CGKS98]

$O(n^{1/(2\ell-1)})$ [Amb97]

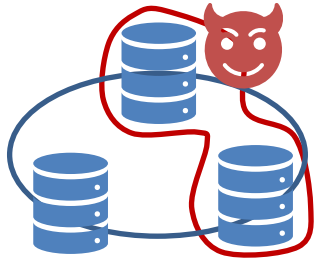
$n^{O\left(\frac{\log\log\ell}{\ell\log\ell}\right)}$ [BIKR02]

劣多項式通信量 $n^{o(1)}$

$\ell = 3$ [Yek08, Efr12, IS10]

$\ell = 2$ [DG16]

t-Private ℓ-server PIR



多項式通信量

通信量 $n^{O(1/d)} = n^{O(t/\ell)}$

サーバ数 $\ell = O(td)$ [WY07]

これまでの研究動向

1-Private ℓ-server PIR

多項式通信量 $n^{\Omega(1)}$

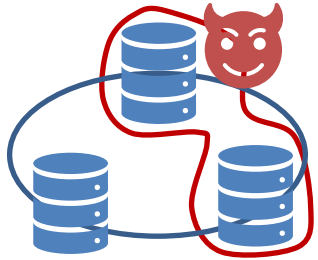
- $O(n^{1/\ell})$ [CGKS98]
- $O(n^{1/(2\ell-1)})$ [Amb97]
- $n^{O\left(\frac{\log\log\ell}{\ell\log\ell}\right)}$ [BIKR02]

劣多項式通信量 $n^{o(1)}$

- $\ell = 3$ [Yek08, Efr12, IS10]
- $\ell = 2$ [DG16]

一般的変換 [BIW10]

t-Private ℓ-server PIR



多項式通信量

- 通信量 $n^{O(1/d)} = n^{O(t/\ell)}$
- サーバ数 $\ell = O(td)$ [WY07]

これまでの研究動向

1-Private ℓ-server PIR

多項式通信量 $n^{\Omega(1)}$

$O(n^{1/\ell})$ [CGKS98]

$O(n^{1/(2\ell-1)})$ [Amb97]

$n^{O\left(\frac{\log\log\ell}{\ell\log\ell}\right)}$ [BIKR02]

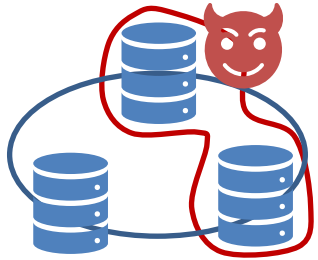
劣多項式通信量 $n^{o(1)}$

$\ell = 3$ [Yek08, Efr12, IS10]

$\ell = 2$ [DG16]

一般的変換 [BIW10]

t-Private ℓ-server PIR



多項式通信量

通信量 $n^{O(1/d)} = n^{O(t/\ell)}$

サーバ数 $\ell = O(td)$ [WY07]

劣多項式通信量

通信量 $n^{o(1)}$

サーバ数 $\ell = 2^{O(t)}$

これまでの研究動向

1-Private ℓ-server PIR

多項式通信量 $n^{\Omega(1)}$

$O(n^{1/\ell})$ [CGKS98]

$O(n^{1/(2\ell-1)})$ [Amb97]

$n^{O\left(\frac{\log\log\ell}{\ell\log\ell}\right)}$ [BIKR02]

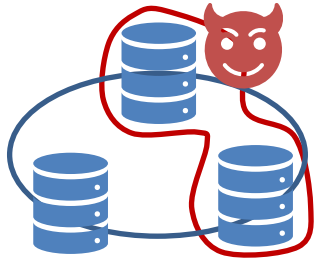
劣多項式通信量 $n^{o(1)}$

$\ell = 3$ [Yek08, Efr12, IS10]

$\ell = 2$ [DG16]

一般的変換 [BIW10]

t-Private ℓ-server PIR



多項式通信量

通信量 $n^{O(1/d)} = n^{O(t/\ell)}$

サーバ数 $\ell = O(td)$ [WY07]

現在最良の方式

劣多項式通信量

通信量 $n^{o(1)}$

サーバ数 $\ell = 2^{O(t)}$

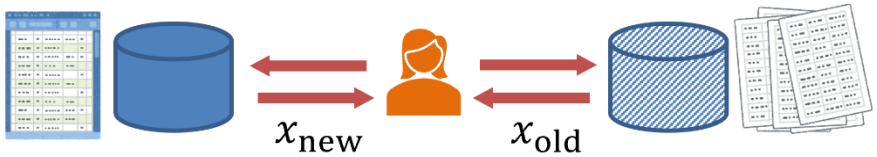
PIR における能動的安全性

能動的安全性

- **Actively secure PIR**

  サーバの応答に誤りが含まれる場合でも正当性を維持する

例. データベースの同期失敗



Active t-security $\Pr[\mathcal{D}(\widetilde{\mathrm{ans}}_1, \ldots, \widetilde{\mathrm{ans}}_\ell) = a_\tau] \geq 1 - 2^{-\lambda}$

t 個の誤りが含まれる

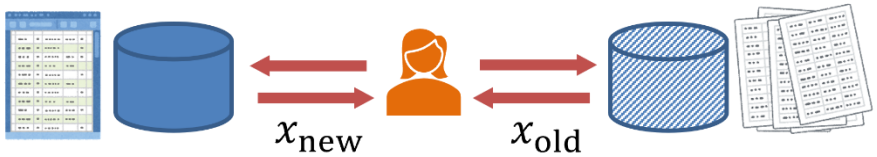
- 1 ラウンド通信を考える場合プライバシへの影響はない。
- 一般的変換 [BS07, EKN@TCC2022, EKN@EUROCRYPT2024]

