GEORGE BROWN PRESS
JOURNALS, BOOKS & PROCEEDINGS

Journal of Computer, Signal, and System Research

Article

Exploration of Intelligent Geographic Information System Development in Big Data Era

Junyu Chen [1,*]

[1] Alexa Music, Amazon, Seattle, WA, 98121, USA
* Correspondence: Junyu Chen, Alexa Music, Amazon, Seattle, WA, 98121, USA

Abstract: With the advent of the big data era, the field of geographic information systems (GIS) has ushered in new opportunities and challenges. An intelligent GIS framework should be built to match the characteristics of big data, covering the collection of diverse geographic information, storage of large-scale geographic data, data preprocessing, supply of service ports following common protocols, and development of customized application software for various industries. Research has shown that big data and GIS technology can enhance the processing speed of geographic information and the efficiency of information services, promote the in-depth application of GIS in smart city construction, disaster warning, and resource optimization, and contribute to the sustainable and balanced development of the social economy.

Keywords: big data; intelligent geographic information system; system development; artificial intelligence; data mining

1. Introduction

Geographic information systems, serving as an integrated platform for managing spatial data, play an important role in areas such as urban planning, natural resource management, and environmental monitoring. However, with the surge in data volume and complexity, conventional GIS has begun to show problems of low efficiency and insufficient adaptability in data collection, processing, and utilization. In this context, the development of big data technology has brought new opportunities for the innovation and improvement of GIS. The research and development of intelligent geographic information systems has become a new trend in the industry, combining advanced technologies such as artificial intelligence, cloud technology, the Internet of Things, and spatial databases. This article explores the core technologies in the development of intelligent GIS from the perspective of big data, and constructs a system architecture that can handle diverse data and provide efficient services, in order to contribute to the technological innovation of GIS in the era of big data.

Received: 14 March 2025
Revised: 20 March 2025
Accepted: 13 April 2025
Published: 18 April 2025

Copyright: © 2025 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

2. Overview of Big Data Technology

Big data technology refers to the use of advanced algorithms and computing architectures to extract key information from massive, rapidly growing, and diverse complex datasets. Its characteristics include massive scale of data, diverse types of data, rapid data processing capabilities, and low concentration of data value. In the application of GIS (Geographic Information System), big data technology is mainly reflected in the effective collection, storage, processing, and interpretation of geographic information. These multi-source data include remote sensing images, sensor data, and socio-economic statistics, and their quantity is growing exponentially, making it difficult for traditional GIS technology to meet modern requirements for timeliness and accuracy. The application of big data technology in GIS mainly involves fields such as distributed storage, parallel computing, and machine learning. Distributed computing platforms such as Hadoop and

Spark have improved the storage efficiency and processing speed of geographic information. At the same time, advanced algorithms such as deep learning greatly enhance the accuracy of spatial data analysis. With the dynamic visualization of big data, geographic data can be vividly displayed, providing clear visual assistance for decision-makers [1].

3. Key Technologies in the Development of Intelligent Geographic Information Systems

3.1. Artificial Intelligence Technology

In intelligent geographic information systems, the role of artificial intelligence technology is mainly demonstrated through fields such as data mining, image recognition, and intelligent decision-making assistance [2]. By utilizing deep learning algorithms that mimic the neural network mechanisms of the human brain, it is possible to efficiently process geographic information datasets. Taking the classification of remote sensing images as an example, the commonly adopted convolutional neural network technology has the following calculation formula:

$$f(x) = ReLU(W * X + b) \quad (1)$$

In formula (1), W represents the convolution kernel weight, x is the input data (such as the image pixel matrix), b is the bias term, $*$ is the convolution operation, $ReLU$ is the activation function used to introduce nonlinear characteristics. With the advanced technology of convolutional neural networks, intelligent GIS can achieve precise sorting of massive remote sensing images, automatically identify various surface features, accelerate the collection speed of geographic information, and improve its accuracy. The application of natural language processing technology enables the system to deeply interpret textual geographic information, such as earthquake rescue reports and news communications, further enriching the system's information sources.

