



ENABLING SPATIAL AND TEMPORAL CORRELATION OF HETEROGENEOUS INFRASTRUCTURE DATA USING KNOWLEDGE GRAPHS

Ina Heise¹, Sebastian Esser¹ and André Borrmann¹

¹Chair of Computing in Civil and Building Engineering, TUM Georg Nemetschek Institute
School of Engineering and Design, Technical University of Munich, Germany

Abstract

Road infrastructure is a complex system with interconnected subsystems, making the assessment of condition correlation crucial for predictive maintenance. However, the substantial heterogeneity of these subsystems complicates the direct correlation of existing data sets. To address this challenge, we propose a conceptual framework representing data sets and their spatial and temporal relations through graphs by enhancing existing approaches with considerations of temporal dependencies. Hence, large-scale correlation analyses between heterogeneous subsystems and their related data are enabled. Validation is performed using real-world data from German infrastructure management, specifically associating traffic count data with bridge conditions based on their spatial and temporal relations.

Introduction

Road infrastructure is vital to society, impacting both social and economic aspects. Thus, ensuring the functionality of road infrastructure is essential, both now and in the future. Achieving this goal requires comprehensive maintenance efforts, whose necessity and effectiveness depend not only on the condition of individual subsystems within the road infrastructure but on their interactions. Consequently, the quality of maintenance can be significantly enhanced by a comprehensive information base that aids in selecting appropriate measures and facilitating informed decision-making. Investigating relations between data sets is particularly important, enabling simulations of potential future condition developments through extrapolation.

A digital twin, as defined by VanDerHorn and Mahadevan (2021), represents the road infrastructure in its various states and could serve as such an information base for analyzing dependencies among subsystems. The challenge of creating a digital twin that effectively supports this analysis stems not only from the need to integrate diverse and heterogeneous subsystems along with their spatial and temporal relations but also from the varying scales of space and time to be considered.

A comprehensive digital twin contrasts sharply with the current methods of road infrastructure management, which are primarily characterized by siloed data storage, failing to acknowledge the relations between different data sets.

Additionally, the heterogeneity of the data complicates straightforward analysis of correlations between these data sets. Previous works introduced approaches to address this limitation through graph-based methods. These methods involve creating graph representations of legacy data and their contextual linking, capturing their spatial relations. We build on these existing strategies for associating spatially related data sets. However, to effectively examine the dependencies of condition developments, it is essential to consider not only spatial relations but also temporal relations. Hence, we propose a conceptual framework enhancing existing approaches for evaluating spatial relations of heterogeneous road infrastructure subsystems by considering temporal relations.

In the first step, we provide a concise overview of the approaches employed to analyze spatial relations, forming the foundation of our work. Subsequently, the general integration of the analysis of temporal relations with spatial relation analysis is discussed. Following this, outlining the various time references present in road infrastructure management data proves the inherent heterogeneity complicating direct correlation analysis. To address this challenge, we apply Allen's interval algebra by associating data based on their relative temporal relation. Knowledge graphs are utilized to enable the extraction of data sets possessing certain temporal relations on a large scale.

Finally, we apply this proposed framework to German infrastructure data, linking bridge condition and traffic count data based on their spatial and temporal relations.

Literature Review

Related Research

The application of overarching digital twins for managing road infrastructure is regarded as highly beneficial Broo and Schooling (2023). However, their implementation remains in its early stages Taherkhani et al. (2024). Research conducted by Ammar et al. (2022), Heise (2023), and Weise et al. (2018) reveals that current road infrastructure management is characterized by heterogeneous data spread across various silos. As a result, comprehensive analysis often necessitates considerable manual effort, if it is feasible at all.

Beck et al. (2021) and Herle et al. (2020) explore the challenges and overarching strategies associated with inte-

grating heterogeneous data models, particularly within the realms of Building Information Modeling (BIM) and Geographic Information Systems (GIS). One solution is developing a comprehensive data model, as applied by Zhang and Hammad (2006) for their spatio-temporal analysis aiming at detecting clashes in the construction scheduling of infrastructure projects. Likewise, Zhang and Hammad (2005) applies this approach to enhance bridge maintenance. However, both studies primarily concentrate on individual assets, neglecting the scalability that allows their application on a broad road infrastructure system.

In contrast, Buuveibaatar et al. (2022) emphasizes the importance of comprehensively capturing infrastructure by proposing extending the Landinfra standard. However, as noted by Herle et al. (2020), integrating multiple heterogeneous data models into a unified data model can create a complex and unwieldy structure that is difficult to manage and expand. These challenges make graph-based linked model approaches seem more promising.

