

平成 28 年 5 月 23 日現在

機関番号：15401

研究種目：基盤研究(C) (一般)

研究期間：2011～2015

課題番号：23500180

研究課題名(和文)分散クラウドコンピューティング環境における個人情報保護データマイニング

研究課題名(英文)Privacy Preserving Data Mining in Distributed Cloud Computing Environment

研究代表者

森本 康彦 (Morimoto, Yasuhiko)

広島大学・工学(系)研究科(研究院)・准教授

研究者番号：00363010

交付決定額(研究期間全体)：(直接経費) 4,000,000円

研究成果の概要(和文)：本課題では、データベースに記録されている大量のデータの中から、何らかの知見を獲得したり、データの全体的な傾向を把握したり、必要なデータを選択する際に利用できる問い合わせ機能の研究を行い、特に「スカイライン問い合わせ」を「MapReduce」上で、効率的にかつ安全にスカイライン問い合わせを計算するためのアルゴリズム、および通信プロトコルを、開発した。他の研究機関の先行研究でも、同様のスカイライン問い合わせの事例が数件報告されているが、我々の方法は、他の研究とは異なり個人情報保護を意識したスカイラインであるという特徴がある。

研究成果の概要(英文)：We have studied methods for querying a large-scale database to find and select some clues or to grasp trends and tendencies. Among them, we have intensively studied “skyline query”, which is one of popular queries used in an early stage of database analysis. We have developed an efficient and secure distributed algorithm for computing “skyline query” that works in MapReduce framework. There are some related works that can compute a skyline query in MapReduce framework. Different from others, ours is designed so that privacy information in a database are preserved.

研究分野：データマイニング

キーワード：Skyline Query Data Mining Privacy Distributed Databases MapReduce Cloud

### 1. 研究開始当初の背景

蓄積されたデータから有用な知見を発掘するデータマイニング技術が様々な方面に応用され、有効利用されている。研究開始当初、クラウドコンピューティングが注目され始め、多くの実システムに浸透しつつあった。クラウドコンピューティング環境上で蓄積される各種のデータからも、有効な知見を発見できることが期待でき、この分野でのデータマイニング技術に着目した。

クラウド環境下においては、データはネットワーク上に分散配置されている。そのため、有用な知見を発掘するためには分散配置されたデータベースでデータマイニングしなくてはならない。

このような状況下では、それまでにない、データ転送を考慮した効率的な分散アルゴリズム、ネットワーク上を流れるデータの機密性を意識したセキュアなアルゴリズムが新たに必要となっていた。

### 2. 研究の目的

本研究は、ネットワーク上の複数のデータベース上に分散配置されたデータからのデータマイニング技術の開発を目指して開始された。具体的にはデータマイニングの過程で、異なるサーバ上にある複数のデータベースをまたがって、必要な情報を収集しなくてはならないが、互いに個人情報了他者に漏らすことなく計算するための新たなアルゴリズムを考案することを目的とした。

データベースの何らかの属性値で他のデータに比べ良いか等しいものを含むデータを列挙する機能を「スカイライン問い合わせ」と呼ぶ。スカイライン問い合わせで得られるデータ集合のなかには、ユーザの多様な要求を満足させるデータが含まれている。そのため、スカイライン問い合わせは、意思決定、情報フィルタリングなど多くのアプリケーションに応用されている。本研究では、個人情報保護の観点から、個々のデータ内容を他者に開示しない新たなスカイライン問い合わせ機能を中心に研究を進めた。

### 3. 研究の方法

研究開始当時、注目度の高かった「スカイライン問い合わせ」の機能を手始めに研究を開始した。この時点ですでにいくつかの先行事例が発表されていたが、地理情報システムのデータを想定した、位置情報を含む「スカイライン問い合わせ」という独自分野で研究開発を進めた。

また、近年、「ビッグデータ」と呼ばれている新しい大規模計算技術やその応用が注目を集めているが、ビッグデータ技術やクラウドアプリケーションの多くが MapReduce や MapReduce 基盤内で利用さ

れる列指向データベースで開発されている。本研究課題でも、この計算基盤への対応が必須であると考え、「MapReduce」上で、効率的にかつ安全にスカイライン問い合わせを計算するためのアルゴリズム、および通信プロトコルを、開発していった。

