H.F.R.I.
Hellenic Foundation for
Research & Innovation

Description of the funded research project
1st Call for H.F.R.I. Research Projects to Support Faculty Members &
Researchers and Procure High-Value Research Equipment

Title of the research project:
GraphTempo: Exploring the History of Temporal Graphs

GRAPHTEMPO

Principal Investigator: Evaggelia Pitoura

Reader-friendly title: Understanding and Predicting the Evolution of Graphs

Scientific Area: Mathematics and Information Sciences

Institution and Country: University of Ioannina, Greece

Host Institution: University of Ioannina

Collaborating Institutions: University of Warwick, UK,
Boston University, USA

Project webpage: http://graphtempo.eu/

Budget: 170,000 euro

Duration: 36 months

Research Project Synopsis

Graphs offer a natural model for representing entities and their interactions and relationships. In collaborative networks, edges capture the cooperation between actors in movies, authors of scientific articles, or co-workers in teams. In social networks, edges express the relationship (e.g., friend, follower) as well as the interactions and reactions (e.g., retweets, likes) between users. In communication networks, edges indicate email and phone exchanges between people, in transportation networks, roads and flights between cities, and in biological networks, interactions between proteins. Other examples include knowledge graphs, program dependency graphs, computer networks, and the web to name just a few.

Most real-world graphs are not static but evolve through time. New interactions and relationships are formed, while existing ones may no longer be valid. In addition, new entities appear, while old ones leave. Furthermore, besides structure, the content such as labels and properties on vertices, and weights on edges, also evolves.

The overall goal of the *GraphTempo* project is to provide a declarative framework for exploring the evolution of a graph that will act as a tool for predicting and potentially affecting its future operation. Towards achieving its overall goal, *GraphTempo* sets the following specific goals:
I.   to offer a general framework for the exploration of the full history of a graph through time including its structure (vertices and edges) and content (e.g., labels, properties, weights) associated with this structure. The framework will offer a declarative approach to temporal graph processing consisting of a set of novel exploration queries.
II.   to reveal hidden properties of real-world networks through the application of the framework, and
III.   to improve the evolution of networks by proposing novel link recommendation algorithms.

Project originality

GraphTempo advances previous research by offering a declarative framework for supporting the exploration of the evolution of graphs through time and the extraction of interesting information from this evolution. Two novel families of queries will be advanced, namely graph evolution and graph journey queries. *Graph evolution queries* will focus on (1) transformations, that is, identifying sets of graph elements that follow specific evolution patterns through time, and (2) change, that is, identifying graph elements based on the volume of change. *Graph journey* queries will concentrate on *ordered set* of graph elements, where a graph element with order i must appear earlier in time than a graph element with order $j > i$. Journeys are interesting since they may reveal dependencies, or causalities. To efficiently support graph evolution and journey queries, **GraphTempo** will use *graph embeddings*.

Graph evolution queries will be used to provide novel solutions to a variety of research problems, including (1) *team formation* in networks by identifying stable cooperation among people, that is, sets of vertices (or, graph elements in general) with a small amount of change, (2) spotting *homophily* in social and cooperation networks, for example, by locating co-evolving content and structural transformations, and (3) *anomaly detection* (such as spam behavior) in social networks, for example, by finding unexpected transformations, or volumes of change. Graph journey queries will be used in (1) identifying *information cascades* in social networks, for example, by finding the most common journeys that a piece of information (e.g., a hashtag) follows, and characterizing *forged information propagation* (e.g., fake news), (2) locating *spurious transaction* in financial networks such as in bitcoin, and (3) finding *interesting routes* in transportation networks.

Finally, **GraphTempo** will introduce new *query-evolution based link recommendation* algorithms.

Expected results & Research Project Impact

Graph data management has been the focus of much current research. There is also a surge on related commercial and research software. Most graphs used to represent real-world entities and their connections are not static but evolve with time. **GraphTempo** is expected to create *scientific impact* through the introduction of a novel declarative framework for the interactive exploration of graphs through time. This framework will go beyond classical graph query processing, by allowing querying the evolution of a graph. Developed software will be open-source and freely available as well as any data collected to maximize impact. The research results of the project will be published and presented in highly competitive international conferences and journals.

GraphTempo has also the potential to create *social* and *economic* impact. Graphs model important real-world networks. Exploring the evolution of such networks through time will improve our understanding of the mechanisms underlying their evolution. Querying the history of a graph will reveal hidden properties, alert us in cases of unexpected changes and help us identify suspicious behavior. In **GraphTempo**, we will apply our framework to several dynamic real-world networks towards capturing, understanding, and addressing specific social and economic problems. We will study homophily, the tendency of individuals to connect with similar individuals or being influenced by their connections. We will also investigate information cascades in networks to address important social problems such as the propagation of fake news, as well as online financial transactions to identify spurious interactions. Taking a step further, **GraphTempo** will investigate predicting and potentially influencing the evolution of real-world networks. To this end, link recommendation algorithms will be developed to recommend links based on their potential to improve desired properties of the network.

The importance of this funding

Many funding opportunities ask for large consortia, put emphasis on development rather than on research, emphasize dissemination impact rather than research impact .The H.F.R.I grant is important, since it funds basic research and focuses mainly on excellence. This fund has given me the opportunity to materialize a novel research idea, support young researchers, and cooperate with colleagues in the US and UK.

H.F.R.I.
Hellenic Foundation for
Research & Innovation