Matrix Reordering: An Overview of its Applications, Challenges and Future Directions

Fatimah Binta Abdullahi
Dept. of Computer Science
Ahmadu Bello University Zaria
zeeh429@gmail.com

Salihu IDI Dishing
Dept. of Computer Science
Ahmadu Bello University Zaria
dishingsalihu@gmail.com

***Abstract*—Matrix reordering, also known as rearrangeable matrices or permutable matrices, are a specialized type of mathematical structure that allows for the rearrangement of rows and columns while maintaining a certain property. This property is typically some form of symmetry or similarity among the matrix's elements. Matrix reordering find applications in various fields, including graph theory, combinatorics, optimization, and information retrieval. It is a generalization of traditional matrices that allow the order of their elements to be changed dynamically during computation. Although, numerous matrix reordering application were reported in the literature however, a survey that describes the current state of research and application is unavailable. This paper presents an overview, applications and challenges of matrix reordering techniques that will promote further advancement in this field. The objectives of this paper are two folds: (i) to understand the applications and challenges of recent work on matrix reordering, and (ii) to identify areas that require additional research. The findings of this work present the limitation of recent works and the future direction of matrix reordering for interested researchers is established.**

Keywords—Matrix reordering, data visualization, binary data.

I. INTRODUCTION

A matrix reordering is the process of rearranging the rows and columns of a matrix to improves its properties. This can be done for a variety of reasons such as to improve the readability and interpretability of the matrix, to reduce the computational cost of performing operations on the matrix or to discover hidden patterns and relationships in the data. Matrix reordering can be used as a visualization technique for exploring and presenting data [1, 2]. A matrix reordering may consist of a rectangular grid of cells, where each cell contains a data value. The user can interactively reorder the rows and columns of the matrix to reveal patterns in the data that might not be visible in a traditional table. Matrix reordering is also a good way to communicate data to a non-technical audience. The core idea behind matrix reordering is to allow for permutations of rows and columns while maintaining some sort of structure. This structure could be related to patterns in the matrix's elements, graph connectivity, eigenvalues, or other algebraic or combinatorial properties. Matrix visualizations are based on the concept of matrix reordering, i.e. a matrix whose rows and columns can be reordered without information loss. This was proposed by Bertin (2010) in 1967 in his first edition of the book "Semiology of Graphics", Since then many researchers have been proposing innovative reordering algorithms that try to reveal patterns that are easy to be visually perceived. As noted, a matrix reordering refers to a matrix whose rows and columns can be rearranged in a specific way to achieve a certain desirable property, often leading to improved efficiency, reduced computational complexity, or better visualization in various applications. Matrix reordering aims to minimize specific measures such as bandwidth, profile, or fill-in, which can affect the performance of algorithms operating on the matrix. The fundamental principle underlying matrix reordering is that the arrangement of rows and columns can significantly impact the matrix's structural properties, and the efficiency of operations performed on it. By strategically reordering the matrix elements, certain patterns can be exploited to reduce the costs associated with various computations, such as matrix-vector multiplication, solving linear systems, and graph analysis. Some key examples of matrix reordering areas are as follows:

1. In archaeology, matrix reordering technique have been used to order artifacts and archaeological sites where the technique have been used to group artifacts together based on their styles, materials or location.
2. In anthropology, the technique has been used to analyze social networks and kinship relationships. Anthropologist can used matrix reordering technique to group people together based on their family ties, social status or religious beliefs.
3. In psychology, matrix reordering can be used to analyze personal traits and cognitive abilities. The technique can group people together based on their scores on personality test or intelligence tests.
4. In biology. He technique have been used to analyze gene expression data and other biological data.

II. RELATED WORK

The history of matrix reordering can be traced back to the 19th century when archaeologists and anthropologists begin using it to order artifacts and archaeological sites chronologically. One of the first systematic methods for matrix reordering was developed by the English Egyptologist Petrie in the 1890s. mid-1960s. In the early 20th century matrix reordering began to be used in fields such as sociology, psychology and biology.