能動的安全性

- **Actively secure PIR**

  サーバの応答に誤りが含まれる場合でも正当性を維持する

例. データベースの同期失敗



Active t-security $\Pr[\mathcal{D}(\widetilde{\mathrm{ans}}_1, \ldots, \widetilde{\mathrm{ans}}_\ell) = a_\tau] \geq 1 - 2^{-\lambda}$

t 個の誤りが含まれる

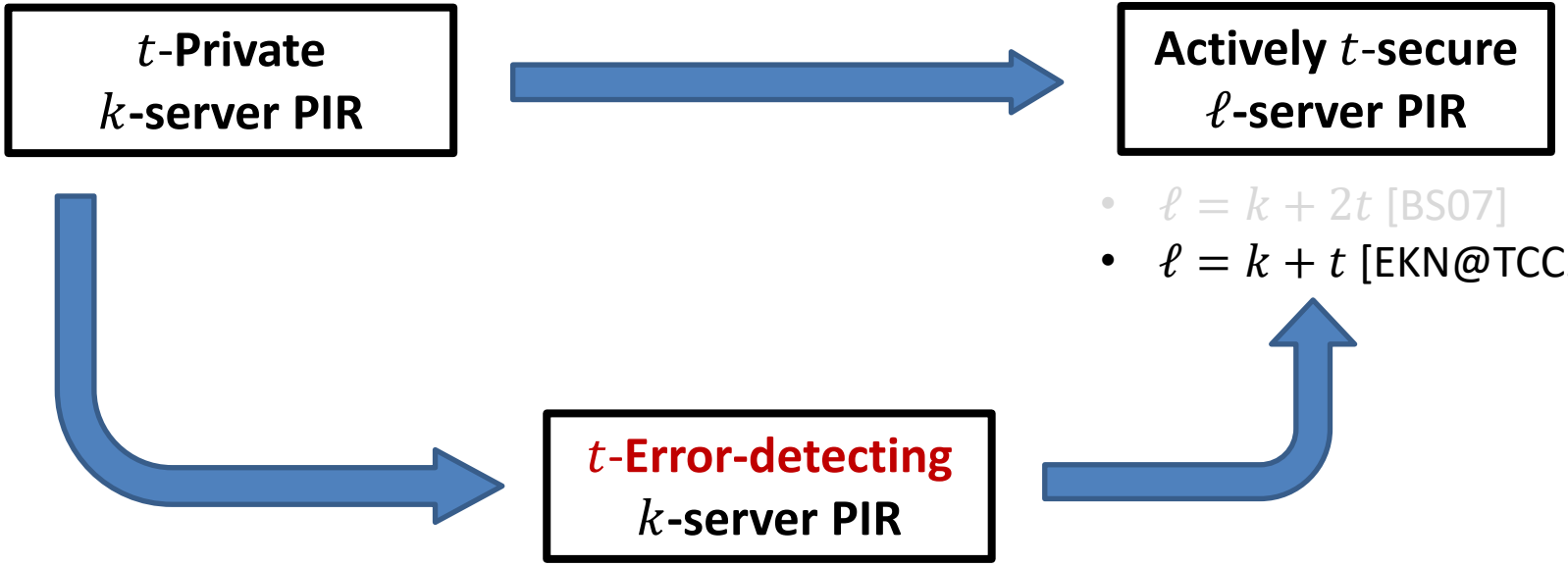
  - 1 ラウンド通信を考える場合プライバシへの影響はない。
  - 一般的変換 [BS07, EKN@TCC2022, EKN@EUROCRYPT2024]

これから説明

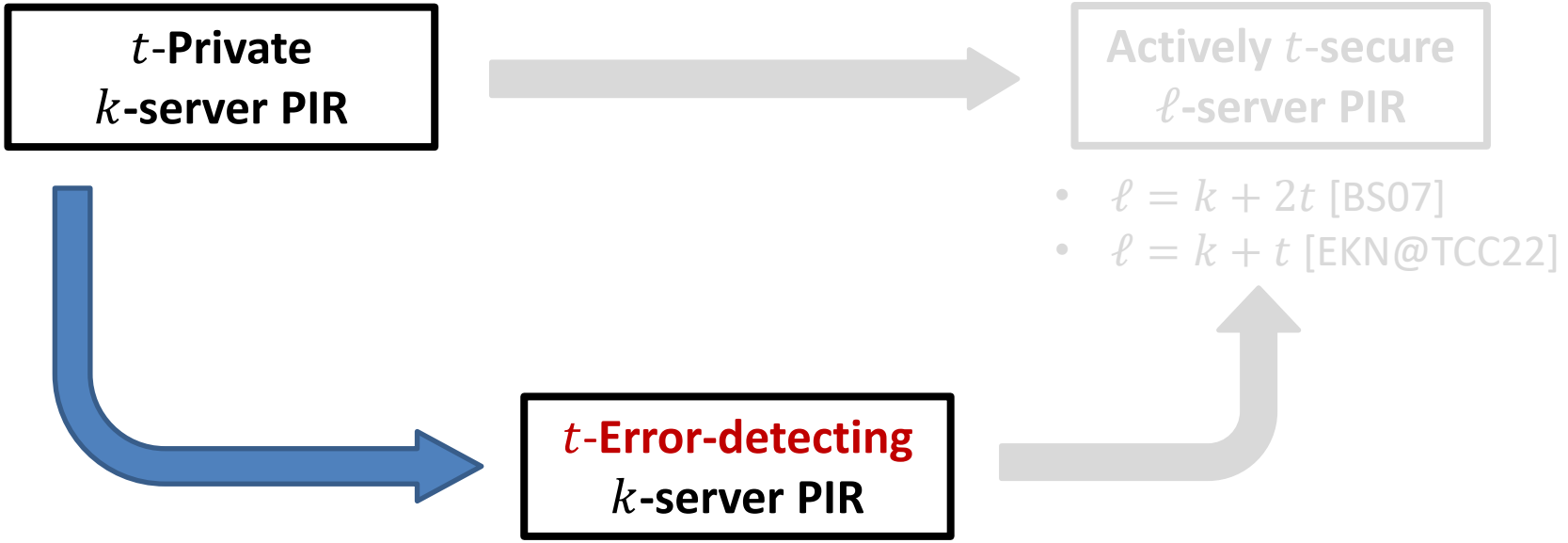
一般的変換

- $\ell = k + 2t$ [BS07]
- $\ell = k + t$ [EKN@TCC22]

一般的変換


t-Private
k-server PIR
Actively t-secure
ℓ-server PIR
• $\ell = k + 2t$ [BS07]
• $\ell = k + t$ [EKN@TCC22]
t-Error-detecting
k-server PIR
$\Pr[\mathcal{D}(\widetilde{\mathrm{ans}}_1, \ldots, \widetilde{\mathrm{ans}}_k) \in \{a_\tau, \perp\}] \geq 1 - 2^{-\lambda}$
t 個の誤りが含まれる
誤りを検知

