

# k-Most Demanding Products Discovery with Maximum Expected Customers

Sofiya S. Mujawar  
Dr.D.Y.Patil College Of Engg, Ambi, Pune  
Pune University

Santosh Biradar  
Dr.D.Y.Patil College Of Engg,Ambi, Pune  
Pune University

## ABSTRACT

Product planning is a basic stage, figuring variables affecting the conclusions needs to made to the center to see the coveted possibilities, we tended to an issue of production arrangements, named  $k$ -most demanding products ( $k$ -MDP) finding. Given a set of customers demanding a certain kind of products with different traits, a set of existing products of the sort, a set of candidate products that can be offered by a company, and a positive integer number  $k$ , we need to help the company to choose  $k$  products from the candidate products such that the normal number of the aggregate customers for the  $k$  products is boosted. We demonstrate the issue is NP hard when the quantity of characteristics for a product is 3 or more. One greedy algorithm is proposed to discover surmised answer for the issue. We additionally endeavor to discover the optimal arrangement of the issue by assessing the upper bound of the normal number of the aggregate customers for a set of  $k$  candidate products for lessening the hunt space of the optimal arrangement. A precise algorithm is then given to discover the optimal arrangement of the issue by utilizing this pruning method. The investigation results exhibit that both the productivity and memory prerequisite of the careful algorithm are similar to those for the voracious algorithm, and the insatiable algorithm is well versatile concerning  $k$ .

## General Terms

Candidate Products, NP Hard, Product Planning.

## Keywords

Algorithms for data and knowledge management, decision support, query processing,  $k$ -MDP.

## 1. INTRODUCTION

Microeconomics is an extension of economics, which ponders how customers and producers settle on choices and how they associate in business sectors [8]. As of late, there has been critical enthusiasm toward creating successful procedures for specially appointed pursuit and recovery in unstructured and organized data repositories, for example, text accumulations and social databases. Specifically, a substantial number of rising applications oblige exploratory questioning on such databases; samples incorporate clients wishing to query databases and lists of items, for example, homes, autos, cams, restaurants, or articles, for example, news and employment promotions. Clients searching these databases ordinarily execute look questions through open front-end interfaces to these databases. Ordinary questions may determine sets of essential words if there should be an occurrence of text databases, or the craved estimations of specific characteristics in the event of organized social databases. The query noting framework answers such queries by either giving back all data questions that fulfill the question conditions, or may rank and furnish a proportional payback  $k$  data objects. Client inclination is a critical element in settling on choices of item

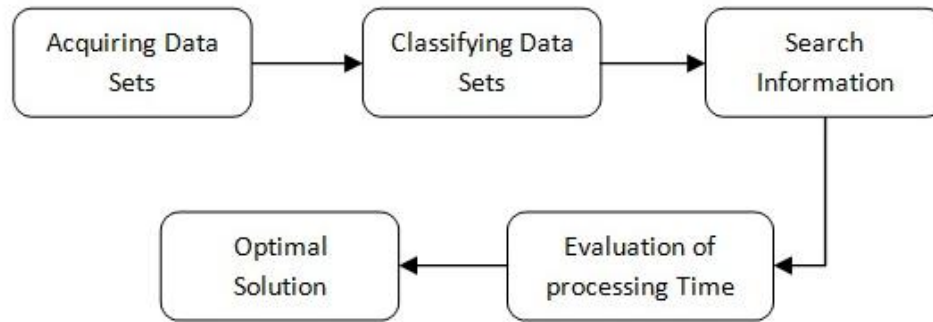
deals, which therefore turns into one noteworthy concern in microeconomics. Kleinberg et al. [6] brought up that, when making creation arrangements or showcasing methods, organizations typically need to recognize one with the most noteworthy utility or worth. They guaranteed that the utility or estimation of a creation arrangement can be displayed as an issue that reflects the association of the organization with different operators, for example, customers and contenders. Propelled by this, to look into rivalry, the issue considered in this paper is to recognize the creation plan with the most elevated utility for an organization, where the utility of a generation arrangement is assessed as indicated by the normal number of the aggregate customers for the chose items in the arrangement. Hence, we have characterized the issue of  $k$ -Most Demanding Products ( $k$ -MDP) finding. Given a set of customers requesting a certain kind of items with numerous traits, a set of existing results of the sort, and a set of hopeful items that can be offered by an organization, we need to help the organization to choose  $k$  items from these applicant items such that the normal number of the aggregate customers for the  $k$  items is expanded.

