**Using Importance Flooding to Identify Interesting Patterns of Criminal Activity**

**Criminal Activity Networks (CAN) are Node-Link Representations of Investigative Intelligence**
- Widely accepted investigative technique
- Visualizable for Understanding
- Shareable Despite Policy and Privacy Concerns
- Analyzable Using Graph-based Techniques

**Design Goals For A CAN Methodology**
- Incorporate domain-appropriate heuristics
- Combine shared data with investigation specific data
- Tolerate missing and ambiguous data
- Be target focused

**Link Chart Challenges**
- Current off-the-shelf software lacks criminal database connectivity
- Assignments are distributed
- “Fishing” in available data for potential criminals is discouraged

**Cross-Jurisdictional Integration Challenges**
- Different schemas and classification taxonomies
- Entity matching
- Privacy and security policies need to protect subpoenaed records, sensitive details, and private personal information in local agency records

**Path-Based Importance Heuristics**
- Simple activity-based group rules identify people who play a particular role, in a particular kind of incident, in a particular date range
- Multi-group rules identify people who have participated in two or more specified associations: a link-node-link network path.
- Path rules identify individuals involved in specified short network paths. For example a person who is in the fraud group, connected in a recent suspect-to-suspect association to someone in the drug sales group, connected in a recent suspect-to-suspect association to a member of the aggrieved assault group. Rules may be node-link-node-link-node or node-link-node.

**Experimentation**
- Several methodologies were compared:
  - BFS randomly selected direct associates
  - Closest Associate used association closeness rules
  - Importance Flooding combined path-based importance and spreading activation

**The Algorithm**
- Assign link weights
- Estimate initial importance using activity records
- Iteratively pass importance along associational links
- Domain-Appropriate, User-Specified Weights
- Weight can be adjusted based on experience
- Rules are stored in XML files separate from the data
- Two types of weights:
  - Link (relationship) weights support spreading activation
  - Initial importance weights for nodes in the network

**Conclusions, Implications, and Future Directions**

**Conclusion:**
- Spreading activation and path-based importance rules seemed to improve results

**Implications:**
- Computer-supported selection techniques could improve analyst efficiency
- Intelligence analysis tools might benefit from spreading activation modules
- Promising association network analysis methodologies can be developed despite imprecise relations, uncertain entity mappings, inconsistent classification, and heuristic analysis rules

**Future Directions:**
- Can users effectively express heuristics? Will the findings be consistent in a larger study?
- Use importance flooding with concept maps and networks of biomedical relationships

---

**Acknowledgement**

This work was supported in part by the NSF, Knowledge Discovery and Dissemination (KDD) # 9883304

NSF, ITR: “COPLINK Center for Intelligence and Security Informatics Research - A Crime Data Mining Approach to Developing Border Safe Research”


We are also grateful to Kathy Martinjak, Tim Petersen, and Chuck Violette from the Tucson Police Department for their input

---

**Cross-Jurisdictional Law Enforcement Link Charts**

- Criminal association networks are useful for:
  - discovering conspiracy,
  - building investigative leads, and
  - illustrating the conspiracy to the jury in a court of law.
- Organized and narcotic crimes can involve hundreds of people, vehicles, and businesses in production, transportation, selling and consumption.

- Creating Investigational Link Charts:
  - Begin with investigational targets
  - Look through incident records for known associates
  - Employ heuristics and judgment to select interesting additions
  - Use drawing tools to depict the results

- Link Chart Challenges:
  - Current off-the-shelf software lacks criminal database connectivity
  - Drawing a link chart is a tedious, manual, and time-consuming (expensive) process - tools filter but they do not analyze
  - Assignments are distributed
  - Investigations cross jurisdictional boundaries
  - Some investigational data can’t be shared
  - “Fishing” in available data for potential criminals is discouraged

**Thus link charts are used in only a few cases and the scope of analysis is limited**

**Cross-Jurisdictional Integration Challenges**

- Different schemas and classification taxonomies
- Entity matching
- Privacy and security policies need to protect subpoenaed records, sensitive details, and private personal information in local agency records

**CANs capture associations but omit sensitive details to provide value despite ambiguous data representations**

---

**Importance Flooding**

- Importance Flooding aims to support applications with:
  - Networks of associations between identifiable entities
  - Ambiguously specified relationships
  - Heuristics used by experts to guide analysis
  - Path-based definitions of importance
  - Different queries using different heuristics

- In this link chart the analyst depicted key people involved with methamphetamines and fraud.
- Records came from two jurisdictions.
- Evidence of conspiracy: Leg-breakers, associated with check washers, associated with drug traffickers.

---

**Conclusions, Implications, and Future Directions**

**Conclusion:**
- Spreading activation and path-based importance rules seemed to improve results

**Implications:**
- Computer-supported selection techniques could improve analyst efficiency
- Intelligence analysis tools might benefit from spreading activation modules
- Promising association network analysis methodologies can be developed despite imprecise relations, uncertain entity mappings, inconsistent classification, and heuristic analysis rules

**Future Directions:**
- Can users effectively express heuristics? Will the findings be consistent in a larger study?
- Use importance flooding with concept maps and networks of biomedical relationships