Graph-based Rare Category Analysis: Exploration, Exploitation, and Interpretation

Dawei Zhou

dzhou21@Illinois.edu

University of Illinois at Urbana-Champaign

https://sites.google.com/view/dawei-zhou/home?authuser=0

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

SAIL Lab

Rare Categories in the Real Word…

1. Insider Threat

Self-Similarity

3. Lone Wolf Terrorism

4. Gene Disease

5. Identity Theft

6. Emerging Trend

- Workshop on Challenges and Opportunities for AI in Financial Services: the Impact of Fairness, Explainability, Accuracy, and Privacy, Co-located with NIPS 2018.
- Workshop on Anomaly Detection in Finance, Co-located with KDD 2017.
- Spaaij, R.. "The enigma of lone wolf terrorism: An assessment." *Studies in Conflict & Terrorism* 33.9 (2010): 854-870.
- Ngai, E. W., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. Decision support systems, 50(3), 559-569.

Malicious Insiders

Malicious Insider

Definition

Current or former employee or contractor who

- ***intentionally*** exceeded or misused an authorized level of network, system or data access in such a way that,
- ***affected*** the security of the organizations' data, systems, or daily business operations

[Cappelli et al, *RSA*2008]

Hard to identify → Insiders often camouflage themselves…

DARPA-BAA-10-84, DARPA-BAA-11-04, DARPA-BAA-11-64, IARPA-BAA-12-01, …

Malicious Insiders: Fort Hood Shooting

Post Inv.

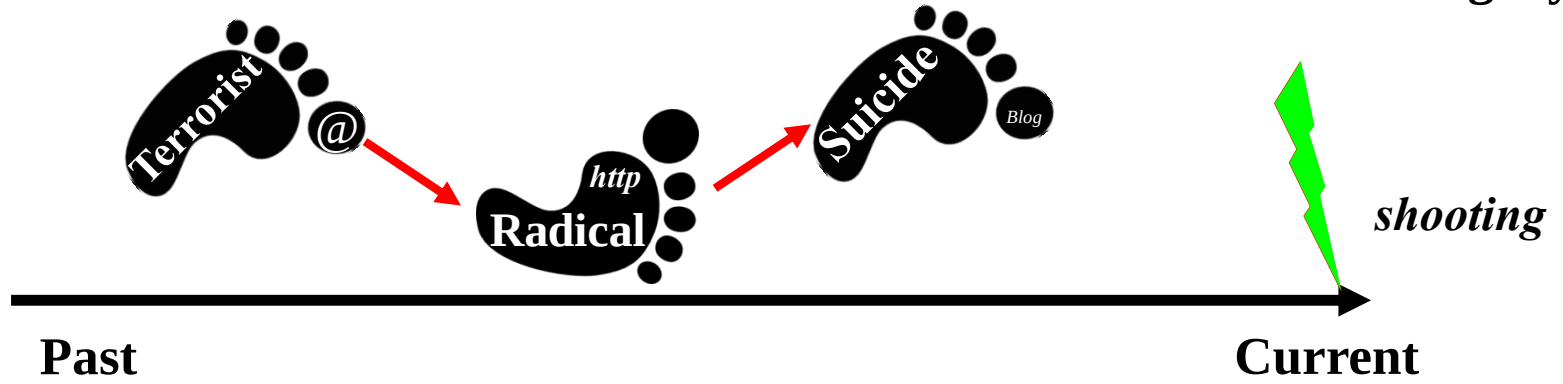
?

‘Obvious’ Clues:

…

Q: Can we reverse it?

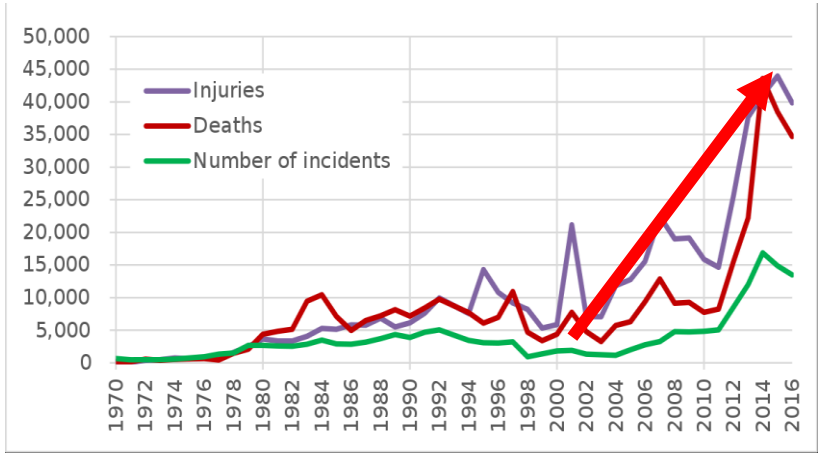
***G1*. Find individual traces** ***G2*. Connect the dots** ***G3*. Prevent the tragedy**



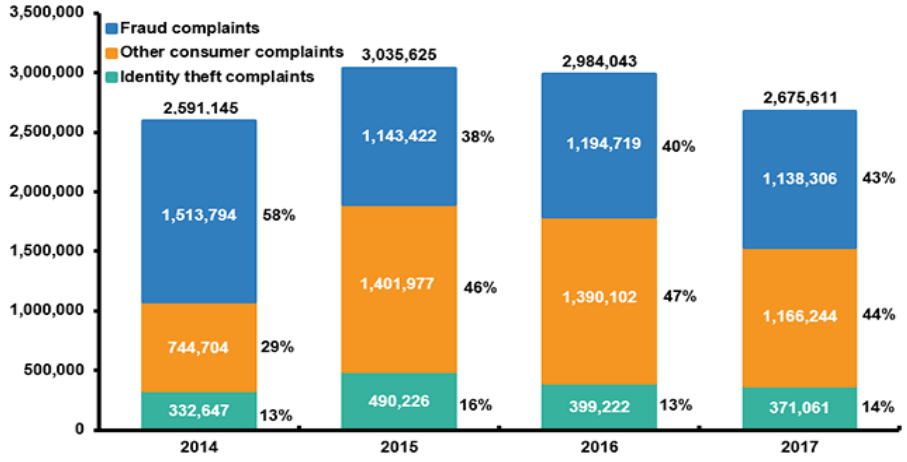
How to find more insiders of the same type?

Why Do We Care?

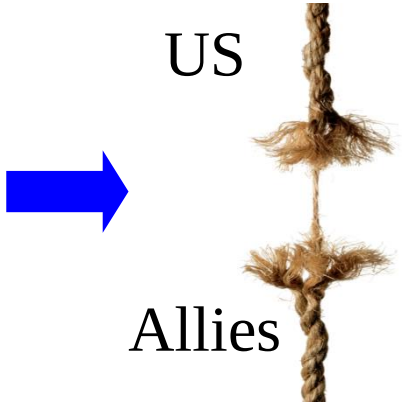
Terrorist Events Worldwide



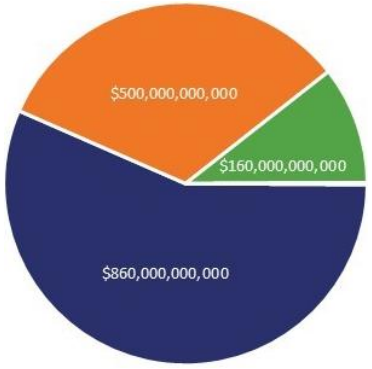
Financial Fraud [Federal Trade Commission, 2017]



PR Crisis



Cybercrime [M., McGuire, RSA2018]

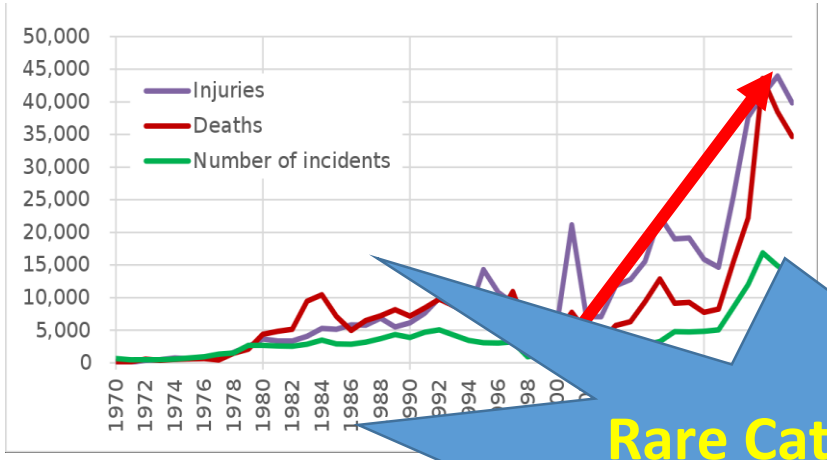


- Illegal online markets
- Trade secret, IP theft
- Data Trading

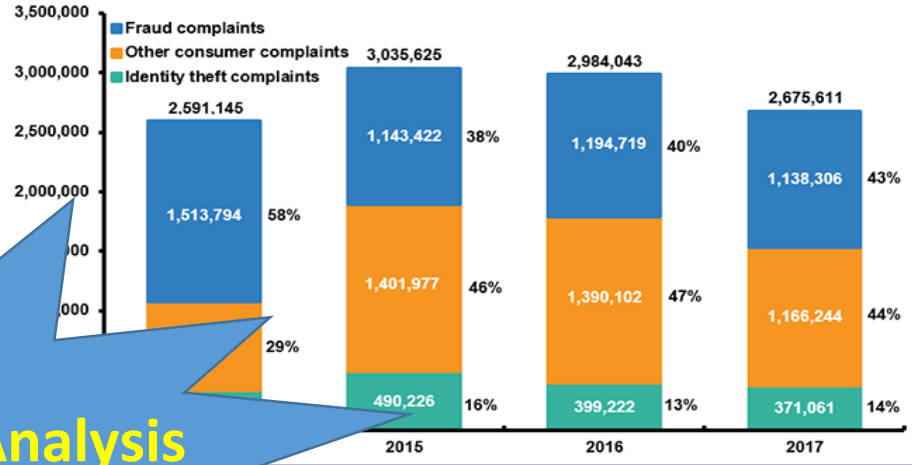
Cost over 1.5 trillion in 2018

Why Do We Care?

- Terrorist Incidents Worldwide



- Financial Fraud [Federal Trade Commission, 2017]



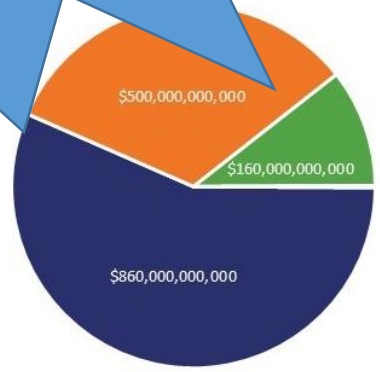
Rare Category Analysis (RCA)

- PR Crisis

US

Allies

- [..., McGuire, RSA2018]



- Illegal online markets
- Trade secret, IP theft
- Data Trading

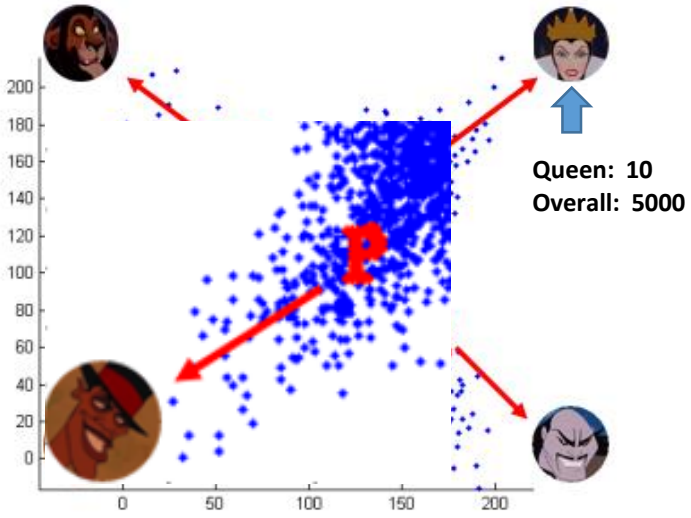
Cost over 1.5 trillion in 2018

Rare Category Analysis

- Problem Definition

  Rare category analysis (**RCA**) refers to the problem of detecting, characterizing, representing and interpreting rare examples from *underrepresented* minority classes in *an imbalanced data set*.