3.2. Cloud Computing Technology

Cloud computing provides a solid underlying architecture for building intelligent geographic information systems through distributed computing, storage, and network resources. Its advantages are reflected in the flexible scalability of computing resources and the dynamic allocation of resources according to actual needs, which improves the efficiency of data processing tasks in geographic information systems and overcomes long-standing performance barriers. On cloud computing platforms, the data processing process of intelligent GIS can rely on the MapReduce architecture. This architecture mainly includes two major steps: "mapping" and "reduction". The mathematical model is as follows:

$$Map: (k_1, v_1) \to [(k_2, v_2)] \; Reduce: (k_2, [v_2]) \to [(k_3, v_3)] \quad (2)$$

In formula (2), $k_1$and v_1 are the key value pairs of the input data, while k_2, v_2 and k_3, v_3 are the intermediate key value pairs and output key value pairs, respectively. This architecture refines the processing of massive geographic information data into multiple small units, which are parallelized and then aggregated to enhance the system's computing power. The cloud storage component in cloud computing technology, with the help of distributed file management systems (such as HDFS), provides efficient storage and management for massive amounts of geographic information. Thanks to virtualization technology on cloud computing platforms, intelligent geographic information systems can deploy computing units designed for complex spatial data analysis and provide real-time geographic information support [3].

3.3. Internet of Things Technology

The development of intelligent geographic information systems relies on the support of Internet of Things (IoT) technology, which connects various hardware devices and sensors to complete the task of real-time information collection and transmission. In the field of intelligent GIS, the application of IoT technology focuses on sensor network layout, determination of data transmission methods, and integration of real-time information, ensuring that intelligent GIS can obtain diversified data resources. As the foundation of IoT technology, sensor networks are widely used in various occasions, such as environmental monitoring, traffic monitoring, and disaster warning. For example, in urban air quality monitoring systems, distributed sensors are responsible for collecting data on pollutants such as PM2.5 and carbon dioxide, and sending this data in real-time to an intelligent GIS platform for analysis. The following displays IoT data in urban traffic monitoring scenarios (see Table 1).

Table 1. Collection of Traffic Monitoring Data.

Equipment number	Collection location	Collection time	Traffic flow (vehicles/minute)	Average vehicle speed (km/h)	Transmission status
A 001	Intersection of Area 1	202x-xx-xx 08:30	45	30	normal
A 002	Intersection of Area 2	202x-xx-xx 08:31	60	25	normal
A 003	Intersection of Area 3	202x-xx-xx 08:40	75	20	Data loss

The processing of real-time data requires an intelligent GIS platform to integrate and analyze massive and heterogeneous data. Using cutting-edge edge computing means, some links of data processing can be implemented in the Internet of Things terminals in a decentralized manner, which helps to reduce the burden of the central processor and speed up the response speed of the system. With the deepening application of IoT technology, intelligent GIS systems have evolved from static analytical tools into dynamic platforms capable of providing real-time decision support [4].

3.4. Spatial Database Technology

Intelligent geographic information systems cannot do without the support of spatial database technology. It is a technical system specifically designed for the storage, processing, and analysis of geographic information. Compared to conventional databases, spatial databases can effectively manage various spatial elements such as points, lines, and surfaces in terms of data organization, retrieval methods, and data processing efficiency, and can perform various complex spatial analysis operations. In intelligent geographic information systems, commonly used spatial database software includes PostGIS, Oracle Spatial, and MongoDB, which utilize spatial indexing techniques such as R-trees and quadtrees to quickly search for large amounts of spatial data. Here is a standard expression for spatial query operation:

$$ST\_Within(geometry1, geometry2) \quad (3)$$

Formula (3) can detect whether a geometric entity ($geometry\,1$) is contained within another geometric entity ($geometry\,2$). For example, in the field of disaster warning, the total number of buildings affected by earthquakes can be quickly calculated. In addition, spatial databases also play a key role in processing topological relationships of spatial data and supporting multidimensional spatial data. In the management of urban underground pipeline systems, spatial databases can accommodate three-dimensional data and analyze the topological structure of pipeline intersections to prevent design conflicts. After adding the time dimension, spatial databases can handle dynamic data, including vehicle trajectories and meteorological changes. The integration of spatial databases and big data technology is particularly crucial in the development process of intelligent geographic information systems.

