Advanced Techniques for Spatio-Temporal Data Management in Graph Databases—A Systematic Review

Farah Ilyana Hairuddin, Suhaibah Azri *, and Uznir Ujang

3D GIS Research Lab, Faculty of Built Environment and Surveying, Universiti Teknologi Malaysia, Johor, Malaysia
Email: farahilyanawork@gmail.com (F.I.H.); suhaibah@utm.my (S.A.); mduznir@utm.my (U.U.)
*Corresponding author

***Abstract*—A high-dimensional representation is required to represent connected information that reflects real events and caters to spatio-temporal dimension. Graph data structures have shown potential for integration into smart city data management frameworks and have evolved to handle spatio-temporal data. To investigate the advanced techniques used in managing spatio-temporal data in graph databases, a systematic literature review of related research papers published from 2019 to 2024 was conducted. The review examines the evolution from basic graphs to specialized structures like dynamic attributed graphs and fuzzy spatio-temporal Resource Description Framework (RDF) and also summarizes algorithms used—including graph representation learning, pattern matching, clustering, and centrality algorithms—that enable sophisticated multi-domain analyses. The research provides five key contributions: (1) the state of graph data structure development and algorithms across various fields; (2) insights on spatio-temporal data inputs used in graph structures; (3) algorithms for spatio-temporal data management and analytics; (4) spatio-temporal analyses conducted using graph-structured databases; and (5) future research trajectories. From the review, we identify challenges in graph-based implementation with spatio-temporal data such as structural graph complexity, temporal representation, semantics, and data quality, while outlining future directions in graph representation techniques, temporal-semantic innovations, scalability solutions, and comprehensive data management.**

***Keywords*—graph data structure, graph database, spatio-temporal data, algorithms, spatio-temporal analysis**

I. Introduction

Spatio-temporal data includes spatial, temporal, and attribute components. Recording an object's temporal state and positions over time provides information about real-world events [1]. In this era of advanced technology, spatio-temporal data can be generated from diverse sources, including mobile devices, social media, sensors, satellites, and remote sensing technologies, with daily logs in the form of geotags and timestamps [2]. Additionally, serial ordered timestamped locations can create spatio-temporal trajectories. Spatio-temporal data can be categorised into events, georeferenced time-series, and trajectories. Analysis of this data enables understanding of object distribution and monitoring of attribute changes over time. Time series data produced from these technologies provide research opportunities in environmental data management, forecasting, tracking, predicting movements, and identifying trajectory patterns at different scales.

However, storing various temporal data sources within the same database presents challenges due to the continuity of temporal data. Furthermore, spatio-temporal data can have a high-dimensional structure that encompasses depth in addition to location and time. Another challenge is the inherent fuzziness in spatio-temporal data, which can degrade the accuracy of data retrieval and analysis. Addressing these challenges is crucial to enhance spatio-temporal analysis. This will enable reasoning with spatio-temporal data by extending temporal query capabilities beyond simple date-based retrieval. Examples of such reasoning include identifying congestion causes at specific times and locations, tracing virus spread patterns, and incorporating depth dimensions into analyses.

This highlights graph databases as platforms for managing spatio-temporal data. The efficiency of data analytics has been enhanced by the increased application of graph databases, such as Neo4j, which utilise the Labelled Property Graph (LPG) data model and graph-structured algorithms. These algorithms include graph embedding algorithms, graph neural networks, graph convolutional networks, and graph attention networks [1, 3–5]. Storing network-like data in graph databases leverages the real-world features of the data's interconnectivity to produce valuable insights and perform complex queries that other databases struggle with. In addition to LPG, knowledge graph development heavily utilises the Resource Description Framework (RDF). Graph data structures, either in respect of databases or in the form of algorithms, are gaining recognition in industry and academia for data management applications.

Manuscript received May 5, 2025; revised August 1, 2025; accepted August 11, 2025; published January 8, 2026.

In our interconnected world, comprehensive insights often require connecting information from diverse sources. This is a concept well-aligned with graph databases and graph-structured algorithms, which apply the same network concept found in real-world information flows. Colarusso *et al.* [6] demonstrated this by developing a graph-based smart city platform to manage complex networks with dynamic graphs, viewing smart cities as "systems of systems" with each system consisting of networked entities and relationships. Spatio-temporal evolution information that assists in analysis, prediction, and forecasting is obtained through the ability to query and analyse complex relationships across different transportation modes, ocean depths [7], and maritime trajectory networks at various spatio-temporal scales [8]. Directed, undirected, and weighted graphs, composed of nodes, edges, and edge weights, are widely recognised types of graphs. However, to accommodate real-world complexity, basic graph types have evolved to better represent complex networks. Das and Soylu [9] refer to these as "special graphs," including heterogeneous graphs, multidimensional graphs, signed graphs, dynamic graphs, and hypergraphs. Their special quality lies in their ability to assign multiple cardinalities to nodes, support different relationship types, and accommodate continuously changing nodes and edges.

This systematic literature review has been motivated by the evolution of graph data structures to address information complexity with temporal dimensions. The objective is to explore graph database applications in managing spatio-temporal data and the algorithms used to support data analytics. This research aims to investigate graph data structures and algorithms in managing spatio-temporal data in graph databases over the past six years. This will be achieved by addressing spatio-temporal data management issues, gaining an understanding of how such data is stored in graph data structures, and providing future directions for graph data structure applications that involve spatio-temporal data.

This paper is organised into six sections. Section II reviewed the related works; Section III describes the methodology; Section IV provides the finding that include statistical analysis of selected publications such as number of publications by year, list of journals' title by year, keyword co-occurrence map summaries of the graph data structures and algorithms used in previous research to store and leverage spatio-temporal data. Section V follows and discusses the suitability of graph data structures for managing spatio-temporal data across different environments, introduces the concept of "spatio-temporal graphs", compiles spatio-temporal analyses from previous research, and addresses challenges and future directions. Section VI concludes the paper.

II. Related Works

Relational databases have dominated as data management platforms for decades, particularly for geospatial information at industrial levels due to their data integrity, consistency, security, and spatial data support capabilities. However, alternative database capabilities should be considered to handle aspects such as heterogeneity, connectivity, and high-dimensional data, like spatio-temporal information, as data volume increases and geographical areas necessitate integration of multiple data sources.

A review by Das and Ghosh [10] provided a comprehensive survey of data-driven approaches for spatio-temporal analysis, acknowledging that spatio-temporal data has a highly interconnected nature, where values at one location and time are influenced by neighboring locations and previous time periods, requiring handling of "complex relationships". This perspective is further supported by Breunig *et al.* [11], whose review identified five key milestones in geospatial data management research. Particularly relevant to our research, they highlighted the revival of graph databases as a future direction for supporting big geospatial data analysis. However, they also acknowledged that "one of the major topics within this research will be how to integrate known geospatial-, spatio-temporal- or nD-access methods into the property graph system," indicating a significant research gap in understanding how graph databases can be practically applied to spatio-temporal data management. This gap motivates our systematic review, which investigates the current state of graph database applications specifically for spatio-temporal data across various domains and identifies the techniques and challenges involved in this emerging field.

Several empirical studies on managing spatio-temporal data have provided insights into the advantages of using graph databases for this purpose. Effendi *et al.* [2] conducted an empirical study comparing graph databases (TigerGraph and JanusGraph) with relational databases (PostgreSQL) for managing spatio-temporal data. The results indicated that TigerGraph provided faster response times and notable horizontal scaling advantages that improved querying of large data volumes. These findings demonstrate the efficacy of graph databases to manage spatio-temporal data through efficient query traversal rather than the multi-join queries required by relational databases. Sun and Sarwat [12] further demonstrated that graph databases can be applied to geospatial data by developing GEOEXPAND, a query operator involving spatial predicates (e.g., within, range, and spatial join). This operator benefits applications such as geospatial knowledge base queries, point of interest recommendations, and GeoSocial advertisements. Although this research did not include spatio-temporal aspects, it indicates the potential for extension with temporal dimensions to advance graph-based geospatial analytics.

While recent reviews have begun exploring graph-based approaches for spatio-temporal data, existing reviews remain limited in scope. Del Mondo *et al.* [13] provided a prospective study of spatio-temporal graphs and knowledge graphs for geographical phenomena, specifically using maritime transportation as a case study, focusing primarily on theoretical modeling principles and integration frameworks. Rakhmangulov *et al.* [14] examined spatio-temporal graphs specifically for transportation, revealing significant growth in graph-based deep learning approaches for traffic forecasting.

The growing recognition of graph-based approaches for spatio-temporal data is further evidenced by domain-specific algorithmic reviews. Bui *et al.* [15] provided a comprehensive taxonomy of Spatio-Temporal Graph Neural Networks (ST-GNNs) specifically for traffic forecasting. While their focus was on predictive algorithms rather than data storage and management, their work demonstrates the inherent compatibility between graph-theoretic approaches and spatio-temporal data analysis. Additionally, Ma *et al.* [16] advanced the state of research on spatio-temporal graphs by focusing on computational efficiency in spatio-temporal prediction algorithms. However, these studies primarily focus on algorithm optimization rather than examining the fundamental data management infrastructure.

These reviews adopt domain-specific approaches and have not examined the unique requirements and opportunities present in Geographical Information System (GIS) and spatio-temporal data management contexts across diverse application domains. This limitation necessitates further investigation of spatio-temporal data management techniques in graph databases, comprehensively examining both algorithms and graph database data structures.

To accommodate various complex network needs, graph data structures are constantly evolving. Previous reviews have documented important aspects of graph database evolution and applications, though from perspectives outside the GIS domain. Das and Soylu [9] provided a comprehensive review of "special graphs" from a general data science perspective, cataloguing the structural evolution from basic graphs to heterogeneous, dynamic, and multidimensional variants that better accommodate real-world complexity. While their focus was on general complex network applications rather than geospatial systems, their work established theoretical foundations for understanding how graph structures can assign multiple cardinalities to nodes, support different relationship types, and accommodate continuously changing nodes and edges. These principles are directly applicable to spatio-temporal data management.

Similarly, Xia *et al.* [5] extensively surveyed graph-structured algorithms from an industrial asset maintenance perspective, organizing applications according to maintenance workflows spanning from anomaly detection to decision-making. Although their domain focus was manufacturing systems rather than geospatial applications, they demonstrated graph-structured algorithms' capacity to leverage complex network data and temporal dependencies, providing valuable insights for spatio-temporal analysis approaches. Additionally, Wang *et al.* [17] conducted an investigation on addressing spatio-temporal data management in evolving graph networks, highlighting effective continual learning approaches for traffic prediction using pattern-based frameworks. Their work establishes important foundations for managing evolving spatio-temporal networks through pattern expansion and consolidation mechanisms.

Building upon these domain-specific advances, this systematic literature review extends the investigation to encompass a broader range of spatio-temporal applications and graph database technologies. This systematic examination of existing literature reveals a critical gap in spatio-temporal data management. While individual components exist, such as spatio-temporal data as interconnected networks [10, 11], graph database spatial capabilities [12], temporal handling approaches [2], spatio-temporal analytics techniques [13–16] and structural evolution frameworks [9, 17], no comprehensive synthesis exists of how graph data structures specifically address the multidimensional challenges of spatio-temporal data management across diverse application domains. Current reviews either address spatial OR temporal aspects separately with limited relation to spatio-temporal scenarios.