In this paper, we have briefly studied the problem of  $k$ -MDP. Also, we have optimized the problem of the objective functions. We have also proved that, if the number of attributes of the product is more than 3, then  $k$ -MDP discovery is a NP Hard problem. We have proposed 2 greedy algorithms to find the approximate solutions for the problem of  $k$ -MDP discovery;

1. Single-Product-Based Greedy (SPG) Algorithm,
2. Incremental-Based Greedy (IG) Algorithm.

We also attempted to get the finest solution to this problem by simply estimating the lower bound and upper bound of the number of expected customers for set of the candidate products. This helps in reducing the search space of the best solution. By using the pruning strategies, we have also proposed two identical algorithms, to find the optimal solution for the problem. To verify the effectiveness and the efficiency of these proposed algorithms, a systematic performance study is performed.

The probability of a product purchased by a customer may be influenced by the values of the quality describers of the product. In addition, in some applications, nominal attributes are used to describe the characteristics of a product in some aspects, whose orderings depend on the preferences of the users. How to extend our framework for these additional issues is under our investigation. The remaining paper is summarized as: Section II gives the quick look on the previous approaches proposed in the same field and tries to understand them. The section III concludes the paper with the future works in this field.



**Fig. 1 Proposed system architecture**

## 2. LITERATURE SURVEY

The dominant part of exploration applicable to microeconomic issues has concentrated on the potential customers finding. This is to bail an organization discover the potential customers who may be occupied with its defined item, and afterward the organization can publicize the item to the potential customers. Thusly, the organization may increase benefit just from the distinguished potential customers. To draw in more customers' consideration, the issue of the item points of interest disclosure tended to in [7] and [11] is to discover the benefits of the indicated result of an organization. The organization can advance the item by utilizing the discovered benefits, and further development the aggressiveness of the organization in the business sector. In any case, the consideration of the item preferences disclosure is focused on a current item whose qualities have been known, and thusly the item may not fulfill the customers despite the fact that its benefits are known.

Numerous studies like, [5], [12], [14] have managed the potential customers finding, for example, the opposite k-closest neighbor query [3], the converse horizon question [4], and the opposite top-k query [1]. The ideas of these works are comparable. Given a set of client inclination and a pointed out item, the queries give back where it's due whose most loved items contain the defined item as per their client inclination. The visibility of the item is, accordingly, restricted to the potential customers. Then again, the objectives of the studies in [7] and [11] are to expand the visibility of the pointed out item by finding the item preferences. The study in [11] goes for finding gimmicks of an item by which the rank of the detailed item is the highest point of every last one of items as per a given scoring capacity. Since it doesn't mull over client necessities, thusly the customers may not be occupied with the found item points of interest. Considering the client prerequisites, Miah et al. [7] propose a calculation to pick k gimmicks of the particular item, which fulfill the most extreme number of customers. Utilizing the discovered benefits to advance the item ought to have the higher chance to pull in more customers' consideration.

To help organizations create the items which are well known with the customers, the point of the studies in [2], [9], and [15] is to focus the right situating for items in the generation plan. Given a set of existing items with numerous segments, Wan et al. [9] consider the issue of creating preferred items over existing ones with agreeable organizations. On the other hand, the client prerequisites are not mulled over, which is one main consideration in microeconomics. Moreover, the quantity of the new items can be greatly substantial. As an issue, the chief of the organization may be overpowered when he/she needs to choose a few new items physically to

distinguish the ones that will in the long run be viewed as aggressive with the current items.

Li et al. [2] develop the idea of strength, utilized as a part of the horizon administrator [10], [13], for business examination. Given a set of client prerequisites and the benefit stipulation of an organization, the issue tended to in [2] is to distinguish the item ruling the biggest client necessities, which fulfills the benefit obligation of the organization. Stretched out from [2], assume there are various organizations with their separate benefit obligations and a set of client prerequisites, by looking into rivalry, the objective of [15] is to discover one item with the most extreme expected number of the customers for each one organization, which fulfills the benefit demand of the organization. The discovered items in [2] and [15] needs to fulfill the benefit requirements of the organizations, which may be much hard to specify. Additionally, to pull in more customers, an organization may decide to offer numerous items in the meantime. The studies in [2] and [15] consider stand out item for an organization, and subsequently can't reflect the need in this present reality. In this paper, taking both item rivalry and client prerequisites into attention, we propose systems to discover k items from all applicant items an organization can offer such that the normal number of the aggregate customers for the k items is augmented.

