

Graph Databases as a Paradigm Shift in Behavior Science

Juan C. Correa^{1,2*}

^{1*}Departamento de Estudios Empresariales, Universidad Iberoamericana,
Mexico City, Mexico, 1219.

^{2*}Research and Development Unit, Critical Centrality Institute.

Corresponding author(s). E-mail(s): juan.correa@ibero.mx;

Abstract

Behavior scientists have multiple options for managing research data. While relational databases offer robust tools for storage and analysis, they impose structural constraints that limit the representation of complex behavioral phenomena. This article argues that graph databases provide a rigorous and conceptually richer alternative, enabling the modeling of human behavior as an interconnected system rather than a set of isolated variables. Beyond a technological shift, adopting graph-based approaches invites a paradigm change in behavior science—one that embraces complexity, dynamic relationships, and multi-level contingencies. Practical implications are illustrated through examples from clinical, consumer, and industrial/organizational psychology.

Keywords: Graph database, Network modeling, complex behavior

1 Introduction

In a recent paper, [Soto \(2025\)](#) introduced relational databases for behavior science and used real-world examples to illustrate how relational databases have been used by behavior scientists. Although relational databases remain the dominant paradigm in research and industry, alternative approaches are gaining traction. The so-called ‘Not Only SQL’ (NoSQL) category includes database systems that employ non-tabular data models, such as key-value pairs, documents, wide columns, matrices, and graphs. Whereas relational databases store structured data in tables (i.e., columns as variables,

rows as cases, and cells containing specific values), NoSQL systems accommodate flexible formats that better support evolving and highly connected data.

In this article, I examine graphs as a different paradigm from the traditional paradigm of relational databases, where nodes and edges (instead of tables and joins) represent the basic elements of any behavior that can be represented as a network or complex system. Networks have a long history in mathematics as “*graph theory*” (Estrada, 2011). In sociology and social sciences, graph theory is known as “social network analysis” (Wasserman & Faust, 1994), and psychologists have leveraged this framework to analyze the structure of psychopathology (Borsboom & Cramer, 2013), estimate the correct number of dimensions in psychological and educational instruments (Golino & Epskamp, 2017), or understand the measurement of organizational climate (Menezes, Menezes, Moraes, & Pires, 2021).

2 Network as a collection nodes and edges

One of the easiest way to grasp the idea of a network is by looking its visual representation (see Figure 1). According to Estrada (2011), a network is a collection of points (called nodes) joined together in pairs by lines (called arcs or edges). Scientists can model different kinds of networks from physical networks (e.g., flights between airports) to biological networks (e.g., protein-protein interactions), and social networks (e.g., who follows whom in LinkedIn or X). In this context, the term “social network” should not be confused with online platforms such as Facebook or Instagram, as they are technological implementations that do not necessarily represent all aspects of social networks as an academic discipline.

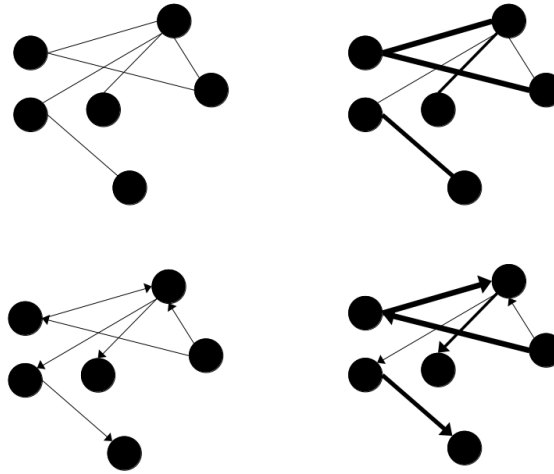


Fig. 1 A visual representation of four types of simple networks: non-directed unweighted network (upper left), non-directed weighted network (upper right), directed unweighted network (bottom left), and directed weighted network (bottom right).

Graphs offer fundamental concepts for understanding how entities (nodes) and their relationships (edges) form interconnected structures. Network structures underpin a wide range of behavioral phenomena —from disease transmission and social clustering to the spread of information and misinformation. As we have mentioned above, the recognition of these patterns requires tools that go beyond the rigid tabular constraints of relational databases. As the concepts of graphs are thoroughly covered in introductory texts (Newman, 2010), these will not be revisited here. Instead, this article aims to illustrate how graph-based databases can enrich the methodological toolbox of behavior scientists, enabling analyses that embrace complexity, dynamic relationships, and multi-level contingencies (Robinson, Webber, & Eifrem, 2015).

3 Graph databases: A gentle introduction

Robinson et al. (2015) define a graph database as a system that implements Create, Read, Update, and Delete (CRUD) operations on a graph data model, where entities are represented as nodes and relationships as edges. Unlike relational databases, which organize data in tables, graph databases treat relationships as first-class elements rather than secondary links between tables. This design enables efficient traversal and pattern matching across highly connected data, making it ideal for modeling complex networks such as behavioral contingencies or social interactions.

