

Understanding the graph databases and power grid systems

Xuming Lv¹, Shuo Chen¹, Shanqi Zheng¹, Jingzhao Luan², Yonghui Guo³

¹State Grid Liaoning Electric Power Co., Ltd., Shenyang, Liaoning 110006, China

²Dalian Power Supply Company of State Grid, Dalian, Liaoning, 116001, China

³State Grid Liaoning Electric Power Company Limited Maintenance Branch Company, No.21, Guangrong Avenue, Shenyang, Liaoning, 110003, China

xuming_lv@126.com

Abstract. This paper serves as a survey of the utilisation of graph database in power grid system. The main idea here is to provide an in-depth insight into graph databases, analytic techniques used in these databases as well as discussing a real-world case around their usage. As the increase of communication and information asset management in power grid, the importance of database management is growing. Compared to the conventional table-based database, graph database shows distinct performance in both flexibility and data visualization. In the following parts, the relationship between graph database and power grid will be discussed in three perspectives: 1) Nodes as Grid Points; 2) Relationships and Grid Lines; and 3) Highly Flexible, Rapidly Indexed and Visually Clear. Also, this paper will talk about the limitations of it.

1. Introduction

As one of the most common sustainable recourses, electricity plays a more and more important role in modern society. Relative researches are proposed to provide effective and efficient electric power supply, including electric power generation, voltage transformation, electric power transmission, etc. Gradually, these researches lead to the study on power grid systems or smart grid, which is known as the central structure of the electric system in nations. This paper serves as the survey study of power grid system from the perspective of database management. The aim of the following discussion is associated with the overall usage as well as key elements related to graph databases. This combination would enable an overall cohesive argument to understand and create a clear discussion around graph databases when compared with traditional databases. The utilisation of graph databases allows organisations to look beyond the usual set of parameters, often tied to traditional databases. This approach hence allows the development of a highly dynamic and adaptable model for data visualisation and data analytics.

From a structure point of view, the discussion is divided into two main sections: the initial part highlights an overall platform as well as key aspects related to graph databases. Key benefits and limitations of the platform are highlighted here. This is followed by analysis of how graph databases can be used in the power grid sector, their limitations in this market and the value add of such databases in a power grid environment. Finally, a conclusion summarizing the overall contents of the discussion is presented.



2. Graph Databases

Robinson and Webber [1] state that the graph database approach involves tying together multiple pieces of information to form a single exhaustive model in order to both utilise and analyse data sets. The main focus here is to build multiple links in the form of relationships between the datasets, hence creating vast and expanded data pool. The figure below showcases a normal graph database and how various nodes are tied/linked together.

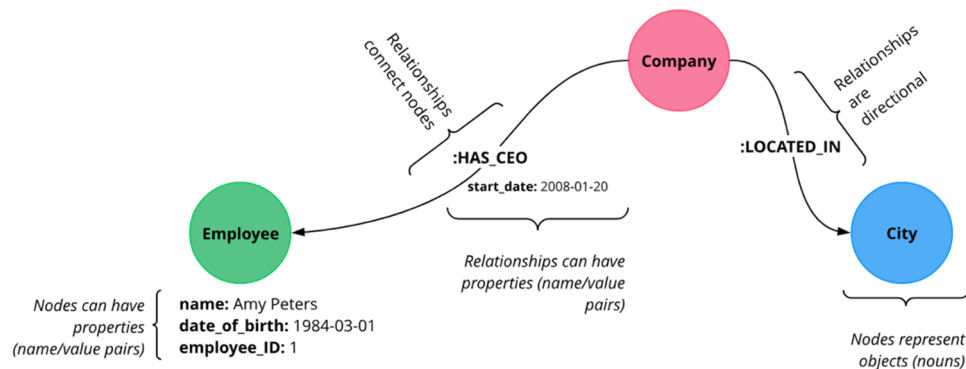


Figure 1. an example of graph database [2]

As shown in the Figure 1, each node in the graph database represents a logical entity. Rodriguez and Neubauer [3] state that a graph database allows the creation of almost infinite relationships, unlike traditional databases, graphs are easily relatable, forming a neural net of information.

On top of a graph database, it is important that a platform to analyze and then utilize this data is developed. Mike, Poly and Luis [4] state that over the past few years, multiple graph database platforms have been developed which take into account the relationship modelling of the nodes and then create meaningful data outputs from these nodes. This allows the development of bespoke and specific platforms that analyze data and hence provide meaningful outputs based on the analysis. It further enables the creation of a technical model which is in line with modules that are data-centric and native to the properties of data nodes.

