BACKGROUND

• Graph querying: locate entities with specific relationships among them
  • financial transaction networks
  • flag "near cliques" formed among company insiders
  • money-laundering
  • online auctions
  • uncover fraudsters and their accomplices
  • Bioinformatics
  • Social network analysis

DATA TO VIS AND DERIVED RESULTS

• DBLP Dataset.
  • DBLP is a computer science bibliography website.
  • Co-authorship network of DBLP’s computer science bibliography data, focusing on the data mining and information visualization communities
  • 59,655 authors; 48,677 papers; 7,236 sessions
  • 4,417 proceedings; 21 conferences; 1,634,742 relations

• Derived results
  • a novel interactive visual analytics system, for exploring and making sense of query results

METHODS & ARCHITECTURE

• Extract Features
  • Calculate the topological- and node-features.

• Vectorize
  • Merge the common features into per-result vectors.
  • Aggregated & Normalize into Signature - Reduce the large input vectors into uniform signatures.

• Reduce & Cluster
  • Reduce the signatures using dimensionality reduction.

• Reorder, realign, hovering highlight

METHODS & ARCHITECTURE (CONT’D)

• Extract Features
  • Structure features
    • Subgraph neighborhood and egonet information
    • An egonet of node i is (i) the neighbor nodes of i, (ii) the edges to these neighbors and (iii) all the edges among neighbors.

• Vectorize
  • Nodes feature
    • Author name
  • Number of co-authors
  • Number of conferences
  • Merge common feature
EVALUATION

- User Study
  - 12 participants from computing related majors.
  - 7 female, 5 male
  - age 21 to 31
  - Paid $10 for 70 minutes test.
  - Dataset: DBLP co-authorship network
  - Real World Application: Discovering Cybersecurity Blindspots

USER STUDY

- Tasks
  1. Find the count of ICDM conference papers by Daniel Keim.
  2. From the last two years of KDD publications, find and list the authors who are on more than one paper with 'entity' in the name.
  3. Find the number of distinct groups of researchers that Tobias Shreck is in from INFOVIS publications.
  4. Among coauthors of at least two papers together at INFOVIS and KDD, who has the most publications.

- Quantitative Results
  - Tasks: find out the software affected by executing four tasks and examine the average task time, and average # of errors.

- Observations and Subjective Results
  - Participants rate various aspects comparing both systems

CONTRIBUTIONS OF VIGOR

- Novel visual analytics system, VIGOR
- Exploring and making sense of graph querying results
- Exemplar-based interactive exploration
  - bottom-up: how many similar values are matched to each query-node
  - top-down: how a particular node value filters the results from the whole structure
- Novel result summarization through feature-aware subgraph result embedding and clustering.
  - VIGOR provides a top-down, high-level overview
  - Clustering node-feature and structural result similarity
  - An integrated system fusing multiple coordinated views
  - Brushable linked views among Exemplar View, Subgraph Embedding View, and the Fusion Graph

CRITIQUE

- The number of people for user study might not be enough and they are all professional users.
- Query sentence is hard to generate for non-professionals.
- The co-authorship is limited to one-hop

Thank you!