



SYSTEMATIC LITERATURE REVIEW UNTUK IDENTIFIKASI METODE FRAGMENTASI PADA PERANCANGAN SISTEM DATABASE TERDISTRIBUSI

Erba Lutfina¹ ✉, Ahmad Nugroho², Muhammad Zakki Abdillah³

1,2,3 Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Nasional Karangturi, Jl. Raden Patah 182-192, Kota Semarang, 50227, Indonesia

DOI: <http://dx.doi.org/10.26623/jtphp.v13i1.1845.kodeartikel>

Info Artikel

Sejarah Artikel:

Disubmit 05-07-2022

Direvisi 12-07-2022

Disetujui 14-07-2022

Keywords:

*Systematic literature review;
distributed database;
fragmentation*

Abstrak

Kompleks dan besarnya suatu database yang ada pada suatu sistem dapat mengakibatkan performa menjadi buruk seperti membesarnya waktu dan biaya akses pada suatu sistem. Terdapat berbagai cara untuk mengurangi waktu dan biaya akses data serta meningkatkan performa suatu database, salah satunya dengan merancang sebuah Database Terdistribusi. Belum diketahui berapa banyak metode yang dapat digunakan untuk merancang sebuah database terdistribusi tersebut. Tujuan dari penelitian ini adalah mengidentifikasi metode perancangan database terdistribusi dengan metode Systematic Literature Review (SLR) yang datanya diperoleh dari jurnal yang terkait pada tahun 2012–2022. Review ini berguna bagi peneliti maupun praktisi, untuk para peneliti penelitian ini akan membantu mereka untuk melihat tren dalam bidang fragmentasi vertikal pada database terdistribusi.

Abstract

Large and complex databases that exist on a system can result in poor performance such as increasing access time and costs on a system. There are various ways to reduce the time and cost of data access and improve the performance of a database, one of them is by designing a Distributed Database. It is not known how many methods can be used to design a distributed database. This study aims to identify a distributed database design method using the Systematic Literature Review (SLR) method whose data were obtained from related journals from 2012–2022. This review is helpful for researchers as well as practitioners, for researchers this research will help them to see trends in the field of vertical fragmentation in distributed databases.

✉ Alamat Korespondensi:
E-mail: erbalutfina@gmail.com

1. Pendahuluan

Kompleks dan besarnya suatu database yang ada pada suatu sistem dapat mengakibatkan performa menjadi buruk (Jawad & Khan, 2013) seperti membesarnya waktu dan biaya akses pada suatu sistem (Lutfina, 2018). Performa yang menurun berdampak pada membengkaknya waktu dan biaya akses yang disebabkan karena fungsi *query* akan mengakses dan melakukan pencarian data pada berbagai atribut pada tabel *database* (Nashat & Amer, 2018), meskipun tidak semua atribut dibutuhkan (A. Suganya & Kalaiselv, 2013). Terdapat berbagai cara untuk mengurangi waktu dan biaya akses data serta meningkatkan performa suatu database, salah satunya dengan merancang sebuah Database Terdistribusi (Al-Sayyed et al., 2014). Merancang sebuah Database Terdistribusi sangatlah kompleks (Goli & Rankoochi, 2012). Untuk menangani masalah tersebut dibutuhkan suatu skema yang memudahkan perancangan Database Terdistribusi, dan skema yang dipilih yaitu dengan menggunakan skema fragmentasi atau partisi data (Singh, 2020). Fragmentasi Vertikal adalah bagian dasar dari proses pembangunan Database Terdistribusi (Le et al., 2018). Ketika suatu database mengalami penurunan performa atau ingin dilakukan suatu proses penurunan biaya dan waktu akses (Rahimi et al., 2018), database terdistribusi dapat digunakan untuk menangani masalah tersebut (Rodríguez-Mazahua et al., 2014).

