VIIQ: Auto-Suggestion Enabled Visual Interface for Interactive Graph Query Formulation

Nandish Jayaram, Sidharth Goyal*, Chengkai Li
University of Texas at Arlington

41st International Conference on Very Large Data Bases, Kohala Coast, Hawaii, Aug. 31st - Sept. 4th, 2015

idir.uta.edu/orion

Motivation

Graphical query interfaces are widely recognized to be important for improving the usability of data management systems.

Limitations of Existing Visual Query Systems

- They allow users to draw nodes and edges of query graphs, but do not rank relevant suggestions regarding what to add.

- Difficult for schema-agnostic users to sift through and search for the exact option.
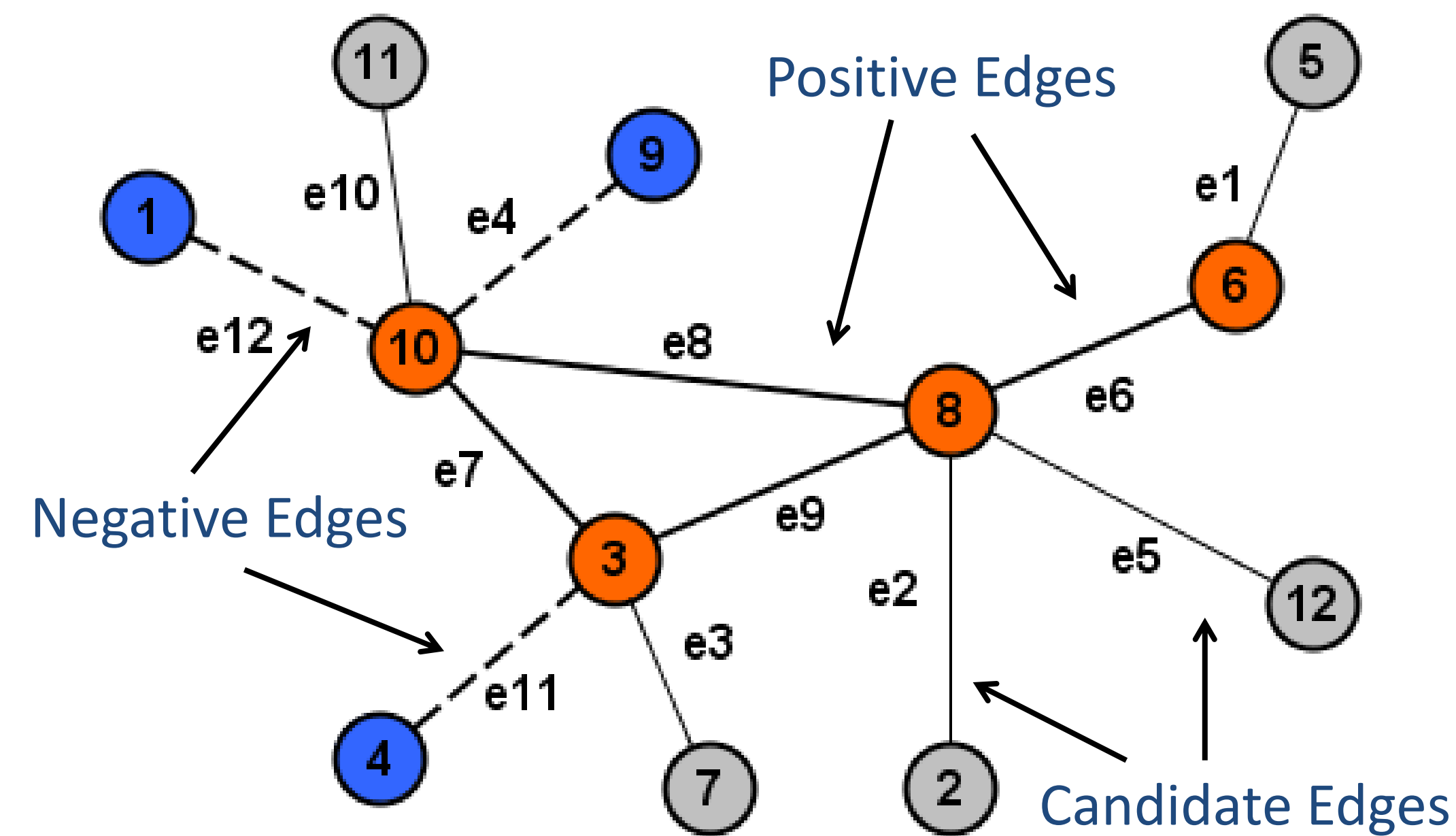
Challenges

- An interactive query interface that seamlessly integrates automatic edge recommendation with query formulation.