The application of the linked models approach is exemplified by Beetz et al. (2018), seeking to connect the data models employed in the Netherlands and Germany for managing road infrastructure. Similarly, Vilgertshofer et al. (2017) explore the integration of BIM models and GIS data in the context of tunnel construction. Direct linking is attempted in both instances. It becomes evident that direct linking is often not sufficient due to the diversity of the data models, leading to the suggestion of contextual linking to enable comprehensive analysis.

An approach for contextual linking in road infrastructure has been proposed by Heise and Borrmann (2024). However, the authors concentrate exclusively on spatial relations, neglecting any further exploration of temporal dependencies. In our work, we build upon the approach presented in Heise and Borrmann (2024), enhancing it to consider both temporal and spatial relations.

Background

Knowledge graphs

Applying knowledge graphs has become a widely accepted approach when it comes to integrating and comprehensively evaluating heterogeneous data sets. Hogan (2020) defines knowledge graphs as graphs representing real-world phenomena and containing either deductive or inductive knowledge. In these graphs, nodes represent elements, while edges represent the relations between those elements.

Various graph models exist, including the Resource Description Framework (RDF) and Labeled Property Graph (LPG) model. Compared to RDF, the LPG model offers the advantage of assigning properties and labels to both nodes and edges, providing greater flexibility in modeling complex relations. Hogan (2020)

As noted by Purohit et al. (2020), this more compact structure significantly enhances scalability for graph analysis. Thus, Purohit et al. (2020) propose mapping RDF graphs onto LPGs, aiming to combine the efficiency of LPGs with

the knowledge afforded by the underlying ontologies of RDF graphs.

In our approach, we use LPGs directly to benefit from their better scalability. To ensure a specific structure in the graph that enables the derivation of further knowledge, we adopt an object-oriented approach for graph generation.

Spatial Relations

A methodology for capturing spatial relations using knowledge graphs at the asset level is outlined in Heise and Borrmann (2024). The presented approach employs a reference system to explicitly model the relations of elements in relation to that system, thereby implicitly conveying information about the spatial relations among the respective elements. Linear reference elements serve as reference systems. The evaluation of spatial relations between assets then entails searching for the occurrence of the subgraph representing the implicitly contained spatial relation in question. The LPG graph structure implemented is described in Heise et al. (2024a): Both linear reference elements and assets are represented as nodes within a graph, each node assigned a specific label. Additional nodes define the spatial relation between the asset and the linear reference element, indicating both longitudinal and transversal directions.

Göbels et al. (2023) and Hamdan and Scherer (2020) propose methodologies for modeling spatial relations within assets, especially bridges. Meanwhile, Pan et al. (2024) introduces a graph structure for representing road sections. Additionally, Heise et al. (2024b) integrates the approach for modeling spatial relations at the asset level outlined in Heise and Borrmann (2024) with the spatial structuring approach from Göbels et al. (2023), thereby presenting a method for cross-scale analysis of spatial relations among different elements of road infrastructure.

For the scope of this paper, we built upon these presented approaches. We assume their application for extracting data sets in a spatial relation, which we enhance considering temporal dependencies.

Temporal Reasoning

A comprehensive review of existing approaches for formalizing the concept of time can be found in Ermolayev et al. (2014). Essential features necessary for effective time representation are elaborated, and existing approaches are evaluated against these criteria. Even though none of the presented approaches meet all the requirements defined, the approach presented by Allen (1983), using time intervals and the relative relations among them, appears particularly suitable for our case as it enables temporal reasoning between heterogeneous data and/or data lacking a precise timestamp. With regard to existing research dealing with graph structures for implementing Allen's interval algebra, Grüninger and Li (2017) compare various existing ontologies. However, the authors conclude that none of the existing ontologies entirely captures Allen's interval algebra.

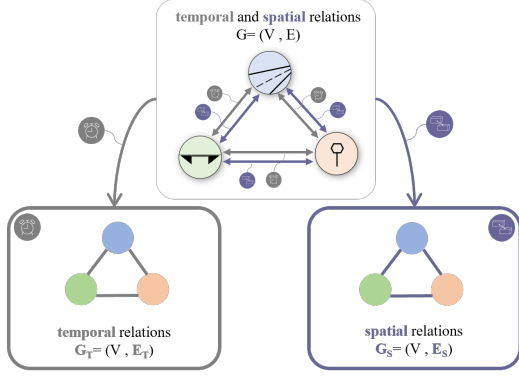


Figure 1: Spatial and temporal relations as independent subsets of relations between the same objects

Combination of Temporal and Spatial Relations

According to Ackoff (1971), objects in relation can be understood as a system, and a system itself can encompass multiple subsystems. Road infrastructure contains multiple elements exhibiting different types of relations. Consequently, road infrastructure can be understood as a system comprising multiple subsystems, each capturing different relation types.