In terms of using a graph platform for information asset management, Borislav [5] states that the creation of the graph database was aimed towards the development of a database that is both modular and flexible. This effectively means that organizations using a graph database as a core platform have the ability to create data sets that can be managed by a set of rules to populate the data repository. Hence, the utilization of graph databases is a logical step in information management that allows databases to scale and manipulate data paths as required. Another important reason for the use of graph nodes and databases is to facilitate communication. IEEE Quad Review [6] states that the dynamic nature of graph nodes allows organizations to use these databases as a communication backend. The ability of such database models to scale and adapt based on rapidly changing communication paths makes this a viable platform for such an operation.

3. Power Grid and Graph Databases

Tan and Novosel [7] state that power grids have some of the most complex data sets in operation, due to the vast scale of these networks and the large number of datapoints they generate.

Graph database modelling allows for significant flexibility in the management of a power grid, as it allows the organization using the platform to map each node as a grid point. What this means is that the graph database development would tie in directly with the mapping of a power grid into different nodes, allowing real time mapping and monitoring of the entire grid in the process. The following benefits are linked to the utilization of a graph database in a power grid:

3.1. Nodes as Grid Points

Graph databases use nodes as the core entities in the database store. Keeping this under consideration, the usefulness for power grids is that each node constitutes a unique grid point, hence forming a relationship with the main or parent grid in question. Mike et al [4] states that a node-based model in the graph data structure allows the formation of a highly flexible database that can be easily expanded or adapted. The utilization of nodes as grid points further allows a visualization platform for the entire deployment and hence creates relationships that can be easily filtered. Bin, Wang and Yatao [8] state that unlike traditional table databases, a graph database allows organizations to create a unique visual representation of a data model. This is a critical factor for power grids, as having such an approach allows the development of a multi-tiered data model that takes into account the complexities of a power grid deployment. Another important factor of a nodes-based model from a power grid deployment is that the overall power flow that can be visualised from the data set representations. Bin et al [8] states that traditional table-based databases are limited in how well they can show data paths efficiently, while with a graph based database and a node structure, the power grid data set can be easily overlaid to a network of data points. Borislav [5] states that power grids have multiple tiers of both transmission lines as well as transmission areas and that using a nodes-based model in a graph database, the grid system could have a clear and distinct representation of each tier. This is an aspect that is not traditionally possible in regular database modelling. Overall, nodes-based graph databases have a clear and distinct advantage if utilised effectively in the power grid and transmission sector.

3.2. Relationships and Grid Lines

The second component of a graph database is tied to the overall relationship structure of a dataset. From a power grid point of view, Tan and Novosel [7] states that the identification and clear distinction of transmission lines is vital. Using a relationship-based mapping, a graph database allows for the creation of a power grid dataset that clearly and distinctly ties different data points together with highly informative data lines in the forms of relationships. One of the core limitations of traditional databases is the fact the relationship formation in these is limited in terms of modularity as well as information handling. The use of relationship lines that show more than just connections within a graph database platform would allow the creation of a dynamic grid system, hence creating a real world, real time image of the transmission handling. This approach is of critical importance for power grid mapping and cannot be easily handled by traditional databases. Aguero, Takayesu, Novosel and Masiello [9] state that graph databases have rich information embedded within them, allowing relation lines to show relevant info, unlike static relationship trees in a table. Furthermore, due to the dynamic nature of graph database platforms, relationships can show in real time status updates and trigger different node statuses, hence allowing a high level of built-in functionality for the entire database. This is a core strength of the platform in the power grid use case.

3.3. Highly Flexible, Rapidly Indexed and Visually Clear

While the previously highlighted benefits aid in the implementation and mapping of a power grid to a graph database, generic benefits are also critical to the operation of a graph dataset. One of the key features of a graph dataset platform is the underlying infrastructure that allows the database to be rapidly indexed. Rudd [10] states that having an indexed database that uses relationship as the naturally indexed module allows graph databases to provide data relations rapidly when compared with alternate options. For a power grid system, having a highly indexed database would allow clear formation of links between different nodes – one that is easily and efficiently searchable. This approach allows the overall grid structure to be indexed and hence searchable using not just the parent node names but also the relationships formed as part of the process. Neo4j [2] shows that the core strength of the graph platform is its ability to quickly index relationships, which is highly relevant in the case of power grids.

