

УДК 004.655.3:004.02

Shablia D.¹, Hunko I.¹, Polska O.², Kudermetov R.³

¹ студ. гр. КНТ-514 НУ «Запорізька політехніка»

² старш. викл. НУ «Запорізька політехніка»

³ доц. НУ «Запорізька політехніка»

EFFICIENCY OF SKYLINE ALGORITHMS USING THE EXAMPLE OF THE SAKILA DATABASE

In modern information systems, multidimensional data processing is critical for decision-making in business, e-commerce, and analytics. In relational databases, there is often a need to select optimal objects based on several criteria simultaneously. Skyline queries enable the selection of such Pareto-optimal objects without prior aggregation or ranking of criteria.

The problem lies in choosing an effective Skyline algorithm for multidimensional data, as classical methods (Naive, BNL, SFS, D&C) demonstrate different performance depending on the size of the dataset and the number of criteria [1]. The simple Naive approach is understandable and easy to implement, but its complexity, $O(n^2)$ makes it unsuitable for large datasets. The BNL and SFS algorithms reduce the number of comparisons and utilise sorting or buffering; however, their efficiency depends on the data distribution and the choice of sorting functions. The D&C algorithm is more scalable, but its implementation in SQL is complex.

To study the effectiveness of these algorithms, the Sakila database can be used. The Sakila database is based on a realistic business model of the video rental industry [2]. Sakila tables (e.g., film and customer) contain both numeric and categorical attributes, enabling the simulation of multidimensional queries. The database reflects real business scenarios and allows for the evaluation of query performance, the impact of indexing, and optimisation techniques. The size of the tables (film ~1000 records, customer ~600) makes Sakila convenient for educational experiments; however, for conclusions regarding scalability, it is necessary to extend the dataset synthetically.

The results of this research will help determine optimal strategies for processing multidimensional queries in relational databases and facilitate the understanding of analytical methods and SQL query optimisation in practical scenarios.

REFERENCES

1. Zhang X., Cavusoglu H. Skyline operators in multi-criteria decision making: A review of characterization, comparison, and perspectives. *Expert Systems with Applications*. 2026. Vol. 306, 130889. DOI: 10.1016/j.eswa.2025.130889
2. Sakila Sample Database. *MySQL*. URL: <https://dev.mysql.com/doc/sakila/en/> (date of access: 15.03.2026)