### 4. 研究成果

研究期間前半では主に、地理情報システムのデータを想定した、位置情報を含む「スカイライン問い合わせ」を完成させることができた。

後半は、「MapReduce」上で、「スカイライン問い合わせ」を中心に、情報検索機能、データマイニング機能を数多く開発した。

他の研究機関の先行研究でも、同様のスカイライン問い合わせの事例が数件報告されているが、本研究で開発した方法は、計算結果・計算過程ともに個人情報保護を意識したものであるという点が特徴である。

本研究に関連して、査読付き学術雑誌 10 編に、著書(分担執筆)1 編に本課題の研究成果を出版した。また、国際会議で 20 件の口頭発表を行った。このうちイタリア、ローマで開催された DBKDA 国際会議で発表した論文は「Best Paper Award」を受賞した。

### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 10 件)

1. Md. Anisuzzaman Siddique, Asif Zaman, Yasuhiko Morimoto, "Efficient Selection of Representative Combinations of Objects from Large Database," International Journal on Advances in Software, Vol. 8, No. 3&4, pp. 481-490, 2015. 査読有  
<http://www.ariajournals.org/software/>
2. Md. Anisuzzaman Siddique, Hao Tian, Yasuhiko Morimoto, "k-dominant Skyline Query Computation in MapReduce Environment," The Institute of Electronics, Information and Communication Engineers (IEICE) Transactions on Information and Systems, Vol. E98-D, No. 5, pp. 1027-1034, 2015. 査読有

- DOI 10.1587/transinf.2014DAP0010
3. Md. Shamsul Arefin and Yasuhiko Morimoto, "Privacy Aware Parallel Computations of Skyline Sets Queries from Distributed Databases ", Computing and Informatics, Vol. 33, No. 4, pp. 831–856, 2014. 査読有  
<http://www.cai.sk/ojs/index.php/cai/issue/view/CAI-33-2014-4>
  4. Md. Shamsul Arefin, Geng Ma, Yasuhiko Morimoto, "A Spatial Skyline Query for a Group of Users," Journal of Software, Academy Publisher, Vol. 9, No. 11, pp. 2938-2947, 2014. 査読有  
DOI:10.4304/jsw.9.11.2938-2947
  5. Md. Anisuzzaman Siddique, Asif Zaman, Md. Mahbulul Islam, and Yasuhiko Morimoto, "Distributed Spatial k-dominant Skyline Maintenance Using Computational Object," International Journal of Networking and Computing, Vol. 3, No. 2, pp. 244-263, 2013. 査読有  
<http://www.ijnc.org/index.php/ijnc/issue/view/9>
  6. Hongbo Chen, Md. Shamsul Arefin, Zhiming Chen, and Yasuhiko Morimoto, "Place Recommendation Based On Users Check-in History for Location-based Services," International Journal of Networking and Computing, Vol. 3, No. 2, pp. 228–243, 2013. 査読有  
<http://www.ijnc.org/index.php/ijnc/issue/view/9>
  7. Md. Shamsul Arefin, Jinhao Xu, Zhiming Chen, and Yasuhiko Morimoto, "Skyline Queries for Selecting Spatial Objects by Utilizing Surrounding Environment", Journal of Computers, Vol 8, No 7 (2013), 1742-1749, Academy Publisher, 2013. 査読有
  8. Md. Shamsul Arefin, Yasuhiko Morimoto, "Skyline Sets Queries for Incomplete Data," International Journal of Computer Sciences and Information Technology (IJCSIT), Vol. 4, No. 5, pp. 67-80, October 2012. 査読有  
DOI: 10.5121/ijcsit.2012.4506
  9. Md. Anisuzzaman Siddique and Yasuhiko Morimoto, "Efficient k-dominant Skyline Computation for High Dimensional Space with Domination Power Index", Journal of Computers, Vol. 7, No. 3, pp. 608-615, Academy Publisher, 2012. 査読有  
DOI:10.4304/jcp.7.3.608-615
  10. Yasuhiko Morimoto, Md. Anisuzzaman Siddique, Md. Shamsul Arefin, "Agent-based Anonymous Skyline Set Computation in Cloud Databases," International Journal of Computational Science and Engineering, Vol. 7, No. 1, pp. 73-81, Inderscience, Publishers 2012. 査読有  
DOI: 10.1504/IJCSE.2012.046182
- 〔学会発表〕(計 20 件)
1. Asif Zaman, Md. Anisuzzaman Siddique, Annisa, Yasuhiko Morimoto, "Selecting Key Person of Social Network Using Skyline Query in MapReduce Framework," International Symposium on Computing and Networking (CANDAR2015), pp. 213-219, 札幌市産業振興センター(北海道札幌市), 2015.12.9. 査読有  
DOI 10.1109/CANDAR.2015.84
  2. Annisa, Md. Anisuzzaman Siddique, Asif Zaman, Yasuhiko Morimoto, "A Method for Selecting Desirable Unfixed Shape Areas from Integrated Geographic