In this paper, we concentrate on the examination of spatial and temporal relations. Spatial relations encompass the functional interdependencies between elements, while temporal relations facilitate an analysis of the interconnections among various states of the elements. Although spatial and temporal dependencies exist concurrently, we consider them independent, as functional relations between assets remain constant over their lifetime, and temporal relations can be assessed without regard to spatial contexts. This distinction allows for representing these different relation types within two separate subsystems, as depicted in Figure 1. The same elements are represented across multiple subsystems, each addressing a distinct type of relation among those elements. This concept aligns with Ackoff’s definition, which asserts that an element can indeed belong to multiple subsystems.

Leveraging knowledge graphs to depict road infrastructure results in viewing data sets, collections of data sets, or subsystems as a set of vertices V interconnected by edges capturing their relations. Given multiple potential relation types between elements, an array of edge types exists accordingly, such as spatial S and temporal T in our case. Since, in principle, each element holds a spatial or temporal relation with every other element, S and T are derived from the Cartesian product of V :

$$T : V \times V \rightarrow \mathbb{N} \quad (\text{temporal relations})$$

$$S : V \times V \rightarrow \mathbb{N} \quad (\text{spatial relations})$$

However, in practice, certainly only specific subsets may be relevant for implementing particular use cases. Using

these relations to define the corresponding graphs:

$$G_T = (V, E_T), \quad \text{where } E_T = \{(u, v) \mid T(u, v) > 0\},$$

$$G_S = (V, E_S), \quad \text{where } E_S = \{(u, v) \mid S(u, v) > 0\}.$$

the relations between T and S can be characterized as follows:

- Independent: The edge sets E_T and E_S satisfy $E_T \cap E_S = \emptyset$
- Parallel: Both graphs G_T and G_S exist simultaneously over the same vertex set V .

These assumptions enable a two-stage process with different approaches for providing and evaluating spatial and temporal relations, each tailored to their respective characteristics.

Regarding temporal relations, each data set can be viewed as an element of a broader super-system providing context through a shared reference system—the time axis. This perspective allows for the uniform treatment of all elements contained, without considering distinctions arising from varying characteristics of multiple reference systems. Such simplification reduces the diversity of potential temporal relation types. However, this reduced number of relation types also results in many data sets exhibiting similar temporal relations with one another. As a result, searching for elements with specific temporal relations within the comprehensive representation of the road infrastructure graph is likely to yield a substantial number of results. This characteristic distinguishes temporal relations from spatial relations, which are tailored to specific use cases and greatly influenced by the characteristics and scales of the related elements. The diversity inherent in spatial relations requires a more complex approach, incorporating multiple reference systems and a more intricate graph structure. However, this complexity also suggests that filtering based on these spatial relations will considerably reduce the number of elements under consideration.

This is particularly relevant when attempting to correlate data sets such as bridge condition recordings and traffic counts on a large scale. The spatial relation between these data sets indicates which traffic flows over which bridge. The temporal relation provides insight into how specific traffic values may have influenced the development of bridge conditions. The data set is likely to contain a substantial amount of traffic data records that can be located in time before a documented bridge condition. In contrast, the analysis of spatial relations will assign only one traffic counting point to each bridge. Therefore, it is most efficient to first evaluate spatial relations and subsequently examine temporal relations of those elements.

Accordingly, we implement a two-step process in which we first analyze spatial dependencies between datasets. The resulting matches form the basis for the investigation of temporal dependencies in the second step. As mentioned above, we assume the application of already existing approaches outlined in the background section when it

comes to representing and extracting information on spatial relations within a graph, and focus on the representation and analysis of temporal dependencies in the following.

Temporal Relations in Road Infrastructure Data

The motivation for considering temporal relations lies in enabling correlation analysis between various condition developments. Such a correlation analysis assesses the dependence of one value (for instance, a bridge condition grade) on another (for instance, traffic volume determined by traffic counts). A significant amount of research, particularly in the medical field, focuses on examining the correlations between different data sets. Statistical methods, especially various types of regression analysis, are typically employed for this purpose. For instance, in Janse et al. (2021), the authors highlight the calculation of the correlation coefficient ρ , which refers to Pearson’s correlation Pearson and Henrici (1896), as a straightforward method for assessing the correlation between two data sets.