4. Architecture Design of Intelligent Geographic Information Systems in the Era of Big Data

4.1. Multi Source Geographic Data Collection

In the era of big data, the types of geographic information are extremely rich, covering multiple aspects such as remote sensing images, surface observation data, social platform information, and IoT sensor data. As shown in Figure 1, this information can be broadly classified into three categories: spatial information, attribute information, and real-time information. *(*Note: Figure 1 illustrates the typical structure of multivariate geographic data sources.*)*. Remote sensing data, as a key component of diverse geographic information, is mainly obtained through methods such as satellites, drones, and aerial photography. Thanks to high-definition imaging technology, remote sensing data can cover a wide area and display detailed information about the earth's surface. Multispectral and hyperspectral remote sensing technologies provide collected data with richer spectral characteristics, which is crucial for land use classification and ecological environment monitoring. Ground monitoring data is mainly collected by geographic survey instruments, involving data information from multiple fields such as terrain, geology, and environmental monitoring [5].

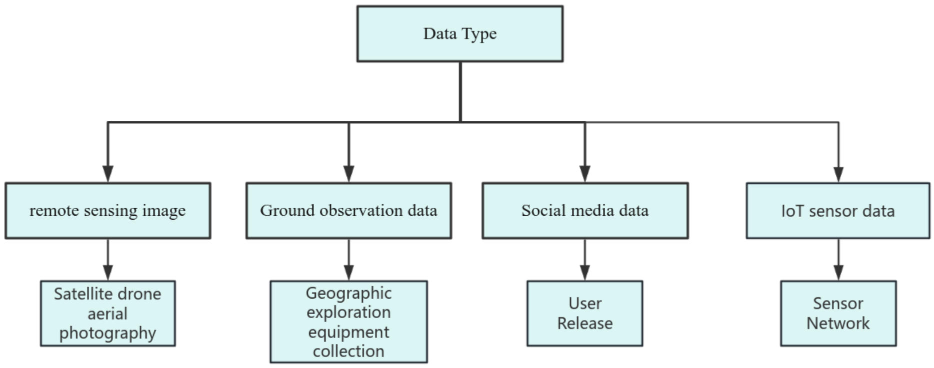

Figure 1. Types of Multivariate Data.

Conventional ground measurement methods mainly rely on total stations and satellite positioning systems, but with the widespread application of laser scanning technology, the speed and accuracy of data acquisition have been greatly improved. The integration of IoT technology has expanded the diversity of geographic information acquisition. Real time monitoring of environmental factors such as air quality, temperature, and humidity through sensor networks enables intelligent geographic information systems to achieve real-time data updates. In the process of building smart cities, street lights, monitoring equipment, and various environmental sensors have become collection points for real-time geographic information. Social networking platforms have accumulated a large amount of unstructured geographic data, including geotagged images, user comments, and various forms of location-related information.

4.2. Geographic Big Data Storage

Geographic information data exhibits distinct spatial correlations and spatiotemporal dynamic characteristics, and its storage requirements involve large capacity data storage, rapid data queries, and multi-dimensional information processing capabilities. The traditional single machine storage method is no longer sufficient to cope with the huge amount of geographic information data. The distributed storage structure enhances storage capacity and system stability by dispersing data across numerous nodes for storage. The widely used Hadoop Distributed File System (HDFS) is a common solution for processing geographic big data, which can effectively manage remote sensing images, sensor data, and vector data. HDFS ensures reliable access and long-term storage of data through specific data block distribution strategies and replica management mechanisms. The rapid retrieval of geographic information relies on advanced spatial indexing techniques, such as R-tree, quadtree, and Hilbert curve, which can improve the search efficiency of spatial data, including quickly locating neighboring points and filtering data points within specific areas [6].

