Towards Querying Uncertain Graphs

(Tutorial)

Arijit Khan
Nanyang Technological University
arijit.khan@ntu.edu.sg

ABSTRACT

Large-scale, highly-interconnected networks pervade both our society and the natural world around us. Uncertainty, on the other hand, is inherent in the underlying data due to a variety of reasons, such as noisy measurements, lack of precise information needs, inference and prediction models, or explicit manipulation, e.g., for privacy purposes. Therefore, uncertain, or probabilistic, graphs are increasingly used to represent noisy linked data in many emerging application scenarios, and they have recently become a hot topic in the database research community. The challenges in uncertain graph processing are both semantics and computation driven. From the perspective of the semantics, there is no uniform model of uncertain graphs; rather assignment and interpretation of the probabilities are application specific. For example, how can we define the shortest path between two nodes in an uncertain graph? From the computation perspective, while many classical graph algorithms such as reachability and shortest path queries become #P-complete, and hence, more expensive in uncertain graphs; various complex queries are also emerging over uncertain networks, e.g., pattern matching, information diffusion, and influence maximization queries. In this tutorial, we discuss the sources of uncertain graphs and their applications, uncertainty modeling, as well as the complexities and algorithmic advances on uncertain graphs processing in the context of both classical and emerging graph queries. We emphasize the current challenges and highlight some future research directions.

Biography

Arijit Khan is an assistant professor in the School of Computer Engineering at Nanyang Technological University. His research interests span in the area of big-data, big-graphs, and graph systems. He received his PhD from the Department of Computer Science, University of California, Santa Barbara, and did a post-doc in the Systems group at ETH Zurich. Arijit is the recipient of the prestigious IBM PhD Fellowship in 2012-13. He published several papers in premier database and datamining conferences and journals including SIGMOD, VLDB, TKDE, ICDE, SDM, EDBT, and CIKM. Arijit copresented tutorials on emerging graph queries, big-graph systems, and uncertain graphs at ICDE 2012, VLDB 2014, VLDB 2015, and served in the program committee of KDD, SIGMOD, ICDM, EDBT, WWW, and CIKM. Arijit served as the co-chair of Big-O(Q) workshop co-located with VLDB 2015.

International Conference on Management of Data, COMAD,