However, applying statistical methods to investigate correlations between condition descriptions poses challenges due to the heterogeneity of the data involved. This goes in line with Beck et al. (2021), who points out differences in timeliness as one of the crucial challenges of linking heterogeneous data models. In the context of road infrastructure maintenance data, these differences in timeliness stem from the diverse methods employed for recording and documenting conditions:

The use of sensors providing high-frequency state information is widespread, particularly in research related to digital twins in infrastructure management. These sensors generate individual measurement values at specific moments in time, producing time series data characterized by short and regular measurement intervals. This contrasts with the traditional methods primarily employed for recording asset conditions, which are marked by significantly longer intervals due to the manual nature of inspections. Furthermore, the time gaps between recorded condition data points are generally similar but rarely identical. Moreover, within conventional condition assessment methods, varying intervals are applied for different asset types, and inspections of spatially related assets often occur independently. For instance, evaluating a road’s condition leading over a bridge takes place independently of the bridge’s condition assessment and, as seen in Germany, can occur at different frequencies.

Moreover, the time references present in the data vary. For example, in the previously mentioned asset condition assessment, a specific condition is assigned to a defined point in time. Conversely, the use of cumulative condition values is quite common. For example, when evaluating the traffic load on a road, traffic censuses are performed over a designated period, and the findings are cumulated to yield a value representing the traffic load. This condition description then does then not correspond to a specific mo-

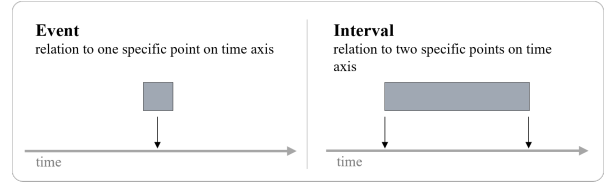


Figure 2: Types of time references

ment in time but rather indicates a state over an extended time frame.

To facilitate temporal reasoning between data, despite their inherent heterogeneity and/or inaccuracies in temporal assignments, Allen (1983) suggests utilizing time intervals and analyzing relative temporal relations. We implement this approach by representing the time references of elements within a subsystem graph, to subsequently extract data sets in a certain particular temporal relation on a large scale.

Proposed Graph Structure

The same principle applied to analyze spatial relations based on a knowledge graph is utilized to evaluate relative temporal relations. The references of the elements to a reference system, in this instance, the time axis, is explicitly modeled in the *Temporal Relation Graph*. The implicitly contained relations among the elements are examined by searching this *Temporal Relation Graph* for a pattern graph representing the specific relation in question.

When representing temporal references in the graph, a basic distinction is made between two types of time references, according to the different types of time-dependent data described above. As shown in Figure 2, data sets or information that can be assigned to a specific point in time are referred to as events, and data sets that can be assigned to a period of time with a start and end point are referred to as intervals.

Events can be used for representing traffic accidents, for instance, or individual data points in a time series. Intervals allow for capturing conditions or condition developments over a specified period. The condition may remain constant throughout the interval, as seen with cumulative data, or it can be described through a time-dependent function or a collection of data points.

An event can be derived from an interval, and conversely, multiple events can be consolidated into an interval, aligning with the reference interval concept introduced by Allen (1983). While utilizing reference intervals may enhance the graph’s structure, the temporal relations among elements of different intervals exist only indirectly through the reference interval. Accordingly, if a distribution of data storage is implemented, such as storing the actual data represented within a reference interval in a time series database, utilizing reference intervals might require further processing.

The proposed structure for the *Temporal Relation Graph* is illustrated in Figure 3. Each data set node is labeled ac-