Combining HDFS's distributed indexing technology, such as GeoSpark, enables efficient searching of large-scale geographic data. Due to the high storage space requirements of geographic information data, data compression technology is particularly critical. Compression formats such as GeoTIFF and JPEG 2000 are commonly used for raster data, reducing storage size while maintaining decoding efficiency. The use of block and incremental storage methods for time series data can accelerate the response time of the storage system. By integrating distributed storage structures, spatial indexing mechanisms, and data compression methods, intelligent geographic information systems can meet the requirements for efficient storage and access, forming a solid foundation for subsequent data analysis in the era of big data, laying a solid foundation for subsequent data processing and analysis work.

4.3. Data Preprocessing

In the construction of intelligent geographic information systems, the core purpose of data preprocessing is to optimize data quality and ensure a solid and trustworthy data foundation for subsequent data analysis and application. In the current context of big data, geographic information data from different channels have problems such as inconsistent formats, duplicate information, and incomplete data, which pose higher requirements for data processing. Preprocessing work typically involves steps such as data cleaning, data transformation, data integration, and data dimensionality reduction. During geographic data collection, data loss, noise, and duplicate records are common issues. For missing data, interpolation techniques can be used to fill in. For abnormal data, statistical analysis methods are used to identify and remove them. Regarding the temperature information collected by sensors, abnormal values may originate from equipment failures or network transmission delays, and it is necessary to rely on consistency checks of time and space for discrimination. Due to differences in format, accuracy, and coordinate systems among data from different sources, standardized conversion methods must be adopted to achieve data consistency. Vector data must be adjusted to a common projection coordinate system. For raster data, a resampling process must be performed to maintain consistency in resolution. For data without a fixed format structure, such as remote sensing images, they must be parsed and converted into a universal format (such as GeoTIFF) for system processing [7].

Integrating geographic information from different sources is a key step in data preparation for intelligent geographic information systems. For example, combining remote sensing images with crustal structure and population distribution data can provide a comprehensive data foundation for the construction of smart cities. When integrating these data, spatial registration techniques (such as nearest neighbor matching) must be used to adjust the spatial correspondence of the data to ensure the accuracy of the merged data. Due to the high-dimensional nature of geographic information big data, computational efficiency may be affected. The use of dimensionality reduction methods such as principal component analysis can reduce data redundancy while preserving core information. The quality of data preprocessing plays a decisive role in determining the performance of intelligent GIS and the reliability of its analytical outcomes. A complete data preprocessing process enables intelligent GIS to extract key information from complex and multi terminal datasets, laying a solid foundation for the era of geographic big data analysis.

4.4. Develop Specialized Geographic Information Application Systems for Different Fields

In the information age, the dependence of various industries on geographic information has significantly increased, leading to more personalized and diversified service demands, and the requirements for geographic information services are also becoming more personalized and diversified. Developing geographic information application systems tailored to specific industry needs, intelligent GIS technology provides strong technical support for multiple industries such as urban planning, agricultural supervision, traffic regulation, and ecological environment protection. Intelligent GIS in urban planning integrates spatial data from various sources, providing scientific decision-making basis for the rational use of urban land and the optimization of infrastructure configuration. With the help of the Multi Criteria Decision Analysis (MCDA) algorithm, it is possible to evaluate the effectiveness of various land use schemes. The basic calculation formula of this algorithm is:

$$Si = \sum_{j=1}^{n} w_j * v_{ij} \tag{4}$$

In formula (4), S_i represents the comprehensive score of schemes i, w_j represents the weight of criterion j, and V_{ij} represents the score of schemes i under criterion j. After integrating intelligent GIS technology into the system, it can assist in land resource planning and important facility layout decision-making. In terms of agricultural production, intelligent GIS combined with remote sensing methods tracks crop growth, estimates yield, and scientifically adjusts planting layout. For example, using $NDVI$ (Normalized Difference Vegetation Index) to evaluate the growth status of crops, the calculation formula is:

