



HAL
open science

Toward a generic approach to capture the temporal evolution in graphs

Landy Andriamampianina

► **To cite this version:**

Landy Andriamampianina. Toward a generic approach to capture the temporal evolution in graphs. Forum jeunes chercheuses, jeunes chercheurs d'INFORSID (JCJC 2020), Jun 2020, Dijon, France. hal-03292892

HAL Id: hal-03292892

<https://hal.science/hal-03292892>

Submitted on 20 Jul 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Toward a generic approach to capture the temporal evolution in graphs

Landy Andriamampianina¹²

1. *Institut de Recherche en Informatique de Toulouse (IRIT) - Université Toulouse 1 Capitole (UT1)*

2 *Rue du Doyen-Gabriel-Marty 31042 Toulouse*

landy.andria@irit.fr

2. *Activus Group*

1 *Chemin du Pigeonnier de la Cépière, 31100 Toulouse*

landy.andriamampianina@activus-group.fr

KEYWORDS: temporal graphs, historization mechanism, evolution types.

MOTS-CLÉS : graphes temporels, mécanisme d'historisation, types d'évolution.

DIRECTEUR DE THÈSE: Franck Ravat.

1. Introduction

Many domains, such as - computer networks, social networks, communication networks, biological networks etc. - are modelled as a graph as its flexible structure allows naturally representing complex relations among interconnected entities. In its most basic form, a graph is a collection of nodes with directed or undirected edges connecting them. The basic graph form is static, that is, it only reflects the snapshot¹ of a graph at a given time. However, real-world graph-based applications evolve constantly over time. Static graphs fail to represent, manage and trace their changes through time. Some studies have introduced the concept of temporal graphs to model the temporal evolution of graphs. Compared to static graphs, temporal graphs allow to model some evolution types in some specific applications. The purpose of our work is to propose a generic solution to manage any type of temporal evolution for any graph-based application.

1. an image of the entire graph at a given time.

2. State of the Art

A temporal graph is a graph representing nodes and edges, as well as all other parameters that the graph can have according to a model, that can be modified over time. The temporal evolution of a graph refers to the changes on its parameters over time, namely (i) *its topology*² and (ii) *data*³ (Zaki *et al.*, 2016).

The temporal evolution of graph topology is discussed in the literature through the addition and deletion of nodes and edges over time (Kostakos, 2009). Regarding the temporal evolution of graph data, it is discussed in the literature through changes in the attributes value of nodes (Desmier *et al.*, 2012) or edges (Zhao *et al.*, 2020) over time.

The standard temporal graph model in the literature is the sequence of graph snapshots⁴. It consists in splitting time into slices and then create a snapshot of the graph for each time slice (Fard *et al.*, 2012). Other studies model the temporal evolution of the graph by attaching to nodes and edges their presence times instead of attaching time to the whole graph (Latapy *et al.*, 2017).

In a nutshell, according to their applications, current modelling solutions include the temporal evolution of topology or data or both to track the temporal evolution of a graph. Most of them are based on the snapshot-based approach to keep temporal evolution traces.

3. Research questions

The first problem we want to address in our research is the representation of multiple types of the temporal evolution of a graph-based application. The state of the art shows that existing temporal graph models capture partially the temporal evolution of a graph. To the best of our knowledge, they do not capture the temporal evolution of graph data structure. Yet, this evolution type is a new information source that can enrich the evolution analysis of an application. As an example, we consider a social network where users and their relationships are respectively represented by nodes and edges in a temporal graph. Nodes and edges have attributes to describe users and their relationships characteristics. When a user adds a new friend, it creates a new edge in the graph modifying the graph topology and it updates the value of the attribute "friends list" of the user. When a user gets a job for the first time, a new attribute called "job status" should be added to its attributes set. This information can be new knowledge to understand evolution trends in the social network. For instance, the time when an important increase in the number of friends of a user may be correlated with the time when he added its job status, meaning that he possibly made new professional

2. the way in which the nodes and edges are arranged within a graph.

3. the attributes value of nodes and edges.

4. denoted G_1, G_2, \dots, G_T where G_i is an image of the entire graph at the time i and $[1; T]$ is the lifetime of the graph.

relationships. Therefore, some applications may require the complete aspect of temporal evolution i.e. the temporal evolution of graph topology, graph data and graph data structure. In this context, the first research question we want to answer is how to extend and generalize existing temporal graph concepts to be generic enough for capturing any evolution type and be reusable in any temporal graph based application.

