



IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

e-ISSN: 2319-8753 | p-ISSN: 2347-6710



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 11, Issue 11, November 2022

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.118



9940 572 462



6381 907 438



ijirset@gmail.com



www.ijirset.com

A NoSQL-Based Architectural Model for Multimedia Data Warehouse of Geographic Data

Dr. Mital Vora

Assistant Professor, Department of Computer Science, T. N. Rao College, Rajkot, Gujarat, India

ABSTRACT: There is an increasing demand of satellite and remote sensing generated data in terms of geographical images. These images are large in size, heterogeneous by nature and continuously changing. To efficiently store, process, access and analyze such a large volume of geographical image data relational data warehouse systems were used. These traditional relational architectures suffer from rigid schemas, limited scalability and poor performance while handling multimedia and spatial data. To overcome these limitations and to achieve better performance, this paper proposes a NoSQL based architectural model for building a multimedia data warehouse specifically for geographic data. To achieve better performance certain techniques are incorporated. Storage efficiency is improved by using already exists compression techniques while query and analytical performance are enhanced using indexing, partitioning, and multilevel feature representation.

KEYWORDS: NoSQL Data Warehouse, Multimedia Data Warehouse, Geographic Data, Satellite Images, Spatial Analytics, Big Data

I. INTRODUCTION

A multimedia data warehouse is a collection of large volume of complex data such as text, images, audio and video data. It is built by integrating large volume of multimedia data from multiple heterogeneous or homogeneous sources. Traditional data warehouses that mainly handle numeric and structured data, multimedia data warehouses must support unstructured and semi-structured data along with advanced analytical capabilities.

Geographic data collected from satellite imagery and remote sensing systems plays an important role in applications such as environmental monitoring, agriculture analysis, urban planning, and disaster management. These images represent real-world objects that continuously change over time and therefore require efficient storage and frequent analysis. Geographic data generally includes raster images, spatial coordinates, temporal attributes, and semantic information such as vegetation or land-use patterns.

Traditional relational multimedia data warehouse architectures face several challenges while managing geographic multimedia data. These traditional relational data warehouse architectures suffer from rigid schemas, limited scalability and poor performance while handling multimedia and spatial data. In contrast, NoSQL databases provide flexible schema design, horizontal scalability, and efficient handling of large-scale data. Hence, adopting a NoSQL-based approach for geographic multimedia data warehousing has become essential for modern geospatial applications [6].

In this paper, a NoSQL-based architectural model for building a multimedia data warehouse for geographic data is proposed. The primary objective is to improve Storage efficiency and enhance query performance.

II. LITERATURE SURVEY

Traditional multimedia data warehousing focused on relational and dimensional models to analyze complex data such as images, audio, and video. Mahboubi et al. proposed an XML-based approach for complex data and highlighted the limitations of traditional relational systems when working with heterogeneous data structures [1]. Their work highlights the need for flexible models capable of handling multimedia data. Rusu et al. proposed a methodology for building XML data warehouses to manage semi-structured data. Their study also pointed out performance challenges when such systems are applied to large datasets [2].

Kaur and Kaur surveyed various lossless and lossy image compression techniques and concluded that lossy compression is more suitable for large multimedia repositories where minor quality loss is acceptable [3]. These techniques are widely used in satellite and geographic image storage systems to reduce storage and input/output overhead.

Research on multimedia data warehouse architectures has increasingly emphasized performance improvement and effective data modeling strategies. Vora et al. presented a modeling architecture for multimedia data warehouses and showed that techniques such as indexing and dimensional modeling can greatly enhance query performance when working with multimedia data [4]. Their findings highlight the importance of well-designed data models in achieving efficient analytical processing.

Boussaid et al. explored methods for integrating complex data into data warehouses using dimensional modeling and pointed out that flexible and extensible architectures are essential for managing non-traditional data types, including images and spatial information [5]. In the context of big data growth, Cattell examined scalable SQL and NoSQL data storage solutions and found that NoSQL systems generally provide superior scalability and performance for handling unstructured and semi-structured data compared to conventional relational databases [6].

Although existing studies address multimedia data warehousing, image processing, and NoSQL data management individually, limited research focuses on a unified NoSQL-based multimedia data warehouse architecture specifically designed for geographic data. This paper addresses this gap.

III. PROPOSED NOSQL ARCHITECTURAL FRAMEWORK

The proposed NoSQL-based multimedia data warehouse architectural model is illustrated by considering geographical data. The proposed NoSQL-based multimedia data warehouse architecture consists of three phases: geographic data extraction and integration, NoSQL-based multimedia data modeling, and geographic data analysis.

3.1 Geographic Data Extraction and Integration

In the phase, geographic images are extracted from the operational sources and remote sensing systems. Feature extraction techniques are applied on the extracted data to extract meaningful information from images. Low-level features such as average Red, Green, and Blue pixel values are calculated, after that HSV are calculated. High-level semantic features such as vegetation area calculated using a program.

After feature extraction, geographic images are compressed using lossy compression techniques to reduce storage requirements. The extracted features and metadata are processed in a staging area before being loaded into the NoSQL data warehouse.

3.2 NoSQL Multimedia Data Model

To design multimedia data warehouse for Geographic data logical and physical dimensional model has been developed. A document-oriented NoSQL data model is used in the proposed architectural framework. Each geographic image is stored as a document containing metadata, extracted low-level features, semantic information, and references to compressed image files. This flexible schema allows easy modification and extension without restructuring the entire database. Indexes are created on frequently used attributes for query such as year, region, and vegetation area to improve query performance.