This systematic literature review addresses this gap by investigating the state of graph data structures and algorithms in managing spatio-temporal data in graph databases over the past six years. This will be achieved by addressing spatio-temporal data management issues, gaining an understanding of how such data is stored in graph structures, and providing future directions for graph data structure applications that involve spatio-temporal data. The contributions of this paper include:

(1) Providing the state of research on graph data structure development and graph-structured algorithms utilised in managing spatio-temporal data across various fields and applications.
(2) Providing insights on the types of spatio-temporal data inputs used in previous research for storage in graph data structures.
(3) Providing a list of algorithms used in previous research for managing spatio-temporal data and performing data analytics.
(4) Providing insight on spatio-temporal analyses conducted using graph-structured databases and algorithms.
(5) Providing the future research trajectories of the graph database and graph-based algorithms with spatio-temporal data.

III. METHOD

A systematic literature review was implemented to synthesise findings based on the four objectives defined above. The review encompasses research conducted from 2019 to 2024. This six-year timeframe was selected because it represents a critical period of maturity for graph database technologies. Three credible databases, namely Scopus, Science Direct, and Web of Science, were used to search for papers using the keywords "Graph" and ("Data Structure" or "Algorithm") and ("spatio-temporal" or "spatial-temporal" or "temporal") and "Graph Database".

This comprehensive keyword approach facilitated acquiring research on graph data structures for databases, graph-structured algorithms, algorithms used in graph databases, and applications of graph databases in spatio-temporal contexts. The three databases yielded 120 articles as a result of the keyword search. Three filtering phases were designed to obtain the most relevant papers.

The first phase resulted in the selection of 40 papers based on the presence of keywords in the title and abstract. In the second phase, articles were ranked using the Ordinatio formula, which considers publication year, citation count, and impact factor. This screening yielded 29 papers, which underwent full-text screening in the third phase, resulting in 23 papers for synthesis. The remaining six papers were excluded due to their lack of relevance to the research scope. Most subject areas involved computer science, engineering, earth and planetary sciences, and mathematics. All selected publications were English-language, full-access journal articles. Table I summarises the filtration and review strategies, and Fig. 1 illustrates the systematic review procedure.

TABLE I. FILTRATION AND REVIEW STRATEGIES

Initial Identification		**Filtration**		
		First phase: Identification and screening by title and abstract	**Second phase: Identify eligibility through Oridnatio method**	**Third Phase: Full-Text Screening & Final Inclusion**
Conducted searching on the databases using defined keywords within the 6 years	Review Strategies	Reviewed title and abstract	Computed ranking using Ordinatio formula considering publication year, citations, and impact factor	Reviewed full articles
	Filtration Details	Included papers with defined keywords in title and abstract, removed duplicates	Excluded papers not satisfying Ordinatio parameters	Identified papers within context of managing spatio-temporal data with graph data structure and algorithm in a graph database
Number of articles	120 identified	40 identified	29 identified	23 identified

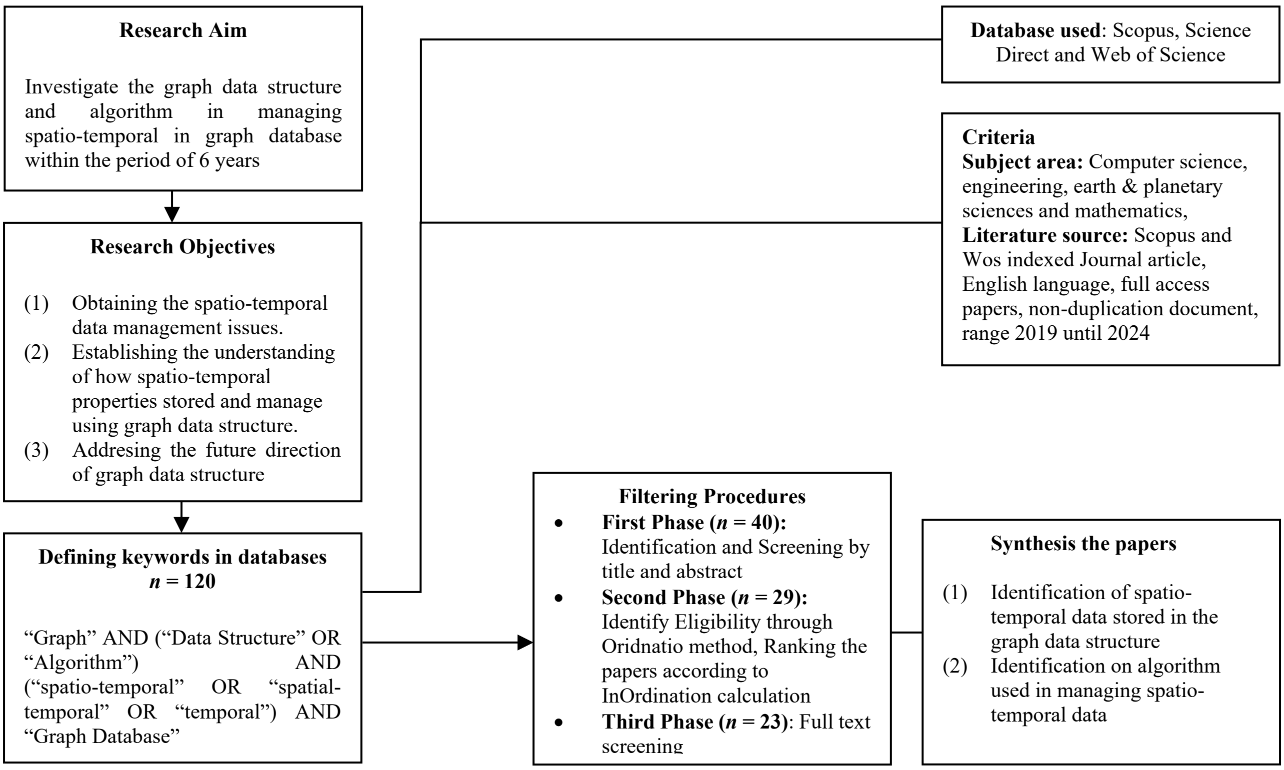

Fig. 1. Systematic review procedure.

IV. FINDINGS

The six-year timeframe is considered to be the crucial emergence of graph database technologies. Neo4j, a graph database, underwent significant advancement during this period. Spatial search trees (R-Trees) were employed as spatial indexing, which is essential for geographic data. This approach is valuable for managing complex geospatial data [18]. The spatial functionality in the Neo4j graph database was available at least by 2018. Research during this period also focussed on improving the performance and scalability of graph databases and reached a significant milestone, where the development of high-performance graph databases could scale to hundreds of thousands of cores [19], making graph databases suitable for a wide range of applications and temporal data.

From an industry perspective, the graph database landscape has evolved significantly. The rapid growth of social networks and other graph data has created a high demand for graph technologies in the market, which led to the emergence of various graph databases, systems, and solution implementations. The global market for graph databases is expected to grow by 21.7% from 2019 to 2027, reaching $4.6 billion by 2027 [20]. Furthermore, this period has seen significant research and development initiatives aimed at advancing graph database technologies. For example, the development of methodologies for knowledge discovery in labelled and heterogenous graphs has demonstrated the ability to effectively extract insights from graph-structured data across various domains. This research was initiated with considerable attention due to the exponential growth of graph-modelled data, resulting in the expansion of the scope of graph database applications [21].

The selected papers were sorted by year to analyse research trends. Fig. 2 shows an alternate rise and fall in publication counts, with 2022 having the highest count. Early 2024 publications suggest a promising start to the year. The initial search yielded 120 articles, but the final inclusion of 23 shows that the intersection of graph databases, spatio-temporal data, and GIS is still a specialised research area in its early stages. Table II illustrates this by showing how the publications are spread across different journals, rather than concentrating them in a specific area. The result suggests that interest is coming from various fields rather than from an established research community.

To understand research trends in managing spatio-temporal data in graph databases, a keyword co-occurrence map was generated using VOSViewer, evaluating 23 papers. Fig. 3 shows five keyword clusters, with “graphic methods” being the most common, followed by “semantics”, “data mining”, “graph database”, and “complex networks”. “Graphic methods” refer to representing and visualising data using graph structures composed of nodes and edges. Semantic data, often managed in graph databases, enable complex queries and reasoning. Network-like or complex network structures are ideal for graph data management. Data mining is a key technique for analysing network data.

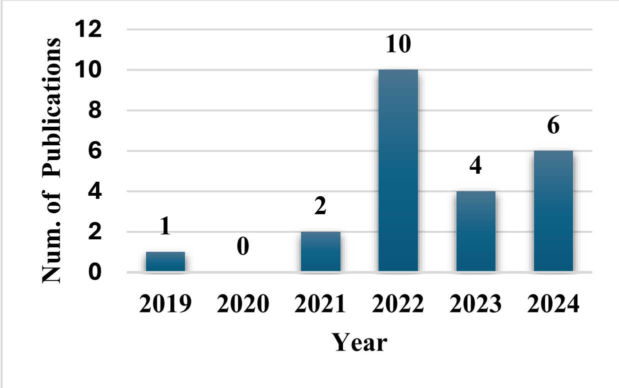

Fig. 2. Number of publications by year.

TABLE II. LIST OF JOURNAL TITLES BY YEAR OF PUBLICATION

Year	Journal Title
2019	Geoinformatica
2021	Computers in Industry Renewable and Sustainable Energy Reviews
2022	Journal of Manufacturing Systems Information Fusion Expert Systems with Applications Building and Environment Transportation Research Part C: Emerging Technologies Future Generation Computer Systems Remote Sensing Data & Knowledge Engineering Computers & Geosciences
2023	International Journal of Digital Earth Chemometrics and Intelligent Laboratory Systems Neurocomputing Applied Soft Computing
2024	SoftwareX World Wide Web Applied Soft Computing Engineering Expert Systems with Applications Automation in Construction Data & Knowledge Engineering Information Systems

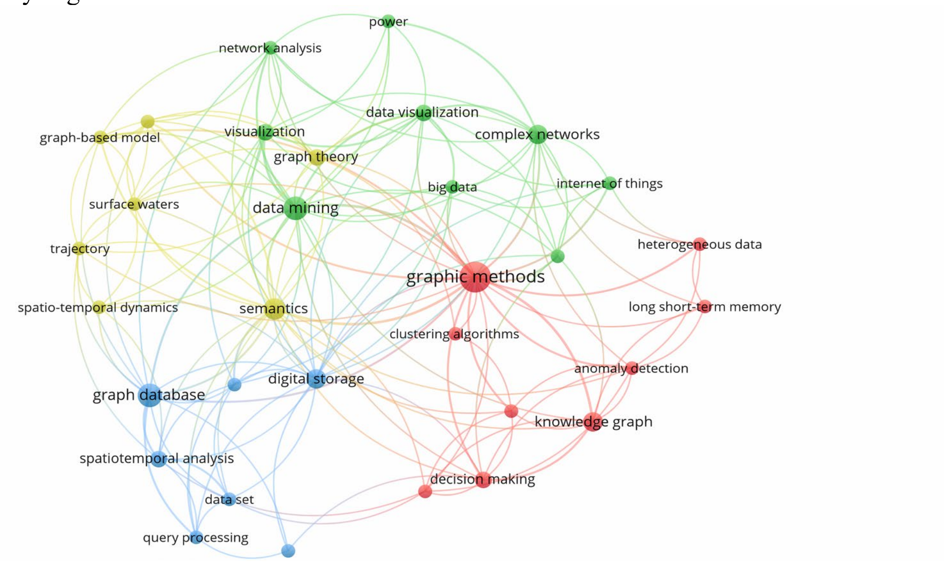

Fig. 3. Keyword co-occurrence map of managing spatio-temporal data using graph data structure and algorithm in graph database.