In the 1960s a French cartographer Jacques Bertin developed a new method for matrix reordering that used color and pattern coding to improve visual perception. Bertin was interested in developing new ways to visualize tabular data and believed that Matrix reordering could be a powerful tool for this purpose. In the early 1970s, Bertin developed a physical device called the Dominos system, which was one of the first implementations of a matrix reordering. The Dominos system consist of a rectangular grid of tiles, where each tile could be flipped over to reveal a different data value. The user could interactively flip the tiles to reorder the rows and columns of the matrix. In the 1980s, matrix reordering began to be implemented in software systems. One of the first software implementations was the Table Lens system [62], developed by researchers at the University of California, Berkeley. The Table Lens system allowed users to interactively reorder the rows and columns of a table using a mouse or touchscreen. In the 1990s, matrix reordering became more widely used in data visualization research. Researchers began to explore new ways to use matrix reordering to visualize different types of data. Bertin [16,17], in 1999, used the terms "construction" and "reconstruction" to describe respectively, the process of generating a matrix reordering from tabular data and the process of moving rows and columns. For example, researchers developed matrix reordering for visualizing financial data, social network data, and gene expression data. As noted above, Petrie applied a reordering technique to study archaeological data [18, 19]. Since then, a number of reordering methods have been used with respect to a variety of applications. For example, Forsyth and Katz [25] were the first to introduce the idea of rearranging the rows and columns of "sociomatrices", tabular representations of data collected as part of some sociometric study, so as to obtain a better presentation of the results of sociometric tests. Brainerd and Robinson in [25], proposed a form of matrix; where the highest values in the matrix were located along the prime diagonal and decreased monotonically when moving away from this diagonal. This matrix become known as the "Robinson Matrix" or
(R-Matrix). In 2000, matrix reordering began to be used in commercial products. One of the first commercial products to use matrix reordering was the Tableau software suite. Tableau allows users to create interactive matrix reordering to visualize data from a variety of sources. Today, matrix reordering can be a visualization technique for exploring and presenting tabular data. They are used in a variety of industries, including finance, healthcare, and marketing. Another technique is seriation, an exploratory data analysis technique used to reorder objects into a sequence along a one-dimensional continuum so that it best reveals regularity and patterning among the whole series [18].

Figures 1 and 2, presents Bertin device for matrix construction and reconstruction and the matrix reordering interface.

Figure 1: Bertin device for matrix construction and reconstruction [24]

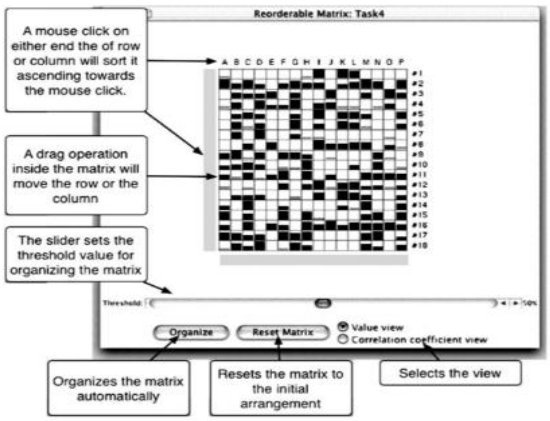

Figure 2. The matrix reordering interface [24].

Some Advantages of matrix reordering are as follows:

1 Enhanced efficiency: Matrix reordering can lead to more efficient algorithms and processes in various domains. For example, in graph theory, reordering matrices can simplify graph isomorphism testing, and in information retrieval, they can speed up search algorithms.
2 Improved Clustering and Analysis: In spectral clustering and data analysis, reordering matrices can enhance the separation of clusters and uncover hidden patterns in data, leading to more accurate insights and conclusions.
3 Flexibility in representations: Matrix reordering provide flexibility on how data is represented. By rearranging rows and columns, thus making it possible to emphasize certain relationships or properties, which can be advantageous in different analysis scenarios.
4 Simplification of problems: In some cases, matrix reordering can transform complex problems into simpler ones, making them more approachable for analysis and solution.

5 Effective data preprocessing: In machine learning and data mining, matrix reordering can be used in data preprocessing to improve the performance of algorithms by transforming the data to a more suitable representation.

III. Material and Methods

In this session, we present a taxonomy that provides a comprehensive overview of matrix reordering, categorizing them based on various dimensions, including their properties, applications, and characteristics.