The second problem we want to address in our research is the historization mechanism of temporal graph models. The snapshots-based model is the most used approach but it does not provide a straightforward solution to track and keep changes in a graph. To ensure the extraction of significant information on temporal evolution, it requires on the one hand to choose relevant time slice durations, which is a research topic itself (Ribeiro *et al.*, 2013). On the other hand, it requires observing differences between two snapshots of the graph. Then it is difficult to track the individual temporal evolution of nodes and edges. The drawbacks of snapshots-based approaches are due to the fact that the temporal evolution is treated at the level of the entire graph. In this context, the second research question we want to answer is how to track and keep evolution traces at different graph levels (node, edge and graph) contrary to existing snapshot-based approaches.

4. Proposition and future actions

The objective of this thesis is to design concepts and formalisms for the development of applications based on temporal graphs. To do so, we study two research axis: the modelling and the manipulation of temporal graphs.

First of all, we focused on the modelling of temporal graphs. To meet the shortcomings of the state of the art, our objective is to propose a modelling solution based on the concepts of *temporal entities*, *temporal relationships*, *instances* and *states*. We give an outline of our model in the following paragraph.

We define a temporal graph as representing evolving entities and relationships in time called respectively *temporal entities* and *temporal relationships*. To track and keep evolution traces of the latter, we model them at two levels of abstraction: the *state level* to describe their attributes set called *schema*, and the *instance level* to describe their attributes value. At each time there is a change in the schema of a temporal entity or relationship, a new state is created. At each time there is change in the attributes value, a new instance of the state is created. We affect to each state and each instance a time value depicting the time when it is valid. Thus, a temporal entity or relationship is composed by a set of states and each state is composed by a set of instances.

Compared to existing temporal graph models, our model has several advantages. First, it manages temporal evolution at a finer granularity level by affecting time values to entities and relationships rather than the entire graph as in snapshot-based approaches. Then, it embeds a complete information about the temporal evolution of entities and relationships. Indeed, the affected time values to temporal entities and relationships capture when the addition of new entity or relationship and new attribute(s)

in the schema of an entity or relationship, when the deletion of an entity or relationship and attribute(s) in the schema of an entity or relationship and, when the change in the attributes value of an entity or relationship are made in the graph.

Our next plan is the manipulation of our previously defined temporal graph model. Our objective for this part is to define a high-level language for querying our temporal graph combining different principles: extension and adaptation to our context of (i) database query languages based on versions (Ozsoyoglu, Snodgrass, 1995), (ii) specific operators of temporal graphs (Moffitt, Stoyanovich, 2017) and (iii) inference of new components in the graph (Fang *et al.*, 2011).

We will validate our modelling as well as its manipulation using data of real-world applications provided by the Activus Group company. We have already identified two implementation alternatives of our model in two data stores based on property graph model: Neo4j and OrientDB. We have defined translation rules to convert our model concepts into the property graph model.

References

- Desmier E., Plantevit M., Robardet C., Boulicaut J.-F. (2012). Cohesive co-evolution patterns in dynamic attributed graphs. In *International Conference on Discovery Science*, pp. 110–124. Springer.
- Fang C., Kohram M., Meng X., Ralescu A. (2011). Graph embedding framework for link prediction and vertex behavior modeling in temporal social networks. In *Proceedings of the SIGKDD Workshop on Social Network Mining and Analysis*.
- Fard A., Abdolrashidi A., Ramaswamy L., Miller J. (2012). Towards Efficient Query Processing on Massive Time-Evolving Graphs. In *Proceedings of the 8th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing*. IEEE.
- Kostakos V. (2009). Temporal graphs. *Physica A: Statistical Mechanics and its Applications*, Vol. 388, No. 6, pp. 1007–1023.
- Latapy M., Viard T., Magnien C. (2017). Stream Graphs and Link Streams for the Modeling of Interactions over Time. *arXiv:1710.04073 [physics, stat]*.
- Moffitt V. Z., Stoyanovich J. (2017). Temporal graph algebra. In *Proceedings of The 16th International Symposium on Database Programming Languages*, pp. 1–12.
- Ozsoyoglu G., Snodgrass R. (1995). Temporal and real-time databases: a survey. *IEEE Transactions on Knowledge and Data Engineering*, Vol. 7, No. 4, pp. 513–532.
- Ribeiro B., Perra N., Baronchelli A. (2013, December). Quantifying the effect of temporal resolution on time-varying networks. *Scientific Reports*, Vol. 3, No. 1, pp. 3006.
- Zaki A., Attia M., Hegazy D., Amin S. (2016). Comprehensive Survey on Dynamic Graph Models. *International Journal of Advanced Computer Science and Applications*, Vol. 7, No. 2.
- Zhao A., Liu G., Zheng B., Zhao Y., Zheng K. (2020). Temporal paths discovery with multiple constraints in attributed dynamic graphs. *World Wide Web*, Vol. 23, pp. 313–336.