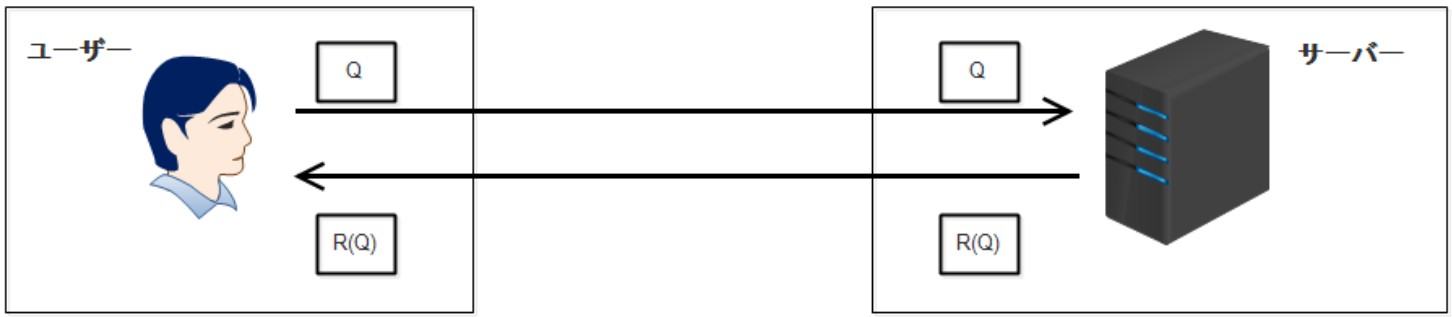
Private Information Retrieval

胡瀚林

May 17, 2016

1. 背景
2. 守り方
3. 攻撃方
4. 参考文献

Private Information Retrieval



- Q:**検索質問**
- R(Q):**質問Qの検索結果**

Location vs Keyword

- Location
  - **地図**
  - **乗換案内**
  - **近くのラストラン**
- Keyword
  - **ウェブ検索**
  - **データベース検索**
  - **クラウドストア検索**

AOL 事件

AOL 質問ログ

AnonID	Query	QueryTime	ItemRank	ClickURL
4417749	care packages	2006-03-02 09:19:32	10	http://booksforsoldiers.com
4417749	care packages	2006-03-02 09:19:32	9	http://www.brandonblog.com
4417749	movies for dogs	2006-03-02 09:24:14		
4417749	blue book	2006-03-03 11:48:52	1	http://www.kbb.com
4417749	best dog for older owner	2006-03-06 11:48:24	1	http://www.canismajor.com
4417749	best dog for older owner	2006-03-06 11:48:24	5	http://dogs.about.com

- 2006 年 8 月 4 日、AOL(American OnLine) が 650,000 人以上のユーザーの匿名化された検索質問ログを研究目的でリリースした。

AOL 事件

AOL 質問ログ

AnonID	Query	QueryTime	ItemRank	ClickURL
4417749	care packages	2006-03-02 09:19:32	10	http://booksforsoldiers.com
4417749	care packages	2006-03-02 09:19:32	9	http://www.brandonblog.com
4417749	movies for dogs	2006-03-02 09:24:14		
4417749	blue book	2006-03-03 11:48:52	1	http://www.kbb.com
4417749	best dog for older owner	2006-03-06 11:48:24	1	http://www.canismajor.com
4417749	best dog for older owner	2006-03-06 11:48:24	5	http://dogs.about.com

- 2006 年 8 月 4 日、AOL(American OnLine) が 650,000 人以上のユーザーの匿名化された検索質問ログを研究目的でリリースした。
- 2006 年 8 月 9 日、ID 4417749 の名前、年齢、住所などが特定された。(Bar)