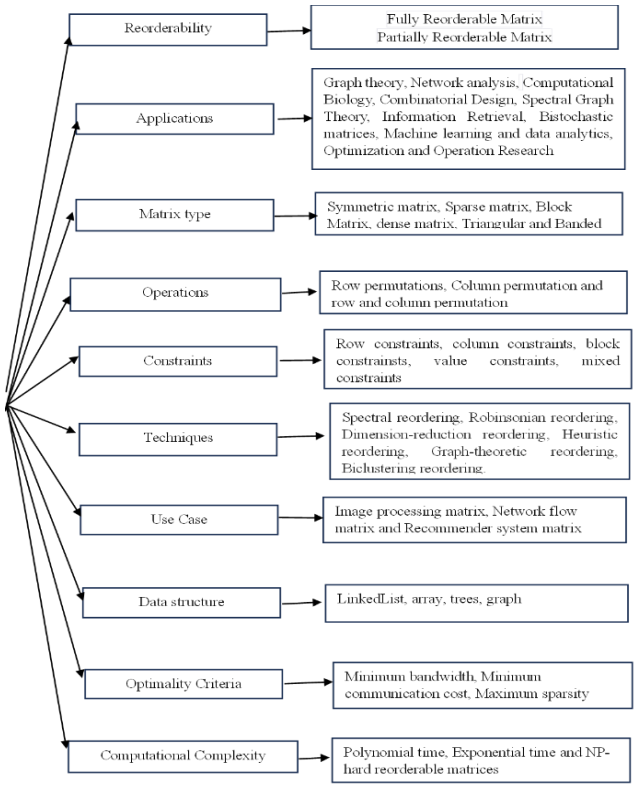

Figure 3: Taxonomy of matrix reordering

Figure 4 present an example of townships dataset where after reordering using different techniques one can find visual answer to local and global patterns found in the data as shown in figures 5,6 and 7 respectively.

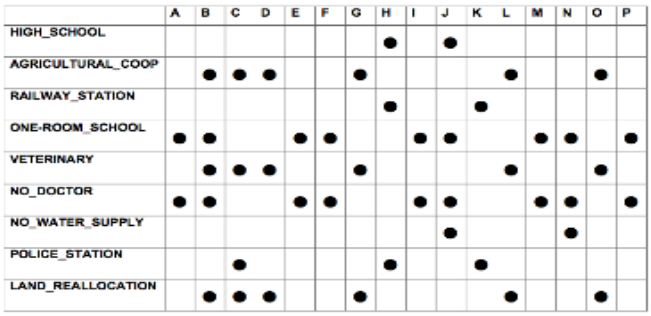

Figure 4: Original township matrix [24]

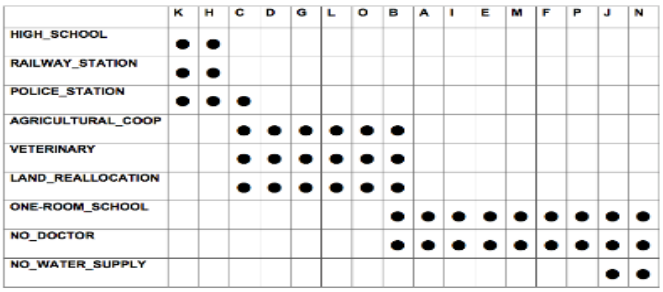

Figure 5: matrix reordered using manual method [24]

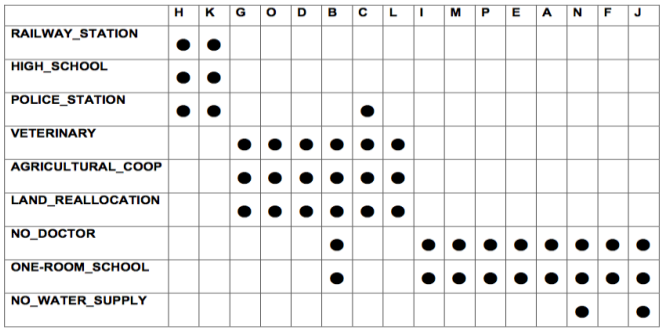

Figure 5: matrix reordered using Sugiyama method []

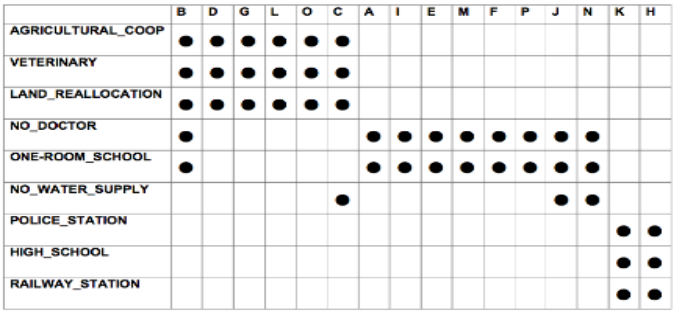

Figure 5: matrix reordered using 2D Sort method [24]