- An Illustrative Example



- ***Imbalanced*** data set
- Minority classes are ***not separable*** from majority class

Comparison with Imbalanced Classification

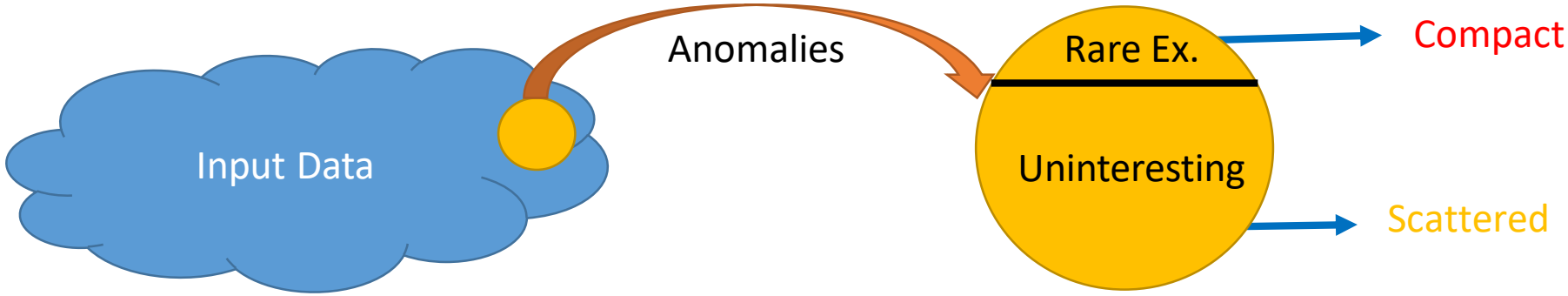
- Imbalanced Classification
  - Problem definition
    - Labeled examples from all the classes
    - Models for imbalanced classification focus on overall accuracy of all classes
  - Methodology: [Kubat & Matwin, ICML1997]; [Chawla et al, JAIR2002]; [Wu & Chang, ICML2003]; [Huang et al., CVPR2016]; …
- Rare Category Analysis
  - Problem definition
    - None/one/few-shot labeled example from the rare categories
    - Models for rare category analysis put heavy emphasis on learning minority classes with a good performance.
  - Methodology: [Fine & Mansour, COLT2006]; [Dasgupta & Hsu, ICML2008]; [Vatturi & Wong, KDD2009]; [Zhou et al., KDD2018]; …

- Chawla, Nitesh V. "Data mining for imbalanced datasets: An overview." Data mining and knowledge discovery handbook. Springer, Boston, MA, 2009. 875-886.
- Workshop on Learning with Imbalanced Domains: Theory and Applications, Co-located with ECML/PKDD 2018.
- Workshop on Active Learning for Decision-Making from Imbalanced Observational Data, Co-located with ICML 2019.

Comparison with Outlier/Anomaly Detection

- ***Outliers/Anomalies***
  - Patterns in data that do not conform to expected behavior [Chandola, Banerjee & Kumar, 2009]
  - Typically separable from normal examples and scattered in the feature space
- ***Rare Categories***
  - Compact in the feature space, may overlap with the majority class [He, 2010]
- ***Outliers/Anomalies vs. Rare Categories***

  "Most of the objects (i.e., majority classes) are well explained by current theories and …. remainder are anomalies, but 99% of these anomalies are uninteresting (i.e., boring anomalies), and only 1% of them (i.e., rare categories) are useful" [Pelleg & Moore, 2004]



- V. Chandola, A. Banerjee, V. Kumar: Anomaly Detection: A Survey. ACM Comput. Surv. 41(3): 15:1-15:58 (2009).
- J. He (2010). Rare category analysis (Doctoral dissertation, Carnegie Mellon University, School of Computer Science, Machine Learning Department).
- D. Pelleg, A. W. Moore: Active Learning for Anomaly and Rare-Category Detection. NIPS 2004.

Example: Sloan Digital Sky

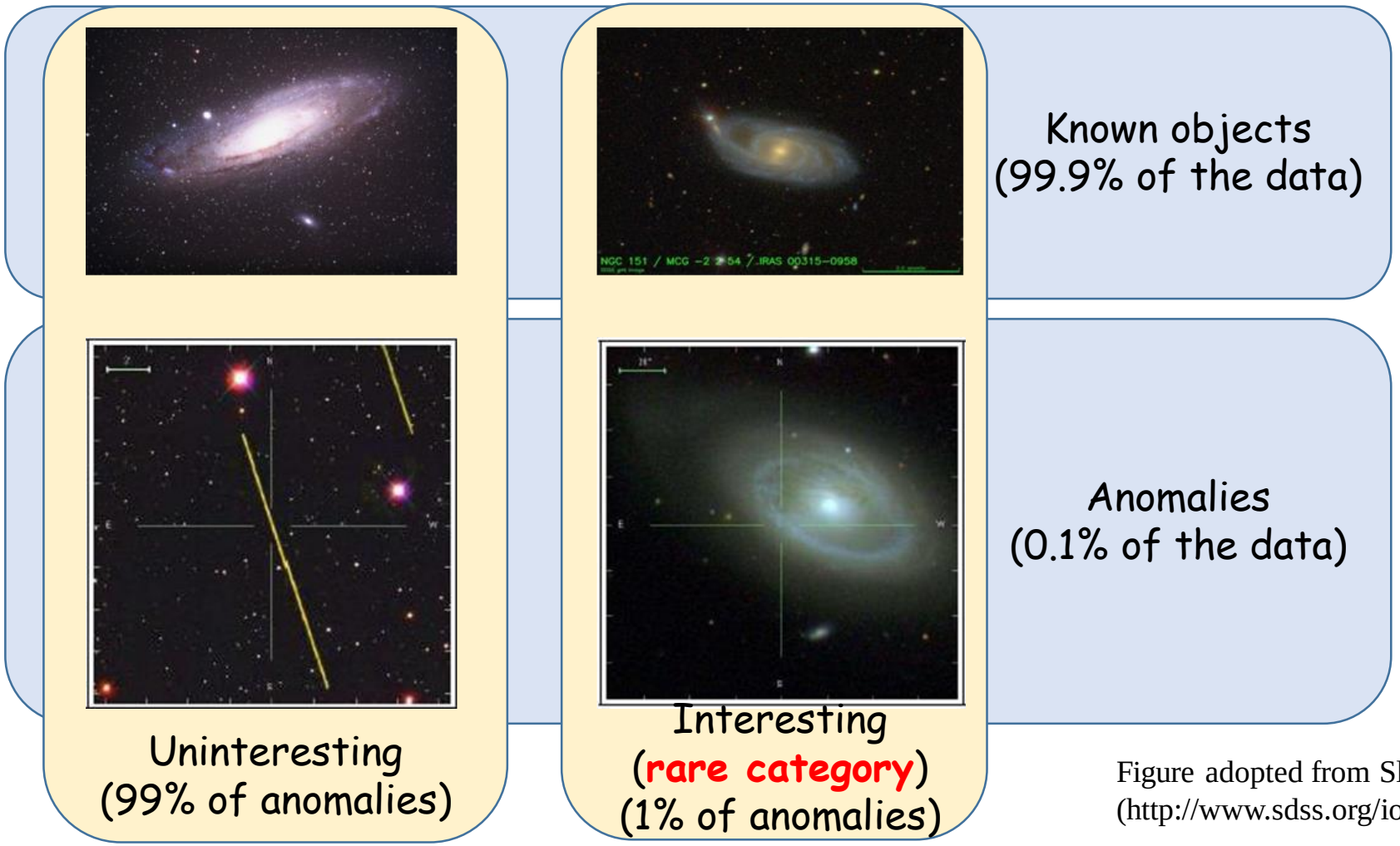


Figure adopted from Sloan Digital Sky Survey
(http://www.sdss.org/iotw/archive.html)

Key Challenge #1: **Rarity**

Highly-skewed distribution

- *0.1%* of any given population of users is malicious
- *99.9%* are normal users

[DARPA-BAA-11-04]

Non-separable nature

- Malicious users often camouflage their abnormal behaviors
- Malicious users try to bypass the detection systems

Commonality of insiders

- Detecting insiders of each type
- Identifying relevant features
- Characterizing each type of insiders

QI: How to detect and characterize insiders of each type, using as little cost as possible?

Key Challenge #2: **Dynamics**

Low-footprint of insiders

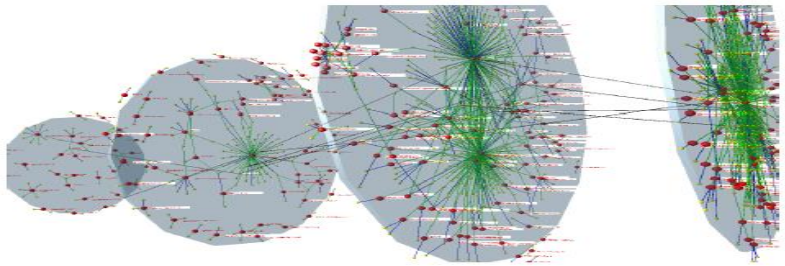
- *0.1%* of any given population of users is malicious
- *20%* of malicious insiders are active on any given day

[DARPA-BAA-11-04]

Costly access to oracle

- [Fort Hood Shooting]
  - ✓ *65K* personnel at Fort Hood
    → 4.7B emails in 2 years
  - ✓ ~*60* initial activity reviews per day per operator

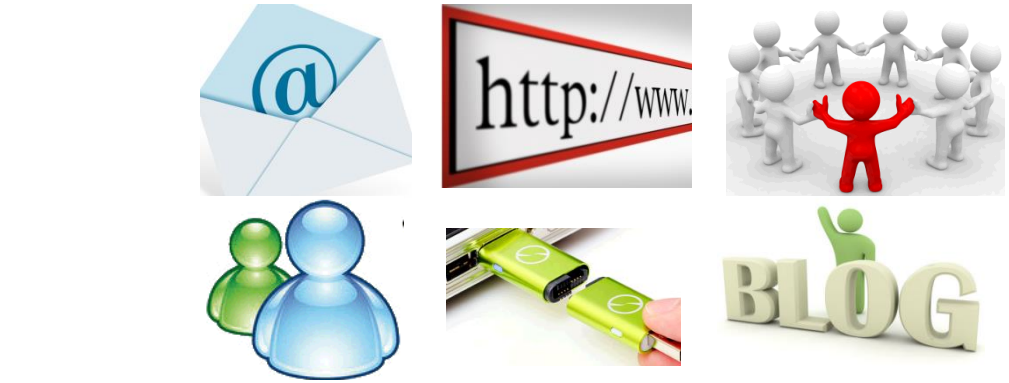
Fine-grained dynamics



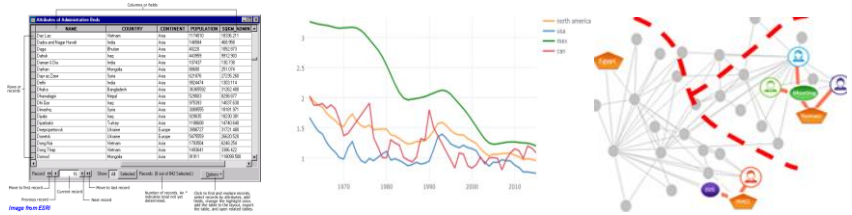
Q2: How to predict and track insiders of each type over time at the proper granularity?

Key Challenge #3: Heterogeneity

Various information sources



Various data types



Attributed Data

Sequential Data

Network Data

Various types of insiders

blackmailed

psychological vulnerability

greedy

Q3: How to reliably analyze malicious insiders using heterogeneous data?