Location vs Keyword

猫　？　犬

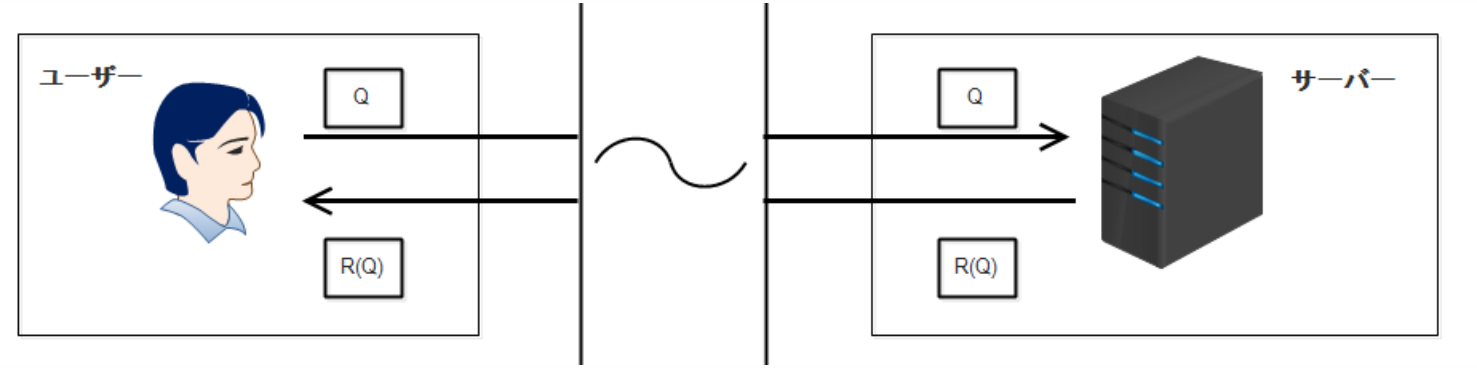
- 位置間の距離は簡単に計算できるが、
  単語間の距離は計算しにくい

猫？

- ノイズを加えにくい

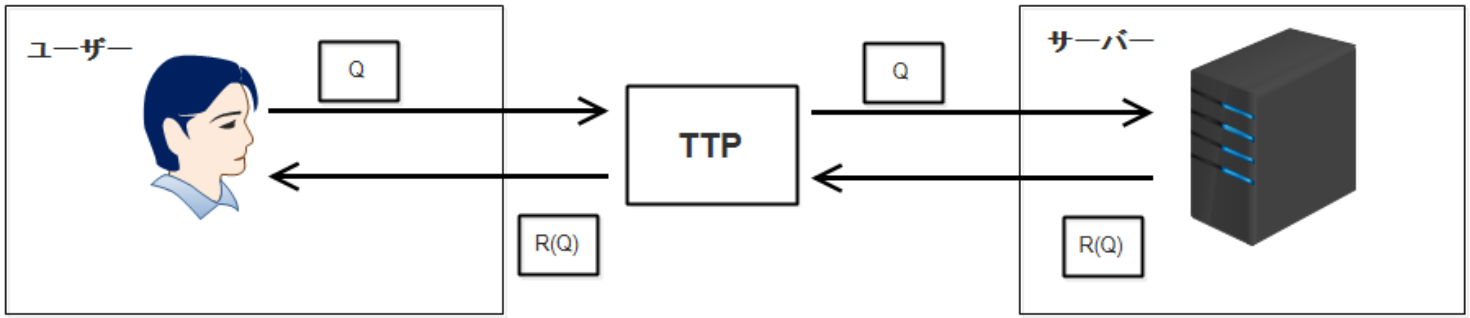
1. 背景
2. 守り方
3. 攻撃方
4. 参考文献

Anonymity



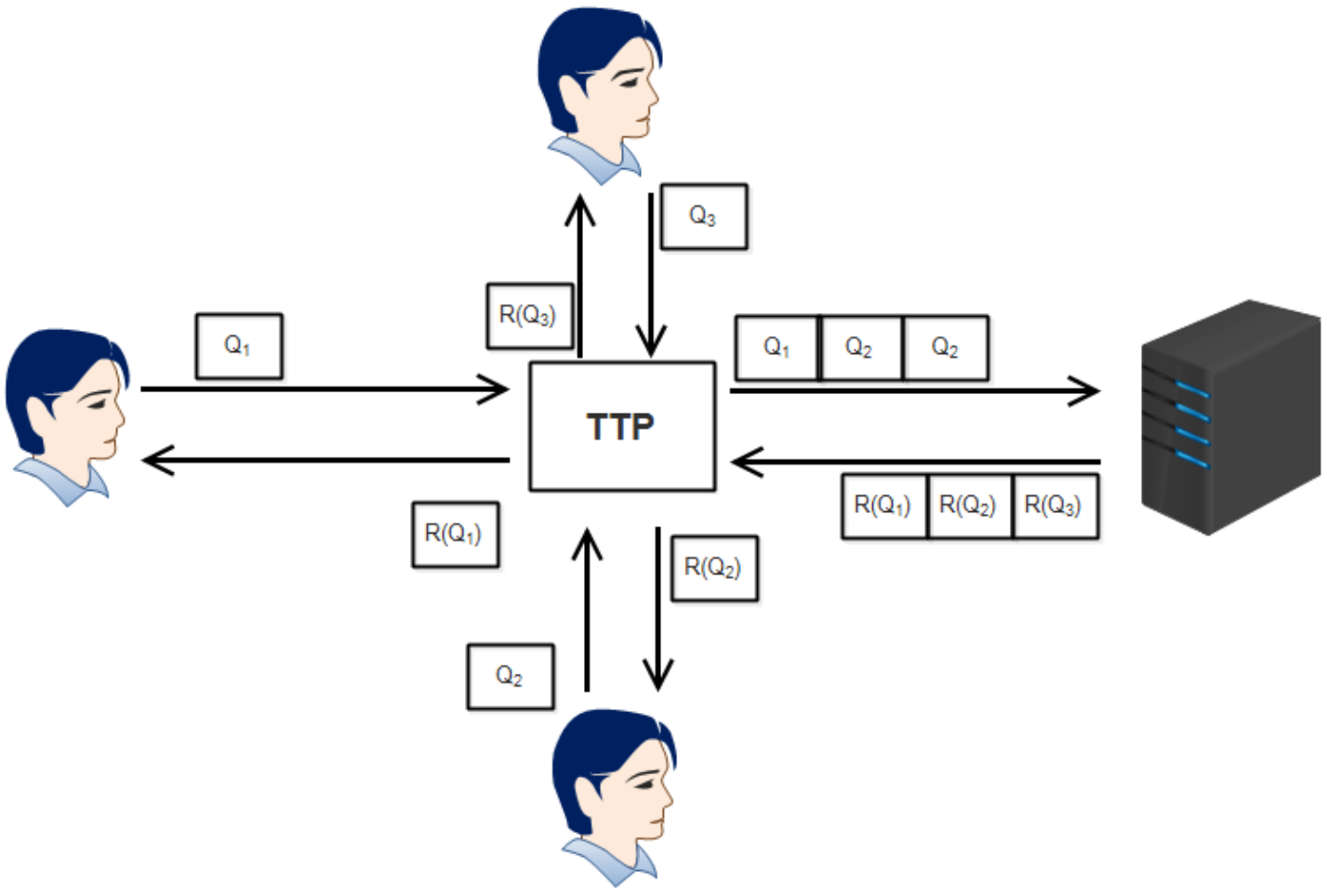