一般的変換

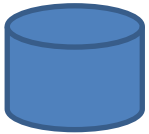

t-Private
k-server PIR
Actively t-secure
ℓ-server PIR
• $\ell = k + 2t$ [BS07]
• $\ell = k + t$ [EKN@TCC22]
t-Error-detecting
k-server PIR
$\Pr[\mathcal{D}(\widetilde{\mathrm{ans}}_1, \ldots, \widetilde{\mathrm{ans}}_k) \in \{a_\tau, \perp\}] \geq 1 - 2^{-\lambda}$
t 個の誤りが含まれる
誤りを検知

Error-detecting PIR への変換

ここでは S_1 が honest (悪者ではない)と分かっているとする。

Honest

S_1



Malicious

S_2

Error-detecting PIR への変換

ここでは S_1 が honest (悪者ではない)と分かっているとする。

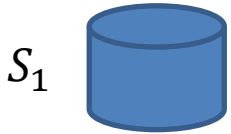
計算用のクエリ

$$\left(\mathrm{que}_1(\mathrm{C}), \mathrm{que}_2(\mathrm{C})\right) \leftarrow Q(\tau)$$

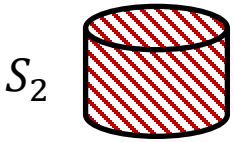
検証用のクエリ

$$\left(\mathrm{que}_1(\mathrm{V}), \mathrm{que}_2(\mathrm{V})\right) \leftarrow Q(\tau)$$

Honest

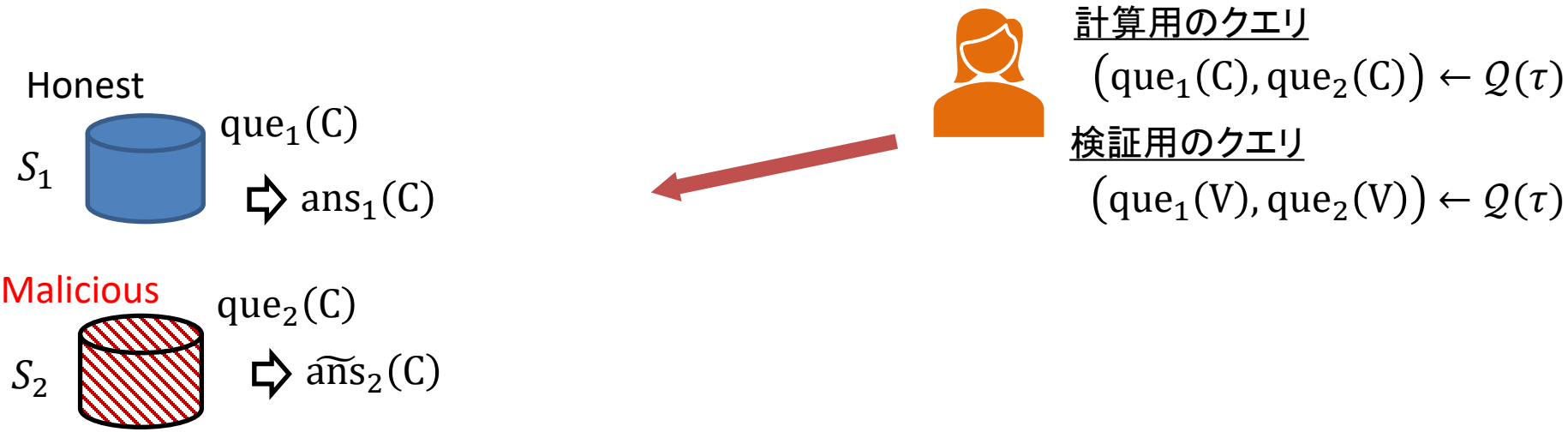


Malicious



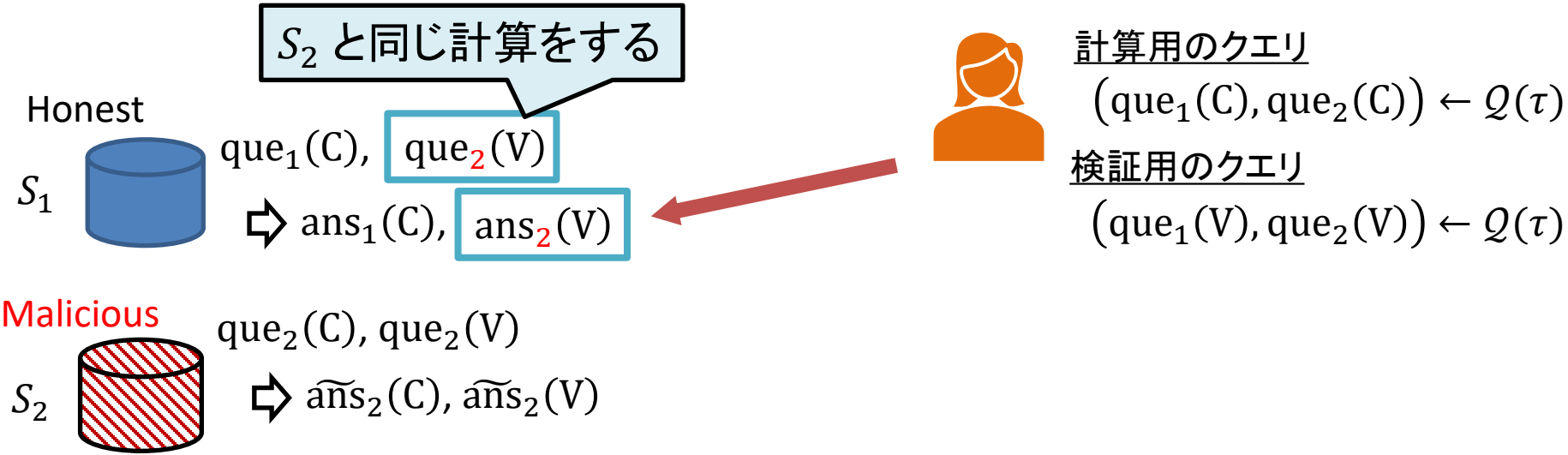
Error-detecting PIR への変換

ここでは S_1 が honest (悪者ではない)と分かっているとする。



Error-detecting PIR への変換

ここでは S_1 が honest (悪者ではない)と分かっているとする。



Error-detecting PIR への変換

1-**Private**
2-**server PIR**

→

1-**Error-detecting**
2-**server PIR**

ここでは S_1 が honest (悪者ではない)と分かっているとする。

S_2 と同じ計算をする

Honest

S_1

$\mathrm{que}_1(\mathrm{C})$, $\mathrm{que}_2(\mathrm{V})$

⇨ $\mathrm{ans}_1(\mathrm{C})$, $\mathrm{ans}_2(\mathrm{V})$

計算用のクエリ

$\left(\mathrm{que}_1(\mathrm{C}), \mathrm{que}_2(\mathrm{C})\right) \leftarrow \mathcal{Q}(\tau)$