- Large ultra-heterogeneous graphs contain thousands of node/edge types, and millions of node/edge instances as candidates.

- Ranking must reflect the relevance to user's query intent.
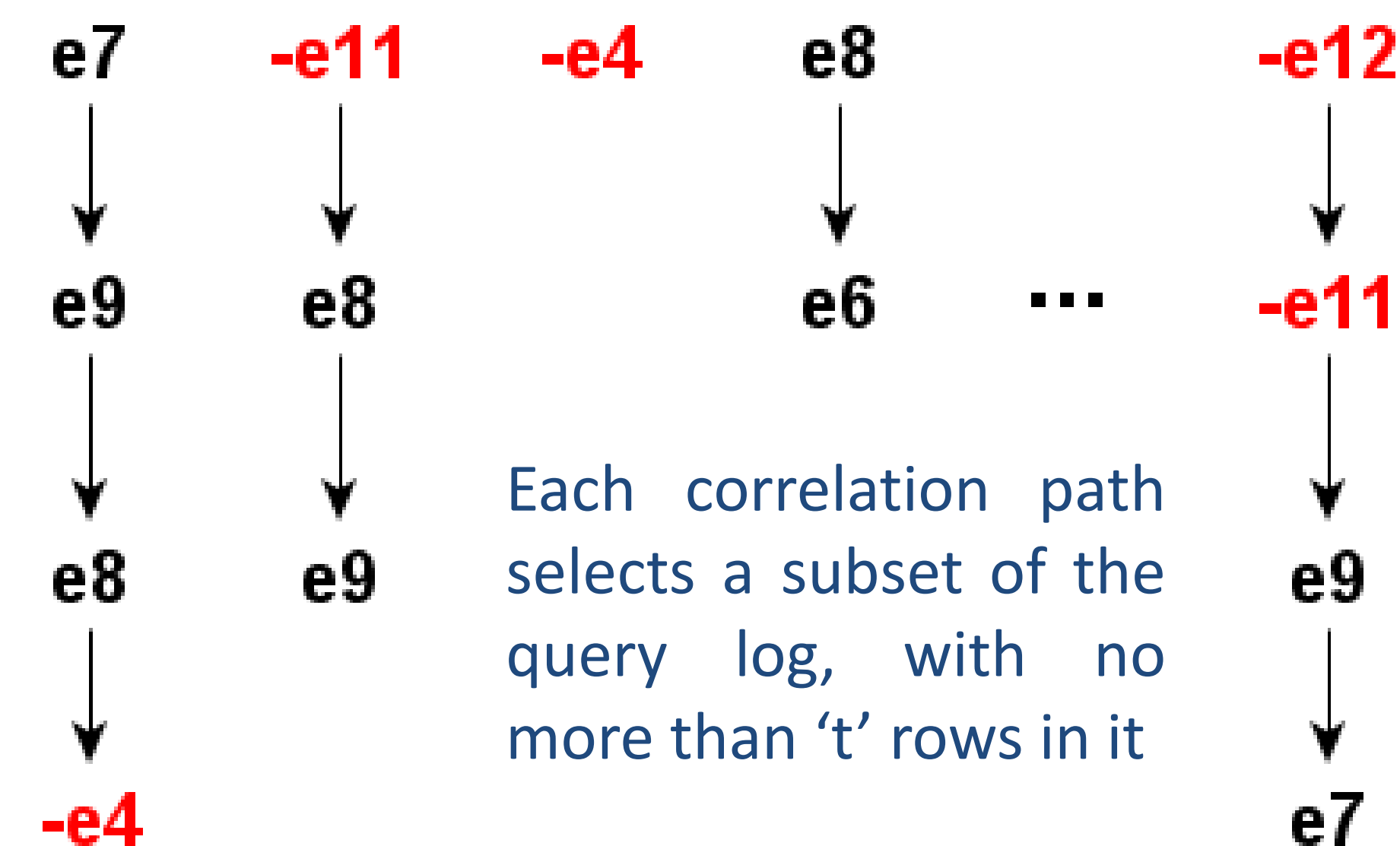
Orion Architecture