A. Matrix reordering Algorithms

Several techniques have been developed to reorder matrices effectively, each catering for specific applications and scenarios. Here are some common implementations:

i. *Reverse Cuthill-McKee (RCM):* This algorithm focuses on reducing the bandwidth of a matrix, which can accelerate the solution of sparse linear systems. It starts from a seed node and performs a breadth-first search to build a reordered matrix with minimized bandwidth. [3, 27]
ii. *Banded Pattern Mining (BPM):* The algorithm focuses on rearranging the elements in a given zero-one datasets to identify a pattern that start from the top left corner and end at the bottom right corner. [10,11]
iii. *Minimum Degree Algorithm:* This approach aims to reduce the degree (number of non-zero elements) of nodes in a graph representation of the matrix. By eliminating nodes with the smallest degrees, the fill-in during factorization is minimized, improving computational efficiency.
iv. *Nested Dissection:* Often used in finite element analysis, this technique recursively partitions the matrix into smaller submatrices. This process creates a hierarchical structure that can lead to reduced fill-in during factorization.
v. *Spectral Methods:* Leveraging spectral properties of matrices, these methods use eigenvectors or eigenvalues to reorder the matrix. Techniques like spectral bisection and spectral clustering focus on partitioning the matrix into submatrices with reduced bandwidth or fill-in. [7]
vi. *2D Sort Method:* The 2D Sort banding algorithm is concerned with arranging a two-dimensional matrix iteratively so as to reveal a pattern where the non-zero entries, called “black areas” [24] are arranged along the leading diagonal.
vii. *Sugiyama Method:* The Sugiyama algorithm was originally developed to support graph drawing where the number of edges “cross-overs” are minimized. The graphs of interest were translated into adjacency matrices where a “1” entry, indicated an edge between the corresponding vertices, represented by the indicated row and column [14], for bipartite graph layout.
viii. *Heuristic Approaches:* Various heuristic strategies, such as genetic algorithms, simulated annealing, and greedy algorithms, have been proposed for matrix reordering. These methods explore different permutations of row and column orders to find locally optimal solutions.
ix. *Graph-Based Approaches***:** Matrices can often be represented as graphs, where rows and columns are nodes and non-zero entries indicate edges. Graph algorithms like Minimum Spanning Trees or Graph Coloring can be adapted to reorder matrices [13].
x. *Simulated annealing*: This algorithm is a probabilistic technique that searches for the optimal order of rows and columns by iteratively swapping pairs of rows and columns [].
xi. *Greedy algorithms*: This algorithm iteratively selects the best row or column to add to the reordered matrix based on a given criterion.

B. Evaluation Metrics for matrix reordering algorithms

i. *Execution time*: The time taken by the algorithm to reorder the matrix. Faster algorithms are generally preferred.
ii. *Memory usage*: The amount of memory used by the algorithm to reorder the matrix.
iii. *Quality of the reordering*: This refers to how well the algorithm reorders the matrix to reveal patterns and relationships in the data.
iv. *Scalability*: This refers to how well the algorithm performs on large datasets.

V. RESULTS AND DISCUSSION

This section presents overview of some areas where research on matrix reordering has been conducted, their challenges and potential future directions.

A. Potential Applications of Matrix reordering

The following are overview of some areas where research on matrix reordering have been conducted:

i. Network Analysis: In network analysis, a common problem is to find the most efficient way to route data through a network of interconnected nodes. This can be represented by a matrix, where each entry represents the distance or cost to travel between two nodes. By reordering the rows and columns of this matrix, we can determine the optimal routing of data through the network, such that the overall cost or distance is minimized. This is known as the "Traveling Salesman Problem" and can be solved using matrix reordering. [28]
ii. Computational Biology: Another application of matrix reordering is in computational biology, specifically in sequence alignment. Sequence alignment is the process of arranging DNA or protein sequences in a way that maximizes the similarity between them. This is often represented by a matrix, where each element represents the similarity score between two sequences. By reordering the rows and columns of this matrix, we can find the optimal alignment of sequences, which can have important implications for understanding genetic.
iii. Graph theory: The research in graph theory explores the properties of matrix reordering, by focusing on the preservation of graph structures under row and column permutations. This includes the study of adjacency matrices, incidence matrices, and permutation matrices, for example, studies have established relationships between graph structures and their matrix representations. In addition, theoretical research results have shown that certain classes of matrix reordering can help determine graph isomorphism, connectivity and automorphism more efficiently. [13]