$$NDVI = \frac{NIR-RED}{NIR+RED} \tag{5}$$

In formula (5), NIR and RED represent the reflectance in the near-infrared and red-light bands, respectively. By utilizing $NDVI$ indicators, the agricultural supervision system integrates climate data and soil information to provide precise cultivation guidance to growers. In ecological environment maintenance, intelligent GIS integrates geographic information with ecological models. This enables effective monitoring of environmental changes and assessment of biological communities. The core of developing geographic information systems that meet the needs of different industries is to closely integrate geographic data parsing with specific industry requirements. With the help of professional model import, algorithm optimization, and data standardization, intelligent GIS tailors' efficient solutions for various complex application scenarios.

5. Conclusion

With the rapid development of information technology, geographic information systems have ushered in a new chapter of intelligence with the help of big data. Intelligent GIS integrates advanced technologies such as intelligent algorithms, cloud computing, network IoT, and spatial data management, demonstrating outstanding performance in processing diverse data and achieving efficient data storage and services. After in-depth research and framework design of the core technology of intelligent GIS, a practical and feasible development strategy has been constructed, which optimizes the processing speed and accuracy of geographic information and provides the possibility for customizing professional application systems for different industries. Looking ahead to the future, the development of intelligent GIS will greatly promote the construction of smart cities, the rational utilization of resources, and the protection of the ecological environment. In the face of the challenges of data privacy and security, as well as the integration of technology and application requirements, further in-depth research is needed. This study provides solid theoretical support and practical reference for the further development and application of intelligent geographic information systems, and it is expected that related technologies can be applied in more fields and demonstrate their potential.

References

1. C. Liang, X. Wang, D. Gu, P. Li, H. Chen, and Z. Xu, "Smart management information systems (Smis): concept, evolution, research hotspots and applications," *Data Intell.*, vol. 5, no. 4, pp. 857–884, 2023, doi: 10.1162/dint_a_00231.
2. P. Emami and A. Marzban, "The synergy of artificial intelligence (AI) and geographic information systems (GIS) for enhanced disaster management: opportunities and challenges," *Disaster Med. Public Health Prep.*, vol. 17, p. e507, 2023, doi: 10.1017/dmp.2023.174.
3. A. Chaiprasert, N. Taweekarn, and J. Chongwatpol, "Applications of a geographic information system–based business intelligence system for decision making at Coffee Refresh," *J. Inf. Technol. Teach. Cases*, vol. 13, no. 1, pp. 16–29, 2023, doi: 10.1177/20438869211040514.
4. V. Lytvyn, O. Lozynska, D. Uhryn, M. Vovk, Y. Ushenko, and Z. Hu, "Information technologies for decision support in industry-specific geographic information systems based on swarm intelligence," *Int. J. Mod. Educ. Comput. Sci.*, vol. 15, no. 2, pp. 62–69, 2023, doi: 10.5815/ijmecs.2023.02.06.
5. B. Wu, X. Lv, A. Alghamdi, H. Abosaq, and M. Alrizq, "Advancement of management information system for discovering fraud in master card based intelligent supervised machine learning and deep learning during SARS-CoV2," *Inf. Process. Manag.*, vol. 60, no. 2, p. 103231, 2023, doi: 10.1016/j.ipm.2022.103231.
6. P. Zhang, C. Lv, Q. Li, B. Cong, and J. Liu, "Research on intelligent platform construction and pavement management of expressway operation and maintenance based on BIM+ GIS technology," *J. Cases Inf. Technol.*, vol. 26, no. 1, pp. 1–19, 2024, doi: 10.4018/JCIT.332879.
7. M. M. Abd-AlHai, S. A. S. Araffa, M. M. Mekkawi, and A. A. M. ElGalladi, "A reconnaissance study for tracing and ordering new mineralisation zones using integrated remote sensing, GIS, and aeromagnetic techniques, west Allaqi-Heiani-Suture, Egypt," *NRIAG J. Astron. Geophys.*, vol. 12, no. 1, pp. 19–44, 2023, doi: 10.1080/20909977.2022.2156068.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of GBP and/or the editor(s). GBP and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.