Rank candidate edges   Generate candidate edges
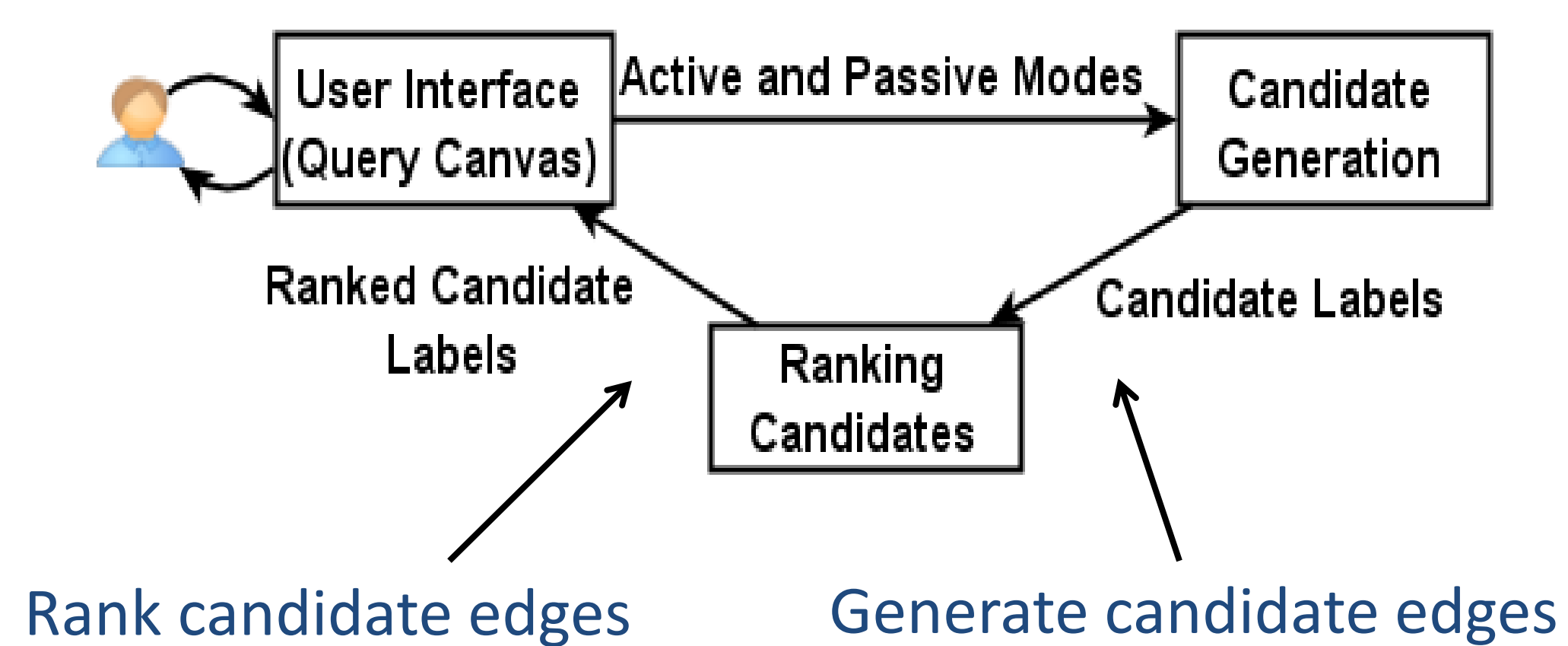
Ranking Candidates

- New Candidate Edges: <e1, e2, e3, e5, e10>

- Positive Edges: <e6, e7, e8, e9>

- Negative Edges: <e4, e11, e12>

- Query Session: <e6, e7, e8, e9, -e4, -e11, -e12>

- Query Log: A collection of other query sessions

Random Correlation Paths (RCP)

- RCP: A sequence of randomly chosen edges from the query session, grown incrementally, until its support drops below a threshold (t).

Each correlation path selects a subset of the query log, with no more than 't' rows in it

Compute support for each candidate, in each correlation path ➡ Find average score over all RCPs for each candidate edge

Orion User Interface

Query Canvas

Newly suggested nodes and edges in grey

Orange nodes with black edges form the partial query graph

Option to view the types of an edge's ends, and its example instances

Dynamic help listing possible user actions at every step

Other useful tips

Click on a suggested grey node to add it to the query graph (color turns blue)

Information Panel

Passive Mode of operation

Passive Mode of operation

Click on an automatically suggested grey edge to view other instances of that edge

Active Mode of operation   Adding a new node manually

Active Mode of operation

Already selected newly suggested node

Adding a new edge manually

This work is partially supported by NSF grants IIS-1018865, CCF-1117369 and IIS-1408928. Any opinions, findings, and conclusions or recommendations expressed in this paper do not necessarily reflect the views of the funding agency.

* Work done as a student at UT-Arlington