It highlights the frequent use of "graphic methods" in applications like surface waters network analysis and anomaly detection, which exemplify complex networks. The data types involved include trajectory, semantics, and heterogeneous data. These data are used for digital storage, visualisation, and creating knowledge graphs to facilitate decision-making. Graph databases are an ideal platform for managing such complex data structures, which aligns with the rise of big data and the Internet of Things. Additionally, the research explores utilising stored data for further analysis through data mining and algorithms like clustering and long short-term memory.

Fig. 4 shows the graph database relation map, which demonstrates how spatio-temporal dynamics applies graph databases with trajectory and semantic data inputs for spatio-temporal analysis. We have observed alternating trends in publication frequency with potential growth in 2024, as a result of these findings. The diverse journal titles indicate multidisciplinary interest in this research area. The five top keywords—graphic methods, semantics, data mining, graph database, and complex networks—provide additional insight into areas related to graph applications. The final included papers underwent synthesis to identify spatio-temporal data issues and compile the ways in which graph data structures and algorithms address these issues, as summarised in the following subsections.

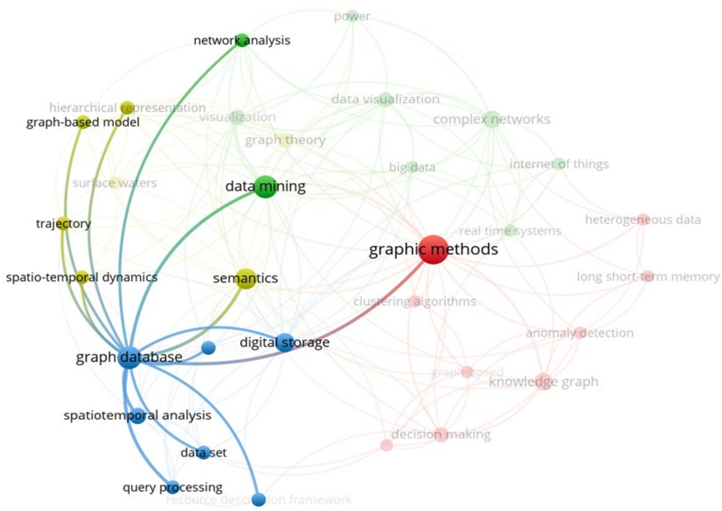

Fig. 4. Graph database keyword co-occurrence map.

A. Spatio-temporal Data Issues

Spatio-temporal data management challenges arise from high data volumes and heterogeneity due to the variety of data sources mentioned earlier. Elayam *et al*. [8] emphasise that spatio-temporal data management complexity stems from incoming data and associated constraints spanning different dimensions, leading to large data volumes. Despite the challenges this high-volume scenario presents it should be viewed as an opportunity for better analysis through data aggregation and fusion, which enable data enrichment. However, achieving data enrichment requires stable, highly scalable data handling and storage infrastructure [2].

Temporal graph traversals face challenges in analysing information diffusion in temporal networks due to the limitation that the temporal networks are only valid at specific times [22]. Since spatio-temporal data is organised in sequential order, this limitation hinders the study of spatio-temporal patterns, such as disease spread and traffic forecasting. Additionally, the highly specific user query selections make spatio-temporal data difficult to query in RDF-structured graph databases. These challenges are compounded when RDF-structured graph databases encounter fuzzy aspects of spatio-temporal data, which are typically the result of inherent vagueness or uncertainty in spatio-temporal applications [23].

These challenges have led to attempts to leverage commonly used Labelled Property Graph data structures and RDF with various modelling techniques to store spatio-temporal data. Several research efforts have extended basic graph data structures by adding new elements to address spatio-temporal aspects. The next section discusses these, including dynamic attributed graphs and fuzzy spatio-temporal RDF data structures.

B. Graph Data Structure in Managing Spatio-temporal Data

Basic graph types include directed, undirected, and weighted graphs. However, graph data structures have adapted in response to the evolution of information environments, from simple unidirectional relationships to complex interconnections between entities, as well as the development of technologies that provide capabilities to gather more information over extended time periods. For example, in oceanographic studies, Li *et al.* [7] stored sea surface temperature datasets, remote sensing images gathered at different temporal intervals and spatial resolutions, and historical climate records using Labelled Property Graph structures in Neo4j, enabling queries of the ocean's evolutionary state. In disease contact tracing, Chen *et al.* [24] utilised information from smart cards, Automatic Vehicle Location systems, shift records, and route schedules in the Resource Description Framework

(RDF) format to identify disease spread sources in public transportation.

When diverse resources are well-extracted and translated into graph data structures, comprehensive information connectivity becomes possible. To accommodate various data sources and network complexity, researchers have adapted components of graph data structure to meet specific application needs. Several studies adopted Neo4j as their graph-based platform [6, 7, 25]. Others adopted RDF data models to establish knowledge graphs, as demonstrated by Zhu *et al.* [23] and Chen *et al.* [24]. While there are also other research that developed custom graph data structures such as to store various data inputs such as temporal reachability graph [3], hierarchical labelled property graph [8], and risk knowledge graph [26].

The research fields using spatio-temporal and temporal-only data include transportation networks, oceanography, epidemiology, construction projects, information networks, power systems, and railway maintenance. This diversity demonstrates the importance of including spatio-temporal data in analyses to enhance understanding of changes through simulations, predictions, and forecasting, as well as to manage complex systems and phenomena over time. Some research using only temporal data was included due to its potential extension to spatio-temporal support. Table III provides details on spatio-temporal and other data inputs and their management in graph-structured approaches. Table IV explains how spatio-temporal data is stored and managed using graph data structures.

TABLE III. SUMMARY OF THE GRAPH DATA STRUCTURE IN VARIOUS FIELD AND APPLICATIONS

Field	Application	Summary	Graph data structure approach
Railway Maintenance	Risk Assessment [26]	Proposed a multi-dimensional network topology graph (knowledge graph) to manage hazard-related entities	Applies directed graph. Edge/node assigned with weight with variable parameter to represent dynamic conditions in static graph structure
Power System	Power system management in extreme events [4]	Proposed time-varying complex network graph approach to represent power system information for simulation and resilience prediction	Applies undirected graph. Nodes represent power system components, edges represent time dimension, and layers represent operational situations within specific time periods
Information Network	systems for storing and analysing temporal graphs [22]	Provides system architecture to resolve temporal graph traversal challenges involving temporal information diffusion analysis	Developed vertex (node) events and edge events where both maintain key-value properties valid at specific times
	Managing fuzzy aspect in spatiotemporal spatio-temporal data [23]	Developed six tuples from initial RDF data structure to improve efficiency and effectiveness of querying fuzzy spatio-temporal RDF	Six tuples consist of directed edges, vertices (with/without temporal or spatial information), information labels (text or temporal), and fuzzy membership degree assigned to vertex and edge structure
	Social Network [27]	Discussed methodologies for detecting events from social networks by considering multiple dimensions including spatio-temporal	Nodes represent users, locations and events with latitude, longitude and timestamp properties. Edges represent temporal relations with start/end time properties
Construction	Construction project management [28]	Proposed location- and time-dependent graph meta-model to manage schedule deviations and cost overruns using knowledge graph	Project location and timestamp defined as node properties
Logistic and Supply Chain Management	Production Logistic [29]	Proposed dynamic spatio-temporal knowledge graph to manage resource allocation in production logistics.	Heterogeneous graph with nodes representing resource locations and edges representing temporal activities and costs. Spatial coordinates and timestamps dynamically updated
Epidemiology	Disease spread [24]	Using LPG data structure to construct knowledge graph for digital epidemic contact tracing on large-scale network	Directed graph with nodes representing passengers or vehicles; edges representing riding acts. Uses composite relationships for passengers with multiple trips
Construction	Construction project management [20]	Proposed location- and time-dependent graph meta-model to manage schedule deviations and cost overruns using knowledge graph	Project location and timestamp defined as node properties
Transportation Network	Traffic congestion identification [23]	Proposed event-process-centred dynamic model for urban traffic congestion using LPG data structure	Nodes represent taxi trajectory states, processes, points of interest. Temporal attributes stored as relationships labelled ‘Next/precedence’ between states and processes to represent sequential order.
	Traffic forecasting [3]	Proposed spatio-temporal reachability graph of road network to forecast dynamic traffic flow	Nodes represent road segments, edges represent reachability between segments within given time periods
	Traffic monitoring [6]	Modelled complex smart city problems into graph data structure	One-layer-directed multiattributed graph with static and dynamic nodes. Static nodes contain static attributes (location/altitude) and dynamic attributes (average travel speed). Dynamic nodes contain ID and timestamp
	Maritime mobility patterns [8]	Used LPG data structure to represent moving objects, trajectory, and activity patterns at various scales	Set of directed graphs with hierarchical properties. Two node types: port (vessel dock location) and Significant Turning Point (trajectory direction change position). Temporal attributes assigned as edges between nodes
Oceanography	Oceanographic research [24]	Proposed ocean current-oriented graph model using LPG data structure	Six node types representing spatial locations at different times/depths and four edge types representing spatio-temporal information trajectories
	Ocean monitoring [7]	Proposed process-oriented graph model using LPG data structure to obtain marine knowledge from time series raster data	Four node types (process, sequence, linked, state) representing marine objects and two edge types (inclusionary and evolutionary) representing relationships between objects

TABLE IV. SUMMARY OF TEMPORAL AND SPATIO-TEMPORAL DATA INPUT AND ITS GRAPH-STRUCTURED DATA MANAGEMENT

Ref.	Data Input	Graph-based Data Management
[26]	Historical accidents / incidents reports. Matter extracted: incident cause, description, consequences, time, location and running speed	Time and location are stored as an entity in the knowledge graph by the Named Entity Recognition (NER) algorithm to identify the location of an entity from the unstructured text information and assign it in correct category. However, the algorithm has shown quite an error percentage when extracting time information. The nodes indicate hazards-related entities and edges denotes linking among entities. The degree of node represents the number of links the node has, and the weighted degree of the node considers the weight of the edge
[4]	Data network: power system, traffic system and communication system	Construct time-varying networks to capture implicit changes in the operational situation, which is helpful to enhance the situational awareness of power systems.
[22]	Temporal Information Network such as bitcoin, college Messege, and e-mail from standford network analysis project	Construct an edge event and vertex event that enable information to be stored in a valid at a specific time. Construct a gamma table that updates frequently whenever there are new nodes added that related to the source (A), uses graph traversal to query the gamma table instead of the whole graph scan when want to query for specific vertix event. Like to see how one information spreads from one person to another
[23]	Fuzzy spatio-temporal data	In dealing with vagueness and uncertainty in spatio-temporal data, the RDF data extend its basis structure by including the fuzzy membership degree to indicate the likelihood of the existence of the vertex in the relation.
[1]	Temporal data consists of the perception of thermal comfort of the occupant, heart rate, and near body temperature. Spatial data consist of a BIM model of the building converted into graph structure	Uses graph embedding to mesh element into a lower-dimensional vector and Graph Neural Networks to extract comfort similarities between different locations in the building
[30]	Data from the event history data of taxi trajectory	The taxi trajectory breaks down into several nodes that represent processes, states, and point of interest (spatial) stored as nodes. Temporal attributes stored as relationships between nodes. It is labelled “Next/precedence” in the relationship (edge) between states and processes to represent sequential order from the location of the taxis to the movement of the taxis.
[3]	Vehicle trajectory, road network	Using vehicle trajectory data to calculate spatio-temporal reachability between road segments. The results are then sorted in a matrix, and the road structure is reconstructed based on the time reachability matrix, which results in spatio-temporal reachability graph. Then, uses the spatio-temporal reachability graph with clustering algorithm to partition road segments into cluster. This clustering approach helps capture regional information and organise road segments into groups, which can improve the accuracy of traffic flow prediction by considering the spatial distribution of road segments within the road network
[6]	Road intersections. Road segments	The nodes represent homogeneous or heterogeneous city entities, and the edges represent the existing relationship between the city entities. The properties of the graph data structure is a multi-modal, multi-layer attributed time-varying network with hierarchical organisation of the nodes. The nodes represent road intersections, edges represent road segments, the attributes of the road intersections consists of a static attribute: unique identifier (osmID16), a latitude-longitude value, and their area name, while the static edges attribute includes linkID, fromNode, toNode, street name, street length, speed limit, estimated free-flow speed, area name, and coordinates; fromNode and toNode are the osmID of the adjacent nodes of an edge, and coordinates is its geometry. The dynamic attribute consists of the link ID, the timestamp of the data aggregation time, and the average travel time. The single layer of the graph composed of multiple subnetworks representing different geographical area
[7]	Data set of sea surface temperature, remote sensing images, and historical climate records. Gathered at different intervals, ranging from daily to annual, with spatial resolutions varying from metres to kilometres and even to global scales	From the time series of the raster formatted dataset, the snapshot objects were extracted, then the process objects were reconstructed from the sequence object, and the evolutionary process was identified from the process objects