検証用のクエリ

$\left(\mathrm{que}_1(\mathrm{V}), \mathrm{que}_2(\mathrm{V})\right) \leftarrow \mathcal{Q}(\tau)$

Malicious

S_2

$\mathrm{que}_2(\mathrm{C}), \mathrm{que}_2(\mathrm{V})$

⇨ $\widetilde{\mathrm{ans}}_2(\mathrm{C}), \widetilde{\mathrm{ans}}_2(\mathrm{V})$

もし $\widetilde{\mathrm{ans}}_2(\mathrm{V}) = \mathrm{ans}_2(\mathrm{V})$ なら
$\mathcal{D}\left(\mathrm{ans}_1(\mathrm{C}), \widetilde{\mathrm{ans}}_2(\mathrm{C})\right)$ を出力する

Error-detecting PIR への変換

1-**Private**
2-**server PIR**

→

1-**Error-detecting**
2-**server PIR**

ここでは S_1 が honest (悪者ではない)と分かっているとする。

S_2 と同じ計算をする

Honest

S_1

$\mathrm{que}_1(\mathrm{C}),\ \mathrm{que}_2(\mathrm{V})$

$\Rightarrow \mathrm{ans}_1(\mathrm{C}),\ \mathrm{ans}_2(\mathrm{V})$

計算用のクエリ

$\left(\mathrm{que}_1(\mathrm{C}), \mathrm{que}_2(\mathrm{C})\right) \leftarrow \mathcal{Q}(\tau)$

検証用のクエリ

$\left(\mathrm{que}_1(\mathrm{V}), \mathrm{que}_2(\mathrm{V})\right) \leftarrow \mathcal{Q}(\tau)$

二つのクエリをシャッフル

Malicious

S_2

$\mathrm{que}_2(\mathrm{C}),\ \mathrm{que}_2(\mathrm{V})$

$\Rightarrow \widetilde{\mathrm{ans}}_2(\mathrm{C}),\ \widetilde{\mathrm{ans}}_2(\mathrm{V})$

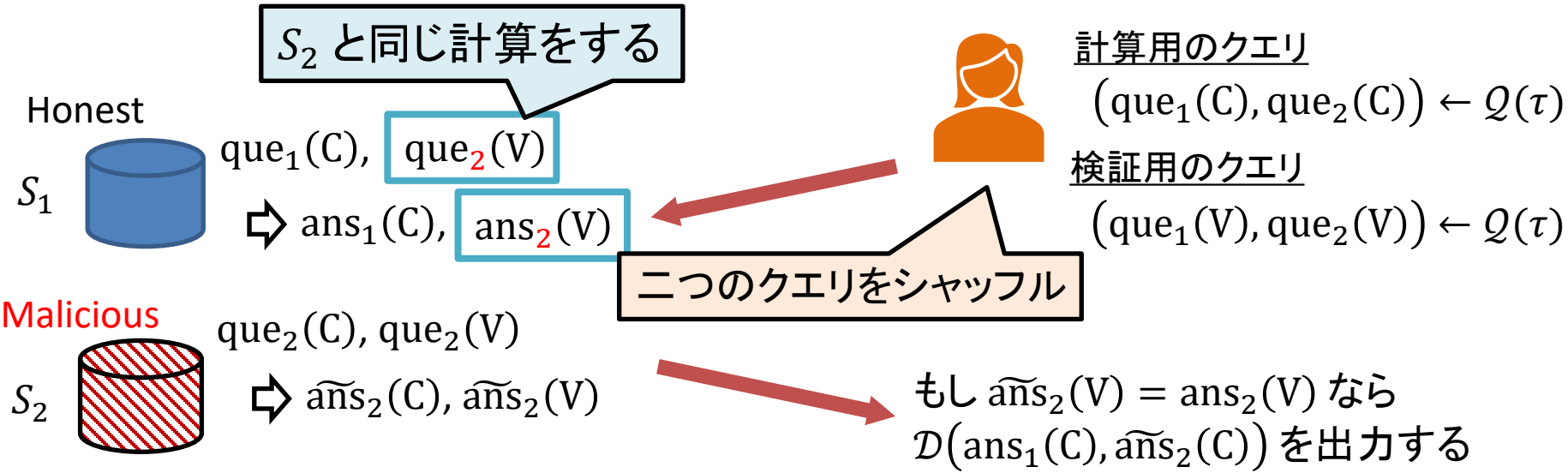
もし $\widetilde{\mathrm{ans}}_2(\mathrm{V}) = \mathrm{ans}_2(\mathrm{V})$ なら

$\mathcal{D}\left(\mathrm{ans}_1(\mathrm{C}), \widetilde{\mathrm{ans}}_2(\mathrm{C})\right)$ を出力する

Error-detecting PIR への変換

1-Private 2-server PIR → 1-Error-detecting 2-server PIR

ここでは S_1 が honest (悪者ではない)と分かっているとする。



S_2 はどちらが計算用のクエリか当てなければならない → 検証失敗確率 $\leq 1/2$

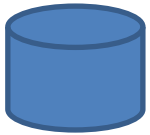
Error-detecting PIR への変換

1-Private
2-server PIR

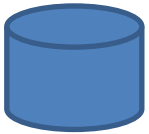
1-Error-detecting
2-server PIR

If $S_i = S_1$:

S_1



S_2



$S_i \leftarrow_{\$} \{S_1, S_2\}$

S_i が honest だと（当てずっぽうで）推測

Error-detecting PIR への変換

1-**Private**
2-**server PIR**

1-**Error-detecting**
2-**server PIR**

If $S_i = S_1$:

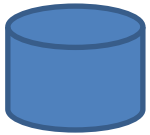
$\left(\mathrm{que}_1(\mathrm{C}), \mathrm{que}_2(\mathrm{C})\right) \leftarrow Q(\tau)$

$\left(\mathrm{que}_1(\mathrm{V}), \mathrm{que}_2(\mathrm{V})\right) \leftarrow Q(\tau)$

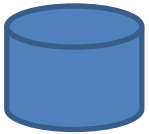
$S_i \leftarrow_{\$} \{S_1, S_2\}$

S_i が honest だと（当てずっぽうで）推測

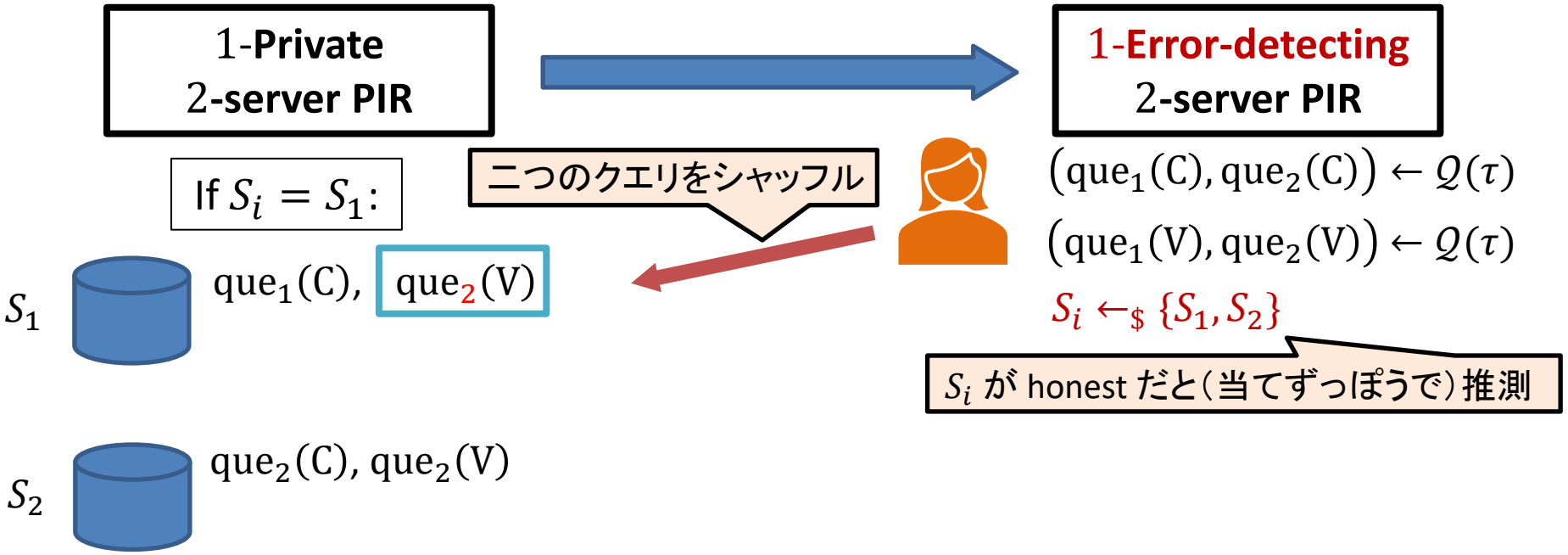
S_1



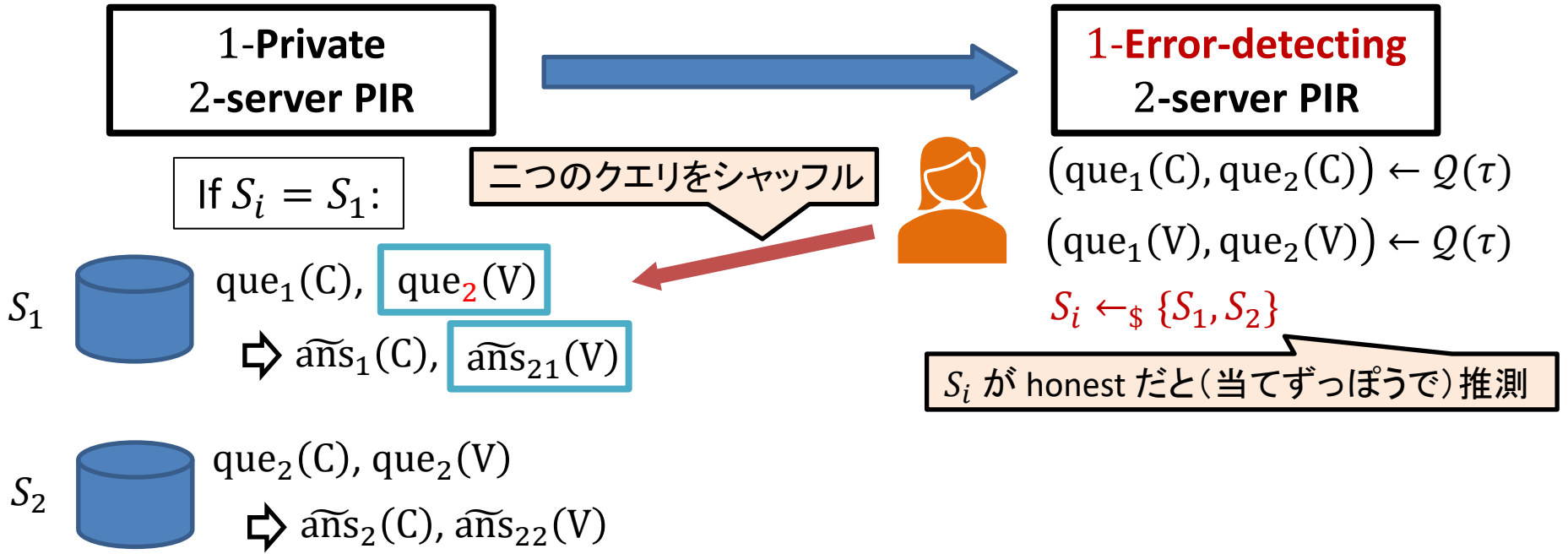
S_2



Error-detecting PIR への変換



Error-detecting PIR への変換



Error-detecting PIR への変換

1-Private
2-server PIR

1-Error-detecting
2-server PIR

$\left(\mathrm{que}_1(\mathrm{C}), \mathrm{que}_2(\mathrm{C})\right) \leftarrow \mathcal{Q}(\tau)$