Overview of Complex Rare Category Analysis

Applications

Malicious Insiders

Emerging Trend

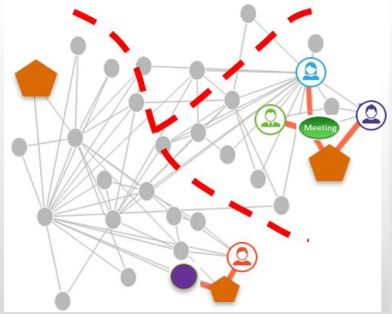
Spam Email

Rare Disease

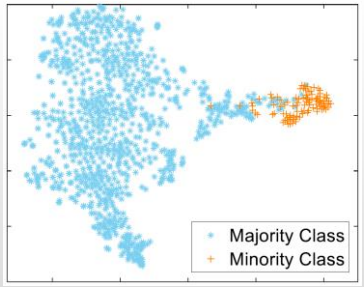
⋮ ⋮

Part I: Exploration

Part II: Exploitation



Part III: Interpretation



Embedding

User Interface

End User

Feedback

Annotation

Diagnosis

Interaction

Part I:
Rare Category Exploration

Overview of Complex Rare Category Analysis

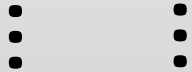
Applications

Malicious Insiders

Emerging Trend

Spam Email

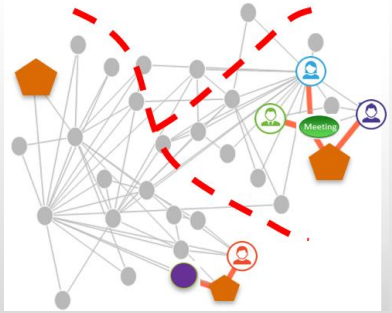
Rare Disease



Part I: Exploration

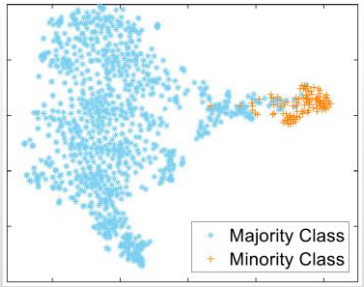
Identified Rare Examples

Part II: Exploitation

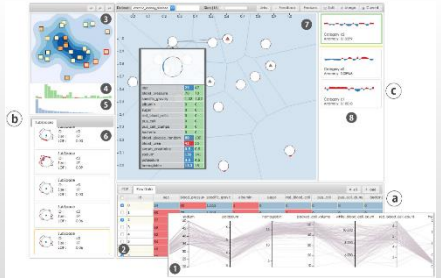


Identified Rare Examples

Part III: Interpretation



Embedding



User Interface

End User

Feedback

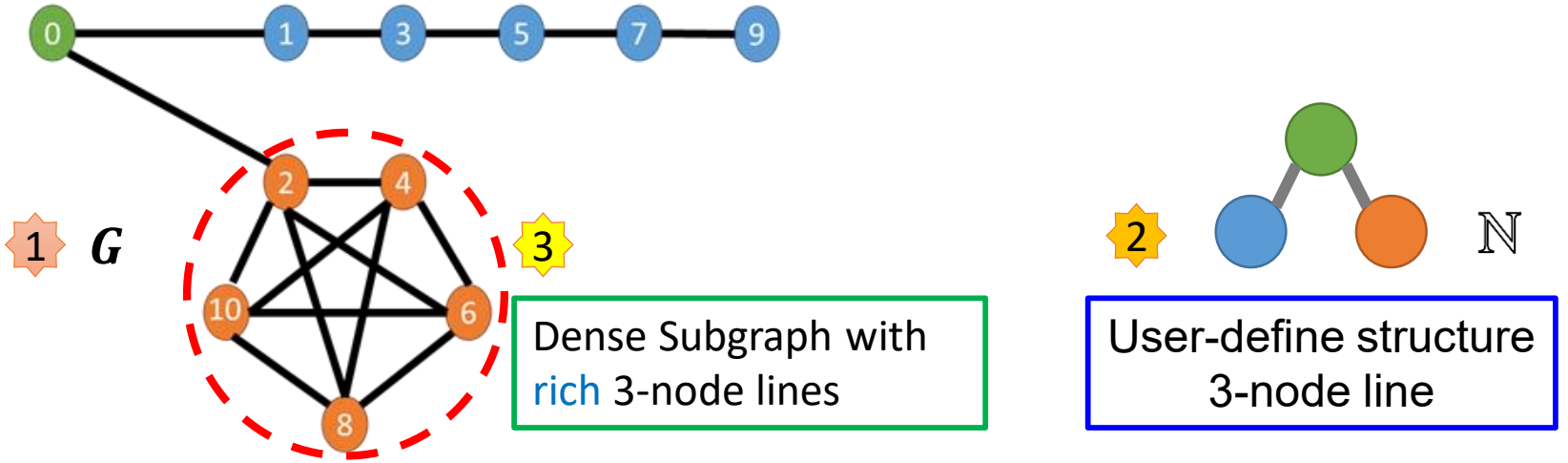
Annotation

Diagnosis

Interaction

Rare Category with High-Order Structures

- Specific Problem Definition
  - Given: Graph $G=(V, E)$, user-defined structure $\mathbb{N}$.
  - Find: A structure-rich dense subgraph that largely preserves the user-defined structures.

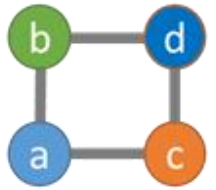


• D. Zhou, S. Zhang, M. Y. Yildirim, S. Alcorn, H. Tong, H. Davulcu, J. He: A Local Algorithm for Structure-Preserving Graph Cut. KDD 2017: 655-664

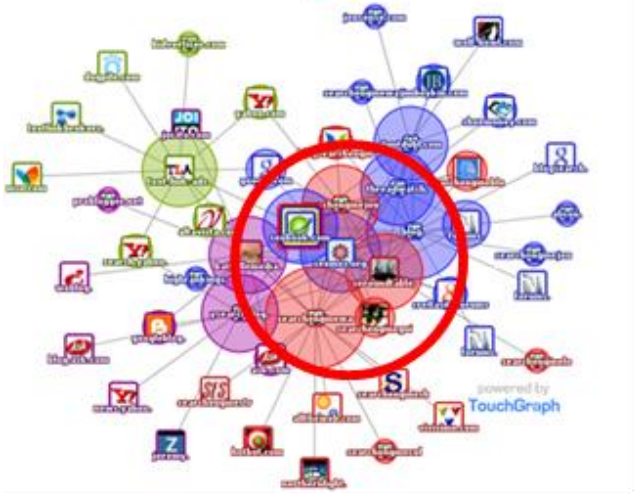
High-Order Structures in Real Applications

Stars

Triangles

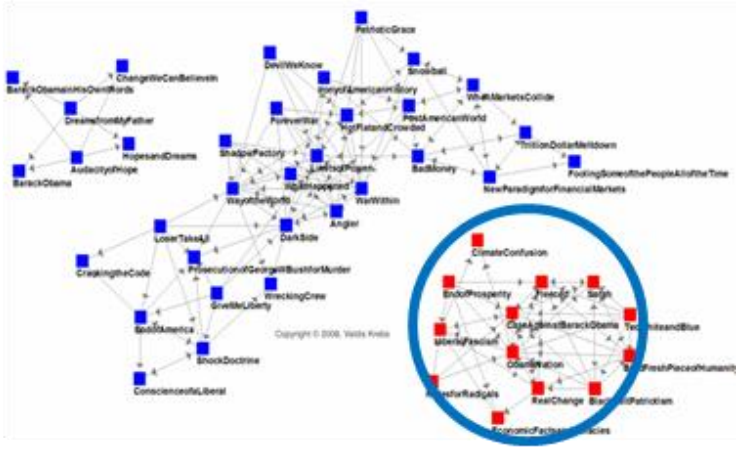


Loops



Internet

Social Network



Online Transaction

• D. Zhou, S. Zhang, M. Y. Yildirim, S. Alcorn, H. Tong, H. Davulcu, J. He: A Local Algorithm for Structure-Preserving Graph Cut. KDD 2017: 655-664

High Order Conductance

- **Definition**

  For any cluster C in graph G and the k^{th}-order structure $\mathbb{N}$, the k^{th}-order conductance $\Phi(C, \mathbb{N})$ is defined as

$$\Phi(C, N) = \frac{cut(C, \mathbb{N})}{\min\{\mu(C, \mathbb{N}), \mu(\bar{C}, \mathbb{N})\}}$$

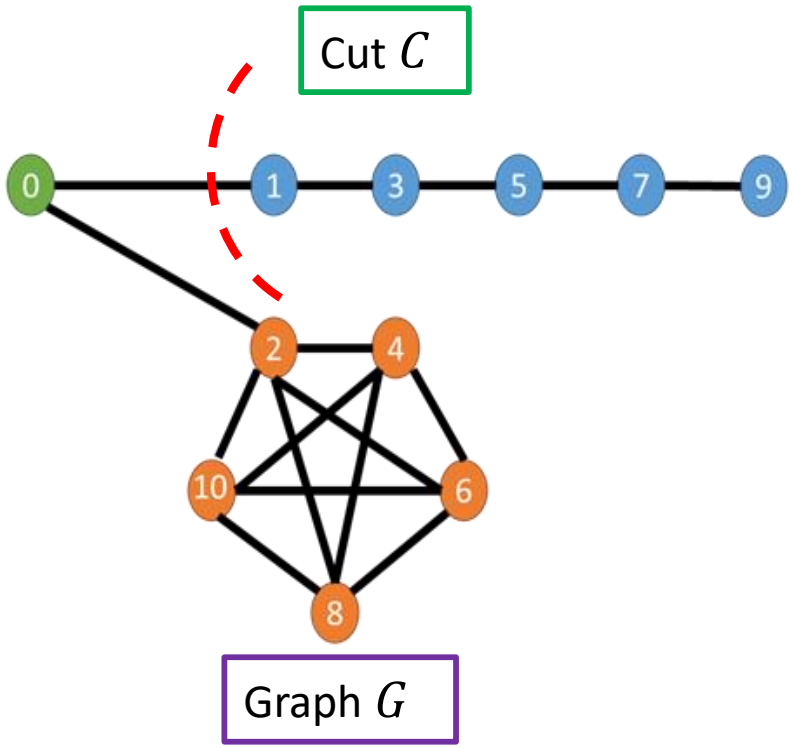
The number of network structures $\mathbb{N}$ broken due to the partition of G into C and $\bar{C}$

The number of network structures $\mathbb{N}$ in C.

The number of network structures $\mathbb{N}$ in $\bar{C}$.

• D. Zhou, S. Zhang, M. Y. Yildirim, S. Alcorn, H. Tong, H. Davulcu, J. He: A Local Algorithm for Structure-Preserving Graph Cut. KDD 2017: 655-664

Conductance vs. High Order Conductance



$$\Phi(C,\mathbb{N}) = \frac{cut(C,\mathbb{N})}{\min\{\mu(C,\mathbb{N}),\mu(\bar{C},\mathbb{N})\}}$$

- **Conductance**

$$\Phi(C,\mathbb{N}) = \frac{1}{\min\{4,11\}} = 1/4$$

- **High-order conductance**

$$\Phi(C,\mathbb{N}) = \frac{2}{\min\{3,34\}} = \frac{2}{3}$$

• D. Zhou, S. Zhang, M. Y. Yildirim, S. Alcorn, H. Tong, H. Davulcu, J. He: A Local Algorithm for Structure-Preserving Graph Cut. KDD 2017: 655-664

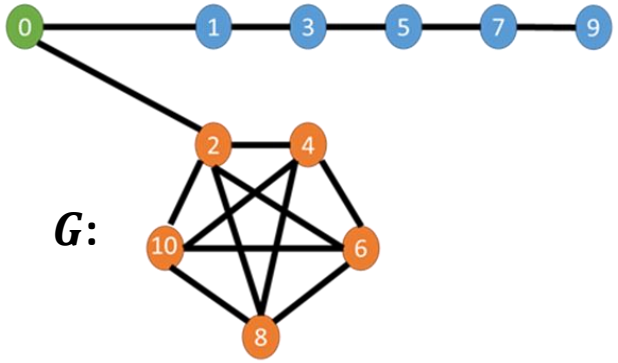
HOSPLOC Algorithm

- **Construct Adjacency Tensor**

  Given a graph $G = (V, E)$, the k^{th}-order network structure $\mathbb{N}$ on G could be represented in a k-dimensional adjacency tensor T as follows

$$T(i_1, i_2, \ldots, i_k) = \begin{cases} 1 & \{i_1, i_2, \ldots, i_k\} \subseteq V \text{ and form } \mathbb{N}. \\ 0 & \text{Otherwise.} \end{cases}$$

Example:



***G*:**

***N*:**

For the set of nodes {2, 4, 1}

$T(1,4,2) = 0$

For the set of nodes {6, 8, 10}

$T(10,8,6) = 1$

• D. Zhou, S. Zhang, M. Y. Yildirim, S. Alcorn, H. Tong, H. Davulcu, J. He: A Local Algorithm for Structure-Preserving Graph Cut. KDD 2017: 655-664

HOSPLOC Algorithm

- **Compute Transition Tensor**

  Given a graph $G = (V, E)$ and the adjacency tensor T for the k^{th}-order network structure $\mathbb{N}$, the corresponding transition tensor P could be computed as

$$P(i_1, i_2, \ldots, i_k) = \frac{T(i_1, i_2, \ldots, i_k)}{\sum_{i_1=1}^{n} T(i_1, i_2, \ldots, i_k)}$$

Example:

• Compute Transition Tensor
Given a graph $G = (V, E)$ and the adjacency tensor T for the k^{th}-order network structure N, the corresponding transition tensor P could be computed as

$\boldsymbol{G}$:

$\mathbb{N}$:

For the set of nodes {2, 4, 1}

$P(1,4,2) = 0$

For the set of nodes {6, 8, 10}

$P(10,8,6) = 1/3$

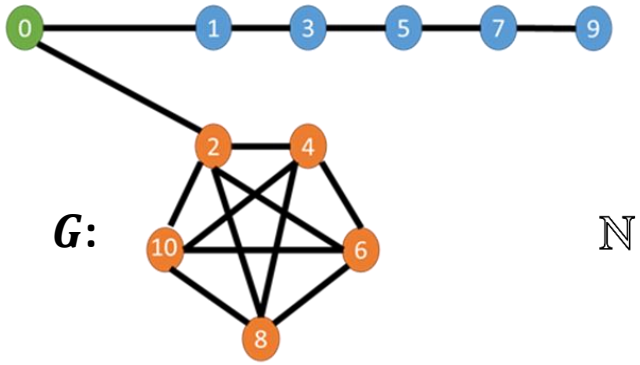
• D. Zhou, S. Zhang, M. Y. Yildirim, S. Alcorn, H. Tong, H. Davulcu, J. He: A Local Algorithm for Structure-Preserving Graph Cut. KDD 2017: 655-664

HOSPLOC Algorithm

- k^{th}-order Markov Chain Interpretation
  - Each vertex represents an individual state
  - Depends on the last k–1 states (nodes)

$$\mathrm{P}(i_1, \ldots, i_k) = \Pr(S_{t+1} = i_1 | S_t = i_2, \ldots, S_{t-k+2} = i_k)$$

Example:



$\boldsymbol{G}$:

$\mathbb{N}$:

For the set of nodes {2, 4, 1}

$P(1,4,2) = P\,r(v_1 | v_4, v_2) = 0$

For the set of nodes {6, 8, 10}

$P(10,8,6) = P\,r(v_{10} | v_6, v_8) = 1/3$

• D. Zhou, S. Zhang, M. Y. Yildirim, S. Alcorn, H. Tong, H. Davulcu, J. He: A Local Algorithm for Structure-Preserving Graph Cut. KDD 2017: 655-664

HOSPLOC Algorithm

- High-Order Random Walk
  - Using the "rank-1" approximation [Li and NG, 2013], the high-order random walks can be formulated as

$$q^{(t)} = Pq^{(t-1)} \dots q^{(t-k+1)}$$

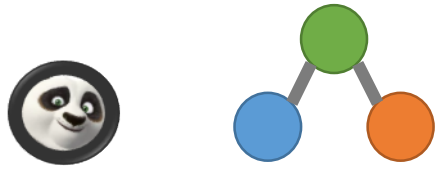
- Vector Based Graph Cut
  - Locally conduct high-order random walks to explore $\mathbb{N}$.
  - Compute the permutation π of the returned HRW distribution q such that:

$$\frac{q(\pi(1))}{d(\pi(1))} \leq \frac{q(\pi(2))}{d(\pi(2))} \leq \dots \leq \frac{q(\pi(n))}{d(\pi(n))}.$$

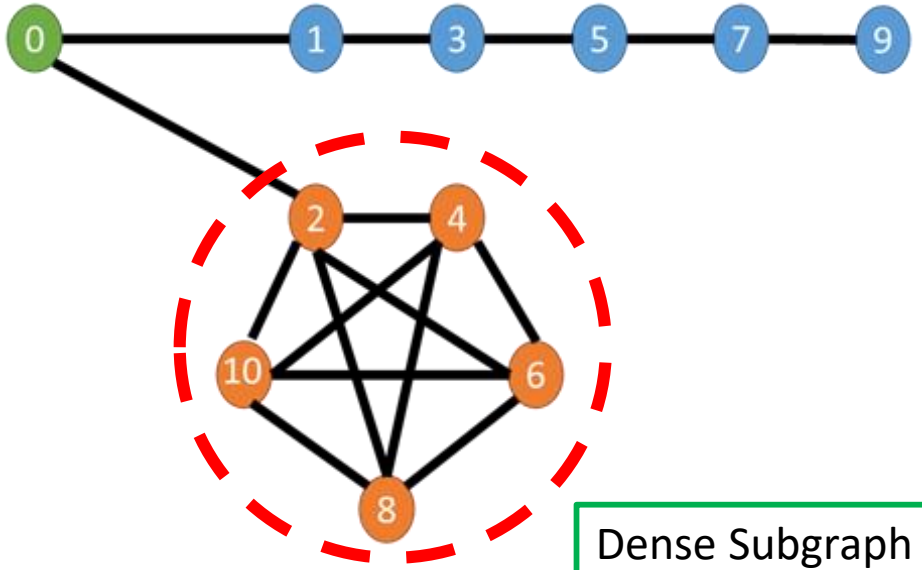
  - Iteratively check the potential cuts $C_1, C_2, \dots, C_{n-1}$, where $C_i = \{\pi(1), \dots, \pi(i)\}$.

• D. Zhou, S. Zhang, M. Y. Yildirim, S. Alcorn, H. Tong, H. Davulcu, J. He: A Local Algorithm for Structure-Preserving Graph Cut. KDD 2017: 655-664

HOSPLOC Algorithm



3-node line



Dense Subgraph with rich 3-node line

Algorithm 1 High-Order Structure-Preserved Local Clustering (*HOSPLOC*)

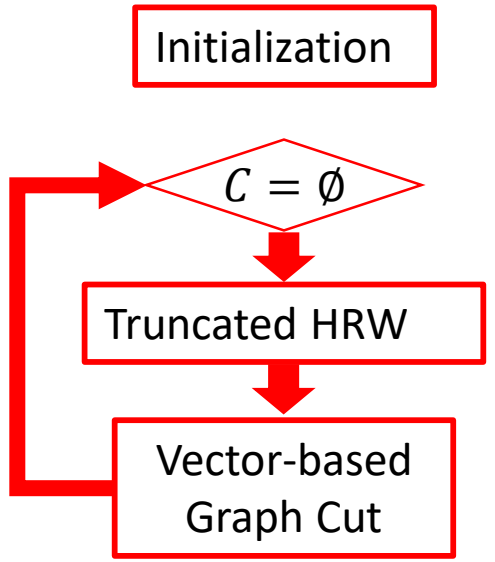
Input:

(1) Transition tensor $\mathbf{P}$ and transition matrix M,
(2) Initial vertex v,
(3) Conductance upper bound ϕ,
(4) Maximum iteration number $t_{\max}$,
(5) Parameters b, c_1, ξ.

Output:

Local cluster C;

1: Construct the unfolding matrix $\bar{P}$ of the transition tensor $\mathbf{P}$.
2: Compute constant ϵ based on Eq. 3.
3: Set initial distribution vectors $q^{(t)} = M^{(t-1)}\chi_v$, where $t = 1, \ldots, k-1$.
4: Compute truncated initial local distribution vectors $r^{(t)} = [q^{(t)}]_\epsilon$, $t = 1, \ldots, k-1$.
5: **for** $t = k : t_{\max}$ **do**
6: Update distribution vector $q^{(t)} = P(r^{(t-1)} \otimes \ldots \otimes r^{(t-k+1)})$.
7: Update truncated distribution vectors $r^{(t)} = [q^{(t)}]_\epsilon$.
8: **if** there exists a j such that: **then**
9: (a)$\Phi(S_j(q^{(t)})) <= \phi$,
10: (b)$2^b <= \lambda_j(q^{(t)})$,
11: (c)$I_x(q^{(t)}, 2^b) >= \frac{\xi}{c_1(l+2)2^b}$.
12: return $C = S_j(q^{(t)})$ and quit.
13: **else**
14: Return $C = \emptyset$.
15: **end if**
16: **end for**



• D. Zhou, S. Zhang, M. Y. Yildirim, S. Alcorn, H. Tong, H. Davulcu, J. He: A Local Algorithm for Structure-Preserving Graph Cut. KDD 2017: 655-664

Experimental Results

- Datasets

Category	Network	Type	Nodes	Edges
Citation	Author	Undirected	61,843	402,074
	Paper	Undirected	62,602	10,904
Infrastructure	Airline	Undirected	2,833	15,204
	Oregon	Undirected	7,352	15,665
	Power	Undirected	4,941	13,188
Social	Epinion	Undirected	75,879	508,837
Review	Rating	Bipartite	8,724	90,962
Financial	PII	Multipartite	375	519

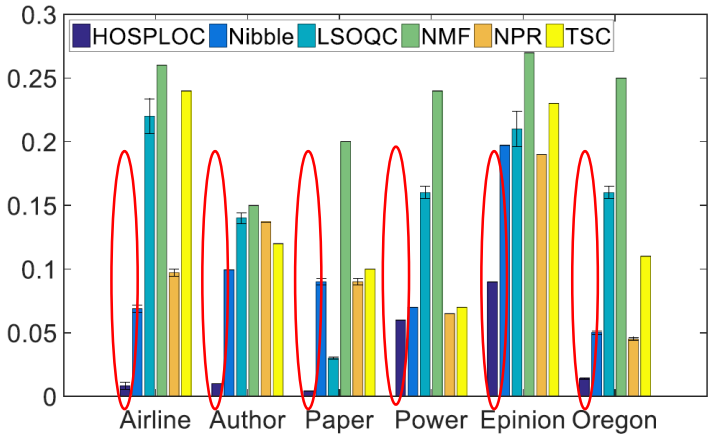
• D. Zhou, S. Zhang, M. Y. Yildirim, S. Alcorn, H. Tong, H. Davulcu, J. He: A Local Algorithm for Structure-Preserving Graph Cut. KDD 2017: 655-664