This section outlines the process of storing networked data, including spatio-temporal data, using a graph data structure. Most of the applications utilised the Neo4j graph database, which uses a Labelled Property Graph (LPG) data structure to store information. Although Neo4j is the most used platform, the technique of modelling the spatio-temporal data into LPG was applied differently by previous researchers based on their respective purpose. The same applies for research that uses RDF in managing spatio-temporal data. However, few studies have focused on implying the spatio-temporal element into part of the graph data structure; several of them are on dynamic graphs [31], on dynamic attributed graphs [32], and on fuzzy spatio-temporal RDF [23], which will be explained in the next section. The term ‘temporal graph’ will be introduced in Section V, with the inclusion of previous research that closely represents temporal graphs and how it enhances spatio-temporal graph data analysis.

C. Algorithms for Managing Spatio-temporal Data

Graph data structures demonstrate how to store and retrieve data, whereas algorithms decipher the stored data to solve specific problems. This section compiles and categorises algorithms associated with applications using spatio-temporal data inputs, as shown in Table V. Some of the categories that assist graph data analytics in retrieving insights from spatio-temporal data are data representation, graph representation learning, graph pattern matching, data transformation and representation learning, pathfinding, and centrality algorithms.

TABLE V. SUMMARY OF GRAPH-STRUCTURED ALGORITHMS AND ITS APPLICATION IN MANAGING SPATIO-TEMPORAL IN GRAPH DATABASES

Type of algorithm	Algorithm	Explanation	Ref
Data representation	Visibility Graph	Constructs time-series data into complex networks and analyses operational data from topology perspective	[4]
Graph representation learning	Graph Convolutional Network	Type of graph neural network that skilfully transfers deep learning to graph data	[3]
	Graph Attention Networks	Leverages attention mechanisms to focus on specific nodes/edges and captures complex relationships and dependencies allowing incorporation of spatio-temporal information in the analysis and prediction tasks.	[4]
	Graph Neural Network	Performs convolution for signal features, generates high-order representation features, and aggregates processed features for prediction	[5]
		Cluster data in spatial area of influence	[1]
Graph pattern matching	Path-based approximate matching	Leverages spatio-temporal data to find query graph occurrences within data graphs, enabling retrieval of information based on spatio-temporal relationships	[23]
Data Transformation and representation learning	Graph embedding	Reduces computational costs during graph data analytics and represents graphs in lower-dimensional space while preserving original data	[1]
		Detects anomalies in dynamic graphs by converting graph structure to vector space to analyze nodes based on structural similarity	[33]
Clustering algorithm	Spectral clustering	Cluster nodes based on graph topology and spatio-temporal connectivity	[3]
	K-means	Cluster nodes based on feature similarity	
	Agglomerative	Hierarchical clustering method building nested clusters reflecting spatial proximity and traffic similarity	
	Graph-based fuzzy clustering algorithm	Partitions graphs into highly connected subgraphs using node betweenness and edge connectivity, then applies fuzzy clustering to determine information membership values	[34]
Pathfinding algorithm	Depth-First-Search	Traversal algorithm focusing on data structure depth, exploring graph network to furthest path before backtracking	[23]
	Breadth-First-Search	Traversal algorithm exploring neighbouring node paths before moving to next level, commonly used for shortest path identification	
Centrality algorithms	Degree centrality	Determines node importance in networks	[4]
	Betweenness centrality	Identifies nodes connecting different network parts	

1) Data transformation and graph learning representation algorithms

In research about building occupant thermal comfort, Abdelrahman *et al.* [1] constructing information from multiple data types (occupant, environmental, and building spatial data) in graph structure form required data discretisation to convert continuous data to sample points for computational analysis. For example, continuous spatial space was converted into cells, then into nodes and edges for graph data structure storage. Graph embedding was necessary to represent the graph in lower-dimensional space because building properties are pairwise related (buildings consist of floors, and floors consist of rooms). The discretisation of different data types resulted in complex graph structures. Graph neural networks then clustered the lower-dimensional data by identifying similarities and grouping them into spatial areas of influence.

2) Data representation and graphical learning algorithms

Modern power systems, Ma *et al.* [4] demonstrate the effective application of graph data structures for data storage. As power systems involve many interconnected entities, breakdowns due to extreme events can cause power flow disruptions leading to productivity and economic losses. Traditional time series data can be converted into graph structures using Visibility Graph algorithms. This enables data simulation and prediction related to extreme events through graph analytics such as Graph Attention Networks, which capture complex relationships and dependencies by integrating spatial and temporal elements for prediction tasks.

3) Pathfinding algorithms

Pathfinding algorithms identify optimal routes between points by exploring and evaluating paths based on criteria like distance and time. They are common in route network applications that use graph structures, and they are also used in digital contact tracing [24] to reconstruct possible infection routes. This process starts with detected cases and traverses backward to potential infection sources, creating directed relationships that indicate infection flow and result in comprehensive contact tracing knowledge graphs.

4) Graph pattern matching algorithms

Graph pattern matching finds subgraphs in a data graph that match query graph patterns. For queries like identifying traffic congestion-prone areas around 8 AM, where "around" represents fuzziness, Zhu *et al.* [23] applied path-based approximate matching. This technique finds paths in the data graph similar to the requested criteria, leveraging depth-first and breadth-first search traversal algorithms to comprehensively explore possible paths within the main graph.

5) Clustering algorithms

Clustering algorithms group graph networks by criteria, distance, or connectivity. For traffic forecasting research [3], it was necessary to investigate route networks using graph structures with spatio-temporal aspects in order to address complex road structures like multi-level highways, where Euclidean distance calculations are inaccurate. For traffic prediction modelling, road segments were grouped based on spatio-temporal reachability using spectral and k-means algorithms. Spectral clustering grouped nodes by spatio-temporal connectivity, while k-means clustered them by feature similarity to reflect real-time traffic conditions. This combined approach integrated spatio-temporal information to enhance the capabilities of the traffic prediction model.

6) Centrality algorithms

Centrality algorithms identify nodes with the highest importance and influence in networks. In modern power system management during extreme events [4], graph data structures enable the identification of which components would be most impacted, using degree centrality and betweenness centrality to prioritise planning and decision-making for critical areas and components.

V. Result and Discussion

This section discusses graph data structure capabilities in storing spatio-temporal data across various data environments, explores the state of research on the development of spatio-temporal graphs, compiles spatio-temporal analyses conducted in previous research using graph structures and algorithms, and addresses challenges and future research directions. Spatio-temporal graph can be defined as a type of graph that captures the temporal patterns and spatial information of entities' usage events, where nodes represent locations, timestamps, and the entity itself while edges represent relationships between locations, timestamps, and specific entities based on similarities and periodic patterns [35].

A. Managing the Versatility of Data Situations with Graph Data Structure

To manage data heterogeneity, information schemas must be converted to graph data models for storage in graph structures. Algorithms expedite the conversion of diverse data sources, resulting in heterogeneous graphs. Many studies in Table IV exhibit heterogeneous graph properties resulting from various data inputs, with several extending to knowledge graphs.

Graph database structures allow the definition of new node and edge types without the need to restructure the entire database. This flexibility stems from the ability to attach attributes to nodes and edges. Data enrichment is facilitated by establishing relationships between nodes with different attributes and contexts, which leads to insightful knowledge graphs.

Data enrichment with temporal data is well demonstrated in analysing power systems' operational resilience during extreme events [4] by converting multiple operational datasets into time-varying networks. These networks use nodes to represent power system components, edges for time dimensions, and layers for operational situations within specific time periods.

This capability has led to the recognition of graph data structures in the management of complex networks. Translating complex networks into graph structures enables the understanding of the ways in which specific nodes impact others by traversing established relationships. Vulnerability points can be identified and used to create intervention strategies and improved operational designs for applications in power systems, transportation networks, and disease spread modelling.

Heterogeneity in geospatial data presents additional challenges, as the dimensions of data sources may vary depending on the instruments used. Some sources may be three-dimensional (X, Y, Z), while others include additional time dimensions, with Z potentially representing altitude or depth. Spatio-temporal information may contain sequences of timestamped locations. To represent dynamic properties of spatio-temporal trajectories, Elayam *et al.* [8] used spatio-temporal directed graphs with sub-graphs composed of spatio-temporal nodes and edges containing location and timestamp information. Spatio-temporal graphs represent entities as nodes and edges with semantic attributes, with relationships reflecting spatial topology and temporal relations.

B. Innovations and Applications in Moving towards Construction of Spatio-Temporal Graphs

1) State of temporal graphs

A Temporal Graph can be defined as a type of graph structure that captures both spatial and temporal dependencies among its features, commonly used in spatio-temporal data analysis for long-term forecasting [36]. Examples of temporal graphs include dynamic graphs [31] and dynamic attributed graphs. In 2007, complex network concerns centred on topology changes. Dynamic graphs were proposed to address weights associated with vertices or edges that change with time, thereby forming time series for each vertex and edge. This weight variability helps identify trends in dynamic graphs, such as in financial market applications.

Dynamic attributed graphs are defined by Fournier-Viger *et al.* [32] as a type of graph structure that have the capability of storing entities in the nodes with multiple attributes and also representing relationships between entities with their evolution over time. These graphs operate using time-ordered sequences of snapshots, where edges and nodes can be inserted or removed, and attribute values may change at each timestamp.

Researchers have proposed their structure for studying attribute evolution and mining frequent patterns [37]. This research establishes a foundational concept for storing evolving temporal attributes in graph structures, with potential extension to spatio-temporal attributes. Fig. 5 displays the definitions related to dynamic attributed graphs as discussed by He *et al.* [37] and Fournier-Viger *et al.* [32]. This involves merging the structures of the dynamic graph (Definition 2) and attributed graph (Definition 3), both of which extend the basic foundation of the graph data structure, originally started with vertices and edges, as shown in Definition 1.

The capability of dynamic attributed graphs resulted in it being proposed for studying attribute evolution, and its structure is valuable for mining frequent patterns, as demonstrated by He *et al.* [37]. This research establishes a foundational concept for storing evolving temporal attributes in a graph data structure, which can be extended to store spatio-temporal attributes.