Paper penelitian ini berisi tentang ulasan atau review dengan teknik *systematic literature review* (SLR) mengenai metode-metode fragmentasi vertikal yang dapat digunakan pada berbagai keadaan tertentu (Kitchenham et al., 2009). Review ini berguna bagi peneliti maupun praktisi, untuk para peneliti penelitian ini akan membantu mereka untuk melihat tren dalam bidang fragmentasi vertikal pada database terdistribusi (Bhuyar et al., 2012). Serta sebagai sarana untuk mengidentifikasi, mengevaluasi, dan menafsirkan keseluruhan hasil penelitian yang telah dilakukan relevan dengan

pertanyaan penelitian atau bidang topik dengan tujuan memberikan jawaban atas pertanyaan (Razavian et al., 2019).

2. Metodologi Penelitian

2.1. Objek Penelitian

Objek pada penelitian ini adalah database terdistribusi. Pengambilan topik fragmentasi vertikal pada database terdistribusi sebagai objek penelitian dilandasi dengan alasan sebagai berikut:

1. Adanya penurunan performa pada database seperti membesarnya waktu dan biaya akses pada suatu proses pencarian database.
2. Perkembangan database terdistribusi yang memiliki metode yang beragam.

2.2. Metode Penelitian

Secara umum *systematic literature review* (SLR) ini dilakukan dalam tiga tahap yaitu tahap perencanaan, tahap pelaksanaan dan yang terakhir adalah tahap pelaporan. Tahap pertama dalam SLR ini adalah mengidentifikasi tujuan dan merumuskan pertanyaan penelitian mengenai topik penelitian yang akan diteliti. Tahap selanjutnya adalah mengidentifikasi literature yang berhubungan dengan topik penelitian dan melakukan analisis mengenai pendekatan apa saja yang berkaitan dengan topik penelitian. Tahapan terakhir adalah penyusunan laporan SLR.

2.2.1 Pertanyaan Penelitian

SLR ini memberikan jawaban atas beberapa *Research Question* seperti berikut:

Tabel 1 Pertanyaan Penelitian (RQ)

ID	Pertanyaan Penelitian	Motivasi
RQ1	Metode apa yang pernah diusulkan peneliti sebelumnya untuk melakukan fragmentasi tabel pada suatu sistem database?	Mengidentifikasi metode yang pernah diusulkan peneliti sebelumnya mengenai fragmentasi pada sistem database
RQ2	Metode apa yang sering digunakan untuk melakukan fragmentasi pada suatu sistem database?	Mengidentifikasi metode yang sering digunakan untuk melakukan fragmentasi vertikal pada suatu database.
RQ3	Metode dari penelitian manakah yang memiliki hasil terbaik dalam meningkatkan performa suatu sistem database?	Mengidentifikasi hasil dari metode penelitian yang memiliki hasil yang terbaik dalam meningkatkan performa suatu database.

2.2.2 Strategi Pencarian

Paper akan menjawab pertanyaan yang diberikan pada bab sebelumnya dengan memenuhi studi *systematic literature* dengan cara yang dijelaskan oleh Kitchenham. SLR ini menggunakan proses dengan langkah-langkah berikut:

- Perencanaan sebuah strategi pencarian paper

- Menentukan kriteria yang digunakan untuk menyeleksi paper
- Mencari paper pada electronic database seperti Science Direct, IEEEExplore, ACM Digital Library dan Springer.
- Mengevaluasi paper berdasarkan criteria yang diinginkan.

- Menganalisis bukti empiris pada paper yang kita pilih.

Pencarian sumber literatur difokuskan pada *electronic database*. Ini secara otomatis berarti bahwa hasil pencarian paper hanya yang dipublikasikan dalam jurnal, konferensi dan workshop.

Peneliti melakukan pencarian dengan menggunakan string (“vertical fragmentation” OR “vertical partitioning”) OR (“horizontal fragmentation” OR “horizontal partitioning”) OR (“mixed fragmentation” OR “mixed partitioning”) AND (distributed database OR database OR method OR approach OR algorithm).

Dengan string pencarian dan batasan kriteria yang dijelaskan diatas, *electronic database* memberi peneliti 72 hasil.