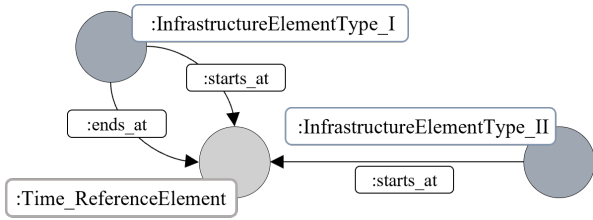


Figure 3: Representation of time references in the graph

ording to its corresponding infrastructure element, while the time axis node is labeled *Time_ReferenceElement*. The relations between the data sets and the time axis are represented through edges, with the edge labels specifying the type of time reference. Additionally, a property *time* associated with each edge defines the position on the time axis. The data itself can be stored as properties directly at the corresponding nodes. This data may consist of either a single value for event data, constant values over an interval, a function capturing a condition's development, or, in the case of time series data, a list of values. Moreover, providing a link to the actual dataset is also feasible if a system utilizing distributed data storage is implemented. In order to link data sets according to their specific temporal relations, the *Temporal Relation Graph* is queried. A pattern graph representing the desired temporal relation is identified to facilitate this. The *Temporal Relation Graph* is then searched for occurrences of this pattern graph. The matches obtained yield the representations of the data sets whose relation adheres to the defined pattern graph structure. Figure 4 illustrates the application of this concept for all possible temporal relations among intervals as defined by Allen (1983). A graphical representation of the pattern graph and the corresponding CYPHER query are provided.

Integration with the Spatial Analysis

The combined analysis of spatial and temporal relations is conducted in a two-stage process, where spatial relations are evaluated first, followed by considering temporal relations (as described above). To evaluate spatial relations, the information necessary for examining temporal relations must be stored in a way that does not affect the analysis of spatial relations. As previously mentioned, time references exhibit significantly less diversity compared to spatial ones, allowing for the utilization of the properties of the corresponding nodes to store the time references during the evaluation of spatial relations. The properties containing time references do not affect the methodology employed for analyzing spatial relations. Instead, they serve as a foundation for creating the *Temporal Relations Graph* to examine temporal relations only for the data sets in relevant spatial relations.

Application of the Concept

Data from the German road infrastructure management is used to demonstrate an implementation of our concept. We facilitate correlation analysis of data from traffic counting,

recording traffic on roads, and bridge condition grades describing the condition of the bridges over which the corresponding roads lead.

Data on bridge conditions is collected in Germany at regular intervals of three years through manual inspections according to DIN 1076 (Deutsches Institut für Normung (1999)), a German standard. During these inspections, damages are documented and interpreted further into a condition grade. The condition grade and the documented damages are assigned to the respective inspection (Bundesministerium für Verkehr (2013)) and stored in *SIB Bauwerke*, a proprietary asset management system based on a relational database.

Traffic data is primarily gathered through automated counts at traffic counting points. This collected data includes information on the types of vehicles counted over specific periods, which is then aggregated to generate metrics representing the traffic load. The frequency of these traffic data collections ranges from quarterly to every five years. The recorded traffic data is linked to the corresponding traffic counting station and stored in *BAYSIS*, the road management system utilized in Bavaria, a federal state in Germany. *BAYSIS* is a Geographic Information System (GIS) that is also built on a relational database system.

Through our concept, we aspire to associate bridge condition grades with measured traffic loads by analyzing the spatial and temporal relations within the data sets. The spatial relations reveal which traffic traversed each bridge, while the temporal relations indicate which traffic loads were recorded prior to specific bridge conditions.

Existing methods are used to create graph representations of legacy data and identify the assets (bridges and traffic counting points) exhibiting a particular spatial relation. Extracting the data sets linked to these assets yields collections of traffic data and bridge condition grades, all in pertinent spatial relations. From these collections, we aim to assign specific bridge condition grades to specific traffic values based on their time references. However, the data sets exhibit the differences in timeliness highlighted by Beck et al. (2021), which stem in part from the differing intervals of data collection and are further compounded by the completely independent gathering of traffic and asset condition data. This heterogeneity inhibits the straightforward assignment of data sets to one another, which we address by employing our proposed method to align the data sets based on their relative temporal relations:

The *Temporal Relations Graph* is created by utilizing the time references contained in the properties of the representations of the data sets stemming from the spatial dependency analysis. In the *Temporal Relations Graph*, the time axis, each traffic count, and every bridge condition data set are represented as nodes. Spatially related data sets possess relations to a shared *Time_ReferenceElement* node, resulting in multiple subgraphs, as illustrated in Figure 5. This graph structure ensures that searching for a specific pattern graph within the *Temporal Relations Graph* retrieves data sets that are temporally and spatially related.

