Landmark indexing for scalable evaluation of label-constrained reachability queries[*]

Lucien Valstar[1] George Fletcher[1] Yuichi Yoshida[2,3]

[1]Eindhoven University of Technology, the Netherlands
[2]National Institute of Informatics, Japan
[3]Preferred Infrastructure, Inc., Japan

Consider a directed edge-labeled graph, such as a social network or a citation network. A fundamental query on such data is to determine if there is a path in the graph from a given source vertex to a given target vertex, using only edges with labels in a restricted subset of the edge labels in the graph. Such label-constrained reachability (LCR) queries play an important role in graph analytics, for example, as a core fragment of the so-called regular path queries which are supported in practical graph query languages such as the W3C's SPARQL 1.1, Neo4j's Cypher, and Oracle's PGQL. Current solutions for LCR evaluation, however, do not scale to large graphs which are increasingly common in a broad range of application domains. In this talk we present the first practical solution for efficient LCR evaluation, leveraging landmark-based indexes for large graphs. We show through extensive experiments that our indexes are significantly smaller than state-of-the-art LCR indexing techniques, while supporting orders of magnitude faster query evaluation times. Our complete C++ codebase is available as open source for further research.[1]

---

[1]`https://github.com/DeLaChance/LCR`