

A Review Paper on Discovering Most Preferential Skyline Product Under Price Promotion

^I Miss. Shilp D. Ravte, ^{II} Prof. A. D. Pathak.

^I Master of Technology, Student, ^{II} Associate Professor

^{I, II} Department of Computer Science & Engineering,

Hindi Seva Mandal's, Shri Sant Gadge Baba College of Engineering & Technology,

Bhusawal - 425203, Dist. – Jalgaon, Maharashtra, India

Email: shilpa.ravte411@gmail.com¹, aniketpathak89@gmail.com²

Abstract: Nowadays, with the development of e-commerce, a growing range of shoppers opt to go shopping online. To seek out enticing product from online searching marketplaces, the skyline query may be a great tool that offers additional attention-grabbing and preferred selections for purchasers. The skyline query and its variants have been extensively investigated. However, to the most effective of our information, they need not taken into consideration the wants of shoppers in sure practical application situations. Recently, online shopping marketplaces typically hold some price promotion campaigns to attract customers and increase their purchase intention. Considering the wants of shoppers during this practical application situation, we tend to be involved concerning product choice under price promotion. We tend to formulate a constrained optimal product combination (COPC) problem. It aims to find out the skyline product combinations that each meet a customer's willingness to pay and bring the most discount rate. The COPC problem is critical to offer powerful decision support for customers under price promotion, that is certified by a customer study. To method the COPC problem effectively, we tend to initial propose a two list exact (TLE) algorithm. The COPC problem is proven to be NP-hard, and also the TLE algorithm isn't scalable as a result of it has to method associate degree exponential range of product combinations. In addition, we tend to style a lower bound approximate (LBA) rule that has guarantee about the accuracy of the results associate degree an incremental greedy (IG) rule that has sensible performance. The experiment results demonstrate the efficiency and effectiveness of our planned algorithms

1. INTRODUCTION

The skyline query could be a common and effective methodology to seek out attractive products. In line with the definition of skyline queries [1], a product that isn't dominated by the other product is claimed to be a skyline product or it's within the skyline. The products within the skyline are the most effective doable trade-offs between all the factors that customers care concerning. Though the skyline question will determine the engaging products, it will very little to assist select product combos with the most discount rate. To handle this drawback, customers typically compare all the engaging product combos in terms of their actual discount rates. With the event of e-commerce, a growing variety of shoppers value more highly to shopping online as a result of it saves time and energy. However, it perpetually contraries to expectations of shoppers. This is often as a result of they'll have to acquire one selection among thousands of products. To assist customers, identify engaging products, a skyline question is true a typical and effective methodology. In line with the definition of the skyline question, a product that isn't dominated by the other product is claimed to be a skyline product or it's within the skyline. The products within the skyline are the most effective doable trade-offs between all the factors that customers care concerning. The skyline question is beneficial in characteristic engaging products. This worth promotion campaigns are often classified into 2 classes thanks to whether or not products can be chosen severally. The primary class, namely, freelance product choice, includes the

campaigns like "buy one product and acquire another product for free" and "25% discount for 2 pics" etc. underneath these campaigns, customers will pick the products meeting their demands severally and directly, and skyline queries might provide powerful decision support. The second class, namely, dependent product choice, consists of the campaigns like "get \$60 off each \$200 purchase" and "\$100 coupon every \$500 purchase" etc. In these eventualities, customers perpetually expect to pick products that are engaging and convey the best profit. Moreover, it has to take into thought the customer's temperament to pay that is a vital issue that affects the customer's buying behavior. The skyline question is powerful to figure the skyline products that have a strong charm to customers. However, it's inadequate to assist customer's choose skyline product combos with the best profit. Considering the wants of customers during this employment situation, we tend to be involved about a new drawback of identifying best product combos under price promotion campaigns. During this paper, we tend to target the dependent-product choice campaigns that are way more common however difficult with comparison to the independent-product selection campaigns. The skyline query aims to return the purposes that aren't dominated by the other point. However, most of the works regarding the skyline query simply analyses individual points, and that they are inappropriate to several applications that decision for analysis of teams of different points. Intended by this, group are developed and paid growing attention.