2.2.3 Seleksi Studi

Setelah tahap pertama dengan cara membatasi paper di electronic database, dilanjutkan dengan tahap kedua. Tahap kedua ini dilakukan dengan mereview

abstrak yang membahas tentang metode fragmentasi vertical sebagai topic utama. Memilih paper yang mengusulkan metode fragmentasi baru atau paper yang mengusulkan perbaikan atau kemajuan dari metode yang sudah ada. Selain itu dalam penelitian ini juga tidak memilih paper yang tidak berbahasa Inggris.

Setelah tahap ini hanya 16 publikasi yang cocok dengan kriteria dari penelitian ini. Tahap terakhir adalah membaca paper secara lengkap dari keenambelas publikasi tersebut. Keenambelas paper telah dibaca secara mendetail dengan menggunakan criteria sebelumnya pada seluruh bagian dari paper. Pada langkah ini dihasilkan 14 paper yang membahas tentang metode penurunan biaya dan waktu pada fragmentasi vertikal.

3. Hasil dan Pembahasan

3.1. Hasil Kualitas Penilaian

Tabel 2 Hasil kualitas penilaian

No	Penulis	Judul	Tahun	RQ1	RQ2	RQ3	Hasil
1	Mehdi Goli, Seyed Mohammad Taghi Rouhani Rankoohi	A new vertical fragmentation algorithm based on ant collective behavior in distributed database systems	2012	Y	X	X	X
2	P. R. Bhuyar, Avinash Gawande, Amar Deshmukh	Horizontal Fragmentation Technique in Distributed Database	2012	Y	X	Y	✓
3	A. Suganya, R. Kalaiselvi	Efficient Fragmentation and Allocation in Distributed Databases	2013	Y	Y	Y	✓
4	Khurram Jawad, Asifullah Khan	Genetic algorithm and difference expansion based reversible watermarking for relational databases	2013	Y	Y	X	✓
5	Lisbeth Rodr'iguez-Mazahua, Giner Alor-Hern'andez, Ma. Antonieta Abud-Figueroa	Horizontal Partitioning of Multimedia Databases Using Hierarchical Agglomerative Clustering	2014	Y	Y	Y	✓
6	Rizik M. H. Al-Sayyed, Fawaz A. Al Zaghoul, Dima Suleiman, Mariam Itriq, Ismail Hababeh2	A New Approach for Database Fragmentation and Allocation to Improve the Distributed Database Management System Performance	2014	Y	Y	Y	✓
7	Hossein Rahimi, Fereshteh-Azadi Parand, Davoud Riahi	Hierarchical simultaneous vertical fragmentation and allocation using modified Bond Energy Algorithm in distributed databases	2015	Y	Y	Y	✓
8	Dalia Nashat, Ali A. Amer	A Comprehensive Taxonomy of Fragmentation and Allocation Techniques in Distributed Database Design	2018	Y	Y	Y	✓
9	Dac-Nhuong Le, Bijeta Seth, Surjeet Dalal	A Hybrid Approach of Secret Sharing with Fragmentation and Encryption in Cloud Environment for Securing Outsourced Medical Database: A Revolutionary Approach	2018	Y	Y	Y	✓

10	Erba Lutfina	A Performance Comparative of Vertical Fragmentation Table Using Bond Energy and Graph Based Vertical Partitioning Algorithm	2018	Y	Y	Y	✓
11	Abhishesh Dahal, Shashidhar Ram Joshi	A Comparative Analysis on Performance of Minimum Spanning Tree and K-Means Clustering Based Vertical Fragmentation Algorithm	2019	Y	X	X	X
12	Arjan Singh	SBBO Based Replicated Data Allocation Approach for Distributed Database Design	2020	Y	Y	Y	✓
13	Nupa Ram, S.P. Tripathi	Beetle Swarm Optimized Priority Heuristic Management for Distributed Real Time Database	2020	Y	Y	Y	✓
14	Parmeet Kaur	Enhanced Bond Energy Algorithm for Vertical Fragmentation of IoT Data in a Fog Environment	2022	Y	X	Y	✓
15	Mukta Aggarwal, Shalini Bhaskar Bajaj, Vivek Jaglan	Performance Analysis of Degree of Redundancy for Replication in Distributed Database System	2022	Y	Y	Y	✓
16	Mukta Aggarwal, Shalini Bhaskar Bajaj, Vivek Jaglan	Vertical Fragmentation using K-Medoid in Distributed Database	2022	Y	X	Y	✓