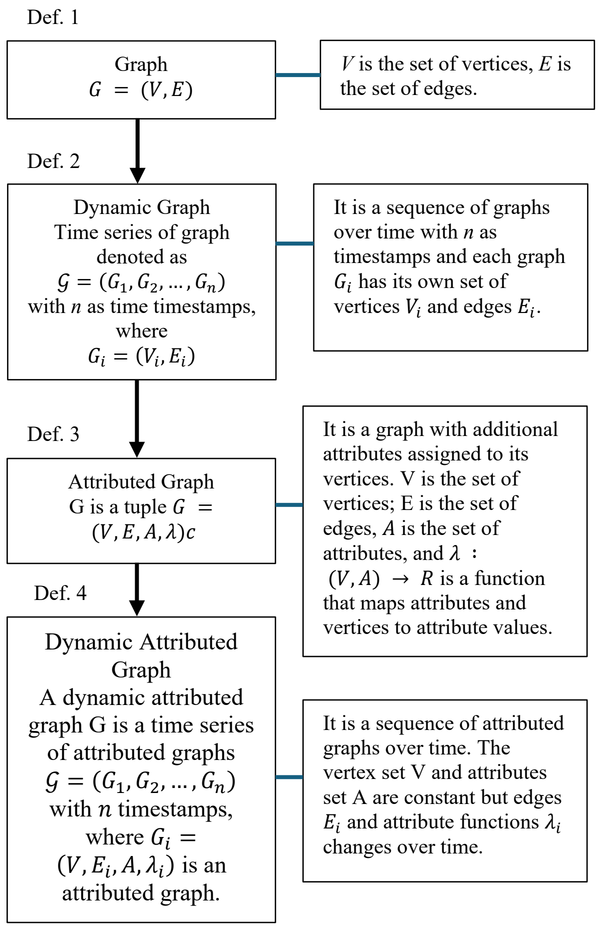

Fig. 5. Constructing the definition of dynamic attributed graph.

2) *Storing dynamic processes in labelled property graph*

Constructing spatio-temporal graph data structures enhances analysis capabilities beyond simple temporal data queries (e.g., specific date or date range queries) by enabling dynamic phenomena analysis to identify the sources and reasons behind events. Dynamic attributed graphs have previously demonstrated the potential to uncover valuable patterns through frequent pattern mining.

In Labelled Property Graphs, He *et al.* [30] proposed a spatio-temporal Event Process-Centred Dynamic Model (EPCDM) to detect and identify causes of traffic congestion. The model considers the taxi's initial location (state), motion (processes), and congestion occurrence (event) structured as a Labelled Property Graph in Neo4j. This research extended the congestion cause analysis by systematically organising information as a dynamic process to imitate real-life traffic congestion scenarios and storing congestion events and taxi motions. It also implemented a hierarchy for traffic events by associating main jam events with subjam events.

Another application involves representing ocean trajectories as hierarchical subtrajectories, sorted by space, time, and depth. The challenge with Argo trajectory data is that the representation is characterised by time-ordered spatial locations that lack a clear distinction between sea surface and parking depth trajectories. Cunjin *et al.* [38] adopted Labelled Property Graph structures to represent Argo trajectories as subtrajectories at different depths, assigning time-ordered relationships between subtrajectories to represent hierarchical organisation. This approach enabled chronological subtrajectory data retrieval and release of trajectories at specific parking depths.

3) *Spatio-temporal reasoning through knowledge graph using resource description framework*

Spatio-temporal reasoning often uses knowledge graphs built with Resource Description Framework (RDF), a standard data interchange model using subject-predicate-object triples. This structure enables complex reasoning by representing entities as interrelated networks [20], making data machine-readable and supporting semantic web and knowledge reasoning applications.

In addition, knowledge graphs have been applied to contact trace networks [24], built from public transportation system data fusion to efficiently organise large-scale transportation data and construct high-resolution contact networks. This enables knowledge reasoning beyond simple data retrieval, focusing on deriving new information. The process of deriving new knowledge from existing data is facilitated by the extraction of edge structures, which helps identify entity relationship patterns, thereby enabling inference development and the creation of new relationships. The knowledge graph is constructed on a trip chaining model, where spatio-temporal correlations between passenger travel trips can be represented through passenger trip chains reflecting continuous travel trajectories.

Knowledge reasoning is commonly associated with RDF data models due to their semantic representation. However, it can also be constructed using property graph structures, as demonstrated by Chen *et al.* [24] who utilised Neo4j to construct public transportation knowledge graphs by considering knowledge graphs as information networks of nodes and edges. Labelled property graphs provide more flexible knowledge graph structures than RDF.

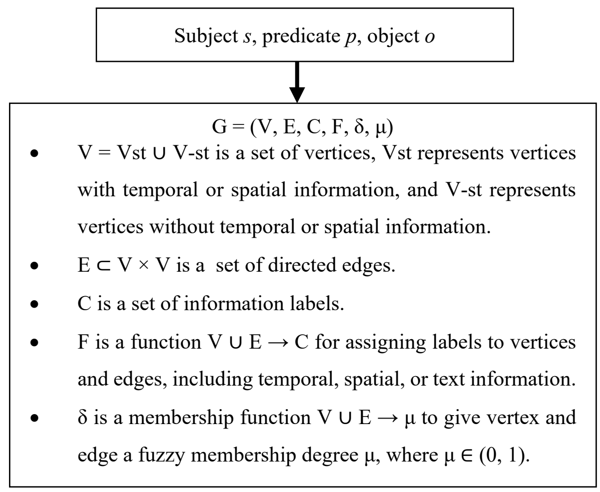

Fig. 6. Construction of the fuzzy spatiotemporal RDF data structure.

RDF faces acknowledged challenges with temporal data, such as maintaining temporal interval consistency, handling multi-temporal data (valid time and transaction time), and integrating contextual information. Wu *et al.* [39] observed that the primary focus of current temporal RDF research is to include valid time, transaction time, and bitemporal information into RDF data structures by employing models like Temporal Knowledge Base (TKB) and multi-temporal RDF. Spatial and spatio-temporal RDF support can be seen in various spatial embedding extensions, such as Multi-Temporal RDF, RDF+, stRDF, stRDFS, and YAGO 1/2, enabling more complex spatial relationship and pattern analysis.

Beyond temporal challenges, the inherent fuzziness in spatio-temporal applications further complicates RDF handling of spatio-temporal data. As shown in Fig. 6, Zhu *et al.* [23] extended the data structure to six tuples and introduced fuzzy spatio-temporal RDF. The fuzzy membership degree μ was uniquely added to indicate the likelihood of a vertex's existence in relationships. This enables spatio-temporal knowledge reasoning queries to include possibility and likelihood temporal queries.

4) Leveraging graph data structure for spatio-temporal analysis

Various insightful spatio-temporal analyses are now possible as a result of the previous discussions of graph data structures for managing versatile data situations and innovations such as dynamic attributed graphs and fuzzy spatio-temporal RDF.

The storage of diverse data sources, including spatio-temporal data, in comprehensive network graphs with algorithms to decipher the stored data enables sophisticated analyses across multiple domains. Table VI summarises key spatio-temporal analyses from the reviewed literature.

TABLE VI. SUMMARY OF SPATIO-TEMPORAL ANALYSIS

Ref	Field	Analysis	Explanation
[4]	Power systems	Forecast wind farm generation to perceive extreme event impacts	Applied graph-based deep learning algorithms (Graph Convolutional Network (GCN), Graph Attention Networks (GAT)) to graph data representing multiple wind turbines with spatio-temporal data.
		Preventive strategies: Develop operational strategies for power systems against extreme events; Reallocate loads to Electric Vehicle (EV) charging stations	Compute generation correlations by multiple wind/ Photovoltaic (PV) stations to obtain temporal and spatial coupling patterns of renewable energy generation. Network data captures coupling characteristics of traffic and power flow; during extreme events, electric flow from traffic can supply EV flow to balance power system demand.
		Restorative Strategies: Gain insight into disruption mechanisms and identify associated failures	Manage spatiotemporal data by modeling coupling mechanisms between extreme events and cascading power system failures.
[22]	Information Network	Generated information insights from temporal data through temporal evolution and information diffusion	Temporal evolution shows network evolution over time; temporal information diffusion shows how information spreads throughout networks over time.
[23]	Spatiotemporal data management	Fuzzy spatiotemporal RDF data	Beneficial for fuzzy spatio-temporal reasoning queries such as likelihood of events at particular locations.
[28]	Construction project	Temporal query of start/end times of building components	Evaluates construction time of specific building elements.
		Query enriched knowledge graph by linking extracted date and location data	Detects modifications on building elements and identifies spatio-temporal working areas and time-sensitive spots to understand construction progress and resource allocation.
[1]	Indoor building environmental quality	Utilisation of spatio-temporal data for predicting similar comfort conditions	Uses graph embedding for lower dimensional vectors and Graph Neural Networks to extract comfort similarities between different building locations.
[24]	Epidemiology	Tracing root causes of virus spreading in transportation networks	Leverages multi-source data with spatiotemporal and temporal elements.
[25]	Transportation network	Identify potential congestion causes through point of interest distribution	Leverages spatio-temporal reasoning in graph databases by querying congestion direction and level.
[3]	Transportation network	Traffic flow forecasting	Optimises prediction by considering temporal reachability to capture regional information
[38]	Oceanography	Obtain complete Argo trajectories including sub-trajectories at different depths	Uses sequence edges to link surface and parking depth trajectory nodes. Enables obtaining sub-trajectories in chronological order, trajectories at specific parking depths, and trajectories within various constraints.
[8]	Maritime transportation network	Historical maritime data analysis	Fine-scale trajectory graph defined as spatio-temporal directed graph of semantic spatio-temporal trajectories. These trajectories form subgraphs with spatio-temporal nodes and edges annotated with semantic information describing evolving properties and context.

C. Challenges and Current Graph Data Structures' Trade-offs

The analysis of the papers identified several challenges. These challenges were categorised into several categories: data representations and modelling, semantic and contextual challenges, and data quality and acquisition challenges. However, these challenges do not imply that they will persist indefinitely; instead, they provide new opportunities for this research and drive it forward towards improvement, which we will discuss further in future directions. Fig. 7 illustrates the challenges and the potential future research.