iv. Combinatorial Design: Matrix reordering are closely related to combinatorial designs and Latin squares. Research in this area aims to construct and analyze designs that remain invariant under row and column permutations. These designs have applications in coding theory, cryptography. Matrix reordering have been utilized in the construction of error-correcting codes and cryptographic systems, these matrices play a role in ensuring error tolerance and security while transmitting and storing information. [9]

v. Spectral Graph Theory: In spectral graph theory, research has been conducted to determine the impact of matrix reordering on eigenvalues and eigenvectors. Spectral clustering algorithms often involve working with similarity matrices that can be reordered to improve clustering results. Here, matrix reordering have been used to enhance spectral clustering algorithms by appropriately reordering the affinity matrix to improve the separation of clusters and the overall clustering accuracy. [10, 14]

vi. Information Retrieval: Research in this area has explored the reordering of term-document matrices to enhance search algorithms and document ranking. This has implications for improving the efficiency and accuracy of search engines and text analysis tasks. [26]

vii. Optimization and Operation Research: Research conducted in optimization and operations research has investigated how matrix reordering can be utilized to solve various combinatorial optimization problems more efficiently. This includes tasks like linear programming, assignment problems, and transportation problems. [22, 23]

B. Challenges of Matrix reordering

In this work, the use of matrix reordering is presented and has shown great promise as well as achieved significant advancements. However, several challenges still exist that researchers needs to address to further improve the accuracy and robustness of the systems.

i. Algorithmic Complexity: While some matrix reordering has shown promising in solving various problems more efficiently, developing algorithms that can effectively exploit the properties of these matrices and take full advantage can be complex and required specialized techniques. The development of efficient algorithms for reordering and utilizing these matrices is an ongoing research area and remains a challenge.

ii. *Generalization of Findings*: Many findings in matrix reordering research have been specific to certain matrix types or problem domains. Generalizing these findings to broader classes of matrices or more diverse applications remains a challenge and requires a deep understanding of both the mathematical properties and the problem context.

iii. Optimal reordering Strategies: Identifying the optimal strategy for reordering matrices to achieve specific goals can be challenging. There might not be a one-size-fits-all solution, and finding effective strategies often requires domain-specific knowledge. In applications like information retrieval, determining the best way to reorder matrices for optimal results is a nontrivial task. Designing systematic and effective reordering strategies that consistently improve performance across diverse datasets remains an open challenge.

iv. Real-world Applicability: While theoretical findings are important, translating these findings into real-world applications with tangible benefits can be challenging. Real-world data often comes with noise, complexities and varying problem characteristics can affect the effectiveness of the reordering techniques in real-world scenarios.

v. Scalability and Big Data: Many matrix reordering techniques need to be scalable to handle large datasets and complex structures. Applying matrix reordering techniques to large datasets or complex structures might pose scalability issues, especially if the algorithms involved are computationally intensive. Therefore, developing an approach that can efficiently handle big data scenarios without sacrificing accuracy or efficiency is a significant challenge.

vi. Interdisciplinary Understanding: Matrix reordering research often involves bridging multiple disciplines and collaboration between researchers from different fields, such as mathematics, computer science, and other domain-specific fields can present challenges in terms of communication and understanding. Building a comprehensive understanding across these disciplines is crucial for developing meaningful applications.

vii. New Matrix Types and Properties: As research progresses, new matrix types and their properties that are amenable to reordering might be discovered. Understanding and characterizing these new matrix structures are important for broader applications.

viii. Benchmarking and Comparison: Evaluating the impact of reordering strategies requires effective benchmarking against existing methods. Developing standardized benchmarks and comparison frameworks is necessary for assessing the true potential of matrix reordering.