3.2.1 MULTIMEDIA DATA MODEL USING STAR SCHEMA (LOGICAL DESIGN)

The proposed multimedia data warehouse follows a logical star schema to support analytical processing and OLAP-style queries which includes fact and dimension documents.

Fact Table

- **Fact_Geographic_Image** : Image_Count, Total_Vegetation_Area, Similar_Image_Count, Time_ID, Location_ID, ImageFeature_ID, Semantic_ID,

Dimension Tables

- **Dim_Time**: Date, Month, Quarter, Year
- **Dim_Location**: Region, District, State, Country, Latitude, Longitude
- **Dim_Image_Features**: RGB and HSV values, texture type
- **Dim_Semantic**: Land cover type, vegetation, water body, urban area

- **Dim_Source:** Satellite name, sensor type, resolution

3.2.2. PHYSICAL NoSQL DATA MODEL

In the physical NoSQL implementation, dimension attributes are embedded within the fact document to eliminate join operations and improve query performance. Each document represents a geographic image along with its associated temporal, spatial, feature, and semantic data. This embedded design supports scalability, fault tolerance, and efficient analytical processing.

PERFORMANCE OPTIMIZATION TECHNIQUE The performance of the proposed NoSQL-based multimedia data warehouse is improved using indexing, partitioning, compression, and distributed storage. Indexing enables faster access to frequently queried attributes, while partitioning restricts query execution to relevant subsets of data, mainly based on year. Compression reduces storage and I/O overhead, and distributed storage ensures horizontal scalability and high availability.

3.3 GEOGRAPHIC DATA ANALYSIS

In this geographic data analysis phase, geographic data can be accessed and analyzed. It allows the retrieval of the total number of images for a given region or year and supports statistical analysis such as total, maximum, minimum, and average vegetation area. Temporal trend analysis is achieved using time-based partitions. Common OLAP operations such as slice, dice, and drill-down are used to access data.

IV. EXPERIMENTAL RESULTS

The proposed NoSQL-based multimedia data warehouse architecture was evaluated on a small-scale NoSQL environment under three configurations: without indexing and partitioning, with indexing only, and with both indexing and partitioning. When no optimization techniques were applied, query execution time was high due to full dataset scans. Indexing significantly reduced execution time for all query types. The best performance was achieved when indexing and partitioning were used together, especially for complex analytical queries on large datasets.

Table 1: Query Execution Time (milliseconds)

Query Type	No Optimization	Indexing Only	Indexing + Partitioning
Simple	820	410	260
Medium	1450	720	430
Complex	2780	1320	760

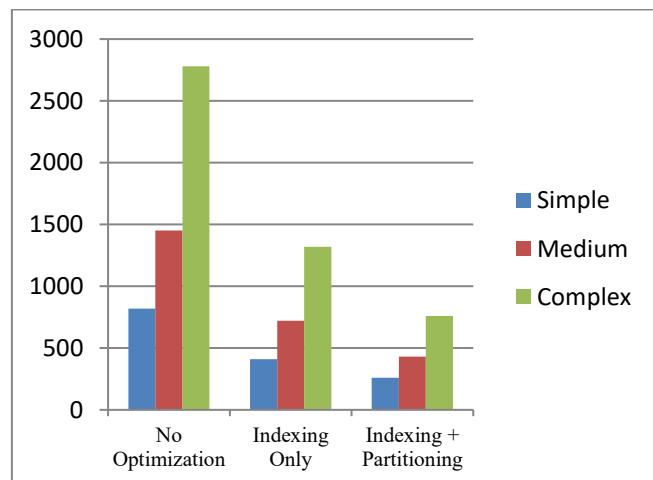


Figure 1: Query Performance without Indexing and Partitioning, with Indexing and with Indexing & Partitioning

V. CONCLUSION

This paper proposed an architectural model that uses the NoSQL database for multimedia data warehousing of geographic information. The proposed methodology overcomes the limitations of traditional data warehouses that use relational databases because of their flexibility, scalability, and high performance compared to traditional methodology. The proposed architecture is very useful for modern geospatial applications that require efficient storage as well as analysis of geographic image information. Experimental results show that indexing technique improves query performance for all the three complexity levels of queries. The query performance further improves by using partitioning approach on indexed data in case of all the three complexity levels of queries. The composite analysis shows that each technique when used with combination of the provided techniques improved results achieved as compared to the use of individual techniques with regards to query performance.

REFERENCES

- [1] H. Mahboubi, J. C. Ralaivao, S. Loudcher, O. Boussaid, F. Bentayeb, and J. Darmont, "X-WACoDa: An XML-Based Approach for Warehousing and Analyzing Complex Data," in *Advances in Data Warehousing and Mining*, IGI Global, 2009.
- [2] L. I. Rusu, W. Rahayu, and D. Taniar, "A Methodology for Building XML Data Warehouses," *International Journal of Data Warehousing and Mining*, vol. 1, no. 2, pp. 23–48, 2005.
- [3] M. Kaur and G. Kaur, "A Survey of Lossless and Lossy Image Compression Techniques," *International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE)*, vol. 3, no. 2, 2013.
- [4] M. Vora, J. Vora, and N. N. Jani, "Modeling Architecture for Multimedia Data Warehouse," *International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)*, vol. 4, no. 1, 2015.
- [5] O. Boussaid, A. Tanasescu, F. Bentayeb, and J. Darmont, *Integration and Dimensional Modeling Approaches for Complex Data Warehousing*, Springer, 2006.
- [6] R. Cattell, "Scalable SQL and NoSQL Data Stores," *ACM SIGMOD Record*, vol. 39, no. 4, 2011.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com



Scan to save the contact details