		$t_{end_1} < t_{start_2}$	<pre>MATCH (d1:Label1)-[e1:ends_at]-()-[s2:starts_at]-(d2:Label2) WHERE e1.time < s2.time RETURN [d1,d2]</pre>
		$t_{end_1} = t_{start_2}$	<pre>MATCH (d1:Label1)-[e1:ends_at]-()-[s2:starts_at]-(d2:Label2) WHERE e1.time = s2.time RETURN [d1,d2]</pre>
		$t_{start_1} < t_{start_2}$ $t_{end_1} > t_{start_2}$	<pre>MATCH (d1:Label1)-[e1:ends_at]-(r)-[s2:starts_at]-(d2:Label2) MATCH (d1)-[s1:starts_at]-(r) WHERE e1.time > s2.time AND s1.time < s2.time RETURN [d1,d2]</pre>
		$t_{start_1} = t_{start_2}$ $t_{end_1} < t_{end_2}$	<pre>MATCH (d1:Label1)-[e1:ends_at]-(r)-[s2:starts_at]-(d2:Label2) MATCH (d1)-[s1:starts_at]-(r) WHERE e1.time > e2.time AND s1.time = s2.time RETURN [d1,d2]</pre>
		$t_{start_1} > t_{start_2}$ $t_{end_1} < t_{end_2}$	<pre>MATCH (d1:Label1)-[e1:ends_at]-(r)-[s2:starts_at]-(d2:Label2) MATCH (d1)-[s1:starts_at]-(r) WHERE e1.time < e2.time AND s1.time < s2.time RETURN [d1,d2]</pre>
		$t_{start_1} < t_{start_2}$ $t_{end_1} = t_{end_2}$	<pre>MATCH (d1:Label1)-[e1:ends_at]-(r)-[s2:starts_at]-(d2:Label2) MATCH (d1)-[s1:starts_at]-(r) WHERE e1.time = e2.time AND s1.time < s2.time RETURN [d1,d2]</pre>
		$t_{start_1} = t_{start_2}$ $t_{end_1} = t_{end_2}$	<pre>MATCH (d1:Label1)-[e1:ends_at]-(r)-[s2:starts_at]-(d2:Label2) MATCH (d1)-[s1:starts_at]-(r) WHERE e1.time = e2.time AND s1.time = s2.time RETURN [d1,d2]</pre>

Figure 4: Querying of possible relative temporal relations between intervals

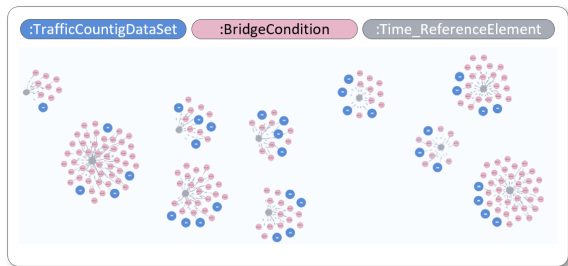


Figure 5: Excerpt of the Temporal Relations Graph

Figure 6 illustrates a created subgraph for a set of data records in spatial relation, with shorter label names used for improved readability. Each data set contains a single time reference, leading them to be initially interpreted as events, which are therefore connected to the time reference element node via an edge labeled *starts_at*. Searching for the pattern graph depicted in the middle of Figure 6 enables determination of the traffic count that occurred immediately before the corresponding bridge condition. Thus, pairs of values are provided based on relative temporal relation, allowing assessment of their correlation. However, as visible in Figure 6 on the right, this approach

does not account for the validity period of a bridge state, potentially neglecting some possible pairs.

Instead, assigning a bridge condition grade to all traffic data collected between the prior inspection and the time frame under consideration for condition assessment can be effectively accomplished by considering the bridge condition grades as an interval representing its validity period. This approach is illustrated in Figure 7 and leads to all contained data records being associated.

A demonstration of the described use case using a very small, anonymized dataset is available online.¹ This demonstration includes the relevant CYPHER queries for creating the *Temporal Relation Graph*, adjusting the graph to include the validity intervals, and performing the pattern matching.

Conclusions and Discussion

In this paper, we propose an approach to evaluate large-scale spatial and temporal relations between different datasets to enable the application of statistical methods for correlation analysis. We focus on condition-describing datasets from the field of road infrastructure management

¹https://github.com/IHeise/EC3_2025_TemporalRelations

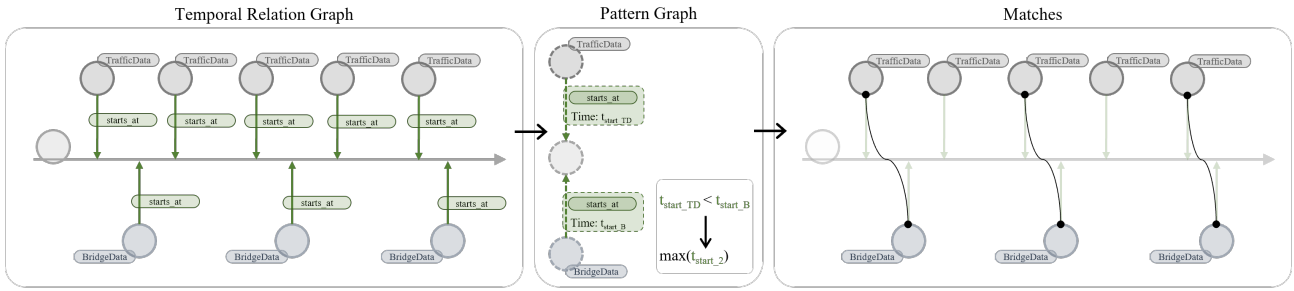


Figure 6: Association of event data sets based on their relative temporal relation