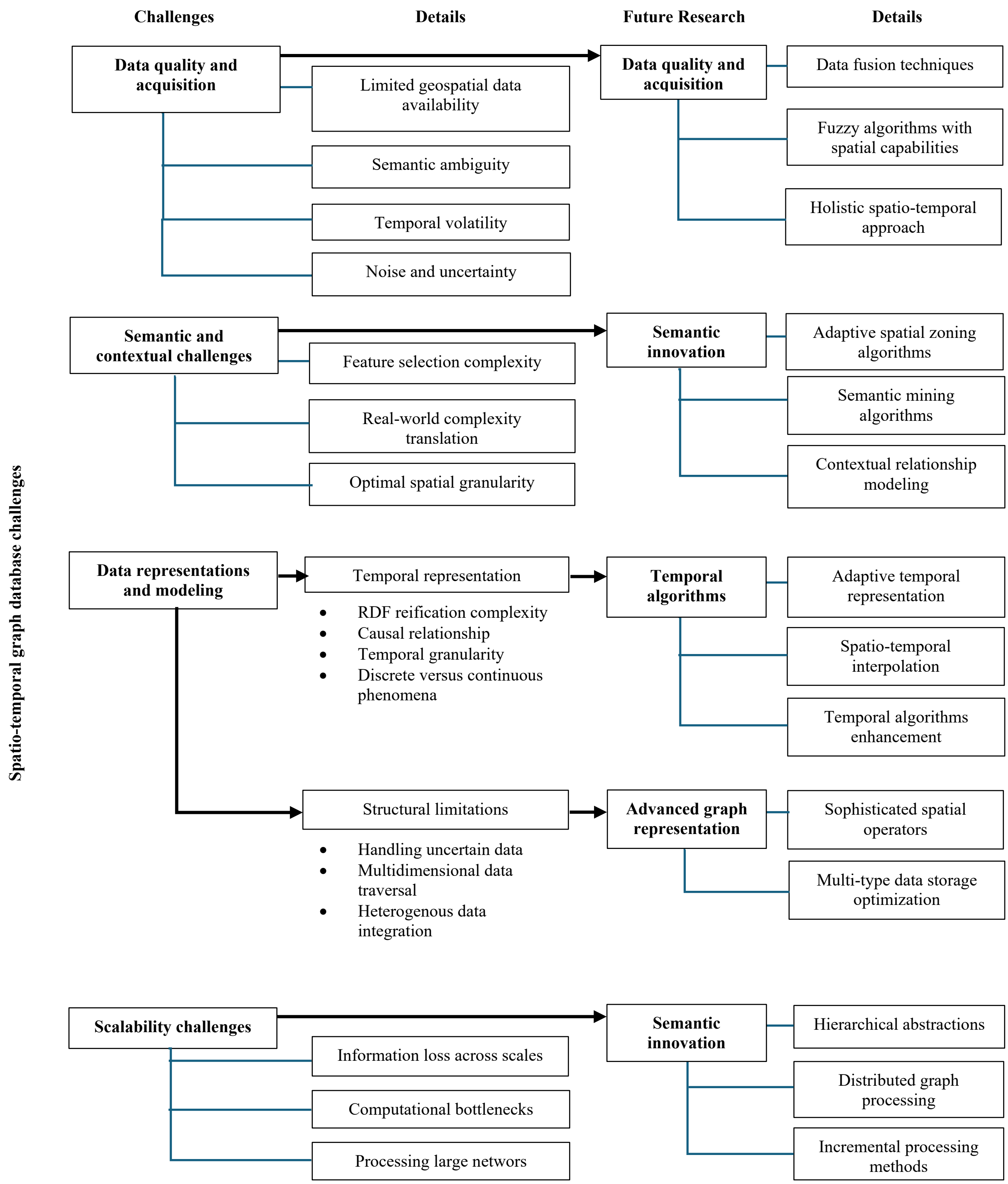

Fig. 7. Summaries of challenges and future research direction of graph database application with spatio-temporal data.

1) Data representation and modelling

This category addresses the complexity of spatio-temporal information that is accurately, efficiently, and meaningfully captured by graph database structures. These challenges emerge from the inherent complexity of representing multidimensional relationships that evolve across space and time. It can further be discussed in the aspect of structural limitations in graph complexity and temporal representation challenges.

2) Structural limitations in graph complexity

Creating graph representations from heterogeneous data sources presents significant challenges [5, 23, 33]. Designing appropriate schemas for heterogeneous spatio-temporal data from multiple sources is complex, particularly when modelling hierarchical structures with temporal aspects. Relationship modelling across time requires domain expertise and consideration of performance implications.

Besides that, the challenges intensify as the data volume increases, which results in the emergence of scalability issues, particularly for temporal pattern matching. This highlights the necessity to enhance the graph database capability; attention should be diverted to the matter of the databases' ability to scale efficiently as temporal data accumulates, which requires focussed strategies for partitioning, archiving, and summarising temporal data that maintain its accessibility during data retrieval. Moreover, another issue that remains a challenge is integrating heterogenous data from different sources in unified graph models due to their varying structures.

However, the technical ability to solve these issues will later bring immense advantages and advance research on graph databases to another significant milestone because of the current interest in integrating data in various domains for the application of knowledge graphs, which requires a stable heterogenous graph. Additionally, there are challenges in handling uncertain and missing data. Zhu *et al.* [23] addressed data integration complexity with a six-tuple model that captured vertices, edges, information labels, labelling functions, membership functions, and fuzzy degrees, though efficient storage remains challenging.

Furthermore, the storage of data with multidimensional structures, such as spatio-temporal, can result in significant performance bottlenecks due to the inherent complexity of multidimensional data traversal. Chen *et al.* [24] and He *et al.* [30] consistently highlighted the computational challenges in processing spatio-temporal relationships. The fundamental complexity of representing real-world relationships in graph computational systems is revealed by the structural challenges in graph databases, which represent more than technical obstacles. Current limitations expose a critical frontier in data science, where the ability to capture multidimensional, heterogeneous information becomes a key research challenge.

3) Temporal representation in graph database

The temporal representation in graph databases exposes the critical point of computational limitation, as the database models are unable to accurately capture the temporal continuity of the phenomena. The existing graph database models struggle to bridge the gap between discrete computational representations and the dynamics of temporal processes. Researchers face an immense challenge in conceptualising and capturing the temporal dimensions of complex systems. In the Labelled Property Graph model, He *et al.* [30] highlighted limitations in representing geographic object change processes; the model relies on discrete state sequences rather than continuous temporal changes. For geographical situations with embedded spatio-temporal and semantic information, graph database modelling increases management complexity, as the challenge in determining appropriate temporal granularity remains unsolved with no established method for adapting temporal representations across different scales.

Additionally, Li *et al.* [7] used a hierarchical graph model to address data representation challenges by modelling oceanic dynamics as process-oriented graph structures rather than discrete observations. Despite the innovative approach, challenges persist in representing spatio-temporal continuity and capturing complex relationships between phenomena. Their research demonstrates the potential of graph databases to model spatio-temporal dynamics, which also highlights the need for further optimisation across diverse conditions.

Next, the algorithm aspect presents additional complexities in managing temporal data. The storage of temporal data in dynamic attributed graphs demands sophisticated algorithmic support. For example, the Credible Attributed Rules (CAR-Miner) for dynamic attributed graphs is capable of identifying patterns in evolving networks; however, it cannot handle time series patterns spanning multiple sequential time points. Its approach is limited to analysing changes between consecutive timestamps, struggling to identify patterns across multiple periods and establish temporal causality chains.

Besides temporal representation in LPG, representing it in RDF graph models presents complexities including reification, which increases triple count and query complexity, the inadequacy of RDF's binary structure to capture higher-order temporal relationships, and the high storage and processing overhead. These challenges have created fragmented temporal RDF extensions with inconsistent semantic interpretations and performance bottlenecks. At its core, the challenge in representing temporal data in a graph database lies in the representation of this continuous phenomenon through discrete nodes and relationships. The discrete observation and storage systems inevitably lead to information loss between temporal observations, which makes establishing causal relationships difficult and creates a mismatch between discrete sampling intervals and continuous process evolution.

4) Semantic and contextual challenges

Semantic representation in graph databases exposes a fundamental challenge of translating complex real-world relationships into computationally meaningful structures. The complicated task of capturing contextual expressions extends beyond simple data connection, demanding sophisticated approaches to understand the underlying meaning and significance of relationships. This results in the graph database facing a critical challenge in extracting meaningful insights from complex datasets where the context itself is as important as the data.

Semantic and contextual challenges emerge in feature selection, where the primary objective is to identify relevant analysis features and extract meaningful contextual information from the complex datasets. Pramanik *et al.* [34] demonstrated this complexity through a graph-based fuzzy clustering algorithm for categorising crime reports. Although their research acknowledged the limitations of cluster labelling and semantic representation, it brings the graph-based algorithm research to a significant intelligent milestone. This highlights the profound difficulties in translating complex information into structured and interpretable formats.

Next, the determination of optimal spatial granularity stands as a critical consideration in graph database design, as it is crucial for correct data retrieval and analysis. Abdelrahman *et al.* [1] explored this challenge through their work on capturing building spatial relationships for occupant thermal comfort using Graph Neural Networks. Their research highlighted the limitations in Neo4j's clustering method, which requires fixed clustering parameters for spatial zoning, underscoring the rigidity of current database computational approaches to contextual representation.

The challenges in semantic and contextual representation within graph databases reveal a fundamental complexity that extends far beyond simple data storage and retrieval. The complex, multidimensional nature of contextual information is difficult to capture using current computational approaches, which results in significant barriers to meaningful data interpretation. The gap between computational representation and real-world complexity demands innovative approaches that can bridge semantic understanding with computational efficiency.

5) Data quality and acquisition challenges

The current research landscape reveals significant barriers to effectively capturing and utilising location-based data, highlighting the gap between the potential of spatial data and the ability of graph databases to comprehensively represent it. Limited geospatial data application within graph databases may result from research focusing on non-spatial relationships, data quality challenges, and manufacturing domain interests prioritising temporal over spatial aspects. Afyouni *et al.* [27] addresses this challenge by noting that only 1–3% of social media posts contain explicit geo-tagging, which reveals the geographical information in current data ecosystems. Besides the data acquisition challenges that are highlighted by the lesser applicability of geographical information in data ecosystems, other challenges such as data quality issues, include significant noise and uncertainty, particularly in social media and IoT-based systems.

Other than that, the heterogenous data sources introduce other complications such as semantic ambiguity, contextual inconsistency, and temporal volatility. Chen *et al.* [24] addressed these challenges through their investigation of spatio-temporal transportation data, revealing limitations in handling noisy spatio-temporal transportation data, including incomplete trip records and lack of uncertainty quantification. The challenges in geospatial data acquisition and quality represent more than technical limitations; they expose fundamental gaps in our approach to understanding complex spatial information. The complex, dynamic nature of location-based data is difficult to capture by current computational systems, creating significant barriers to comprehensive spatial analysis.

6) Comparison of the trade-offs between labelled property graph and resource description framework

From the literature findings, LPG and RDF are popularly used in managing data in graph databases. In general, the flexibility of LPG offers more advantages in organizing complex network data, while RDF is rich with semantic libraries, making it excel in knowledge graph and reasoning queries. However, there are approaches that utilize both of cabilities, such as He *et al.* [25] and He *et al.* [30] The review further synthesizes the capabilities of these graph data structure in terms of spatio-temporal capabilities, examining spatio-temporal representation, dynamic topology handling, query capability, semantic spatial reasoning, and temporal reasoning.

In terms of spatial data representation, LPG demonstrates superior spatial representation capabilities across all dimensions due to its core architectural strengths previously outlined. Specifically, property embedding without reification overhead, flexible hierarchical modeling through native relationship structures, direct coordinate access eliminating complex query patterns, and multi-resolution support through adaptable graph structures enable LPG to excel in spatial data handling. While RDF's spatial limitations directly result from its fundamental architectural constraints, which due to the binary relationship restrictions, it limit the complex spatial representations, reification overhead for multi-attribute spatial properties, and lack of native spatial indexing requiring external extensions.

In terms of temporal data representation reasoning, LPG's operational efficiency advantages enable superior performance in direct timestamp storage, natural sequence modeling, and state transition representation. However, LPG's absence of semantic reasoning frameworks severely limits complex temporal relationships. Conversely, RDF's semantic foundation and standardized ontological frameworks enable sophisticated temporal reasoning through temporal logic support, interval reasoning, and complex temporal constraints. These capabilities directly stem from RDF's formal semantic architecture and access to comprehensive temporal ontologies like Web Ontology Language (OWL)-Time.

Next, in terms of dynamic topology capabilities reflect each paradigm's core architectural characteristics. LPG's flexible schema evolution and native relationship modeling enable real-time topology updates [6, 8, 9], evolutionary change tracking [2, 5, 9], and dynamic relationship creation [2, 6, 9]. RDF's static semantic framework creates inherent challenges for real-time topology modifications [23, 39], with changes requiring complex reification patterns and computational overhead.