Effectiveness Analysis

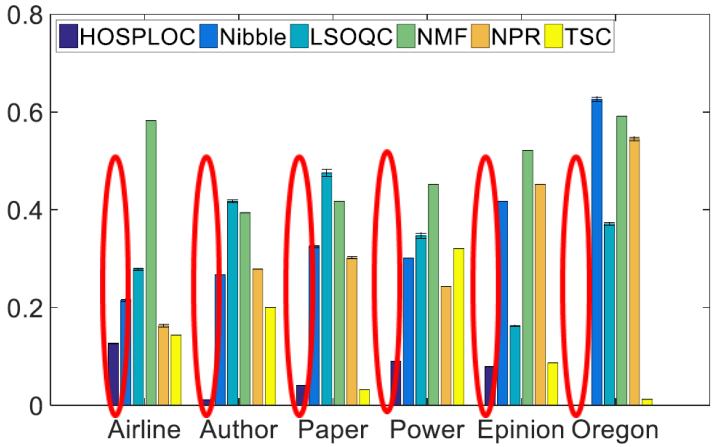
Worse

Better

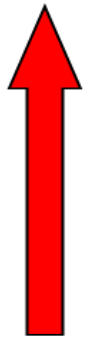
$(a)\ \Phi(C, edge)$



$(b)\ \Phi(C, triangle)$

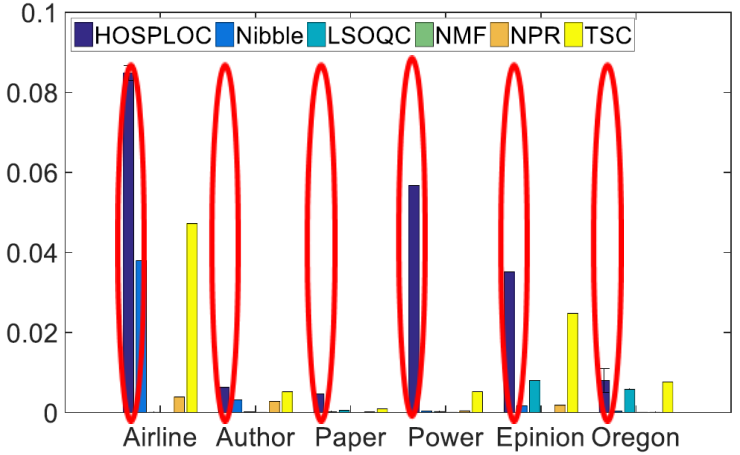


Better



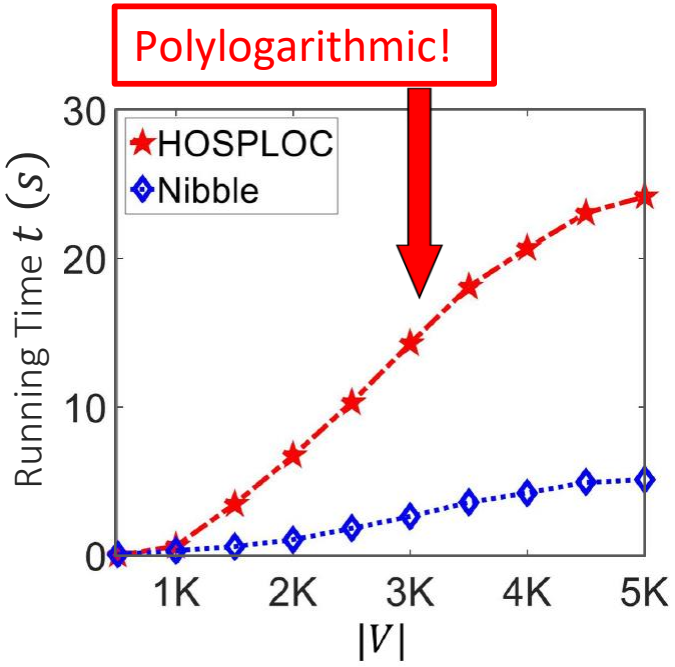
Worse

(c) Triangle Desnsity

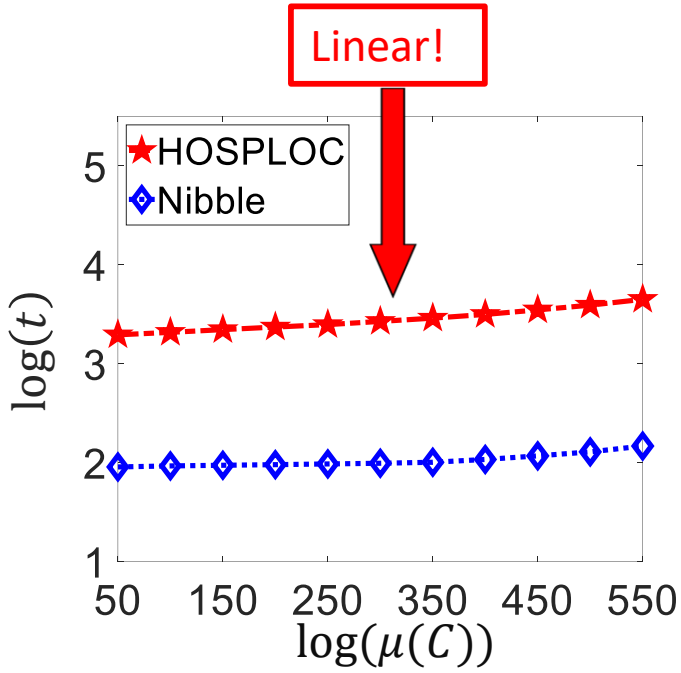


• D. Zhou, S. Zhang, M. Y. Yildirim, S. Alcorn, H. Tong, H. Davulcu, J. He: A Local Algorithm for Structure-Preserving Graph Cut. KDD 2017: 655-664

Scalability Analysis



(a) The number of vertices v.s running time t



(b) The lower bound of $\log(\mu(C))$ v.s $\log(t)$

* Nibble = A special case of HOPLOC

Part III:
Rare Category Representation

Overview of Complex Rare Category Analysis

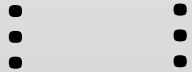
Applications

Malicious Insiders

Emerging Trend

Spam Email

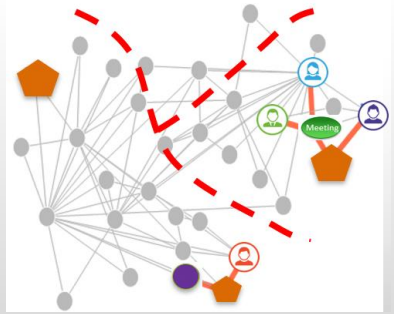
Rare Disease



Part I: Exploration

Identified Rare Examples

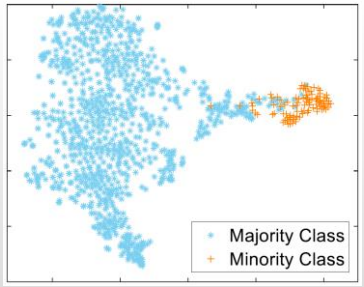
Part II: Exploitation



Identified Rare Examples

Identified Rare Categories

Part III: Interpretation



Embedding

User Interface

End User

Feedback

Annotation

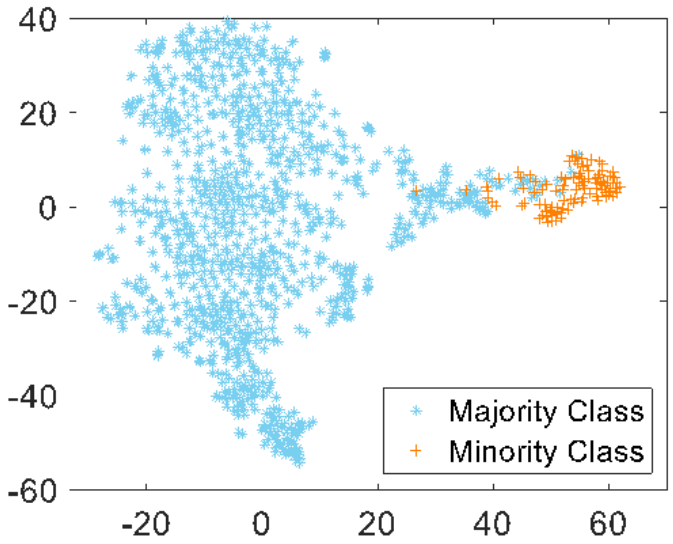
Diagnosis

Interaction

Rare Category Oriented Network Embedding

- Specific Problem Definition
  - Given: An attributed network $G = (V, E, X)$, one-shot or few-shot labeled rare examples $\mathrm{L} = \{x_1, \ldots, x_L\}$, and the desired embedding dimension d.
  - Find: A rare category oriented network embedding $E \in R^{n \times d}$, and a list of predicted rare category examples.

- Challenges
  - C1: Rarity
  - C2: Sparsity of labeled examples
  - C3: Non-separability of rare example from the majority classes



Network Layout of Pubmed

• D. Zhou, J. He, H. Yang, W. Fan. SPARC: Self-Paced Network Representation for Few-Shot Rare Category Characterization, ACM: SIGKDD Conference on Knowledge Discovery and Data Mining (KDD-2018), August 2018.

Rare Category Characterization (RCC)

- Cost-sensitive Learning (Address C1)
  - Focus on rare category examples
  - Learn from the highly-skewed distribution

$$L_s = \sum_{l=1}^{L} c_{y_i,\hat{y}_i} \log \Pr(\hat{y}_i = 1 - y_i | x_i, e_i)$$

Penalize the error of classify minority class examples into majority classes

- Self-paced Learning (Address C2)
  - Start from a handful labeled example
  - Gradually explore more via label propagation

$$L_{RCC} = L_s - \sum_{i=1}^{L+U} v_i^{(1)} \log \Pr(\hat{y}_i = 1 | x_i, e_i) - \sum_{i=1}^{L+U} \lambda^{(1)} v_i^{(1)}$$

Self-paced Vector

Self-paced Regularizers

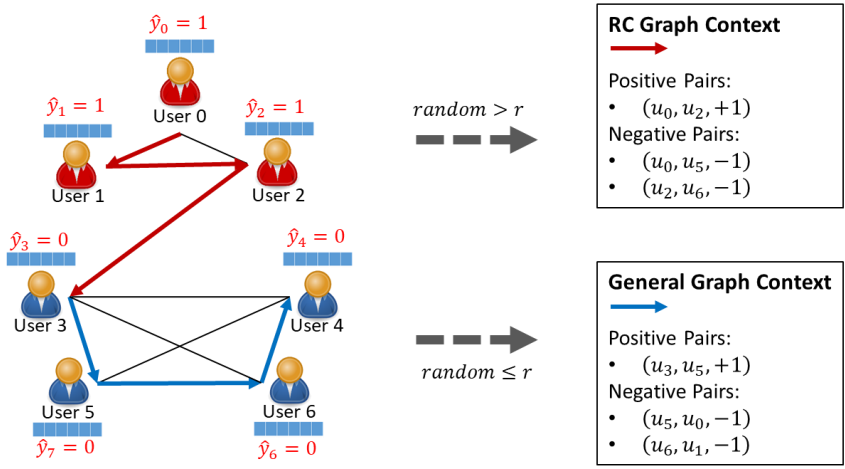
- D. Zhou, J. He, H. Yang, W. Fan. SPARC: Self-Paced Network Representation for Few-Shot Rare Category Characterization, ACM: SIGKDD Conference on Knowledge Discovery and Data Mining (KDD-2018), August 2018.

Rare Category Embedding (RCE)

- Minimize the cross-entropy loss of predicting context pairs (i, c)

$$\min -E_{(i,c,\gamma)} \log \sigma(\gamma \boldsymbol{\theta}_c^T \boldsymbol{e}_i)$$

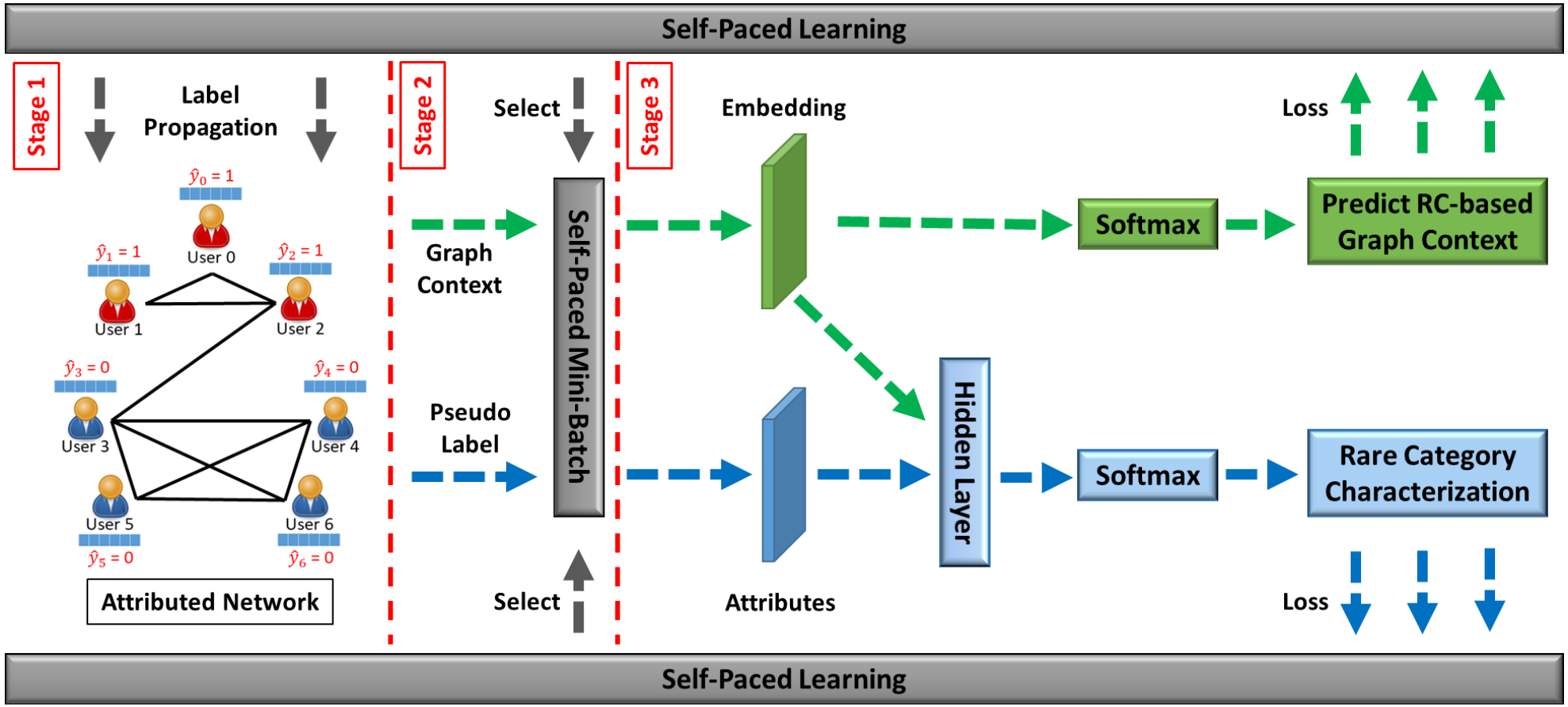
- Rare category oriented context sampling (Address C3)
  - Indicator vector I from RCC.
  - With probability p, extract general network context.
  - With probability $1-p$, extract rare category oriented network context starting from the non-zero elements in I.



- D. Zhou, J. He, H. Yang, W. Fan. SPARC: Self-Paced Network Representation for Few-Shot Rare Category Characterization, ACM: SIGKDD Conference on Knowledge Discovery and Data Mining (KDD-2018), August 2018.