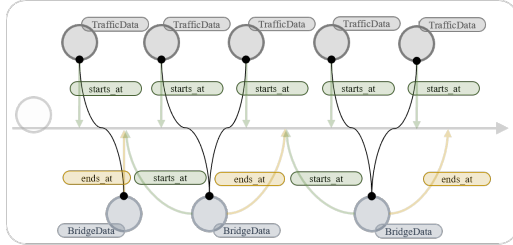


Figure 7: Association of interval and event data sets based on their relative temporal relation

and consider, in particular, the heterogeneity of the underlying data models – a well-known challenge in the context of BIM-GIS integration.

To facilitate a flexible evaluation of relations, we represent the various data sets and their references to reference systems in knowledge graphs. Searching for occurrences of pattern graphs representing specific relations facilitates extracting data sets exhibiting a particular relation. Thereby, contextual linking is implemented.

A sequential process is suggested for the combined analysis of spatial and temporal relations. First, spatial relations are examined using approaches presented in prior research. Subsequently, temporal relations are evaluated, addressing the challenges posed by the inherent heterogeneity of the data through the application of Allen’s interval algebra. Thus, despite their heterogeneity, comprehensive linking of various infrastructure data originating from diverse sub-systems is enabled.

Integrating the consideration of temporal relations with spatial relations makes it possible to associate state-describing data. In the subsequent step, regression analysis can be conducted to reveal and formalize the correlations among these state descriptions. Thereby, a comprehensive digital twin, built on knowledge graphs, contributes significantly to the enhancement of predictive maintenance.

Our approach has limitations concerning the time-dependency of spatial relations. We simplify by assuming spatial relations remain constant over time and can therefore be entirely decoupled from temporal dependencies. However, even if the locations of assets do not change throughout their lifespan, the connecting road network may influence the spatial relations between them. Con-

sequently, alterations to the road network could impact the validity of these spatial relations, making them time-dependent.

Regarding integrating high-frequency data such as time series, the consideration of integrating other database systems, such as relational time series databases, was not considered further. However, the integration of nearly continuous condition descriptions would push the approach toward a Digital Twin.

Furthermore, the approach outlined here emphasizes the verification of presumed correlations. However, there may also be unknown correlations in addition to the expected dependencies among the data sets. Utilizing graph neural networks to explore these unknown relations appears promising.

Acknowledgments

The presented research has been funded by the Bavarian State Ministry for Housing, Construction and Transport in the frame of the project ”Digital twin for operating road infrastructure”.

References

- Ackoff, R. L. (1971). Towards a system of systems concepts. *Management Science*, 17:661–671.
- Allen, J. F. (1983). Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26:832–843.
- Ammar, A., Nassereddine, H., and Dadi, G. (2022). Roadmap to a Holistic Highway Digital Twin: A Why, How, and Why Framework. IntechOpen.
- Beck, S. F., Abualdenien, J., Hijazi, I. H., Borrmann, A., and Kolbe, T. H. (2021). Analyzing Contextual Linking of Heterogeneous Information Models from the Domains BIM and UIM. *ISPRS International Journal of Geo-Information*, 10:807.
- Beetz, J., Amann, J., and Borrmann, A. (2018). Linked Data - Analyse von Einsatzmöglichkeiten von verbundenen Informationen (Linked Data) und Ontologien und damit befassten Technologien (Semantic Web) im Bereich des Straßenwesens.
- Broo, D. G. and Schooling, J. (2023). Digital twins in infrastructure: definitions, current practices, challenges