Considering the necessities of customers during this application scenario, we have a tendency to be concerned about a new drawback of identifying optimal product combinations under price promotion campaigns. we target the dependent-product choice campaigns that are way more popular however complicated with comparison to the independent-product selection campaigns

2. LITERATURE SURVEY

Review of Papers

- **Xu Zhou, Kenli Li(2018)**, with the development of e-commerce, a growing range of customers like better to shopping online. to search out attractive product from online shopping marketplaces, the skyline query may be a great tool that offers a lot of attention-grabbing and preferable selections for customers. The skyline query and its variants are extensively investigated. However, to the simplest of our data, they need not taken into account the necessities of customers in certain usage scenarios. Recently, on-line looking marketplaces typically hold some value promotion campaigns to draw in customers and increase their purchase intention. Considering the necessities of shoppers during this practical application situation, we tend to be involved regarding product selection under-price promotion.
- **Xu Zhou, Kenli Li, Member, Guoqing Xiao, Yantao Zhou, and Keqin Li,(2016)**, The probabilistic dynamic skyline (PDS) query could be a powerful tool for customers to use in choosing product in keeping with their preferences. However, this query suffers many limitations: it needs the specification of a probabilistic threshold, that reports undesirable results and disregards important results; it only focuses on the objects that have massive dynamic skyline probabilities; in addition, the results aren't stable. to deal with this concern, we tend to formulate an unsure dynamic skyline (UDS) query over a probabilistic product set. moreover, we tend to propose effective pruning strategies for the UDS query, and integrate them into effective algorithms. additionally, a unique question kind, specifically the highest k favorite probabilistic products (TFPP) query, is given. The TFPP query is used to pick out k product which might meet the wants of a customer set at the utmost level. To tackle the TFPP question, they propose a TFPP rule and its efficient parallelization. extensive experiments with a range of experimental settings illustrate the potency and effectiveness algorithms. Customer preferences info may be a growing concern in market research. during this paper, we tend to initial propose the UDS question to pick out product that can meet a customer's demands to the greatest extent. Compared to the PDS

query, our UDS query doesn't have to specify a threshold and may come back far better results. additionally, with relation to the preferences of various customers, we have a tendency to formulate the TFPP query, that retrieves the k product with the very best favorite possibilities. Moreover, to method the UDS and TFPP query effectively,

some pruning ways are planned and integrated into many effective algorithms. Finally, the potency and effectiveness of the planned algorithms are verified with intensive experiments. As a part of our future analysis, we'll investigate the UDS and TFPP queries on big information.

- **Xu Zhou, Kenli Li, Member, Yantao Zhou, and Keqin Li, Fellow (2015)**, Query process over unsure data has gained growing attention, as a result of it's necessary to handle uncertain data in many real-life applications. during this paper, we tend to investigate skyline queries over unsure information in distributed environments (DSUD query) whose analysis is simply in AN early stage. The progressive algorithm, referred to as e-DSUD algorithm, is intended for process this question . it's the fascinating characteristics of progressivity and minimum information measure consumption. However, it still must be formed in 3 aspects. (1) progressivity. when it solely returns one query result at the most. (2) potency. There are a big quantity of redundant I/O cost and various iterations that causes an extended total query time. (3) generality. it is restricted to the case wherever native skyline tuples are incomparability. to deal with these considerations, we have a tendency to 1st gift an in depth Analysis of the e-DSUD formula so develop an improved framework for the DSUD query, specifically IDSUD. supported the new framework, we have a tendency to propose AN adaptive formula, referred to as ADSUD, for the DSUD query. within the algorithm, we tend to redefine the approximate international skyline probability and select native representative tuples due to minimum probabilistic bounding rectangle adaptively. furthermore, they style a progressive pruning method and apply the utilize mechanism to improve its efficiency. The results of extensive experiments verify the higher overall performance of our algorithm than the e-DSUD algorithm.
- **Lijiang Chen, Student Member, IEEE, Bin Cui, Senior Member, IEEE, and Hua Lu, Member, IEEE(2011)**,The skyline of a multidimensional purpose set may be a set of attention-grabbing points that are not dominated by others. they investigate constrained skyline queries during some large-scale unstructured distributed surroundings, wherever relevant data are distributed among geographically scattered sites. we tend to initial propose a partition formula that divides all information sites into best