Keterangan symbol:

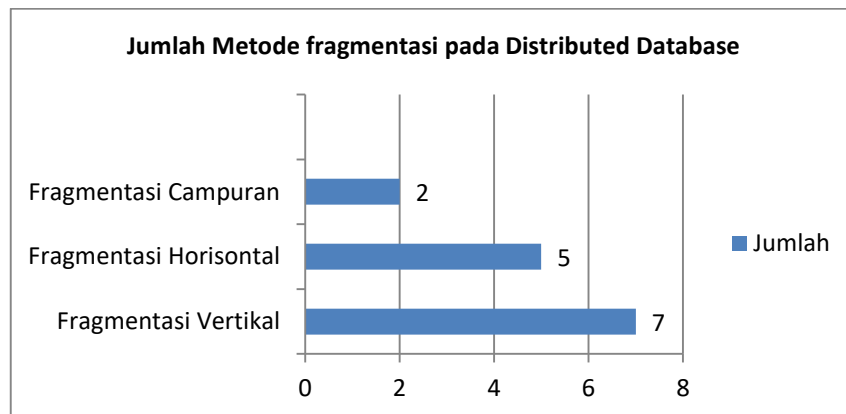
- ✓ : Jurnal tersebut dipilih karena sesuai dengan kualifikasi pertanyaan terkait dengan metode, pendekatan, serta informasi yang telah ditentukan pada *Research Question*.
- X : Jurnal tersebut tidak dipilih karena tidak memenuhi kualifikasi pertanyaan terkait dengan metode, pendekatan, serta informasi yang telah ditentukan pada *Research Question*.

3.2. Pembahasan Hasil

Tahap pembahasan hasil menjawab pertanyaan yang telah ditentukan pada *Research Question*.

RQ1 Metode yang pernah diusulkan peneliti sebelumnya untuk melakukan fragmentasi tabel pada suatu sistem database

Metode penelitian yang sudah pernah dilakukan oleh peneliti terdahulu yang dijadikan referensi terkait topic fragmentasi tabel pada suatu sistem database. Beberapa metode yang pernah diusulkan oleh peneliti terdahulu adalah sebagai berikut:



Gambar 1 Grafik jumlah metode fragmentasi pada Distributed Database

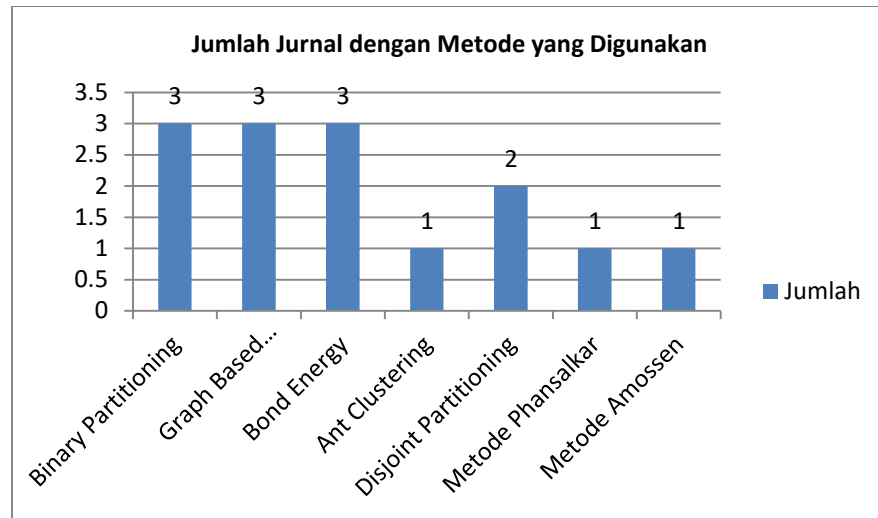
Dari grafik diatas diketahui terdapat tiga metode yang pernah diusulkan oleh peneliti sebelumnya terkait dengan fragmentasi suatu *database*. Metode yang diusulkan adalah metode fragmentasi sebanyak 2 jurnal, metode fragmentasi horizontal sebanyak 5 jurnal, dan metode