- and strategies. *International Journal of Construction Management*, 23:1254–1263.
- Bundesministerium für Verkehr (2013). ASB-ING. Date accessed: 12.02.2024.
- Buuveibaatar, M., Lee, K., and Lee, W. (2022). Development of a Conceptual Data Model for 3D Geospatial Road Management Based on LandInfra Standard: A Case Study of Korea. *ISPRS International Journal of Geo-Information*, 11:316.
- Deutsches Institut für Normung (1999). DIN 1076: Ingenieurbauwerke im Zuge von Straßen und Wegen; Überwachung und Prüfung. Beuth Verlag, Berlin.
- Ermolayev, V., Batsakis, S., Keberle, N., Tatarintseva, O., and Antoniou, G. (2014). Ontologies of time: Review and trends. *Int. J. Comput. Sci. Appl.*, 11:57–115.
- Grüninger, M. and Li, Z. (2017). The Time Ontology of Allen's Interval Algebra. In Schewe, S., Schneider, T., and Wijsen, J., editors, 24th International Symposium on Temporal Representation and Reasoning (TIME 2017), volume 90 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 16:1–16:16, Dagstuhl, Germany. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.
- Göbels, A., Rivadeneyra, F., and Beetz, J. (2023). Transfer of implicit semi-formal textual location descriptions in three-dimensional model contexts.
- Hamdan, A.-H. and Scherer, R. (2020). Integration of BIM-related bridge information in an ontological knowledgebase. *Proceedings of the 8th Linked Data in Architecture and Construction Workshop - LDAC2020*, Dublin, Ireland, pages 77–90.
- Heise, I. (2023). Requirements analysis for a digital road infrastructure twin. *Proceedings of 33rd Forum Bauinformatik*, 2023.
- Heise, I. and Borrmann, A. (2024). *infraspatialOT: An ontology for the representation of spatial relationships in road infrastructure*. *Proceedings of LDAC 2024 - Linked Data in Architecture and Construction*, Bochum, Germany.
- Heise, I., Esser, S., and Borrmann, A. (2024a). Towards a comprehensive digital twin of a road infrastructure system – requirements analysis and system architecture. *Proceedings of the 31st International Workshop on Intelligent Computing in Engineering*, Vigo, Spain, pages 74–83.
- Heise, I., Göbels, A., Borrmann, A., and Beetz, J. (2024b). Enabling Comprehensive Querying of Road and Civil Structure Data using Graph-based Methods. *Proceedings of the 41st International Conference of CIB W78*, Marrakech, Morocco, 2-3 October, ISSN: 2706-6568.
- Herle, S., Becker, R., Wollenberg, R., and Blankenbach, J. (2020). GIM and BIM. *PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 88:33–42.
- Hogan, A. (2020). *Linked Data*, pages 515–625. Springer International Publishing.
- Janse, R. J., Hoekstra, T., Jager, K. J., Zoccali, C., Tripepi, G., Dekker, F. W., and van Diepen, M. (2021). Conducting correlation analysis: important limitations and pitfalls. *Clinical Kidney Journal*, 14:2332–2337.
- Pan, Y., Wang, M., Lu, L., Wei, R., Cavazzi, S., Peck, M., and Brilakis, I. (2024). Scan-to-graph: Automatic generation and representation of highway geometric digital twins from point cloud data. *Automation in Construction*, 166:105654.
- Pearson, K. and Henrici, O. M. F. E. (1896). VII. Mathematical contributions to the theory of evolution.— III. Regression, heredity, and panmixia. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 187:253–318.
- Purohit, S., Van, N., and Chin, G. (2020). Semantic property graph for scalable knowledge graph analytics. *2020 IEEE International Conference on Big Data (Big Data)*, pages 2672–2677.
- Taherkhani, R., Ashtari, M. A., and Aziminezhad, M. (2024). Digital Twin-Enabled Infrastructures: A Bibliometric Analysis-Based Review. *Journal of Infrastructure Systems*, 30.
- VanDerHorn, E. and Mahadevan, S. (2021). Digital Twin: Generalization, characterization and implementation. *Decision Support Systems*, 145.
- Vilgertshofer, S., Amann, J., Willenborg, B., Borrmann, A., and Kolbe, T. H. (2017). Linking BIM and GIS Models in Infrastructure by Example of IFC and CityGML. pages 133–140. *American Society of Civil Engineers*.
- Weise, M., Schmidt, I., and Hettwer, J. (2018). OKSTRA, IFC und BIM: Neue Wege im Umgang mit Straßeninfrastrukturdaten.
- Zhang, C. and Hammad, A. (2005). Spatio-temporal issues in infrastructure lifecycle management systems. In *Proceedings of 1st CSCE Specialty Conference on Infrastructure Technologies, Management and Policy*.
- Zhang, C. and Hammad, A. (2006). Spatio-temporal representation and analysis in infrastructure systems. In *Proc., Joint Int., Conf. on Computing and Decision Making in Civil and Building Engineering*.