In aspect of query capabilities, LPG's direct property access architecture eliminates join overhead, enables efficient graph traversals for spatial proximity operations, provides unified data modeling that avoids impedance mismatch, and supports real-time processing for continuous analytics.

RDF's query limitations stem directly from its semantic processing overhead: quadratic computational complexity for graph operations, reification patterns for temporal operations, complex join requirements for spatio-temporal correlation, and high overhead in temporal extensions [23, 39]. However, RDF's semantic reasoning capabilities enable complex spatial reasoning through standardized vocabularies and semantic validation of query results.

However, semantic reasoning represents RDF's primary domain of superiority, directly enabled by its formal semantic foundation. RDF's comprehensive spatial ontologies, advanced inference engines, standardized vocabularies, and semantic web interoperability all stem from its core semantic architecture. LPG's limitations in this domain are fundamental partly due to absence of built-in ontological frameworks, limited inference capabilities requiring custom implementation, and lack of standardized semantic vocabularies. These limitations directly result from LPG's operational design philosophy that prioritizes performance over semantic sophistication.

The comparison synthesis from existing research reveals that LPG and RDF represent fundamentally different architectural philosophies rather than competing technologies. LPG's operational efficiency philosophy enables superior performance for direct data manipulation, real-time processing, and dynamic topology handling. RDF's semantic sophistication philosophy enables superior reasoning, standardized interoperability, and complex inference capabilities. While hybrid architectures emerge as the optimal solution for comprehensive spatio-temporal applications by strategically combining each paradigm's core strengths.

The literature demonstrates that hybrid implementations using tools like NeoSemantics plugin successfully leverage LPG's operational efficiency for spatial data representation, temporal data handling, dynamic topology management, and query processing, while accessing RDF's semantic reasoning capabilities for complex inference, standardized vocabularies, and sophisticated temporal logic when required. This architectural approach addresses the full spectrum of spatio-temporal data management requirements—from high-performance operational tasks to complex semantic analysis—while minimizing each paradigm's individual limitations. Table VII summarizes the state of the capability level (H = High Capability, M = Medium Capability, and L represent Low Capability) between LPG, RDF and hybrid approach in various range of spatio-temporal aspect.

TABLE VII. COMPARISON OF GRAPH DATA STRUCTURE CAPABILITIES ACROSS SPATIO-TEMPORAL ASPECTS

Spatio-Temporal Aspect		**LPG**	**RDF**	**LPG+ RDF**
Spatial Data Representation	Direct coordinate storage	H	M	H
	Hierarchical spatial organization	H	M	H
	Spatial property embedding	H	L	H
	Multi-resolution spatial indexing	H	L	H
Temporal Data Representation	Timestamp as node properties	H	M	H
	Temporal sequence modeling	H	L	H
	State transition representation	H	M	H
	Complex temporal relationships	L	H	H
Dynamic Topology Handling	Real-Time Topology Updates	H	L	H
	Evolutionary change tracking	H	M	H
	Dynamic relationship creation	H	L	H
	Topological transformation support	M	L	M
Query Capabilities	Spatial proximity queries	H	L	H
	Temporal range queries	H	L	H
	Spatio-temporal joins	H	L	H
	Real-time analytics	H	L	H
Semantic Spatial Reasoning	Spatial ontology support	L	H	H
	Spatial inference capabilities	L	H	H
	Standardized spatial vocabularies	L	H	H
Temporal Reasoning	Temporal logic support	L	H	H
	Temporal interval reasoning	L	H	H
	Complex temporal constraints	L	H	H

D. Future Research Directions in Spatio-Temporal Graph Databases

The challenges explored in previous sections demand a comprehensive, multi-faceted approach to advancing spatio-temporal graph database technologies. The path forward requires innovative strategies that address the fundamental limitations in data representation, computational efficiency, and semantic understanding.

1) Advanced graph representation techniques

Addressing the complex challenges of spatio-temporal data representation requires investigating the capabilities of graph databases, including their data structure, data storage, and indexing. Research opportunities include enabling graph databases to store spatio-temporal data that support multiple data types while maintaining performance. Zhu *et al.* [23] defined five temporal relationship types—meet, overlap, contain, equal, and separate—suggesting the implementation of sophisticated spatial queries using operators like adjacency, intersects, contains, and other proximity-based operators to be used with spatial data that have geometry information.

Next, the development of temporal algorithms in dynamic attributed graphs presents a critical research frontier. Temporal algorithms in dynamic attributed graphs could be enhanced by adapting sequential pattern mining algorithms, and incorporating spatial awareness would allow discovering patterns influenced by geographic proximity for transportation, urban planning, and epidemic monitoring applications. This can extend the graph database capabilities to create intelligent systems that can discover complex patterns influenced by geographic proximity, detect region-specific trends, recognise spatial anomalies, and enable multi-scale analysis.

2) Temporal and semantic representation innovation

Wu *et al.* [39] addressed the ongoing implementation of Bitemporal RDF (BiTRDF) models incorporating valid and transaction times into RDF frameworks without reification. Currently, spatial dimension is not discussed in RDF applications and also BiTRDF. Addressing temporal granularity in graph databases, which is highly crucial for storing spatio-temporal phenomena research, presents new research opportunities for the development of adaptive temporal representation methods with flexible granularity adjustment mechanisms. The challenge lies in creating computational models that can dynamically adjust to the complex, evolving nature of spatiotemporal data.

Furthermore, the research direction further branches out to mining semantic information, such as converting stored data into insights using algorithms like CAR-miner and Named Entity Recognition, which demonstrate a promising direction to mine semantic information. On top of that, addressing spatial granularity challenges offers opportunities for developing data-driven adaptive spatial zoning algorithms that can capture the contextual relationships inherent in complex geographical datasets. For specialised domains that involve spatio-temporal data continuity, like oceanic research, future work should invest in developing spatio-temporal interpolation techniques with robust uncertainty quantification. Adaptive interpolation techniques for graph databases can estimate missing states between observations while quantifying uncertainty assessment.

Besides that, additional future directions include handling dynamic graph structure evolution without performance degradation, implementing time-versioning for graph models, and addressing scalability challenges with increasing attribute numbers.

3) Scalability and processing challenges

Ma *et al.* [4] proposed decomposing large networks through community detection algorithms and leveraging graph computing approaches, such as bulk synchronous parallel computing models for large-scale data processing. This approach presents a promising solution to overcome the computational bottleneck when processing large power networks. It opens up a new research area to integrate this algorithm with graph databases. In addressing the scalability and processing challenges, future research should address ways to improve distributed graph processing through better partitioning algorithms and developing incremental processing methods, as well as avoiding full graph recomputation.

Distributed processing techniques can address spatio-temporal data through targeted partitioning strategies. For temporal dimensions, time-window partitioning can leverage temporal metrics to organize graph structures [3], while spatial partitioning can divide networks geographically to enable parallel processing and optimize data flow for local relevance [6]. Community detection and hierarchical representation approaches can manage spatio-temporal data across multiple abstraction levels [8]. BSP computing models can handle time-varying networks that capture spatio-temporal relations between components [4].

These techniques can be applied to various spatio-temporal scenarios such as traffic flow forecasting where temporal metrics organize road network graphs [3], maritime transportation networks requiring hierarchical mobility data representation [8], city infrastructure management through geographic network partitioning [6], power system resilience analysis using time-varying network models [4], production logistics resource allocation with dynamic spatio-temporal tracking [29], and oceanic dynamics analysis across multiple temporal scales [7].

Other important considerations concerning the scalability and processing challenges include the effect of data detail on query performance. Storing and querying at the most detailed level (micro-scale) leads to prohibitive computational costs, as addressed by Elayam *et al.* [8], which represents the oceanic dynamic representation at various spatio-temporal scales using hierarchical graphs. Representing the hierarchical relationships between different abstraction levels significantly increases the complexity of the graph structure.

Selecting appropriate temporal granularities across abstraction levels is crucial for both performance and meaningful analysis. Consequently, it is necessary to continuously explore methods that integrate spatio-temporal, and semantic dimensions simultaneously and improve the abstraction functions that transform data between hierarchical levels in a graph database, thereby enabling the graph data structures to adapt with different time periods.

Moreover, Li *et al.* [7] addressed a research limitation that involves information loss within spatio-temporal data. They provide insights in developing a spatio-temporal interpolation with uncertainty quantification that is specifically designed for graph databases. This can estimate the missing spatio-temporal information between the dynamic observations and strengthen the capability of graph databases to model evolutionary relationships by developing adaptive graph database structures that can represent the continuous evolution of the spatio-temporal data rather than connecting it as discrete snapshots. This will eventually enable the graph database to accommodate spatio-temporal data across space and various temporal scales.

4) Comprehensive data management strategies

The research trajectory must shift from an isolated temporal or spatial focus to a holistic approach that equally prioritises both dimensions. These findings will also enhance the research area of urban development and spatial-related forecasting. It is crucial to identify spatial queries that leverage the interconnectedness of graph data structures. Subsequently, the research area can progress by determining the best way to store spatial data alongside temporal and create spatio-temporal relationships within the graph data structure. This will create new avenues in discovering, storing, and extracting hierarchical structures in graph databases and graph-based algorithms.

Furthermore, addressing the challenges with distributed graph approaches can optimise spatio-temporal data processing. This draws the attention of future research to exploring the data fusion approach for graph databases. Moreover, spatial proximity must be integrated into analyses involving geospatial aspects, such as epidemiology, to address data coverage issues. Although fuzzy algorithms have been developed in existing research, these approaches require extension through spatial operators to enable reliable data identification in spatial void areas.

Extending fuzzy algorithms with enhanced spatial capabilities offers a comprehensive solution to address data voids. This approach leverages advanced techniques, including probabilistic modelling, spatial interpolation, clustering, multi-resolution adaptability, and temporal-spatial correlations. By doing so, knowledge graphs can effectively represent uncertain spatial relationships, maintain robust tracing capabilities, and clearly communicate confidence levels to decision-makers.

5) Broader technological implementations

Although the graph database is the most suitable platform to cater to the diversity and dynamic nature of the data, it can be too much to handle at some point. This issue prompts another research in developing Multi-Model Databases (MMDB) that are natively capable of storing and accessing data in several models, including relational-based, document-based, and graph-based, to manage multidimensional elements that are compatible with the current big data environment. MMDB was applied by Bimonte *et al.* [40] in a case study on the management of complex phenomena in agroecology, where it enabled the analysis of the spatio-temporal dynamics of diseases, the investigation of field and landscape factors affecting disease propagation, the organisation of observation tasks, and the provision of easily understandable indicators to farmers.

Data mining, which is currently emerging, is a soaring application utilised to assist in data analytics upon completion of the comprehensive database platform. Another area of research that focusses on utilising spatio-temporal data is the effective conduct of data mining when it is stored in a graph data structure. Chen *et al.* [24] proposed that their knowledge graph for epidemic contact tracking could be extended with spatio-temporal correlation analysis by relating it to another knowledge graph about built environments that will help to mine connections between travel sequences in the transportation system with the underlying land use types.

Having an accurate graph data structure is important to achieve causal inference capability within a graph database. The realisation of this capability would require a substantial amount of temporal data, which has resulted in the significant topic of continuously evolving graph structures. A key area of interest in this context is Dynamic Knowledge Graph due to its capability in constructing connections to new entities. This further emphasised the potential direction of research towards solving the evolution of graph structures [5].