- 質問者を隠す

Tursted Third Party



- 質問者の IP アドレスなどを隠す

Tursted Third Party



- 複数の質問者を混ぜて検索する

Perturbation

Location

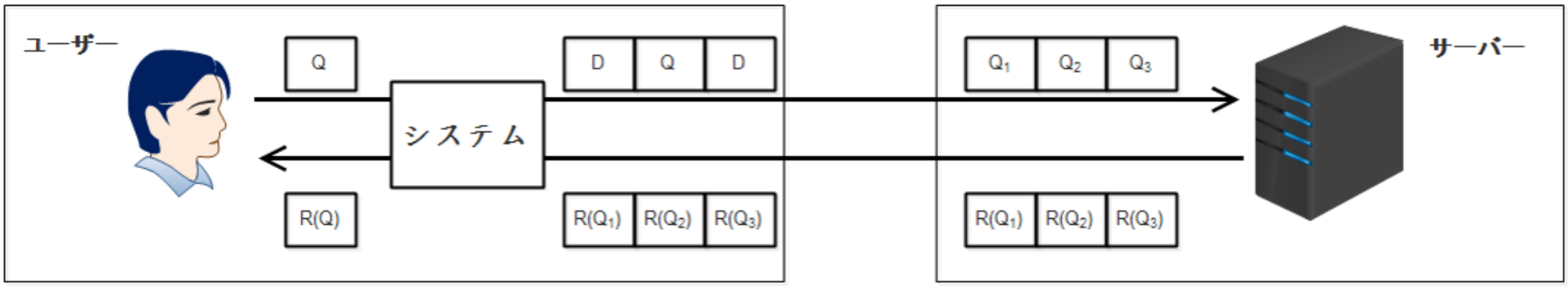
- Geo-indistinguishability (ABCP13)

Keyword

- **質問を一般化して検索する** (AED12)