$\left(\mathrm{que}_1(\mathrm{V}), \mathrm{que}_2(\mathrm{V})\right) \leftarrow \mathcal{Q}(\tau)$

二つのクエリをシャッフル

$S_i \leftarrow_{\$} \{S_1, S_2\}$

S_i が honest だと（当てずっぽうで）推測

If $S_i = S_1$:

S_1

$\mathrm{que}_1(\mathrm{C})$, $\mathrm{que}_2(\mathrm{V})$

⇨ $\widetilde{\mathrm{ans}}_1(\mathrm{C})$, $\widetilde{\mathrm{ans}}_{21}(\mathrm{V})$

S_2

$\mathrm{que}_2(\mathrm{C})$, $\mathrm{que}_2(\mathrm{V})$

⇨ $\widetilde{\mathrm{ans}}_2(\mathrm{C})$, $\widetilde{\mathrm{ans}}_{22}(\mathrm{V})$

もし $\widetilde{\mathrm{ans}}_{21}(\mathrm{V}) = \widetilde{\mathrm{ans}}_{22}(\mathrm{V})$ なら
$\mathcal{D}\left(\widetilde{\mathrm{ans}}_1(\mathrm{C}), \widetilde{\mathrm{ans}}_2(\mathrm{C})\right)$ を出力する

Error-detecting PIR への変換

1-Private
2-server PIR

1-Error-detecting
2-server PIR

$\left(\mathrm{que}_1(\mathrm{C}), \mathrm{que}_2(\mathrm{C})\right) \leftarrow \mathcal{Q}(\tau)$

$\left(\mathrm{que}_1(\mathrm{V}), \mathrm{que}_2(\mathrm{V})\right) \leftarrow \mathcal{Q}(\tau)$

二つのクエリをシャッフル

If $S_i = S_1$:

S_1 $\mathrm{que}_1(\mathrm{C})$, $\mathrm{que}_2(\mathrm{V})$

$\Rightarrow \widetilde{\mathrm{ans}}_1(\mathrm{C})$, $\widetilde{\mathrm{ans}}_{21}(\mathrm{V})$

$S_i \leftarrow_{\$} \{S_1, S_2\}$

S_i が honest だと（当てずっぽうで）推測

S_2 $\mathrm{que}_2(\mathrm{C})$, $\mathrm{que}_2(\mathrm{V})$

$\Rightarrow \widetilde{\mathrm{ans}}_2(\mathrm{C})$, $\widetilde{\mathrm{ans}}_{22}(\mathrm{V})$

もし $\widetilde{\mathrm{ans}}_{21}(\mathrm{V}) = \widetilde{\mathrm{ans}}_{22}(\mathrm{V})$ なら
$\mathcal{D}\left(\widetilde{\mathrm{ans}}_1(\mathrm{C}), \widetilde{\mathrm{ans}}_2(\mathrm{C})\right)$ を出力する

検証失敗確率 $\leq \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} = \frac{3}{4} < 1$

並列実行によって
成功確率を増幅

Honest サーバの推測に失敗

攻撃者が計算用のクエリを当てる

一般的変換

t-Private
k-server PIR

Actively t-secure
ℓ-server PIR

- $\ell = k + 2t$ [BS07]
- $\ell = k + t$ [EKN@TCC22]

t-Error-detecting
k-server PIR

$$\Pr[\mathcal{D}(\underbrace{\widetilde{\mathrm{ans}}_1, \ldots, \widetilde{\mathrm{ans}}_k}) \in \{a_\tau, \perp\}] \geq 1 - 2^{-\lambda}$$

