Optimizing Skyline Query Processing in Incomplete Data

By: Gulzar, Y (Gulzar, Yonis)[1]; Alwan, AA (Alwan, Ali A.)[2]; Turaev, S (Turaev, Sherzod)[3]

View Web of Science ResearcherID and ORCID

IEEE ACCESS
Volume: 7 **Pages:** 178121-178138
DOI: 10.1109/ACCESS.2019.2958202
Published: 2019
Document Type: Article
View Journal Impact

Abstract

Given the significance of skyline queries, they are incorporated in various modern applications including personalized recommendation systems as well as decision-making and decision-support systems. Skyline queries are used to identify superior data items in the database. Most of the previously proposed skyline algorithms work on a complete database where the data are always present (non-missing). However, in many contemporary real-world databases, particularly those databases with large cardinality and high dimensionality, such assumption is not necessarily valid. Hence, missing data pose new challenges if the processing skyline queries cannot easily apply those methods that are designed for complete data. This is due to the fact that imperfect data cause the loss of the transitivity property of the skyline method and cyclic dominance. This paper presents a framework called Optimized Incomplete Skyline (OIS) which utilizes a technique that simplifies the skyline process on a database with missing data and helps prune the data items before performing the skyline process. The proposed strategy assures that the number of the domination tests is significantly reduced. A set of experiments has been accomplished using both real and synthetic datasets aimed at validating the performance of the framework. The experiment results confirm that the OIS framework is indeed superior and steadily outperforms the current approaches in terms of the number of domination tests required to retrieve the skylines.

Keywords

Author Keywords: Algorithms; incomplete data; database; preference queries; query processing; skylines; skyline queries
KeyWords Plus: IDENTIFYING SKYLINES; FRAMEWORK

Author Information

Reprint Address: Gulzar, Y (reprint author)

King Faisal Univ, Coll Business Adm, Dept Management Informat Syst, Al Hasa 31982, Saudi Arabia.

Reprint Address: Alwan, AA (reprint author)

Int Islamic Univ Malaysia, Kulliyyah Informat & Commun Technol, Dept Comp Sci, Gombak 35100, Selangor, Malaysia.

Addresses:

[1] King Faisal Univ, Coll Business Adm, Dept Management Informat Syst, Al Hasa 31982, Saudi Arabia

[2] Int Islamic Univ Malaysia, Kulliyyah Informat & Commun Technol, Dept Comp Sci, Gombak 35100, Selangor, Malaysia

[3] United Arab Emirates Univ, Coll Informat Technol, Dept Comp Sci & Software Engn, Al Ain 15551, U Arab Emirates

E-mail Addresses: ygulzar@kfu.edu.sa; aliamer@iium.edu.my

Funding

Funding Agency	Grant Number
Deanship of Scientific Research in King Faisal University, Saudi Arabia	186304

View funding text

Publisher

IEEE-INST ELECTRICAL ELECTRONICS ENGINEERS INC, 445 HOES LANE, PISCATAWAY, NJ 08855-4141 USA

Journal Information

Impact Factor: Journal Citation Reports

Categories / Classification

Research Areas: Computer Science; Engineering; Telecommunications
Web of Science Categories: Computer Science, Information Systems; Engineering, Electrical & Electronic; Telecommunications

See more data fields

Cited References: 44

1. **A Framework for Identifying Skylines over Incomplete Data** — **Times Cited: 2**
   By: Alwan, Ali A.; Ibrahim, Hamidah; Udzir, Nur Izura
   3RD INTERNATIONAL CONFERENCE ON ADVANCED COMPUTER SCIENCE APPLICATIONS AND TECHNOLOGIES ACSAT 2014 Book Series: International Conference on Advanced Computer Science Applications and Technologies Pages: 79-84 Published: 2014

2. **An Efficient Approach for Processing Skyline Queries in Incomplete Multidimensional Database** — **Times Cited: 7**
   By: Alwan, Ali A.; Ibrahim, Hamidah; Udzir, Nur Izura; et al.
   ARABIAN JOURNAL FOR SCIENCE AND ENGINEERING Volume: 41 Issue: 8 Pages: 2927-2943 Published: AUG 2016

3. **Processing skyline queries in incomplete distributed databases** — **Times Cited: 4**
   By: Alwan, Ali A.; Ibrahim, Hamidah; Udzir, Nur Izura; et al.
   JOURNAL OF INTELLIGENT INFORMATION SYSTEMS Volume: 48 Issue: 2 Pages: 399-420 Published: APR 2017