Another important area of the graph database module is tied to the overall flexibility of the platform. Due to the node-based nature of the database, the implementation is not tied into any specific

formatting requirements, hence creating a highly flexible layer for power grids. This ties in directly with the fact that graph databases are developed to be visually clear. Mike et al [4] highlight that unlike traditional databases, graph models are built to be visually analysed, hence providing a clear platform to both monitor and visualise datasets in the process. This is true for the case of power grids, where data, if visualised properly, would provide the user with clear and concise mapping of different grid data points.

3.4. Limitations of Graph Databases in Power Grids

One of the most integral components, as highlighted earlier, was the fact that graph databases have a multi-tiered model for operational processes that allows them to create a platform that is visually effective. While this is a viable component of the platform, Rudd [10] states that due to the visually intensive nature of the database, the overall platform normally has a very high computing requirements. This means that the database requires significant processing power in order to operate under large scale conditions. When compared with a traditional table-based database, this is a notable drawback, as intensive data requirements have a major cost impact on large scale deployments such as those in a power grid system. Another area of concern when deploying a large scale graph database platform is the level of node-centric search and its potentially slow application. Rudd [10] highlights that while relationship indexing within graph databases is effective, the overall application from a node search point of view is limited and slow, creating issues when implemented in large scale systems such as power grid platforms. Finally, the level of modularity within a graph database can often cause issues in large deployments such as power grid systems, creating a fragmented system over time which is difficult to both maintain and upgrade. Bin et al [8] highlight that modularity is effective only when used efficiently, however graph databases often have third party plugins which, when scaled, hamper the performance of the entire system. These elements present a clear list of issues that may cause problems when utilising a graph database in a power grid system.

4. Conclusion

Overall, the utilisation of graph databases is a trend that is rapidly picking up pace. The discussion initially highlights how graph databases form the basis of developing a clear pathway to effective database visualisation. From an operational point of view, if used efficiently, graph database platforms provide a power grid with a system that is easy to scale, has the ability to show rich relationships between nodes and provides a visual platform that can be clearly monitored. These form the core benefits of using this approach. On the other hand, graph databases as mentioned above are resource intensive and therefore require compute requirements that are much higher than traditional databases. Further, when implemented for a power grid system, it is important that the database considers the specific requirements of the power grid platform and will be builds on top of those requirements. Generic graph databases are limited in their ability to provide a viable solution for a large scale power grid, however if used with the right data points, graph databases have clear advantage over inflexible table-based data models.

References

- [1] Robinson, R. and Webber, J. (2016), Graph databases, O'Reilly Media, Inc
- [2] Neo4j (2018) What is a Graph Database?, <https://neo4j.com/developer/graph-database/> (Accessed on 22nd May 2018)
- [3] Rodriguez, M. and Neubauer, P. (2015) "The graph traversal pattern", arXiv preprint arXiv, vol. 1004.1001
- [4] Mike, B. Poly, C. and Luis, S. (2012), The current state of graph databases, Cal Poly San Luis Obispo: Department of Computer Science
- [5] Borislav, I. (2016) "Hypergraphdb: a generalized graph database" in Web-Age Information Management, Berlin Heidelberg: Springer, pp. 25-36

- [6] IEEE Quadrennial Energy Review Report to DOE (2017), Available: <http://www.ieee-pes.org/qer> (Accessed on 22nd May 2018)
- [7] Tan, D. and Novosel, D. (2017) “Energy challenge, power electronics & systems (PEAS) technology and grid modernization,” CPSS Transactions on Power Electronics and Applications, vol. 2 no. 1
- [8] Bin, S. Wang, H. and Yatao, L. (2013) "Trinity: A distributed graph engine on a memory cloud", Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data, pp. 505-516
- [9] Agüero, J. Takayasu, E. Novosel, D. and Masiello, D. (2017) “Modernizing the grid: Challenges and opportunities for a sustainable future,” IEEE Power & Energy Magazine, vol. 14, pp. 74–83
- [10] Rudd, B. (2017) The Good, The Bad and the Hype about Graph Databases, <https://tdwi.org/articles/2017/03/14/good-bad-and-hype-about-graph-databases-for-mdm.aspx> (Accessed on 22nd May 2018)