fragmentasi vertikal sebanyak 7 jurnal yang menggunakan metode tersebut.

RQ2 Metode yang sering digunakan untuk melakukan fragmentasi pada suatu sistem database

Metode fragmentasi yang sering digunakan untuk melakukan fragmentasi vertikal dijadikan state-of-the-art pada

penelitian SLR ini. Tabel dibawah merupakan rangkuman metode dari Research Qestion 1 dan 2.



Gambar 2 Jumlah jurnal yang menggunakan metode tertentu

Dari grafik diatas diketahui terdapat tiga metode yang paling sering digunakan yaitu metode Binary Partitioning, Graph Based Partitioning, dan Bond Energy Algorithm dengan masing-masing 3 jurnal. Selanjutnya terdapat 2 jurnal yang menggunakan metode Disjoint Partitioning. Metode yang masing-masing satu satu kali dalam jurnal adalah metode Ant Clustering, Metode Phansalkar, dan Metode Amossen.

RQ3 Metode dari penelitian yang memiliki hasil terbaik dalam meningkatkan performa suatu sistem database

Dari hasil review beberapa paper didapatkan hasil yaitu terdapat 3 metode

yang sering digunakan dalam proses fragmentasi vertikal suatu database yaitu metode Bond Energy, Graph-based Vertical Partitioning dan Binary Vertical Partitioning. Dari ketiga metode tersebut dilakukan pembahasan dan review lebih lanjut untuk menentukan metode mana yang menghasilkan hasil yang lebih baik. Dari jawaban pertanyaan penelitian nomor dua mengenai metode manakah yang menghasilkan hasil fragmentasi yang terbaik didapatkan suatu kesimpulan menurut penulis yaitu metode Bond Energy menghasilkan hasil yang lebih baik dibanding metode yang lain terkait dengan penurunan waktu dan biaya akses, meskipun belum optimal.

Tabel 3 Rangkuman Hasil Metode

Dataset	Inputan	Output	Evaluasi	Hasil
Private Dataset	Matriks Afinitas	Fragment	Partition Evaluator	Metode Bond Energy menghasilkan matriks yang meminimalkan biaya fungsi untuk seluruh matriks menggunakan Matriks affinity. Metode ini mengelompokkan atribut berdasarkan sejauh mana atribut tersebut digunakan dengan menghitung afinitas dari tiap-tiap pasangan atribut.
Private Dataset	Matriks Afinitas (graf affinity)	Fragment	Partition Evaluator	Algoritma <i>Graph-Based Vertical Partitioning</i> (GBVP) memiliki kompleksitas komputasi yang lebih sedikit dan menghasilkan fragmen yang semuanya memiliki makna dengan menggunakan metode graf
Private Dataset	Matriks Clustered Affinity	Fragment	Partition Evaluator	Metode Binary Vertical Partitioning menggunakan parameter inputan dalam bentuk matriks penggunaan atribut untuk membentuk matriks afinitas atribut dengan melakukan pengelompokan atau clustering. Setelah clustering, fungsi empiris digunakan untuk

Dataset	Inputan	Output	Evaluasi	Hasil
				melakukan partisi biner iterative. Metode ini memiliki hasil yang subjective tergantung kepada designer untuk memutuskan pengelompokkan atribut untuk membentuk fragmen.

4. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat diambil beberapa kesimpulan sebagai berikut:

1. Metode yang pernah diusulkan oleh peneliti sebelumnya untuk melakukan freagmentasi tabel pada database terdistribusi adalah metode fragmentasi campuran, fragmentasi horizontal, dan fragmentasi vertical. Dengan metode fragmentasi vertikal memiliki frekuensi yang paling banyak digunakan oleh peneliti sebelumnya.
2. Metode yang paling sering digunakan dalam fragmentasi vertikal adalah metode Bond Energy, Graph-based Vertical Partitioning dan Binary Vertical Partitioning.
3. Metode yang memiliki hasil terbaik terkait peningkatan performa penurunan waktu dan biaya akses database adalah metode Bond Energy Algorithm.

5. Daftar Pustaka

- A. Suganya, & Kalaiselv, R. (2013). Efficient Fragmentation and Allocation in Distributed Databases. *International Journal of Engineering Research and Technology (IJERT)*, 2(1).
- Al-Sayyed, R. M. H., Al Zaghoul, F. A., Suleiman, D., Itriq, M., & Hababeh, I. (2014). A New Approach for Database Fragmentation and Allocation to Improve the Distributed Database Management System Performance. *Journal of Software Engineering and Applications*, 07(11), 891–905. <https://doi.org/10.4236/jsea.2014.711080>
- Bhuyar, P. R., Gawande, A. D., & Deshmukh, A. B. (2012). Horizontal Fragmentation Technique in Distributed Database. *International Journal of Scientific and Research Publications*, 2(5), 1–7.
- Goli, M., & Rankoochi, S. M. T. R. (2012). A new vertical fragmentation algorithm based on ant collective behavior in distributed database systems. *Knowledge and Information Systems*, 30(2), 435–455. <https://doi.org/10.1007/s10115-011-0384-6>
- Jawad, K., & Khan, A. (2013). Genetic algorithm and difference expansion based reversible watermarking for relational databases. *Journal of Systems and Software*, 86(11), 2742–2753. <https://doi.org/10.1016/j.jss.2013.06.023>
- Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering - A systematic literature review. *Information and Software Technology*, 51(1), 7–15. <https://doi.org/10.1016/j.infsof.2008.09.009>
- Le, D. N., Seth, B., & Dalal, S. (2018). A hybrid approach of secret sharing with fragmentation and encryption in cloud environment for securing outsourced medical database: A revolutionary approach. *Journal of Cyber Security and Mobility*, 7(4), 379–407. <https://doi.org/10.13052/jcsm2245-1439.742>
- Lutfina, E. (2018). A Performance Comparative of Vertical Fragmentation Table Using Bond Energy and Graph Based Vertical Partitioning Algorithm. *Proceedings - 2018 International Seminar on Application for Technology of Information and Communication: Creative Technology for Human Life, Isemantic 2018*, 377–380. <https://doi.org/10.1109/ISEMANTIC.2018.8549755>
- Nashat, D., & Amer, A. L. I. A. (2018). A Comprehensive Taxonomy of Fragmentation and. *ACM Computing Surveys*, 51(1), 1–25.
- Rahimi, H., Parand, F. A., & Riahi, D. (2018). Hierarchical simultaneous vertical fragmentation and allocation using modified Bond Energy Algorithm in distributed databases. *Applied Computing and Informatics*, 14(2), 127–133. <https://doi.org/10.1016/j.aci.2015.03.001>
- Razavian, M., Paech, B., & Tang, A. (2019). Empirical research for software architecture decision making: An analysis. *Journal of Systems and Software*, 149(2019), 360–381. <https://doi.org/10.1016/j.jss.2018.12.003>
- Rodríguez-Mazahua, L., Alor-Hernández, G., Abud-Figueroa, M. A., & Peláez-Camarena, S. G. (2014). Horizontal partitioning of multimedia databases using hierarchical agglomerative clustering. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8857, 296–309. https://doi.org/10.1007/978-3-319-13650-9_27
- Singh, A. (2020). SBBO Based Replicated Data Allocation Approach for Distributed Database Design. *International Journal of Engineering Research and Technology*, 13(9), 2461–2473. <https://doi.org/10.37624/ijert/13.9.2020.2461-2473>.