4. **Skyline sets queries for incomplete data** — **Times Cited: 6**
   By: Arefin, M.S.; Morimoto, Y.
   International Journal of Computer Science & Information Technology Volume: 4 Issue: 5 Pages: 67-80 Published: Oct. 2012

5. **DERIVING SKYLINE POINTS OVER DYNAMIC AND INCOMPLETE DATABASES** — **Times Cited: 3**
   By: Babanejad, Ghazaleh; Ibrahim, Hamidah; Udzir, Nurl Zura; et al.
   PROCEEDINGS OF THE 6TH INTERNATIONAL CONFERENCE ON COMPUTING AND INFORMATICS: EMBRACING ECO-FRIENDLY COMPUTING Book Series: Proceedings of the International Conference on Computing & Informatics Pages: 77-83 Published: 2017

6. Title: [not available] — **Times Cited: 1**
   By: BALKE WT
   LECT NOTES COMPUTER Volume: 2992 Published: 2004

7. **SaLSa: Computing the Skyline without scanning the whole sky** — **Times Cited: 2**
   By: Bartolini, I.; Ciaccia, P.; Patella, M.
   15 ACM INT C INF KNO Published: 2006

8. **Effect of Base Station Configurations and Complexity on the Accuracy of Ultra Wideband Localisation** — **Times Cited: 4**
   By: Bharadwaj, Richa; Yang, Ke; Alomainy, Akram; et al.
   2013 IEEE ANTENNAS AND PROPAGATION SOCIETY INTERNATIONAL SYMPOSIUM (APSURSI) Book Series: IEEE Antennas and Propagation Society International Symposium Pages: 19-20 Published: 2013

9. **Finding Skylines for incomplete data** — **Times Cited: 1**
   By: Bharuka, R.; Kumar, P. S.
   24 AUSTR DAT C AD AU Volume: 137 Published: 2013

10. **The Skyline operator** — **Times Cited: 906**
    By: Borzsonyi, S; Kossmann, D; Stocker, K
    17TH INTERNATIONAL CONFERENCE ON DATA ENGINEERING, PROCEEDINGS Book Series: IEEE International Conference on Data Engineering Pages: 421-430 Published: 2001

11. **Finding k -dominant Skylines in high dimensional space** — **Times Cited: 2**
    By: Chan, C.-Y.; Jagadish, H. V.; Tan, K.-L.; et al.
    ACM SIGMOD INT C MAN Published: 2006
    [Show additional data]

12. **On high dimensional skylines** — **Times Cited: 87**
    By: Chan, Chee-Yong; Jagadish, H. V.; Tan, Kian-Lee; et al.
    ADVANCES IN DATABASE TECHNOLOGY - EDBT 2006 Book Series: Lecture Notes in Computer Science Volume: 3896 Pages: 478-495 Published: 2006

13. **Skyline with presorting** — **Times Cited: 319**
    By: Chomicki, J; Godfrey, P; Gryz, J; et al.
    19TH INTERNATIONAL CONFERENCE ON DATA ENGINEERING, PROCEEDINGS Book Series: PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON DATA ENGINEERING (SERIES) Pages: 717-719 Published: 2003

14. **Crowdsourcing for Query Processing on Web Data: a Case Study on the Skyline Operator** — **Times Cited: 1**
    By: El Maarry, K.; Lofi, C.; Balke, W.-T.
    CIT. Journal of Computing and Information Technology Volume: 23 Issue: 1 Pages: 43-60 Published: 2015

15. **Processing k-skyband, constrained skyline, and group-by skyline queries on incomplete data** — **Times Cited: 17**
    By: Gao, Yunjun; Miao, Xiaoye; Cui, Huiyong; et al.
    EXPERT SYSTEMS WITH APPLICATIONS Volume: 41 Issue: 10 Pages: 4959-4974 Published: AUG 2014

16. **Maximal vector computation in large data sets** — **Times Cited: 2**

By: Godfrey, P.; Shipley, R.; Gryz, J.
31 INT C VER LARG DA Published: 2005

17. **Processing skyline queries in incomplete database: Issues, challenges and future trends** Times Cited: 4
By: Gulzar, Y.; Alwan, A. A.; Salleh, N.; et al.
Journal of Computer Science Volume: 13 Issue: 11 Pages: 647-658 Published: 2017
[Show additional data]