The literature synthesis process has revealed that there are several areas that are gaining momentum in the study of spatio-temporal applications with graph data structures that would require continuous attention. These include the integration of graph-based components in multimodal databases to accommodate big data environments, enabling spatio-temporal reasoning within these environments, creating data models that can predict and reason with big data, mining spatio-temporal graphs, and effectively storing spatio-temporal data that encompass multiple events. These areas represent key opportunities to advance the field and address the challenges posed by increasingly complex and large-scale spatio-temporal data. Fig 8. illustrates the potential future research in graph databases.

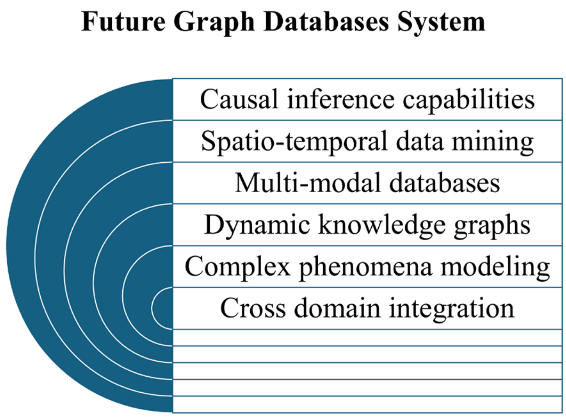

Fig. 8. Potential future graph database application research.

VI. Conclusion

The keyword co-occurrence map reflects the emergence of several key concepts that can be empirically applied to spatio-temporal graph data structures during the synthesis process. First, complex networks originate from situations with time-changing entities interacting through multiple complex time-changing relationships, where both entities and relationships appear, disappear, and change attributes over time. Next, the geographic scene model was extended through the Event-Process-Centred Dynamic Model (EPCDM) for spatio-temporal evolution structure. The geographic scene model organised spatio-temporal dynamics into hierarchical nesting to facilitate management of multiple interacting geographic changes in complex phenomena. These research approaches demonstrate spatio-temporal data's network-like properties and hierarchical representation, making graph databases suitable management platforms.

This research investigated graph data structures and algorithms for managing spatio-temporal data in graph databases over six years, building knowledge about graph database suitability by providing evidence of graph data structure utilisation for storing spatio-temporal data and algorithms for extracting insights and improving management efficiency. The paper provides the state of research on graph data structure development and graph-structured algorithms for managing spatio-temporal data across various applications, lists algorithms used for managing spatio-temporal data and analytics, offers insights on spatio-temporal data input types used for graph data structure storage, and examines spatio-temporal analyses conducted using graph-structured databases and algorithms.

The relevance of this review remains strong in 2025, even though it extends to mid-2024. The fundamental graph-based concepts and algorithms identified in this review form the building blocks for geospatial data management systems. Additionally, the insights and challenges identified provide valuable direction for ongoing research and development in this emerging field.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Farah Ilyana Hairuddin (FIH): Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Writing-original draft, Writing-review & editing. FIH led the systematic literature review process, including developing the paper structure, conducting the review using the Ordinatio method, synthesizing the selected papers, and identifying and constructing the knowledge development framework.

Suhaibah Azri (SA): Methodology supervision, Validation, Writing-review & editing. SA supervised the implementation of the Ordinatio methodology, validated the systematic review process, and provided thorough review and critical feedback on the manuscript.

Uznir Ujang (UU): Methodology supervision, Validation, Writing-review & editing. UU supervised the implementation of the Ordinatio methodology, validated the systematic review process, and provided thorough review and critical feedback on the manuscript.

All authors have approved the final version.

Funding

This work was supported by the Ministry of Malaysia Higher Education through Fundamental Research Grant Scheme (FRGS/1/2022/WAB07/UTM/02/3).

References

[1] M. M. Abdelrahman and C. Miller, "Targeting occupant feedback using digital twins: Adaptive spatial–temporal thermal preference sampling to optimize personal comfort models," *Build. Environ.*, vol. 218, 109090, 2022.

[2] S. B. Effendi, B. v. d. Merwe, and W. T. Balke, "Suitability of graph database technology for the analysis of spatio-temporal data," *Future Internet*, vol. 12, no. 5, 78, 2020.

[3] F. Chen *et al.*, "Temporal metrics based aggregated graph convolution network for traffic forecasting," *Neurocomputing*, vol. 556, 126662, 2023.

[4] X. Ma, H. Zhou, and Z. Li, "On the resilience of modern power systems: A complex network perspective," *Renewable and Sustainable Energy Reviews*, vol. 152, 111646, 2021.

[5] L. Xia *et al.*, "Toward cognitive predictive maintenance: A survey of graph-based approaches," *J. Manuf. Syst.*, vol. 64, pp. 107–120, 2022.

[6] C. Colarusso *et al.*, "PROMENADE: A big data platform for handling city complex networks with dynamic graphs," *Future Generation Computer Systems*, vol. 137, pp. 129–145, 2022.

[7] L. Li *et al.*, "PoSDMS: A mining system for oceanic dynamics with time series of raster-formatted datasets," *Remote Sens.*, vol. 14, no. 13, 2991, 2022.

[8] M. M. Elayam, C. Ray, and C. Claramunt, "A hierarchical graph-based model for mobility data representation and analysis," *Data Knowl. Eng.*, vol. 141, 102054, Sep. 2022.

[9] R. Das and M. Soylu, "A key review on graph data science: The power of graphs in scientific studies," *Chemometrics and Intelligent Laboratory Systems*, vol. 240, 104896, 2023.

[10] M. Das and S. K. Ghosh, "Data-driven approaches for spatio-temporal analysis: A survey of the state-of-the-arts," *J. Comput. Sci. Technol.*, vol. 35, pp. 665–696, May 2020.

[11] M. Breunig *et al.*, "Geospatial data management research: Progress and future directions," *ISPRS Int. J. Geo-Inf.*, vol. 9, no. 2, 95, 2020.

[12] Y. Sun and M. Sarwat, "A spatially-pruned vertex expansion operator in the Neo4j graph database system," *Geoinformatica*, vol. 23, pp. 397–423, 2019.

[13] G. D. Mondo *et al.*, "Leveraging spatio-temporal graphs and knowledge graphs: Perspectives in the field of maritime transportation," *ISPRS Int. J. Geo-Inf.*, vol. 10, no. 8, 541, Aug. 2021.

[14] A. Rakhmangulov, N. Osintsev, and P. Mishkurov, "Spatio-temporal graphs in transportation: Challenges, optimization, and prospects," *Systems*, vol. 13, no. 4, 263, 2025.

[15] K. H. N. Bui, J. Cho, and H. Yi, "Spatial-temporal graph neural network for traffic forecasting: An overview and open research issues," *Applied Intelligence*, vol. 52, pp. 2763–2774, 2021.

[16] J. Ma *et al.*, "BiST: A lightweight and efficient bi-directional model for spatiotemporal prediction PVLDB artifact availability," in *Proc. VLDB Endowment*, 2025, pp. 1663–1676.

[17] B. Wang *et al.*, "Pattern expansion and consolidation on evolving graphs for continual traffic prediction," in *Proc. ACM SIGKDD International Conf. on Knowledge Discovery and Data Mining*, 2023, pp. 2223–2232.

[18] N. Hein and J. Blankenbach. (2021). Evaluation of a NoSQL database for storing big geospatial raster data. [Online]. (1). pp. 76–84. Available: https://publications.rwth-aachen.de/record/826388/files/826388.pdf

[19] M. Besta *et al.*, "The graph database interface: Scaling online transactional and analytical graph workloads to hundreds of thousands of cores," in *Proc. International Conf. for High Performance Computing, Networking, Storage and Analysis*, 2023, 22.

[20] Y. Tian, "The world of graph databases from an industry perspective," *ACM SIGMOD Rec.*, vol. 51, no. 4, pp. 60–67, Jan. 2023.

[21] V. H. Ortega-Guzmán *et al.*, "A methodology for knowledge discovery in labeled and heterogeneous graphs," *Applied Sciences*, vol. 14, no. 2, 838, Jan. 2024.

[22] H. Gaza and J. Byun, "Chronoweb: An open-source platform for analyzing temporal information diffusion on the web," *SoftwareX*, vol. 26, 101738, 2024.

[23] L. Zhu, J. Lu, and L. Bai, "Path-based approximate matching of fuzzy spatiotemporal RDF data," *World Wide Web*, vol. 27, 11, 2024.

[24] T. Chen *et al.*, "A knowledge graph-based method for epidemic contact tracing in public transportation," *Transp. Res. Part C: Emerg. Technol.*, vol. 137, 103587, 2022.

[25] Y. He *et al.*, "Processes and events in the centre: A dynamic data model for representing spatial change," *Int. J. Digit. Earth*, vol. 15, no. 1, pp. 276–295, 2022.

[26] C. Liu and S. Yang, "Using text mining to establish knowledge graph from accident/incident reports in risk assessment," *Expert Syst. Appl.*, vol. 207, 117991, Nov. 2022.

[27] I. Afyouni, Z. A. Aghbari, and R. A. Razack, "Multi-feature, multi-modal, and multi-source social event detection: A comprehensive survey," *Information Fusion*, vol. 79, pp. 279–308, 2022.

[28] F. Pfitzner, A. Braun, and A. Borrmann, "From data to knowledge: Construction process analysis through continuous image capturing, object detection, and knowledge graph creation," *Autom. Constr.*, vol. 164, 105451, 2024.

[29] Z. Zhao *et al.*, "Digital twin-enabled dynamic spatial-temporal knowledge graph for production logistics resource allocation," *Comput. Ind. Eng.*, vol. 171, 108454, 2022.

[30] Y. He *et al.*, "Processes and events in the center: A taxi trajectory-based approach to detecting traffic congestion and analyzing its causes," *Int. J. Digit. Earth*, vol. 16, no. 1, pp. 509–531, 2023.

[31] R. Jin, S. McCallen, and E. Almaas, "Trend motif: A graph mining approach for analysis of dynamic complex networks," in *Proc. Seventh IEEE International Conf. on Data Mining*, 2007, pp. 541–546.

[32] P. Fournier-Viger *et al.*, "Mining significant trend sequences in dynamic attributed graphs," *Knowledge-Based Systems*, vol. 182, 104797, 2019.

[33] B. Lindemann *et al.*, "A survey on anomaly detection for technical systems using LSTM networks," *Comput. Ind.*, vol. 131, 103498, 2021.

[34] A. Pramanik, A. K. Das, and W. Ding, "Graph based fuzzy clustering algorithm for crime report labelling," *Appl. Soft Comput.*, vol. 141, 110261, 2023.

[35] L. Yao *et al.*, "Building entity graphs for the web of things management," in *Managing the Web of Things*, Boston: Morgan Kaufmann, 2017, pp. 275–303.

[36] Z. Chen *et al.*, "Long sequence time-series forecasting with deep learning: A survey," *Information Fusion*, vol. 97, 101819, 2023.

[37] C. He *et al.*, "Mining credible attribute rules in dynamic attributed graphs," *Expert Syst. Appl.*, vol. 246, 123012, 2024.

[38] X. Cunjin *et al.*, "An ocean current-oriented graph-based model for representing Argo trajectories," *Comput. Geosci.*, vol. 166, 105143, 2022.

[39] D. Wu, H. T. Wang, and A. U. Tansel, "A survey for managing temporal data in RDF," *Inf. Syst.*, vol. 122, 102368, 2024.

[40] S. Bimonte, F. A. Coulibaly, and S. Rizzi, "An approach to on-demand extension of multidimensional cubes in multi-model settings: Application to IoT-based agro-ecology," *Data Knowl. Eng.*, vol. 150, 102267, 2024.