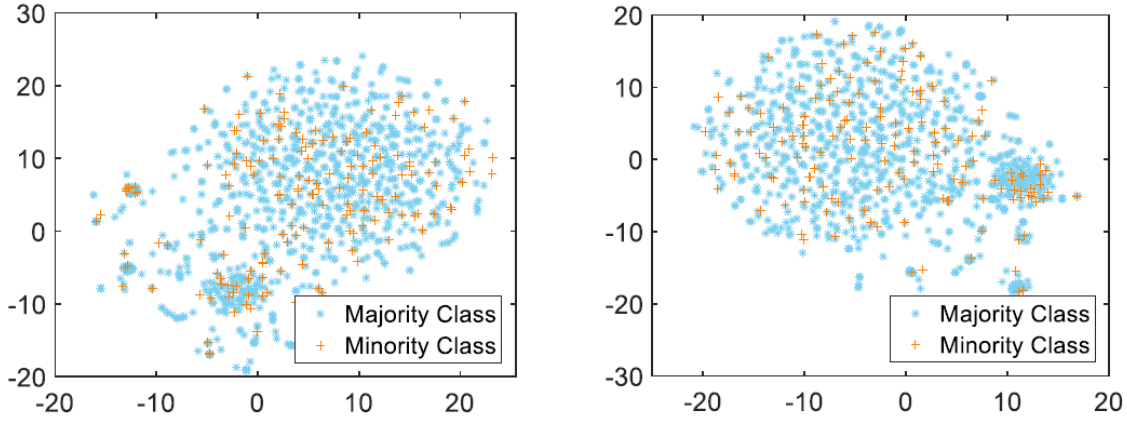
SPARC Algorithm

- Proposed Framework



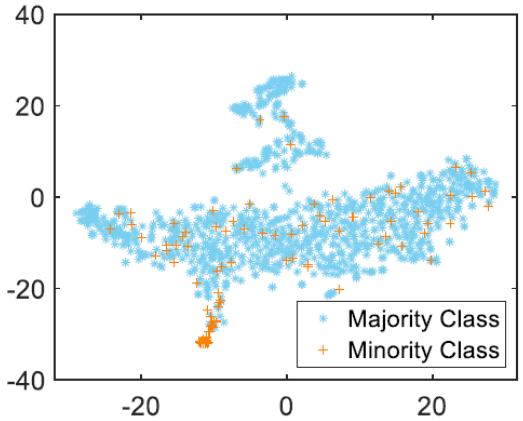
• D. Zhou, J. He, H. Yang, W. Fan. SPARC: Self-Paced Network Representation for Few-Shot Rare Category Characterization, ACM: SIGKDD Conference on Knowledge Discovery and Data Mining (KDD-2018), August 2018.

2-D t-SNE Visualization

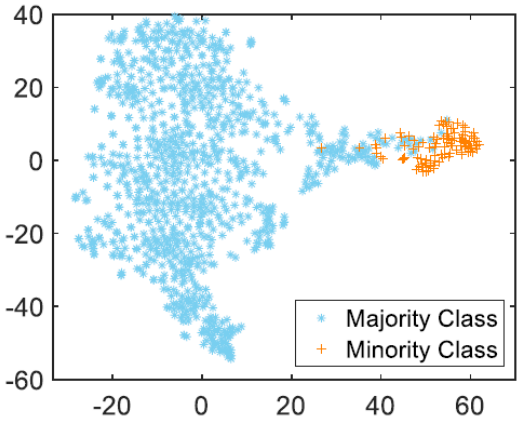


(a) DeepWalk

(b) LINE



(c) PLANETOID



(d) SPARC

- D. Zhou, J. He, H. Yang, W. Fan. SPARC: Self-Paced Network Representation for Few-Shot Rare Category Characterization, ACM: SIGKDD Conference on Knowledge Discovery and Data Mining (KDD-2018), August 2018.

Experimental Results

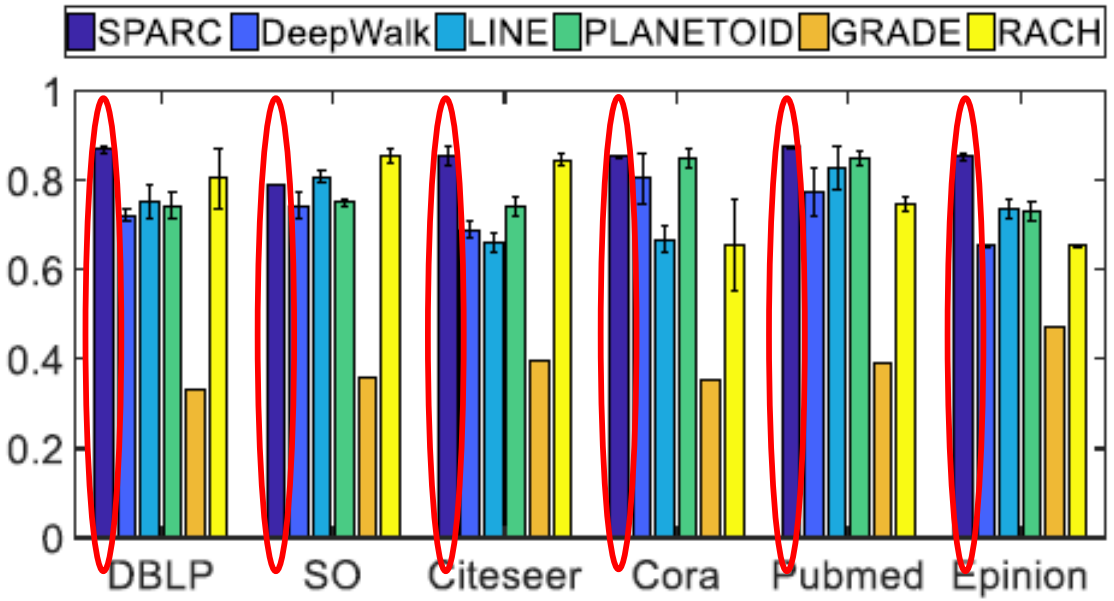
- Datasets

Category	Network	Classes	Smallest Class	Nodes	Edges
Collaboration	DBLP	20	1.91%	2,309	7,913
	SO	2	1.29%	3,262	19,926
NLP	Citeseer	6	3.42%	3,327	4,732
	Cora	7	1.14%	2,708	5,429
	Pubmed	3	4.05%	19,717	44,318
Social	Epinion	19	1.38%	75,879	508,837

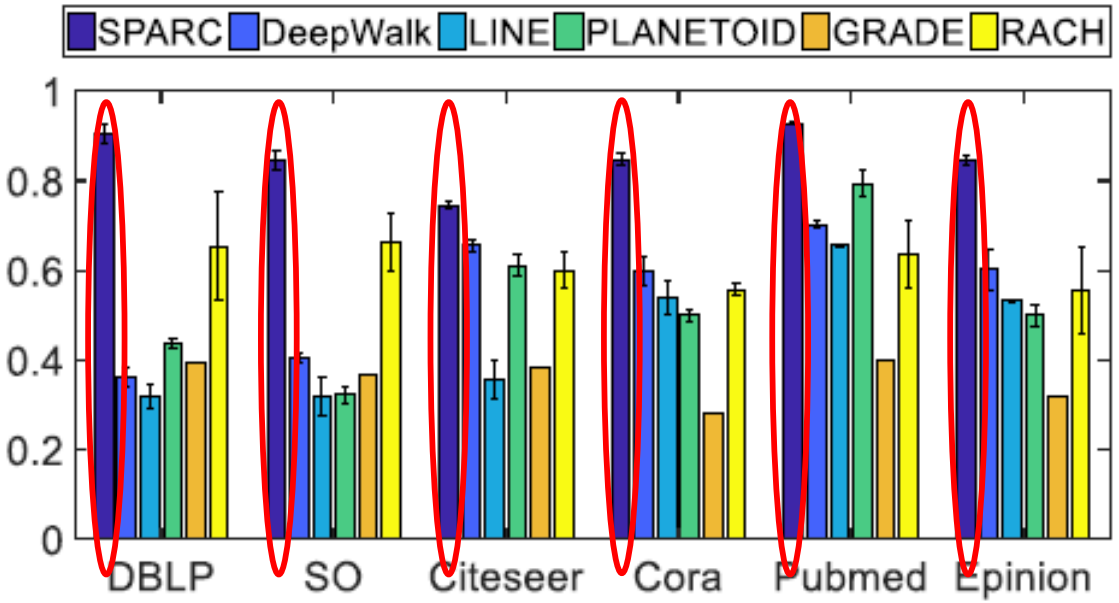
- D. Zhou, J. He, H. Yang, W. Fan. SPARC: Self-Paced Network Representation for Few-Shot Rare Category Characterization, ACM: SIGKDD Conference on Knowledge Discovery and Data Mining (KDD-2018), August 2018.

Experimental Results

- Effectiveness analysis with 5% labeled minority class examples.



(a) Accuracy



(b) Recall

- D. Zhou, J. He, H. Yang, W. Fan. SPARC: Self-Paced Network Representation for Few-Shot Rare Category Characterization, ACM: SIGKDD Conference on Knowledge Discovery and Data Mining (KDD-2018), August 2018.