teams such the skyline computations all told groups may be parallelized while not dynamical the ultimate result. we tend to then develop a completely unique formula framework referred to as PaD Skyline for parallel skyline query process among partitioned off web site teams. we have a tendency to conjointly use internal optimization and multi filtering technique to enhance the skyline query processes at intervals every group. specifically, multiple (local) skyline points are sent along with the query as filtering points, that facilitate establish unqualified native skyline points early a data web site

- **Yufei Tao, Xiaokui Xiao, and Jian Pei(2007)**, Skyline and top-k queries are 2 standard operations for preference retrieval. In practice, applications that need these operations typically offer varied candidate attributes, whereas, depending on their interests, users might issue queries relating to totally different subsets of the scale. the existing algorithms are inadequate for mathematical space skyline/top-k search because they need a minimum of one in all the subsequent defects: 1) They need scanning the complete info at least once, 2) they are optimized for one mathematical space however incur vital overhead for different subspaces, or 3) they demand big-ticket maintenance value or house consumption. they propose a method SUBSKY, that settles each forms of queries by mistreatment strictly relative technologies. The core of SUBSKY could be a transformation that converts two-dimensional knowledge to one-dimensional (1D) values. These values are indexed by an easy B-tree, that permits us to answer mathematical space queries by accessing a fraction of the info. SUBSKY entails low maintenance overhead, that equals the price of change a standard B-tree. intensive experiments with real knowledge make sure that our technique outperforms different solutions considerably in each efficiency and scalability.
- **Nan Zhang ; Chengkai Li ; Naemul Hassan ; SundaresanRajasekaran ; Gautam Das(2014)**, With the development of the economy, product are considerably enriched, and uncertainty has been their inherent quality. The probabilistic dynamic skyline (PDS) query could be a powerful tool for patrons to use in choosing product consistent with their preferences. However, this question suffers many limitations: it needs the specification of a probabilistic threshold, that reports undesirable results and disregards vital results; it only focuses on the objects that have massive dynamic skyline probabilities; and, to boot, the results aren't stable. to handle this concern, during this paper, we tend to formulate associate degree unsure dynamic skyline (UDS) query over a probabilistic product set. moreover, we tend to propose effective pruning ways for the UDS query, and integrate them into effective algorithms. additionally, a

completely unique query kind, particularly the highest k favorite probabilistic product (TFPP) query, is given. The TFPP query is employed to pick k product which may meet the wants of a client set at the utmost level. To tackle the TFPP query, we tend to propose a TFPP formula and its efficient parallelization. intensive experiments with a variety of experimental settings illustrate the potency and effectiveness of our projected algorithms.