t 個の誤りが含まれる

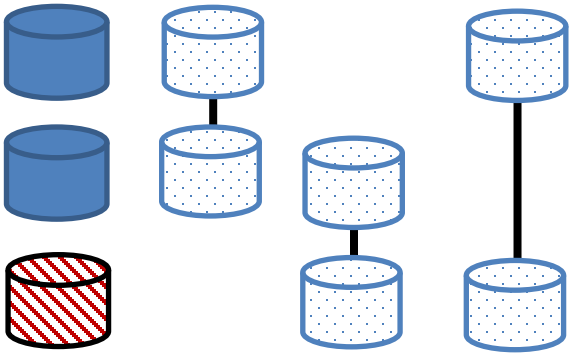
誤りを検知

Actively Secure PIR への変換

t-Error-detecting k-server PIR → **Actively t-secure $(k+t)$-server PIR**

Error-detecting PIR を k 台のサーバからなる全ての部分集合に対して実行する

Honest

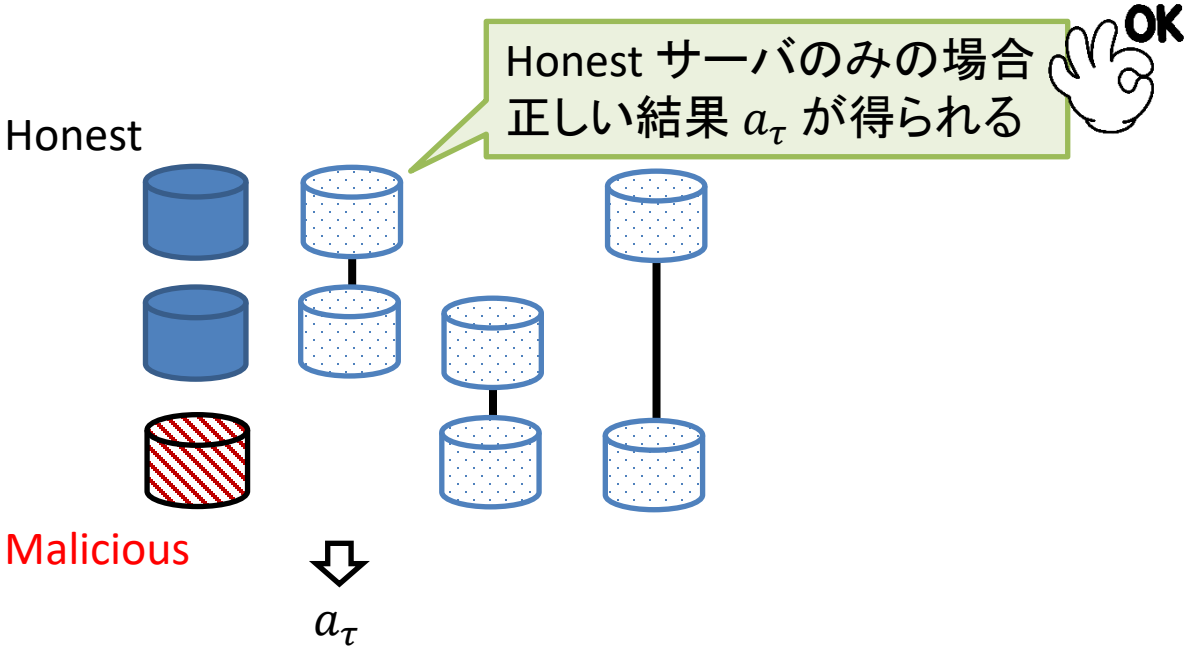


Malicious

Actively Secure PIR への変換

t-Error-detecting k-server PIR → **Actively t-secure $(k+t)$-server PIR**

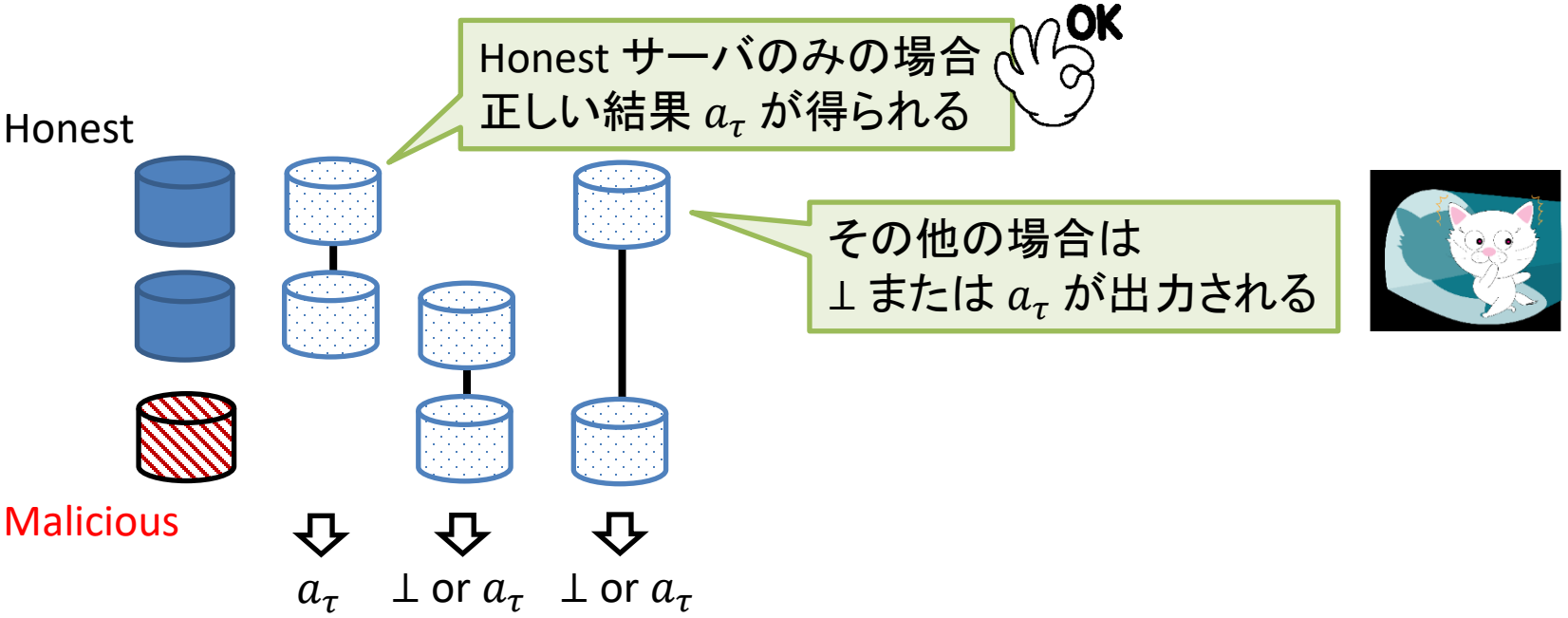
Error-detecting PIR を k 台のサーバからなる全ての部分集合に対して実行する



Actively Secure PIR への変換

t-Error-detecting k-server PIR → **Actively t-secure $(k+t)$-server PIR**

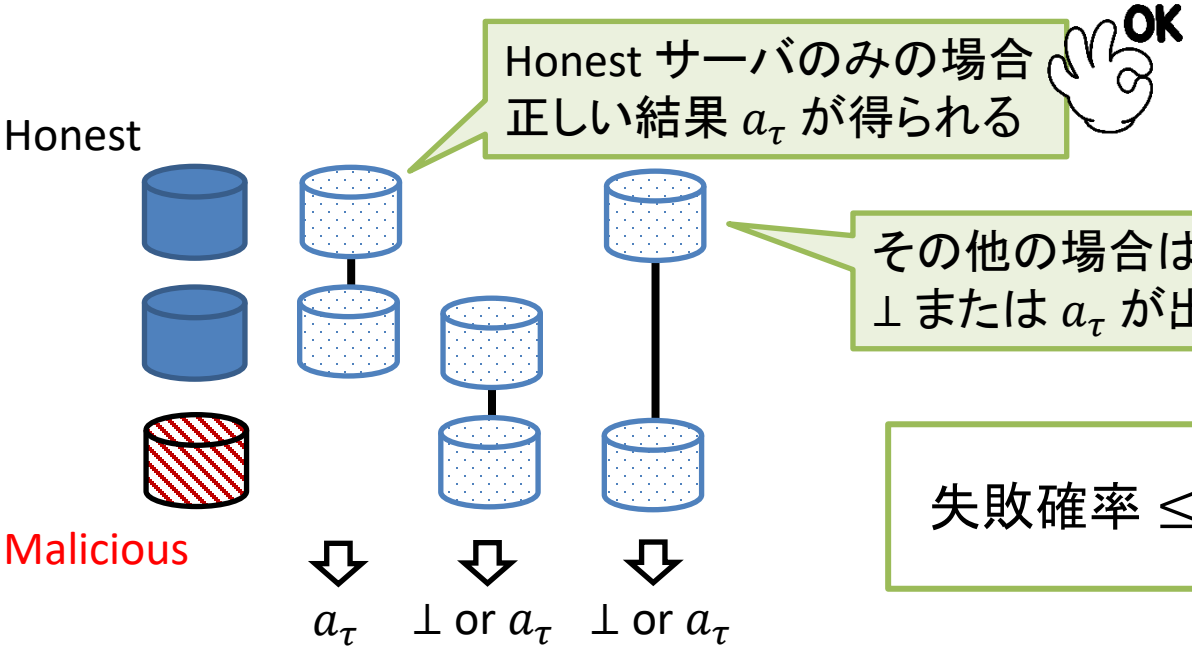
Error-detecting PIR を k 台のサーバからなる全ての部分集合に対して実行する