Part IV:
Rare Category Analysis Interpretation

Overview of Complex Rare Category Analysis

Applications

Malicious Insiders

Emerging Trend

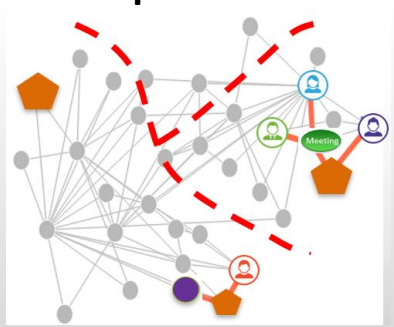
Spam Email

Rare Disease

Part I: Exploration

Identified Rare Examples

Part II: Exploitation

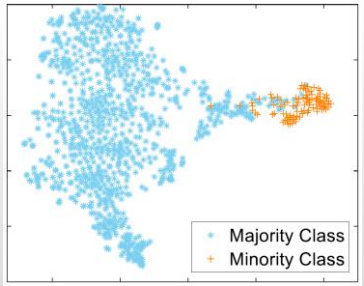


Identified Rare Examples

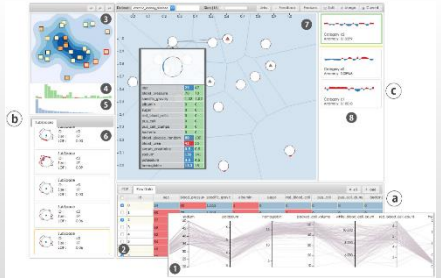
Visualization & Explanation

Identified Rare Categories

Part III: Interpretation



Embedding



User Interface

End User

Feedback

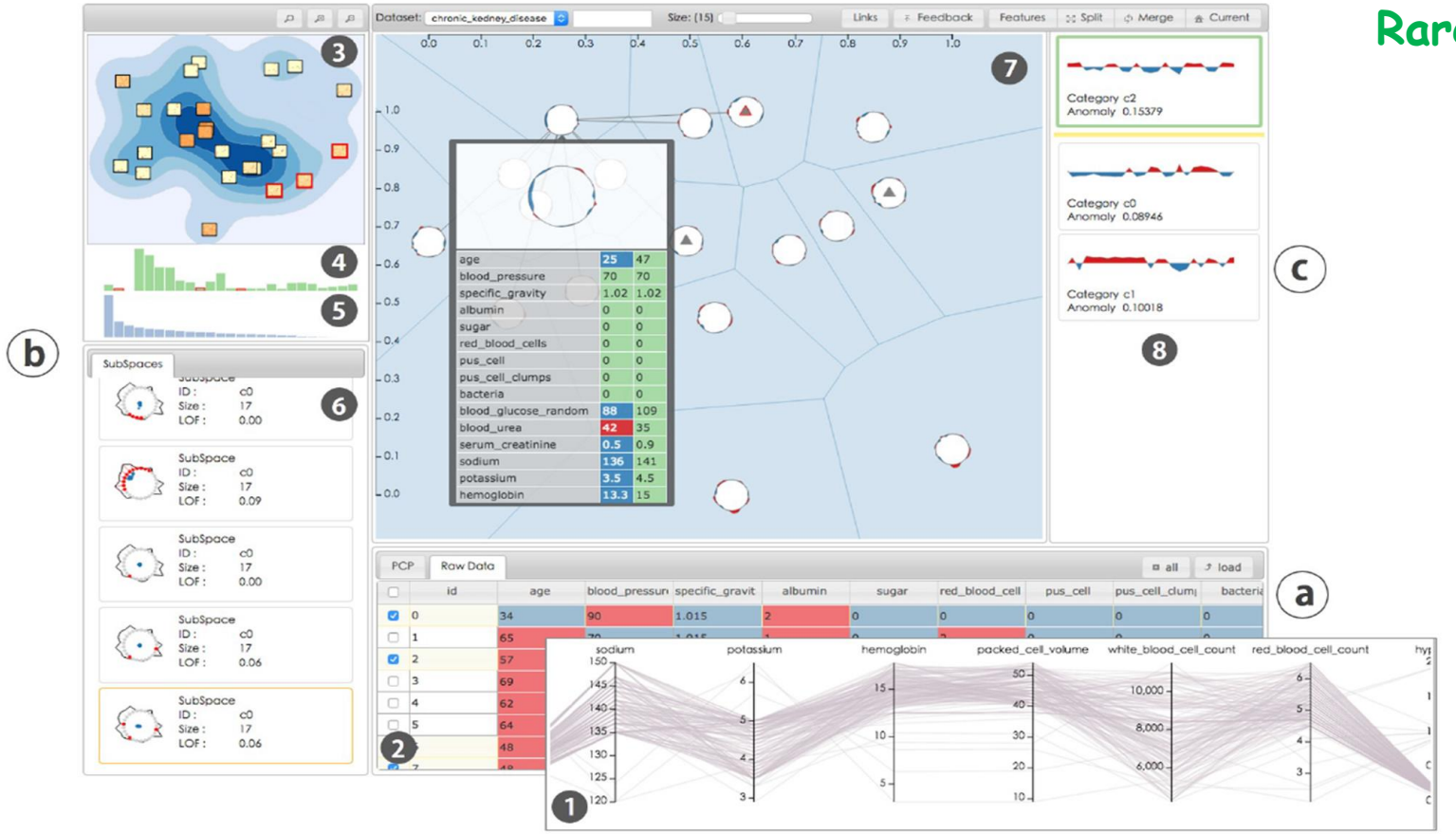
Annotation

Diagnosis

Interaction

The Prototype System

Feature Selection Module

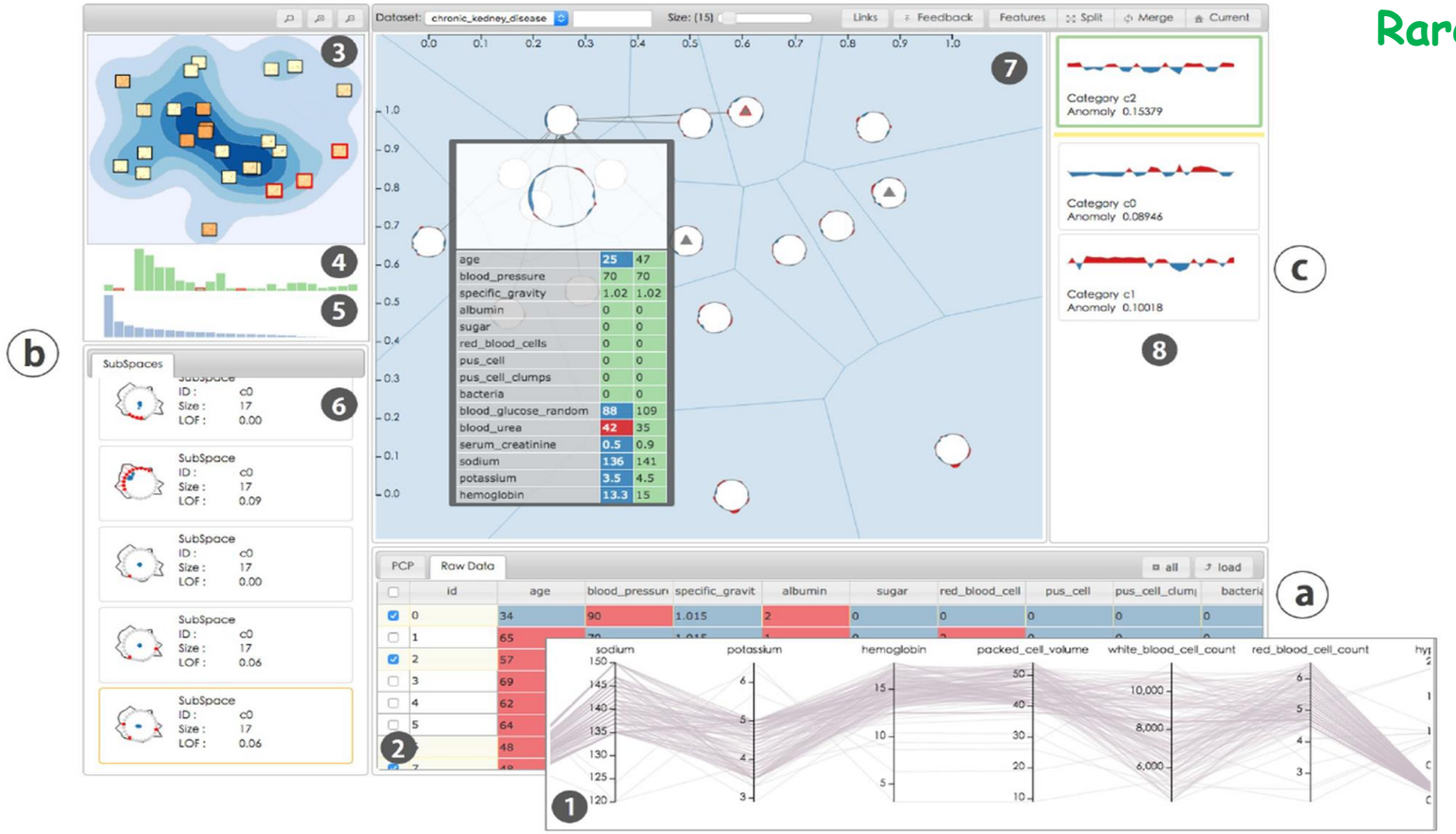


Rare Category Analysis Module

Data Exploration Module

The Prototype System

Feature Selection Module



Rare Category Analysis Module

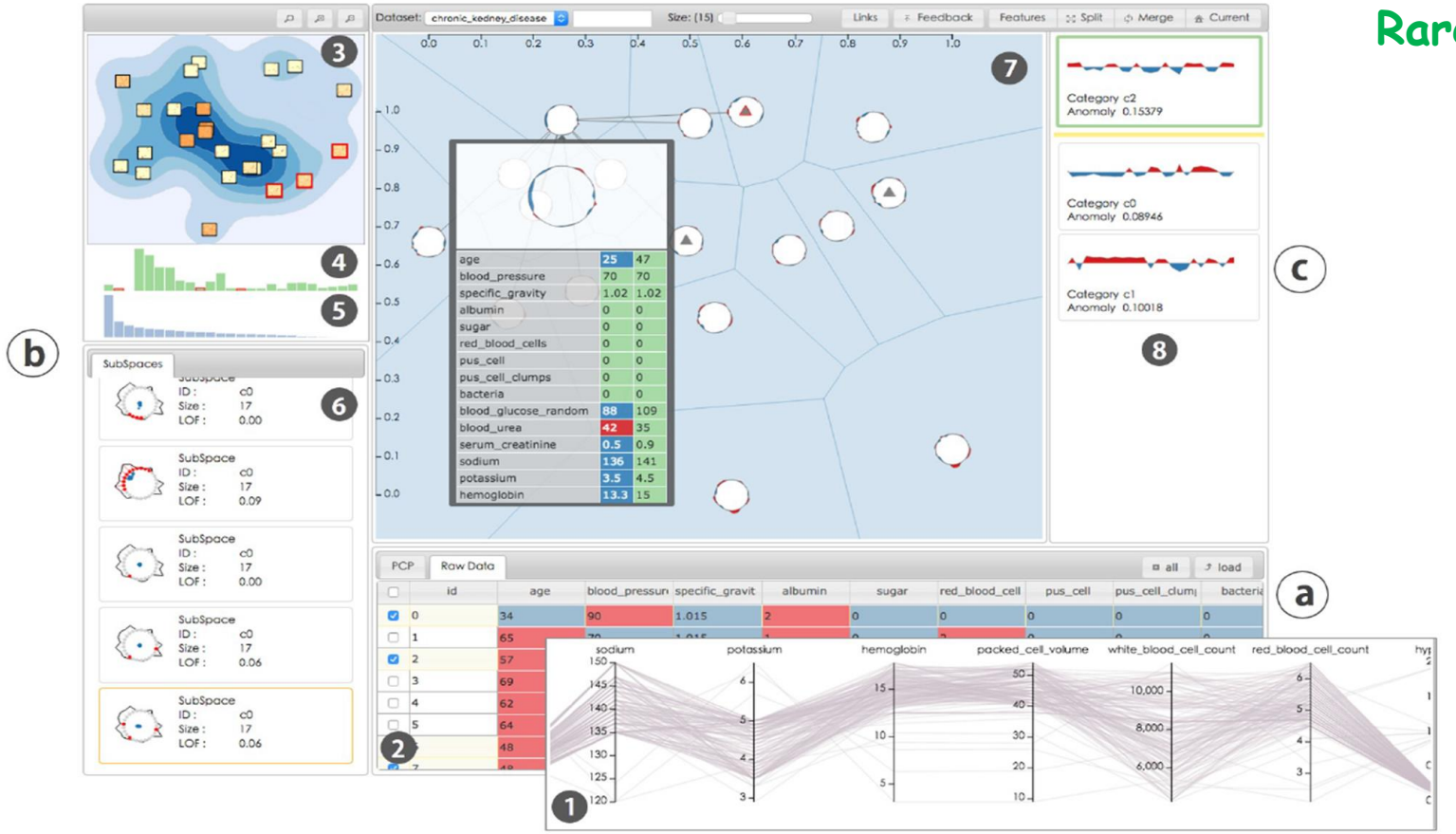
Data Exploration Module

- Represent raw data
- Interactive visualization for data querying
- Support Data filtering

The Prototype System

Feature Selection Module

- Visualize the Variance of data
- Visualize the correlation of data
- Guide the feature selection and subspace investigation process



Rare Category Analysis Module

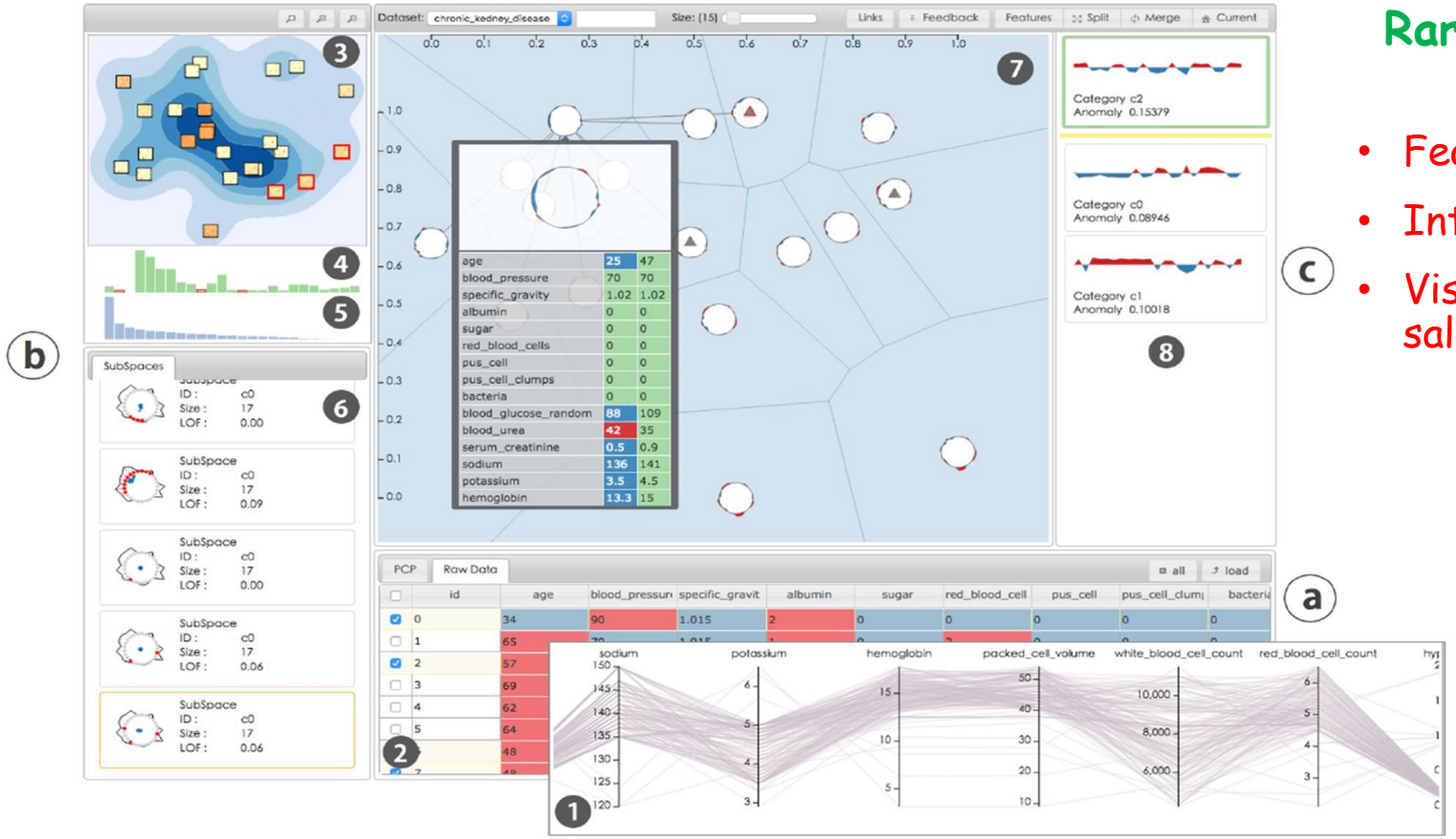
Data Exploration Module

- Represent raw data
- Interactive visualization for data querying
- Support Data filtering