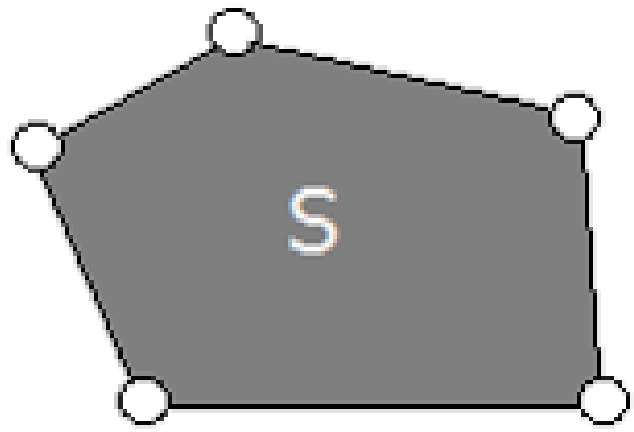
  リンゴ ⇒ 赤 果物
- **事前に標準質問を作って、本当の質問の代わりに使う** (MC09)

Obfuscation



- 複数の質問を混ぜて検索する

Obfuscation-Location



定義 ((k, s) – *privacy* (LJY08))

本当の位置と $k-1$ 個のダミー位置に囲まれた図形の面積が S 以上ある

Obfuscation-Keyword (BTD12)

問題

どうのようなダミー質問がいいダミー質問

Obfuscation-Keyword (BTD12)

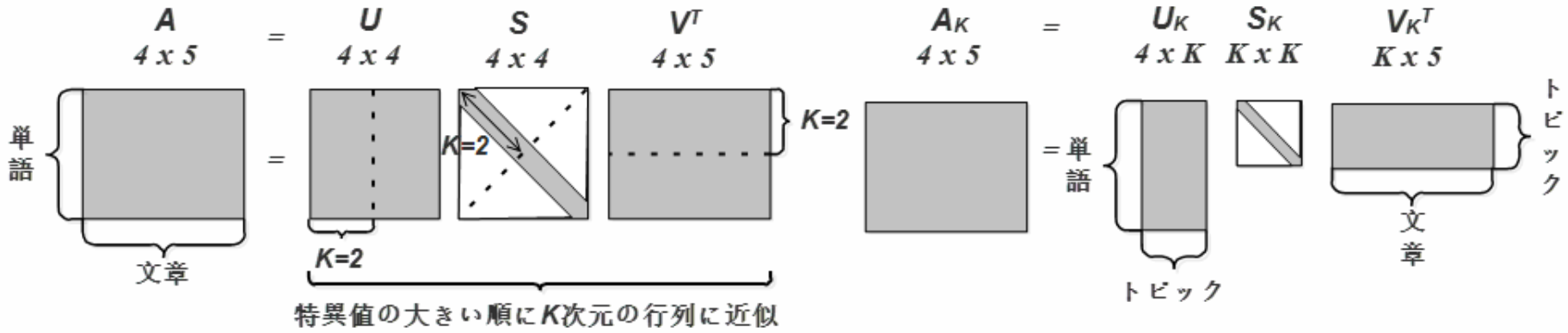
問題

どうのようなダミー質問がいいダミー質問

Plausibly Deniable Search (MC09)