- **Hua Lu Christian S. Jensen ; Zhenjie Zhang(2011)**, a skyline query returns the attention-grabbing points that are not dominated by alternative points. it's been determined that the particular cardinality (s) of a skyline question result could dissent considerably from the specified result cardinality (k), that has prompted studies on the way to scale back s for the case wherever k;s. supported these observations, the paper proposes a replacement approach, referred to as skyline ordering, that forms a skyline-based partitioning of a given information set such Associate in Nursing order exists among the partitions. Then, set-wide maximization techniques could also be applied among every partition. economical algorithms are developed for skyline ordering and for resolving size constraints exploitation the skyline order. The results of intensive experiments show that skyline ordering yields a versatile framework for the efficient and scalable resolution of arbitrary size constraints on skyline queries
- **Xuemin Lin ; Yidong Yuan ; Qing Zhang ; Ying Zhang(2007)**, Skyline computation has many applications together with multi-criteria deciding. they studied the matter of selecting k skyline points so the number of points, that are dominated by a minimum of one amongst these k skyline points, is maximized. we tend to 1st gift an efficient dynamic programming primarily based actual algorithmic rule during a 2d-space. Then, we tend to show that the matter is NP-hard once the spatial property is three or a lot of and it are often around resolved by a polynomial time algorithmic rule with the warranted approximation quantitative relation $1 - 1/e$. To speed-up the computation, Associate in Nursing economical, scalable, index-based randomized algorithm is developed by applying the FM probabilistic counting technique. A comprehensive performance analysis demonstrates that our randomized technique is incredibly efficient, extremely correct, and climbable.

Database entrance logs are the initial point for many forms of database administration, from database performance modification, to safety auditing, to standard plan, and many more. Unfortunately, query logs are also bulky and unwieldy, and it can be complicated for an analyst to remove broad patterns from the set of queries initiate therein. Clustering is a normal first step towards perceptive

the massive query logs. However, many clustering methods rely on the concept of pairwise similarity, which is difficult to compute for SQL queries, particularly when the underlying data and database schema is occupied.

- **Disadvantages**

- Clustering is not be done accurately
- Time consumption was very large

- **2.1. Proposed System**

Nothing of the similarity metrics execute as well as preferred, so suggest and estimate a pre-processing step to create more standard, uniform query representations by leveraging query correspondence rules and data partitioning process. These rules are commonly used by database management systems when parsing and evaluating SQL queries. This process extensively improves the feature of all three distance metrics. Examine and recognize sources of errors in the clustering process. Concretely, the specific contributions of this article are:

- A review of presented SQL query similarity metrics
- An estimation of these metrics on multiple query logs, and
- Applying query standardization techniques to get better query clustering accurateness.

3. SYSTEM DESIGN

3.1. Existing System

Skyline query may be a useful gizmo to seek out enticing products which provide a lot of interesting and preferred choices for customers. However, the dimensions of the skyline query results cannot be controlled flexibly. consequently, several analysis efforts are dedicated to cope with this downside. the present approaches to handle this downside are developed to spot k representative skylines that have the utmost dominant capability or the maximum diversification. The skyline query aims to come the purposes that aren't dominated by the other point. However, most of the works regarding the skyline query simply analyse individual points, and that they are inappropriate to several applications that decision for analysis of teams of various points. motivated by this, cluster are developed and paid growing attention. In most of the cluster skyline queries, optimum teams are computed by the dominance relationship between corresponding aggregate-based points of various groups. Su et al. developed prime k combinatorial skyline query (kCSQ). It returns those combinatorial skyline tuples whose mixture values for a specific attribute are most. Since solely the first k combinations are needed, the k-CSQ query method will be simplified.

- **3.2. Proposed System**

- formulate the COPC problem to retrieve best skyline product combinations that satisfy the customer's payment constraint and produce the utmost discount rate. To tackle the COPC problem, we have a tendency to propose a precise algorithmic program, style AN approximate algorithm with an approximate certain, ANd develop a progressive greedy algorithmic program to spice up the performance. we have a tendency to conduct a client study to verify the numerous of our COPC problem. in addition, the experimental results on each real and synthetic

dataset illustrate the effectiveness and potency of the projected algorithms. This work opens to some promising directions for future work. First, additionally to mixtures of solid product, we are going to concentrate on the COPC problem over product of various classes. After that, in reality, the customer's demands are diversification and individuation, and it's vital and attention-grabbing to figure best product mixtures that meet totally different client demands like save or pay the foremost money under their budgets. Last however not least, we have a tendency to may also analysis high k COPC problem that aims to figure k optimal product combinations because of customer demands based on the works.