- Information System,” IIAI International Congress on Advanced Applied Informatics (IIAI AAI), pp. 195-200, 岡山コンベンションセンター (岡山県岡山市), 2015.7.12. 査読有  
DOI 10.1109/IIAI-AAI.2015.242
3. Md. Anisuzzaman Siddique, Asif Zaman, Yasuhiko Morimoto, “Skyline Objectset: Efficient Selection of Complete Non-dominate Sets from Database,” International Conference on Advances in Databases, Knowledge, and Data Applications (DBKDA), pp. 114-120, Rome (Italy), 2015.5.24. 査読有
  4. Md. Anisuzzaman Siddique, Asif Zaman, and Yasuhiko Morimoto, “Skyband-Set for Answering Top-k Set Queries of Any Users,” International Workshop on Databases in Networked Information Systems (DNIS), pp. 41-55, 会津大学(福島県会津若松市), 2015.3.23. 査読有  
DOI:10.1007/978-3-319-16313-0\_4
  5. Md. Shamsul Arefin, Rahma Bintey Mufiz Mukta, Yasuhiko Morimoto, “Agent-based Privacy Aware Feedback System,” International Conference on Advanced Data Mining and Applications (ADMA 2014), pp. 725-738, Guilin (China), 2014.12.19. 査読有  
DOI:10.1007/978-3-319-14717-8\_58
  6. Md. Anisuzzaman Siddique, Hao Tian, Yasuhiko Morimoto, “Selecting Representative Objects from Large Database by Using K-Skyband and Top-k Dominating Queries in MapReduce Environment,” International Conference on Advanced Data Mining and Applications (ADMA), pp. 560-572, Guilin (China), 2014.12.19. 査読有  
DOI: 10.1007/978-3-319-14717-8\_44
  7. Hao Tian, Md. Anisuzzaman Siddique, Yasuhiko Morimoto, “An Efficient Processing of k-Dominant Skyline Query in MapReduce,” ACM International Workshop on Bringing the Value of Big Data to Users (Data4U), pp. 29-34, Hangzhou (China), 2014.9.1 査読有  
DOI:10.1145/2658840.2658846
  8. Md. Anisuzzaman Siddique, Hao Tian, Yasuhiko Morimoto, “Distributed Skyline Computation of Vertically Splitted Databases by Using MapReduce,” DASFAA International Workshop on Big Data Management and Analytics (BDMA2014), pp. 33-45, Bali (Indonesia), 2014.4.21. 査読有  
DOI:10.1007/978-3-662-43984-5\_3
  9. Md. Anisuzzaman Siddique, Yasuhiko Morimoto, “Efficient Selection of Various k-Objects for a Keyword Query Based on MapReduce Skyline Algorithm,” International Workshop on Databases in Networked Information Systems (DNIS), pp. 40-52, 会津大学(福島県会津若松市), 2014.3.24. 査読有  
10.1007/978-3-319-05693-7\_3
  10. Zhichao Chang, Md. Shamsul Arefin, Yasuhiko Morimoto, “Hotel Recommendation Based On Surrounding Environments,” IIAI International Conference on Advanced Applied Informatics (IIAI AAI), pp. 330-336, くにびきメッセ(島根県松江市), 2013.8.31. 査読有  
DOI:10.1109/IIAI-AAI.2013.39
  11. Md. Shamsul Arefin, Yasuhiko Morimoto, “Skyline Queries for Sets of Spatial Objects by Utilizing Surrounding Environments,” International Workshop on Databases in Networked Information