Actively Secure PIR への変換

t-Error-detecting k-server PIR → **Actively t-secure $(k+t)$-server PIR**

Error-detecting PIR を k 台のサーバからなる全ての部分集合に対して実行する



$$\text{失敗確率} \leq \binom{k+t}{k} 2^{-\lambda} = k^{O(t)} 2^{-\lambda}$$

まとめ

- PIR は最も基本的なインデックスクエリを実現する秘匿検索技術
- 効率的かつ情報理論的安全な方式（抜粋）

1-Private ℓ-server PIR

劣多項式通信量 $n^{o(1)}$

$\ell = 3$ [Yek08, Efr12, IS10]

$\ell = 2$ [DG16]

t-Private ℓ-server PIR

多項式通信量

通信量 $n^{O(1/d)} = n^{O(t/\ell)}$

サーバ数 $\ell = O(td)$ [WY07]

劣多項式通信量

通信量 $n^{o(1)}$

サーバ数 $\ell = 2^{O(t)}$

- 能動的安全性
  - サーバの応答に誤りが含まれる場合（DB の同期失敗等）でも正当性を維持する

今後の課題

- タイトな下限を求められるか
- 複数の通信ラウンドを許すことで通信量を減らすことは可能か
- サーバ数を削減できるか

1-Private ℓ-server PIR

劣多項式通信量 $n^{o(1)}$

$\ell = 3$ [Yek08, Efr12, IS10]

$\ell = 2$ [DG16]

t-Private ℓ-server PIR

多項式通信量

通信量 $n^{O(1/d)} = n^{O(t/\ell)}$

サーバ数 $\ell = O(td)$ [WY07]

劣多項式通信量

通信量 $n^{o(1)}$

サーバ数 $\ell = 2^{O(t)}$

$\ell = \mathrm{poly}(t)$ に削減できるか

参考文献

[CGKS98] Chor, B., Goldreich, O., Kushilevitz, E., Sudan, M.: Private information retrieval. Journal of the ACM 45(6)

[BI05] Barkol, O., Ishai, Y.: Secure computation of constant-depth circuits with applications to database search problems. CRYPTO 2005

[BGI16] Boyle, E., Gilboa, N., Ishai, Y.: Function Secret Sharing: Improvements and Extensions. ACM CCS 2016

[DHRW16] Dodis, Y., Halevi, S., Rothblum, R.D., Wichs, D.: Spooky encryption and its applications. CRYPTO 2016

[KO97] Kushilevitz, E., Ostrovsky, R.: Replication is not needed: single database, computationally-private information retrieval. Proceedings 38th Annual Symposium on Foundations of Computer Science.

[CMS99] Cachin, C., Micali, S., Stadler, M.: Computationally private information retrieval with polylogarithmic communication. EUROCRYPT ’99

[Cha04] Chang, Y.C.: Single database private information retrieval with logarithmic communication. Information Security and Privacy

参考文献

[Lip05] Lipmaa, H.: An oblivious transfer protocol with log-squared communication. Information Security
[GR05] Gentry, C., Ramzan, Z.: Single-database private information retrieval with constant communication rate. Automata, Languages and Programming

[CK20] Corrigan-Gibbs, H., Kogan, D.: Private information retrieval with sublinear online time. EUROCRYPT 2020

[CHK22] Corrigan-Gibbs, H., Henzinger, A., Kogan, D.: Single-server private information retrieval with sublinear amortized time. EUROCRYPT 2022

[ZLTS23] Zhou, M., Lin, W., Tselekounis, Y., Shi, E.: Optimal single-server private information retrieval. EUROCRYPT 2023

[Amb97] Ambainis, A.: Upper bound on the communication complexity of private information retrieval. Automata, Languages and Programming

[BIKR02] Beimel, A., Ishai, Y., Kushilevitz, E., Raymond, J.F.: Breaking the o(n/sup 1/(2k-1)/) barrier for information-theoretic private information retrieval. The 43rd Annual IEEE Symposium on Foundations of Computer Science

[Yek08] Yekhanin, S.: Towards 3-query locally decodable codes of subexponential length. Journal of the ACM (JACM) 55(1)

参考文献

[Efr12] Efremenko, K.: 3-query locally decodable codes of subexponential length. SIAM Journal on Computing 41(6)

[IS10] Itoh, T., Suzuki, Y.: Improved constructions for query-efficient locally decodable codes of subexponential length. IEICE Transactions on Information and Systems E93.D(2)

[DG16] Dvir, Z., Gopi, S.: 2-server PIR with subpolynomial communication. Journal of the ACM 63(4)

[WY07] Woodruff, D., Yekhanin, S.: A geometric approach to information-theoretic private information retrieval. SIAM Journal on Computing 37(4)

[BIW10] Barkol, O., Ishai, Y., Weinreb, E.: On locally decodable codes, self-correctable codes, and tprivate PIR. Algorithmica 58(4)

[BS07] Beimel, A., Stahl, Y.: Robust information-theoretic private information retrieval. Journal of Cryptology 20(3)

[EKN22] Eriguchi, R., Kurosawa, K., Nuida, K.: On the optimal communication complexity of errorcorrecting multi-server PIR. Theory of Cryptography

[EKN24] Eriguchi, R., Kurosawa, K., Nuida, K.: Efficient and generic methods to achieve active security in private information retrieval and more advanced database search. EUROCRYPT 2024

[WdW05] Wehner, S., de Wolf, R.: Improved lower bounds for locally decodable codes and private information retrieval. Automata, Languages and Programming