- 本当の質問との“ 距離 ”が遠い
- 本当の質問と似たような“ 確率 ”で提出される

Latent Semantic Analysis

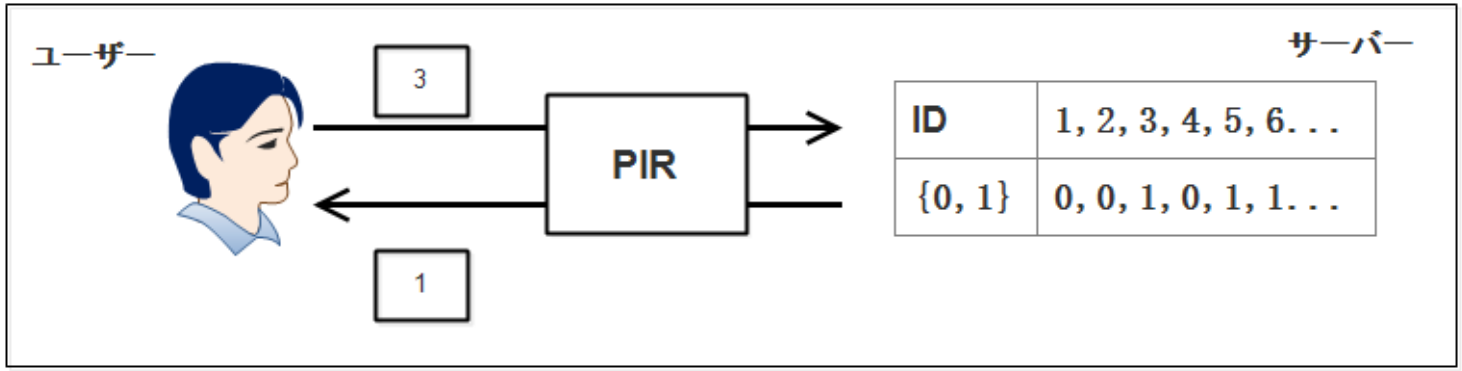


潜在的意味インデキシング

単語・文書行列 A を特異値分解 $A = USV^T$ し、U、S、V の各列ベクトルを特異値が大きい順に K 個用いて A の低ランク近似 $A_K = U_K S_K V_K^T$ を得る。
このように低ランク分解によって、単語とトピックの関係を分析することができる

Plausibly

PIR (OI07)



- 暗号などの手法を用いて質問の内容を完全に隠す

準同型暗号

定義 (準同型暗号)

二つの暗号文 $Enc(m_1), Enc(m_2)$ が与えられた時に、平文や秘密鍵なしで $Enc(m_1 \circ m_2)$ を計算できる暗号

例 (加算ができる準同型暗号)

$Enc(\cdot)$: **暗号化**　$Dec(\cdot)$: **復号**

$Dec(Enc(m_1) \cdot Enc(m_2)) = m_1 + m_2$

準同型暗号

ユーザー

質問生成

1: Input:i^*, n
2: **for** $i = 1, \ldots, n$:
3: **if** $i == i^*$:
4: $q_i = Enc(1)$
5: **else**
6: $q_i = Enc(0)$
7: **return** $Q = \{q_1, \ldots, q_n\}$

復号

1: input:R
2: **return** $Dec(R)$

サーバー

結果計算

1: Input:$Q, \{x_1, \ldots, x_n\}$
2: $R = 0$
3: **for** $i = 1, \ldots, n$:
4: $R = R \cdot q_i^{x_i}$
5: **return** R

Note

$m_1 = m_2 \nRightarrow Enc(m_1) = Enc(m_2)$
$Dec(R) = \sum_{x_i=1} Dec(q_i) = x_{i^*}$

Obfuscation + PIR

Bibliography I

Miguel E. Andrs, Nicols E. Bordenabe, Konstantinos Chatzikokolakis, and Catuscia Palamidessi.
Geo-indistinguishability: Differential Privacy for Location-based Systems.
In *Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security*, CCS '13, pages 901–914, New York, NY, USA, 2013. ACM.

Avi Arampatzis, Pavlos S. Efraimidis, and George Drosatos.
A query scrambler for search privacy on the internet.
Information Retrieval, 16(6):657–679, October 2012.

Zeller Barbaro.
A Face Is Exposed for AOL Searcher No. 4417749 - New York Times.

E. Balsa, C. Troncoso, and C. Diaz.
OB-PWS: Obfuscation-Based Private Web Search.
In *2012 IEEE Symposium on Security and Privacy*, pages 491–505, May 2012.

Hua Lu, Christian S. Jensen, and Man Lung Yiu.
PAD: Privacy-area Aware, Dummy-based Location Privacy in Mobile Services.
In *Proceedings of the Seventh ACM International Workshop on Data Engineering for Wireless and Mobile Access*, MobiDE '08, pages 16–23, New York, NY, USA, 2008. ACM.

M. Murugesan and C. Clifton.
Providing Privacy through Plausibly Deniable Search.
In *Proceedings of the 2009 SIAM International Conference on Data Mining*, Proceedings, pages 768–779. Society for Industrial and Applied Mathematics, April 2009.

Bibliography II

Rafail Ostrovsky and William E. Skeith Iii.
A Survey of Single-Database Private Information Retrieval: Techniques and Applications.
In Tatsuaki Okamoto and Xiaoyun Wang, editors, *Public Key Cryptography PKC 2007*, number 4450 in Lecture Notes in Computer Science, pages 393–411. Springer Berlin Heidelberg, April 2007.
DOI: 10.1007/978-3-540-71677-8_26.