Graph Queries: Generation, Evaluation and Learning

Angela Bonifati

University Lyon 1

angela.bonifati@univ-lyon1.fr

1. INVITED TALK ABSTRACT

Several modern graph query languages are capable of expressing sophisticated graph queries, which return nodes connected by arbitrarily complex paths. Such paths can be synthesized by regular expressions and often involve recursion. Such graph queries are known as Regular Path Queries and correspond to Property Paths in Sparql 1.1. Recently, with my colleagues I have been investigating regular path queries and their combinations by looking at the generation problem [1, 2, 3, 10], the complexity of query evaluation [5] and the learning problem [9, 7, 6, 8]. Precisely, we focused on schema-driven generation of complex and broad graph queries with user-defined features, on the complexity of the evaluation of regular simple path queries and on learning algorithms for regular path queries. In this talk, I will begin with a brief recap of graph queries and their expressive power. I will then provide an overview of a comprehensive query-oriented graph benchmark that we have designed and assessed [1, 2, 3, 10]. I will next discuss the theoretical results of our study on the complexity of regular simple path queries [5]. I will then present a learning framework for regular path queries and discuss its potential along with its practical feasibility [7, 6, 8]. To conclude, I will briefly outline our ongoing work [4] and pinpoint lingering issues and research directions in the study of graph queries.

2. REFERENCES

[1] G. Bagan, A. Bonifati, R. Ciucanu, G. H. L. Fletcher, A. Lemay, and N. Advokaat. Generating Flexible Workloads for Graph Databases. *PVLDB*, 9(13):1447–1460, 2016.

[2] G. Bagan, A. Bonifati, R. Ciucanu, G. H. L. Fletcher, A. Lemay, and N. Advokaat. gMark: Schema-Driven Generation of Graphs and Queries. *IEEE Trans. on Knowl. Data Eng.*, In press, 2017.

[3] G. Bagan, A. Bonifati, R. Ciucanu, G. H. L. Fletcher, A. Lemay, and N. Advokaat. gMark: Schema-Driven Generation of Graphs and Queries (Extended Abstract). In *Proceedings of the International Conference on Data Engineering ICDE 2017, San Diego, USA*, To appear, 2017.

[4] G. Bagan, A. Bonifati, G. H. L. Fletcher, and H. Kheddouci. Labeling Schemes for RPQs and beyond. In preparation, 2017.

[5] G. Bagan, A. Bonifati, and B. Groz. A trichotomy for regular simple path queries on graphs. In *Proceedings of the 32nd ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems, PODS 2013, New York, NY, USA*, pages 261–272, 2013.

[6] A. Bonifati, R. Ciucanu, and A. Lemay. Interactive Path Query Specification on Graph Databases. In *Proceedings of the 18th International Conference on Extending Database Technology, EDBT 2015, Brussels, Belgium*, pages 505–508, 2015.

[7] A. Bonifati, R. Ciucanu, and A. Lemay. Learning Path Queries on Graph Databases. In *Proceedings of the 18th International Conference on Extending Database Technology, EDBT 2015, Brussels, Belgium*, pages 109–120, 2015.

[8] A. Bonifati, R. Ciucanu, A. Lemay, and S. Staworko. A Paradigm for Learning Queries on Big Data. In *Proceedings of the First International Workshop on Bringing the Value of "Big Data" to Users, Data4U@VLDB 2014, Hangzhou, China*, 2014.

[9] A. Bonifati, R. Ciucanu, and S. Staworko. Learning Join Queries from User Examples. *ACM Trans. Database Syst.*, 40(4):24:1–24:38, 2016.

[10] W. V. Leuween, A. Bonifati, and N. Yakovets. Stability notions in synthetic graph generation: a preliminary study. In *Proceedings of the 18th International Conference on Extending Database Technology, EDBT 2017, Venice, Italy*, To appear, 2017.