- Systems (DNIS), pp. 293-309, 会津大学 (福島県会津若松市), 2013.3.25. 査読有 DOI:10.1007/978-3-642-37134-9\_22
12. Asif Zaman, Md. Mahbulul Islam, Md. Anisuzzaman Siddique, Yasuhiko Morimoto, "Distributed k-dominant Skyline Queries," IEEE International Conference on Computer and Information Technology (ICCIT 2012), pp. 569-574, Chittagong (Bangladesh), 2012.12.22. 査読有
  13. Jinhao Xu, Md. Shamsul Arefin, Zhiming Chen, Yasuhiko Morimoto, "Real Estate Recommender: Location Query for Selecting Spatial Objects," IEEE International Conference on Networking and Computing (ICNC 2012), pp. 278-282, 沖縄県男女共同参画センター(沖縄県那覇市), 2012.12.5. 査読有
  14. Geng Ma, Md. Shamsul Arefin, Yasuhiko Morimoto, "A Spatial Skyline Query for a Group of Users Having Different Positions," IEEE International Conference on Networking and Computing (ICNC 2012), pp. 137-142, 沖縄県男女共同参画センター(沖縄県那覇市), 2012.12.5. 査読有
  15. Zhiming Chen, Md. Shamsul Arefin, Yasuhiko Morimoto, "Skyline Queries for Spatial Objects: A Method for Selecting Spatial Objects Based on Surrounding Environments," IEEE International Conference on Networking and Computing (ICNC 2012), pp. 215-220, 沖縄県男女共同参画センター(沖縄県那覇市), 2012.12.5. 査読有
  16. Hongbo Chen, Zhiming Chen, Md. Shamsul Arefin, Yasuhiko Morimoto, "Place Recommendation from Check-in Spots on Location-Based Online Social Networks," IEEE International Conference on Networking and Computing (ICNC 2012), pp. 143-148, 沖縄県男女共同参画センター(沖縄県那覇市), 2012.12.5. 査読有
  17. Md. Anisuzzaman Siddique, Asif Zaman, Md. Mahbulul Islam, Yasuhiko Morimoto, "Multicore based Spatial k-dominant Skyline Computation," IEEE International Conference on Networking and Computing (ICNC 2012), pp. 188-194, 沖縄県男女共同参画センター(沖縄県那覇市), 2012.12.5. 査読有
  18. Md. Shamsul Arefin, Yasuhiko Morimoto, "Skyline Sets Queries from Databases with Missing Values," Proceedings of The 22<sup>nd</sup> IEEE International Conference on Computer Theory and Applications (ICCTA2012), pp. 24-29, Alexandria (Egypt), 2012.10.13. 査読有
  19. Yasuhiko Morimoto, Md. Shamsul Arefin, Md. Anisuzzaman Siddique, "Information Filtering by Using Materialized Skyline Views", International Workshop on Databases in Networked Information System (DNIS), pp. 179-189, 会津大学 (福島県会津若松市), 2011.12.14. 査読有 DOI: 10.1007/978-3-642-25731-5\_15
  20. Md. Shamsul Arefin, Yasuhiko Morimoto, "Privacy Aware Parallel Computation of Skyline Sets Queries from Distributed Databases," Proceedings of The IEEE International Conference on Networking and Computing, pp. 186-192, 大阪大学(大阪府吹田市), 2011.12.1. 査読有
- [図書](計 1件)  
Md. Shamsul Arefin, Zhichao Chang, and Yasuhiko Morimoto, "Recommending

Hotels by Social Conditions of Locations,"  
Tourism Informatics - Towards Novel  
Knowledge Based Approaches, Springer,  
pp.91-106, 2015. (分担執筆)  
DOI 10.1007/978-3-662-47227-9\_7

〔その他〕  
ホームページ等

<http://www.morimo.com/morimo-ken/>  
(森本研究室ホームページ)

<http://www.hiroshima-u.ac.jp/en/sangaku/HUQTN/>  
(Hiroshima Univ Quarterly Technology  
Newsletter) (掲載予定)

## 6. 研究組織

### (1) 研究代表者

森本 康彦 (MORIMOTO YASUHIKO)  
広島大学・大学院工学研究院・准教授  
研究者番号： 00363010