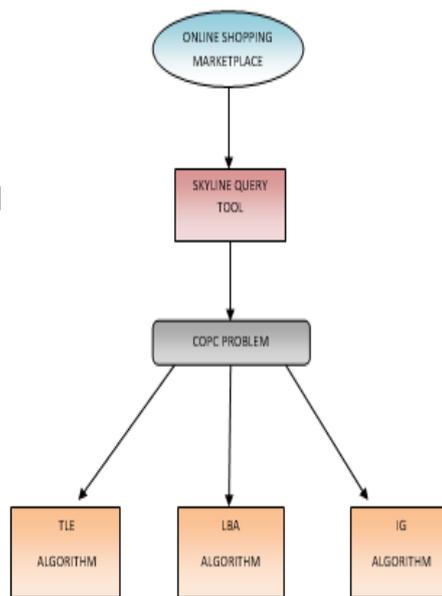


Fig. 1:- Architectural Diagram

3. CONCLUSION

we formulate the COPC problem to retrieve optimum skyline product combos that satisfy the customer's payment constraint and convey the utmost discount rate. To tackle the COPC problem, we tend to propose a precise rule, style associate degree approximate rule with an approximate sure, associate degreeed develop an incremental greedy rule to spice up the performance. we tend to conduct a client study to verify the numerous of our COPC problem. to boot, the experimental results on each real and artificial datasets illustrate the effectiveness and potency of the planned algorithms. This work opens to some promising directions for future work. First, additionally to combos of unvaried product, we are going to specialize in the COPC problem over product of various classes. After that, in reality, the customer's demands are diversification and individuation, and it's important and attention-grabbing to reason optimum product combos that meet completely different client demands like save or pay the foremost cash under their budgets. Last however not least, we tend to might additionally analysis high k COPC

problem that aims to reason k optimum product combos thanks to customer demands based on the works.

ACKNOWLEDGMENT

I would like to thank our honorable Principal, Dr. R. P. Singh and my special thanks to my guide A.D. Pathak and Head of Department Prof. Dinesh. D. Patil, & sincere thanks to all the respected teaching faculties of Department of Computer Science & Engineering of Hindi Seva Mandal's, Shri SantGadge Baba College of Engineering & Technology, Bhusawal. My special gratitude to all the authors of reference paper that are referred by me.

REFERENCES

- [1] S. Borzsönyyi, D. Kossmann, and K. Stocker, "The skyline operator," in Proc. Int'l Conf. Data Eng. (ICDE), pp. 421–430, 2001.
- [2] Q. Wan, R. C.-W. Wong, I. F. Ilyas, M. T. Özsu, and Y. Peng, "Creating competitive products," Proc. of the VLDB Endowment, vol. 2, no. 1, pp. 898–909, 2009.
- [3] I.-F. Su, Y.-C. Chung, and C. Lee, "Top-k combinatorial skyline queries," in Database Systems for Advanced Applications, pp. 79–93, Springer, 2010.
- [4] Y.-C. Chung, I.-F. Su, and C. Lee, "Efficient computation of combinatorial skyline queries," Information Systems, vol. 38, no. 3, pp. 369–387, 2013.
- [5] H. Im and S. Park, "Group skyline computation," Information Sciences, vol. 188, pp. 151–169, 2012.
- [6] M. Magnani and I. Assent, "From stars to galaxies: skyline queries on aggregate data," in Proc. 16th Int'l Conf. on Extending Database Technology, pp. 477–488, ACM, 2013.
- [7] N. Zhang, C. Li, N. Hassan, S. Rajasekaran, and G. Das, "On skyline groups," IEEE Trans. on Knowl. Data Eng., vol. 26, no. 4, pp. 942–956, 2014.
- [8] J. Liu, L. Xiong, J. Pei, J. Luo, and H. Zhang, "Finding pareto optimal groups: Group-based skyline," Proc. of the VLDB Endowment, vol. 8, no. 13, 2015.