## 3. CONCLUSION

In this paper, we form the k-MDP finding issue for deciding k most requesting items with the most astounding expected number of the aggregate customers. We have demonstrated that the issue is NP-hard when the quantity of value describers for an item is 3 or more. As needs be, two greedy calculations, the SPG calculation and the IG calculation, are proposed to discover the results approaching the ideal arrangement. Also, two successful pruning techniques are given to create two calculations, the APR calculation and the UBP calculation, for endeavoring to discover the ideal arrangement of the issue. The execution for all the proposed calculations on productivity is enhanced with the BMI list structure.

The likelihood of an item bought by a client may be impacted by the estimations of the quality describers of the item. What's more, in a few applications, ostensible ascribes are utilized to portray the attributes of an item in a few angles, whose orderings rely on upon the inclination of the clients. Step by step instructions to develop our schema for these extra issues is under our examination.

## 4. REFERENCES

- [1] A. Vlachou, C. Doulkeridis, Y. Kotidis, and K. Norvag, "Reverse Top-k Queries," Proc. 26th Int'l Conf. Data Eng., pp. 365-376, 2010.

- [2] C. Li, B.C. Ooi, A.K.H. Tung, and S. Wang, “DADA: A Data Cube for Dominant Relationship Analysis,” Proc. 25th ACM SIGMOD Int. Conf. Management of Data, pp. 659-670, 2006.
- [3] E. Aichert, C. Bohm, P. Kroger, P. Kunath, A. Pryakhin, and M. Renz, “Efficient Reverse k-Nearest Neighbor Search in Arbitrary Metric Spaces,” Proc. 25<sup>th</sup> ACM SIGMOD Int Conf. Management of Data, pp. 515-526, 2006.
- [4] E. Dellis B. Seeger, “Efficient Computation of Reverse Skyline Queries”, Proc. 33rd Int. Conf. Very Large Data Bases, pp. 291-302, 2007.
- [5] F. Korn, S. Muthukrishnan, “Influence Sets Based on Reverse nearest Neighbor Queries”, Proc. 19th ACM SIGMOD Int. Conf. Management of Data, pp. 201-212, 2000.
- [6] J. Kleinberg, C. Papadimitriou, and P. Raghavan, “A Microeconomic View of Data Mining”, Data Mining and Knowledge Discovery, vol. 2, no. 4, pp. 311-322, 1998.
- [7] M. Miah, G. Das, V. Hristidis, and H. Mannila, “Determining Attributes to Maximize Visibility of Objects”, IEEE Transactions on Knowledge and Data Engineering, v.21 n.7, p.959-973, July 2009.
- [8] N. G. Mankiw, “Principles of Economics”, 5<sup>th</sup> ed. South-Western College Pub, 2008.
- [9] Q. Wan, R. C. -W. Wong, I. F. Ilyas, M. T. Ozsu, Y. Peng, “Creating Competitive Products,” Proc. 35th Int. Conf. Very Large Data Bases, pp. 898-909, 2009.
- [10] S. Borzsonyi, D. Kossmann, and K. Stocker, “The Skyline Operator,” Proc. 17th Int. Conf. Data Eng., pp. 421-430, 2001.
- [11] T. Wu, D. Xin, Q. Mei, and J. Han, “Promotion Analysis in Multi-Dimensional Space”, Proc. 35th Int. Conf. Very Large Data Bases, pp. 109-120, 2009.
- [12] X. Lian and L. Chen, “Monochromatic and Bichromatic Reverse Skyline Search over Uncertain Databases”, Proc. 27th ACM SIGMOD Int. Conf. Management of Data, pp. 213-226, 2008.
- [13] X. Lin, Y. Yuan, Q. Zhang, and Y. Zhang, “Selecting Stars: The k Most Representative Skyline Operator,” Proc. 23rd Int. Conf. Data Eng., pp.86-95, 2007.
- [14] Y. Tao, D. Papadias, and X. Lian, “Reverse kNN Search in Arbitrary Dimensionality”, Proc. 30th Int. Conf. very Large Data Bases, pp. 744-755, 2004.
- [15] Z. Zhang, L. V. S. Lakshmanan, and A. K. H. Tung, “On Domination Game Analysis for Microeconomic Data Mining,” ACM Trans. Knowledge Discovery from Data, vol. 2, no. 4, pp. 18-44, 2009.