18. **D-SKY: A framework for processing Skyline queries in a dynamic and incomplete database** Times Cited: 1
By: Gulzar, Y.; Alwan, A. A.; Ibrahim, H.; et al.
20 INT C INF INT WEB Published: 2018
[Show additional data]

19. **A Framework for Evaluating Skyline Queries over Incomplete Data** Times Cited: 5
By: Gulzar, Yonis; Alwan, Ali A.; Salleh, Norsaremah; et al.
11TH INTERNATIONAL CONFERENCE ON FUTURE NETWORKS AND COMMUNICATIONS (FNC 2016) / THE 13TH INTERNATIONAL CONFERENCE ON MOBILE SYSTEMS AND PERVASIVE COMPUTING (MOBISPC 2016) / AFFILIATED WORKSHOPS Book Series: Procedia Computer Science Volume: 94 Pages: 191-198 Published: 2016

20. **A Model for Skyline Query Processing in a Partially Complete Database** Times Cited: 3
By: Gulzar, Yonis; Alwan, Ali A.; Salleh, Norsaremah; et al.
ADVANCED SCIENCE LETTERS Volume: 24 Issue: 2 Pages: 1339-1343 Published: FEB 2018

21. **SCSA: Evaluating skyline queries in incomplete data** Times Cited: 2
By: Gulzar, Yonis; Alwan, Ali A.; Abdullah, Radhwan Mohamed; et al.
APPLIED INTELLIGENCE Volume: 49 Issue: 5 Pages: 1636-1657 Published: MAY 2019

22. **IDENTIFYING SKYLINES IN CLOUD DATABASES WITH INCOMPLETE DATA** Times Cited: 1
By: Gulzar, Yonis; Aljuboori, Ali Amer Alwan; Salleh, Norsaremah; et al.
JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGY-MALAYSIA Volume: 18 Issue: 1 Pages: 19-34 Published: JAN 2019

23. **SKYLINE QUERY PROCESSING FOR INCOMPLETE DATA IN CLOUD ENVIRONMENT** Times Cited: 2
By: Gulzar, Yonis; Alwan, Ali A.; Salleh, Norsaremah; et al.
PROCEEDINGS OF THE 6TH INTERNATIONAL CONFERENCE ON COMPUTING AND INFORMATICS: EMBRACING ECO-FRIENDLY COMPUTING Book Series: Proceedings of the International Conference on Computing & Informatics Pages: 567-576 Published: 2017

24. **Efficient Processing of the Skyline-CL Query** Times Cited: 1
By: Huang, Zhenhua; Zhang, Juan; Tian, Chunqi
ARABIAN JOURNAL FOR SCIENCE AND ENGINEERING Volume: 41 Issue: 8 Pages: 2801-2811 Published: AUG 2016

25. **Skyline query processing for incomplete data** Times Cited: 54
By: Khalefa, Mohamed E.; Mokbel, Mohamed F.; Levandoski, Justin J.
2008 IEEE 24TH INTERNATIONAL CONFERENCE ON DATA ENGINEERING, VOLS 1-3 Book Series: IEEE International Conference on Data Engineering Pages: 556-565 Published: 2008

26. **Efficient progressive skyline computation** Times Cited: 219
By: Kian-Lee Tan; Pin-Kwang Eng; Beng Chin Ooi
Proceedings of the 27th International Conference on Very Large Data Bases Pages: 301-10 Published: 2001

27. **Continuous Top-k Dominating Queries in Subspaces** Times Cited: 12
By: Kontaki, Maria; Papadopoulos, Apostolos N.; Manolopoulos, Yannis
PCI 2008: 12TH PAN-HELLENIC CONFERENCE ON INFORMATICS, PROCEEDINGS Pages: 31-35 Published: 2008

28. **Shooting stars in the sky: An Online algorithm for skyline queries** Times Cited: 1
By: Kossmann, D.; Ramsak, F.; Rost, S.
28 INT C VER LARG DA Published: 2002

29. **Optimizing skyline queries over incomplete data** Times Cited: 10
By: Lee, Jongwuk; Im, Hyeonseung; You, Gae-won
INFORMATION SCIENCES Volume: 361 Pages: 14-28 Published: SEP 20 2016

30. **Scalable skyline computation using a balanced pivot selection technique** Times Cited: 27
By: Lee, Jongwuk; Hwang, Seung-won
INFORMATION SYSTEMS Volume: 39 Special Issue: SI Pages: 1-21 Published: JAN 2014

Showing 30 of 44 **View All in Cited References page**