C. Potential Future Directions

Currently, the potential future directions for matrix reordering are focusing on addressing the existing challenges. As technology progresses and research in matrix reordering advances, it is expected that these potential future directions will lead to more accurate and robust matrix reordering techniques with widespread application in various domain.

i. Algorithmic Complexity: Researchers can explore the computational complexity of reordering algorithms and develop efficient techniques for various matrix types, investigate how matrix reordering can be used to optimize algorithmic performance in different computer science applications, such as data mining and graph algorithms. Develop approximation algorithms that leverage matrix reordering to solve NP-hard problems more effectively.

ii. Data Analysis and Machine Learning: Researchers can study how matrix reordering can enhance clustering and classification algorithms by promoting better separation of data points. Similarly, explore the use of matrix reordering in feature selection techniques to identify the most relevant attributes for improved model performance as well as investigate how matrix reordering can be integrated into dimensionality reduction method like PCA to enhance data visualization and analysis. Matrix reordering can also be utilized in machine learning algorithms to improve data processing and analysis. By intelligently reordering the feature matrices, it may be possible to optimize the performance of machine learning models, enhance data visualization, and improve algorithm efficiency.

iii. Graph Theory and Network Analysis: Matrix reordering can have implications for solving graph isomorphism and automorphism problems more efficiently by leveraging matrix symmetries. exploring how matrix reordering can further optimize graph algorithms such as shortest path finding, connectivity analysis, and centrality measures.

iv. Information Retrieval and Natural Language Processing: Researcher can study how matrix reordering can improve the efficiency and accuracy of search algorithms in information retrieval systems by investigating how the use of matrix reordering can uncover semantic relationships between terms and documents in text analysis tasks.

v. Practical Applications: Researchers can study how matrix reordering can be applied to real-world optimization problems in areas like supply chain management, logistics, and resource allocation.

vi. Coding and Cryptography: Novel coding schemes and cryptographic protocols based on matrix reordering can be explored to further enhance data security and transmission.

vii. Scalability and Big Data: Researchers can investigate how matrix reordering can be adapted to handle streaming data scenarios, where data arrives sequentially and needs to be processed in real time. Exploring parallel and distributed algorithms for reordering matrices to leverage the power of modern parallel computing architectures.

viii. Benchmarking and Evaluation: Develop standardized benchmarks to evaluate the performance of matrix reordering algorithms and compare them with existing methods.

ix. Network optimization: Matrix reordering can play a significant role in optimizing networks, such as transportation networks, communication networks, or social networks. By rearranging the adjacency matrices or distance matrices of these networks, it may be possible to minimize costs, improve connectivity, or enhance resource allocation.

x. Data compression and storage: Matrix reordering have the potential to be leveraged for efficient data compression and storage techniques. By rearranging data in a matrix form, it may be possible to exploit patterns or redundancies, leading to more compact data representation, faster processing and storage.

VI. CONCLUSION

This study presents an overview of matrix reordering. The work aims at assisting interested researchers to have an up-to-date review of works related to matrix reordering by presenting the applications, challenges and future directions. The findings demonstrate that matrix reordering have been applied to different domain areas. Based on this, a taxonomy is presented. The taxonomy revealed various dimensions, including their properties, applications as well as characteristics. Furthermore, open issues and challenges as recommendation for future research direction are discussed, this includes scalability, network optimization, algorithmic complexity, graph theory, and data compression and storage.

REFERENCES

[1] G.C. Garriga, E. Junttila, H. Mannila, “Banded structures in binary matrices. In: Proceedngs Knowledge Discovery in Data Mining” (KDD 2008), pp. 292–300.

[2] G.C Garriga, E. Junttila, H. Mannila, “Banded structure in binary matrices. Knowledge Information System”, 2001, vol. 28(1), pp.197–226.

[3] Brian Rodrigues, Fei Xiao, and Andrew Lim. A new node centroid algorithm for bandwidth minimisation. 18th International Joint Conference on Artificial Intelligence (IJCAI) Mexico. Research Collection Lee Kong China School Business, 2003

[4] F. Alizadeh, R.M Karp, L. A Newberg, D.K.. Weisser. “Physical mapping of chromosom es: a combinatorial problem” in molecular biology. Algorithmica, vol. 13(1–2), pp. 52–76, 1995

[5] Y. Cheng and G.M. Church. “Biclustering of expression data”. In: ISMB, 2000, vol. 8, pp.93–103