In Figure 2 we can see a graph of three X users (i.e., Tom, Paul, and Anna) and three messages (i.e., 99, 100, and 101). Anna and Tom **FOLLOWS** Paul, Tom also **FOLLOWS** Anna, but Paul does not follow Anna. Anna has a string of messages. Her most recent message can be found by following a relationship marked **CURRENT**, and the **PREVIOUS** relationships then create Anna’s timeline.

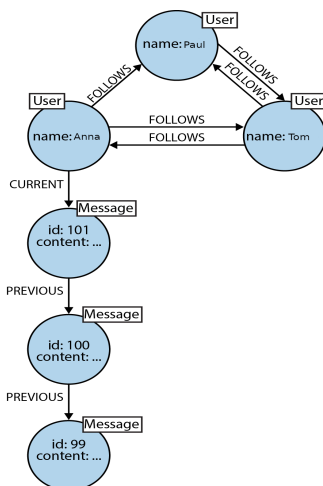


Fig. 2 A visual representation of a graph database that combines users and messages

A graph database like the one depicted in Figure 2 leverages the “labeled property graph” which has the following characteristics: 1) it contains nodes and relationships, 2) Nodes contain properties (key-value pairs), 3) Nodes can be labeled with one or more labels, 4) Relationships are named and directed, and always have a start and end node, 5) Relationships can also contain properties.

3.1 Benefits of graph databases.

According to [Robinson et al. \(2015\)](#), an advantage of graph databases is the sheer performance increase when dealing with connected data versus relational databases. Compared with relational databases, where join-intensive query performance deteriorates as the dataset gets bigger, a graph database performance remains relatively constant, even as the dataset grows with millions of records or variables. This is because queries are localized to a portion of the graph (the specific edge(s) to be consulted). As a result, the execution time for each query is proportional only to the size of the part of the graph traversed to satisfy that query, rather than the size of the overall graph.

Another advantage of graph databases has to do with their flexibility. Scientists aim to connect data in ways that reflect the domain itself, allowing structure and schema to evolve alongside their understanding of the problem rather than being rigidly defined upfront, when knowledge is most limited. This is particularly important when a behavioral phenomenon lacks theoretical background. Graph databases fulfill this need by providing a flexible model that adapts to changing requirements. Because graphs are inherently additive, new nodes, relationships, labels, and subgraphs can be introduced. The modifications introduced to the original graph do not imply a threat to existing queries or application functionality. This flexibility minimizes the need for exhaustive upfront modeling and lowers the frequency of costly migrations, thereby reducing maintenance overhead.

A third benefit of graph databases is the agility that they offer. Modern graph databases enable smooth development and easy system maintenance. Their schema-free design, combined with testable APIs and query languages, allows controlled evolution of applications. While the absence of rigid schemas means traditional governance mechanisms are missing, this is not a drawback. On the contrary, it encourages more transparent and actionable data governance. Typically, governance is enforced programmatically through tests that validate data models, queries, and business rules. This approach aligns well with agile and test-driven development practices, making graph database-based applications adaptable to changing business needs.

It is worth mentioning that relational databases acknowledge relationships, but only during modeling, where they serve as join mechanisms between tables. In graph databases, we often need to clarify the meaning of relationships and even qualify their strength—something relational models cannot do ([Robinson et al., 2015](#)). As datasets grow more complex and less uniform, relational systems like Microsoft Access, PostgreSQL, or SQLite become burdened with large join tables, sparsely populated rows, and extensive null-checking logic. Increased connectedness in relational databases translates into more joins, which degrade performance and make adapting the database to evolving requirements difficult.

These and other limitations of relational databases are highlighted as “challenges to adoption” (Soto, 2025).

References

- Borsboom, D., & Cramer, A. (2013, 03). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, 9(Volume 9, 2013), 91–121, <https://doi.org/10.1146/annurev-clinpsy-050212-185608>
- Estrada, E. (2011). *The structure of complex networks: Theory and applications*. Oxford University Press.
- Golino, H.F., & Epskamp, S. (2017). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PLOS ONE*, 12(6), e0174035, <https://doi.org/10.1371/journal.pone.0174035>
- Menezes, I., Menezes, A.C., Moraes, E., Pires, P.P. (2021). Measuring organizational climate via psychological networks analysis. *International Journal of Organization Theory & Behavior*, 24(3), 229–250, <https://doi.org/10.1108/IJOTB-08-2020-0142>
- Newman, M.E.J. (2010). *Networks: An introduction*. Oxford University Press.
- Robinson, I., Webber, J., Eifrem, E. (2015). *Graph databases: New opportunities for connected data*. Sebastopol, California: O’Reilly.
- Soto, P.L. (2025, 12). Relational databases for behavior science. *Perspectives on Behavior Science*, , <https://doi.org/10.1007/s40614-025-00486-w>
- Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge University Press.