

AN ANALYTICS-AWARE CONCEPTUAL MODEL FOR EVOLVING GRAPHS



Amine Ghrab^{1,2}, Sabri Skhiri¹, Salim Jouili¹ and Esteban Zimányi²

¹Eura Nova R&D, Mont-Saint-Guibert, Belgium

²Department of Computer & Decision Engineering, Faculty of Engineering, Université Libre de Bruxelles, Belgium
{amine.ghrab, salim.jouili, sabri.skhiri}@euranova.eu, ezimanyi@ulb.ac.be



Limitation of current graph models

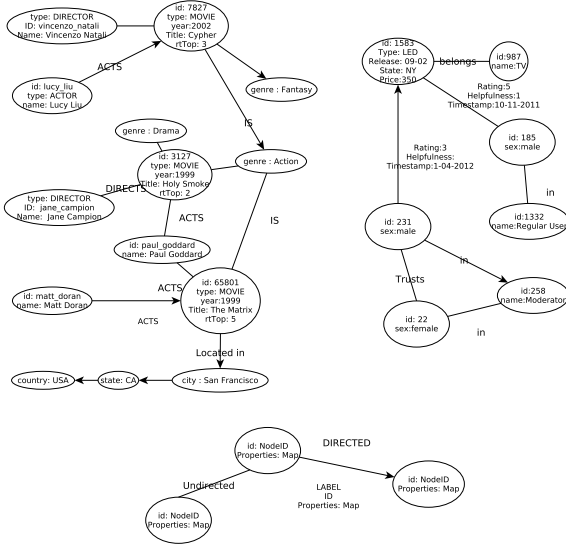


FIGURE 1: Sample networks modeled as a property graphs

In current graph databases, the property graph model is the most adopted. However, there is no commonly agreed conceptual model for graph databases. Moreover, none explicitly suggests tools for graph evolution tracking. On multidimensional graph analysis, graphs are assumed to be homogeneous and static.

Graph Data Modeling

1 The Graph Model

We defined a semi-structured data model for graph data analysis (Ghrab et al., 2013). The model introduces a well defined set of data structures and operators needed for the analysis.

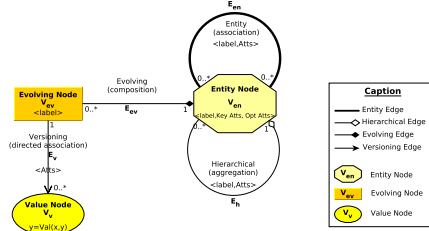


FIGURE 2: Analytics-oriented evolving graph metamodel

The model provides a set tools for modeling graph data and introduces two new data structures oriented for high level analysis, namely **analytics hypernodes** and **classes**. It supports representation and querying of nodes data as it **evolves**. Hence complementing current state of the art focusing more on the overall network structure (Khurana and Deshpande, 2013).

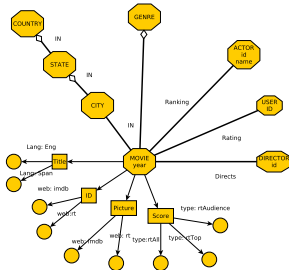


FIGURE 3: Our model applied on MovieLens recommendation website

2 Querying the Graph Data

- **Selection:** $\sigma_{([NLabel, AttVals]^*; [ELabel, AttVals]^*)}(\mathcal{G})$
- **Projection:** $\pi_{(EvLabel, \{ValSet\})}(\mathcal{G}, NLabel)$
- **Traversal:** $\tau_{(Start, Pattern)}$

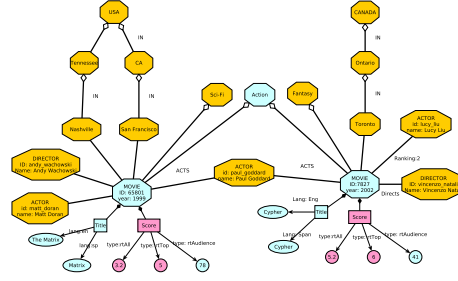


FIGURE 4: A subgraph of the model instance applied on MovieLens

Multidimensional Analysis of Graph Data

We propose to build the stack of multidimensional analysis of graph data, enriching current graph aggregation techniques (Chen et al., 2009) on both graph and cube levels.

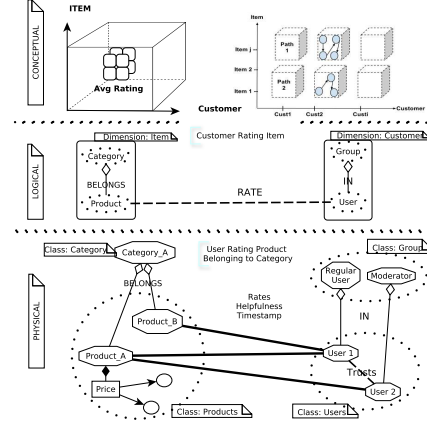


FIGURE 5: Multidimensional analysis on a recommendation network

Future Works: Towards a Graph Data Warehouse

- Algebra verification and completion
- ETL: either from a property graph or an instance of model-compliant graph
- Multidimensional structures and operations extension, algebra definition
- Multidimensional Query language for graph data
- Distributed storage and processing of graph data
- Temporal support for management of graph data

This work has been partially funded by the Wallonia Region in Belgium (Grant FIRST-ENTERPRISES N° 6850).

References

- Chen, C., Yan, X., Zhu, F., Han, J., and Yu, P. S. (2009). Graph OLAP: a multi-dimensional framework for graph data analysis. *Knowl. Inf. Syst.*, 21(1):41–63.
- Ghrab, A., Skhiri, S., Jouili, S., and Zimanyi, E. (2013). An analytics-aware conceptual model for evolving graphs. In *Data Warehousing and Knowledge Discovery*, Lecture Notes in Computer Science.
- Khurana, U. and Deshpande, A. (2013). Efficient snapshot retrieval over historical graph data. *2013 IEEE 29th International Conference on Data Engineering (ICDE)*, 0:997–1008.