[6] S. Myllykangas, J. Himberg, T. Bohling, B. Nagy, J. Hollman, S. Knuutila,“DNA copy number amplification profiling of human neoplasms”. Oncogene 2006, vol.25, pp. 7324–7332

[7] J.E. Atkins, E.G. Boman, B. Hendrickson,“Spectral algorithm for seriation and the consecutive ones problem”. Journal of . Computing. SIAM, 1998, vol. 28, pp. 297–310

[8] M. Fortelius, M.F. Kai Puolamaki, H. Mannila, “Seriation in paleontological data using Markov chain monte carlo methods”. PLoS Comput. Biol. 2006, vol. 2, pp. 2

[9] E. Makinen, H. Siirtola, “The barycenter heuristic and the Matrix reordering”.Informatica, 2005, vol. 29, pp. 357–363.

[10] F.B Abdullahi, F. Coenen, R. Martin, “ Finding banded patterns in data: the banded pattern mining algorithm.” In: Madria, S., Hara, T. (eds.) DaWaK LNCS, 2015, vol. 9263, pp. 95–107. Springer.

[11] F.B. Abdullahi, “Finding Banded Patterns in 2-Dimrnsional Zero-Data Using Weighted Column and Row Scores” *Nigerian* Journal of Scientific Research (NJSR) 2018, .Vol 17(2) pages 186-196.

[12] Ch. H Papadimitriou. “The np-completeness of the bandwidth minimization problem. *Computing”,* 1976 vol. 16: pp. 263-270.

[13] M. Koebe and J. Kn¨ochel. “On the block alignment problem. Elektronische Informationsverarbeitung und Kybernetik”, 1990, vol. 26(7), pp. 377–387.

[14] K. Sugiyama, S. Tagawa, M. Toda. “ Methods for visual understanding of hierarchical system structures”. IEEE Trans. Syst. Man Cybern. 11, 1981, pp.109–125.

[15] K.E. Atkinson, “An Introduction to Numerical Analysis”. John Wiley & Sons, New York, 2008.

[16] J. Bertin. “Graphics and Graphic Information Processing”. Walter de Gruyter, New York. 1981

[17] J. Bertin. “Graphics and graphic information processing. In: Readings in Information Visualization”, pp. 62–65. Morgan Kaufmann Publishers Inc. 1999.

[18] I. Liiv.”Seriation and matrix reordering methods: an historical overview”. Stat. Anal. Data Min 2010, vol.. 3(2), pp. 70–91.

[19] F. Makinen, H. Siirtola."Reordering the Matrix reordering as an algorithmic problem". In: Anderson, M., Cheng, P., Haarslev, V. (eds.) Diagrams 2000. LNCS (LNAI), vol. 1889, pp. 453–468. Springer, Heidelberg (2000). https://doi.org/10. 1007/3-540-44590-0 37.

[20] C. Aykanat, A. Pinar, U. Catalyurek. "Permuting sparse rectangular matrices into block-diagonal form". SIAM J. Sci. Comput.2004, vol. 25, pp.1860–1879.

[21] S.B Deutsch, J.J. Martin. "An ordering algorithm for analysis of data arrays". Operation Research, 1971, vol. 19(6), pp.1350–1362.

[22] K.Y. Cheng. "Minimizing the bandwidth of sparse symmetric matrices. Computing", 1973, vol. 11(2), pp.103–110.

[23] K.Y. Cheng. "Note on minimizing the bandwidth of sparse, symmetric matrices". Computing, 1973, vol. 11(1), pp. 27–30.

[24] H. Siirtola and E. Makinen. "Constructing and reconstructing of Matrix reordering". *Information Visualisation*, 2005, vol. 4, pp.32-48

[25] W. S Robinson. A method for chronologically ordering archaeological deposits. American Antiquity, 16:293–301, 1951

[26] Baeza-Yates, R., RibeiroNeto., B.: Modern Information Retrieval. Addison-Wesley (1999

[27] Cuthill, A.E., McKee, J.: Reducing bandwidth of sparse symmentric matrices. In: Proceedings of the 1969 29th ACM National Conference, pp. 157–172 (1969)

[28] Arindam Banerjee, Chase Krumpelman, Joydeep Ghosh, Sugato Basu, and Rahmand J. Mooney. Model-based overlapping clustering. In Proceedings of the 11th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'05), pages 532–537, 2005.