The Prototype System

Feature Selection Module

- Visualize the Variance of data
- Visualize the correlation of data
- Guide the feature selection and subspace investigation process



Rare Category Analysis Module

- Feature selection
- Interactive active learning
- Visualize rare examples in a salient representation

Data Exploration Module

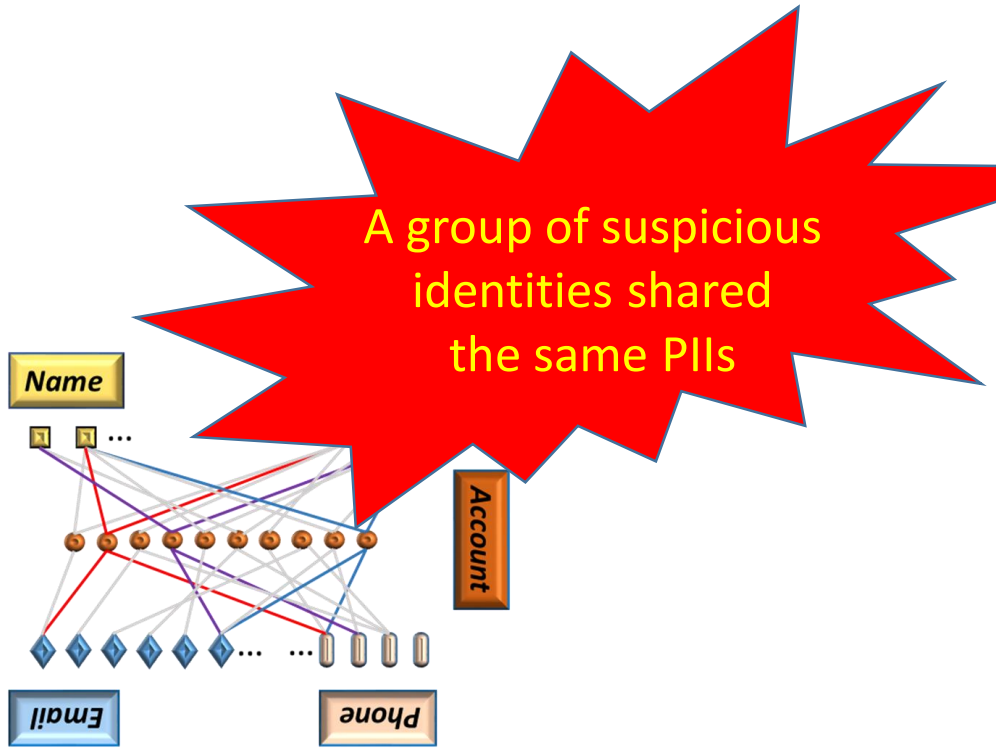
- Represent raw data
- Interactive visualization for data querying
- Support Data filtering

A Case Study in Financial Fraud Detection

- Problem
  - Given: Personal identification information (PII) network of the bank customers.
  - Find: Suspicious synthetic identities.

- Identified Abnormal Patterns

PII Network



Identified Rare Category

Part IV:
Challenges & Future Directions

Next Generation Rare Category Analysis System

- Built Upon Four Key Pillars
  - P1. Comprehension: Theoretical foundation of rare category analysis.
  - P2. Explainability: Understandable rare category prediction models.
  - P3. Robustness: Trustworthy of RCA models.
  - P4. Scalability: Bounded/Constant computational complexity.

- Advancing High-Impact Applications
  - Aviation Safety
  - Healthcare
  - Manufacturing
  - Finance
  - Security

P1. Comprehension

- Observations
  - O1: Different assumptions behind different types of heterogeneity

Heterogeneity	Underlying Assumption
Task	Task Relatedness
View	View Consistency
Instance	Bi-level Label Relationship
Oracle	Better-than-Random

  - O2: As the learning model becomes more complex by encoding more types of heterogeneity in the rare category analysis models
    - The likelihood of model assumption violation increases
    - The likelihood of model over-fitting increases

P2. Explainability

$$\varepsilon(x) = \text{argmin}_{g \in G} \mathcal{L}(f, g, \pi_x) + \Omega(g)$$

$\mathcal{L}(f, g, \pi_x)$

- f: rare category analysis model
- π_x: Local vicinity of example x
- L: Model fidelity function

$\Omega(g)$

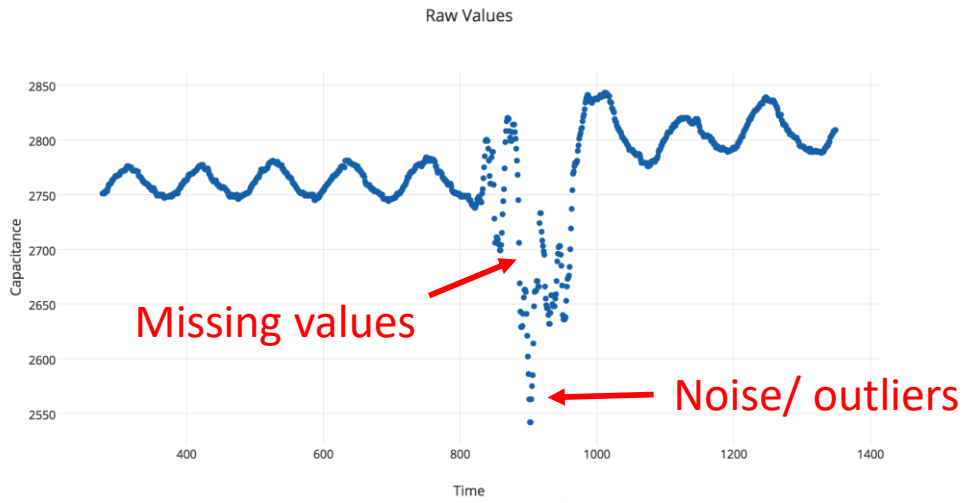
- $g \in G$: Interpretable model
- $\Omega(g)$: Model complexity (inversely proportional to model interpretability)

	Problem	State-of-the-art	Future Directions
Fidelity	Supervision	Output of outlier detection model	• Query values; • Correct wrong predictions
	Model Types	Flat models	• Hierarchical • Multi-resolution
Explainability	Data Sensitivity	NA	• Relevant instances • Relevant features
	Model Diagnosis	NA	• Explaining correct or wrong predictions
	Model Improvement	NA	• Explainable "user in the loop" learning process

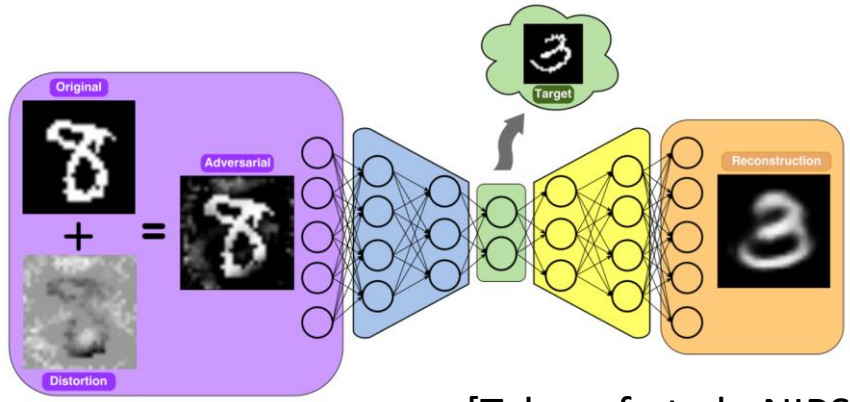
Key Insight: **Data heterogeneity brings both challenges and opportunities for explainability!**

P3. Robustness

- Operational Robustness
  - Ex. Noise, outliers, missing values



- Adversarial Robustness
  - Ex. Adversarial attacks



[Tabacof et al., NIPS 2016]

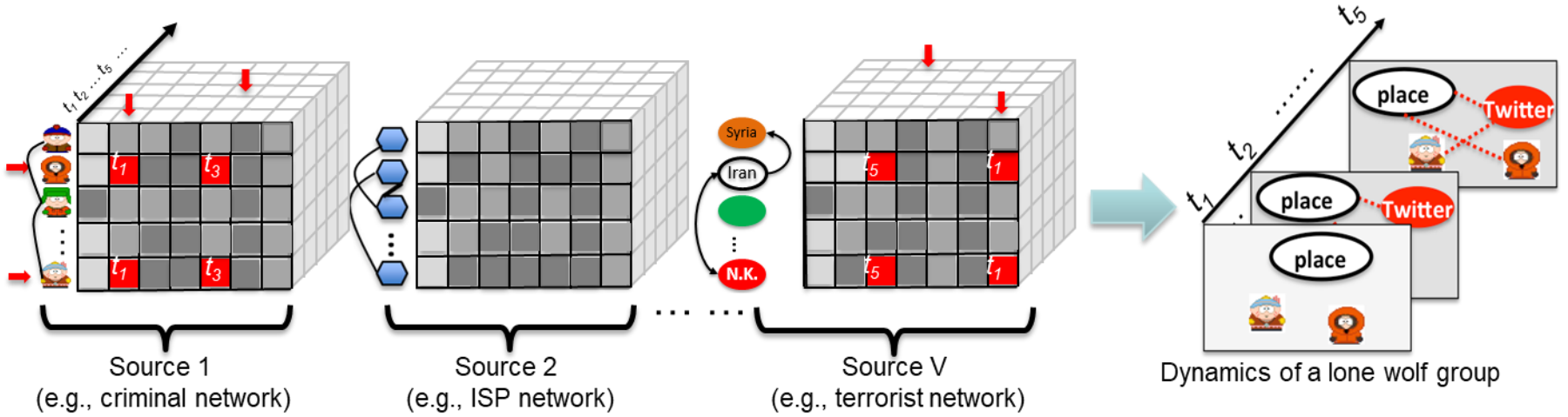
Two competing hypotheses for robust learning in the presence of data heterogeneity

- H1: Heterogeneity makes the learning system more vulnerable (more attacking channels for adversaries, possibly causing **a cascading attack**).
- H2: Heterogeneity makes the learning system more robust (it makes the attacker harder to generate examples that is **universally adversarial**, i.e., it becomes harder to fool ALL the learning tasks).

P4. Scalability

- Key Research Questions
  - Time Complexity
  - Space Complexity

- A Case Study: Long-Wolf Attack Detection
  - Multi-Sourced Data
  - Evolving over time



References

Introduction

- Chen, C.. "CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature." Journal of the Association for Information Science and Technology 57.3 (2006): 359-377.
- Spaaij, R.. "The enigma of lone wolf terrorism: An assessment." Studies in Conflict & Terrorism 33.9 (2010): 854-870.
- Ngai, E. W., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. Decision support systems, 50(3), 559-569.
- D. Pelleg, A. W. Moore: Active Learning for Anomaly and Rare-Category Detection. NIPS 2004.
- He, J. (2010). Rare category analysis (Doctoral dissertation, Carnegie Mellon University, School of Computer Science, Machine Learning Department).
- He, Haibo, and Edwardo A. Garcia. "Learning from imbalanced data." IEEE Transactions on Knowledge & Data Engineering 9 (2008): 1263-1284.
- Chawla, Nitesh V., Nathalie Japkowicz, and Aleksander Kotcz. "Special issue on learning from imbalanced data sets." ACM Sigkdd Explorations Newsletter 6.1 (2004): 1-6.
- Chawla, Nitesh V. "Data mining for imbalanced datasets: An overview." Data mining and knowledge discovery handbook. Springer, Boston, MA, 2009. 875-886.
- Workshop on Learning from Imbalanced Data Sets, Co-located with AAAI 2000.
- Workshop on Learning with Imbalanced Domains: Theory and Applications, Co-located with ECML/PKDD 2018.
- Workshop on Active Learning for Decision-Making from Imbalanced Observational Data, Co-located with ICML 2019.

References

Part I

- D. Pelleg, A. W. Moore: Active Learning for Anomaly and Rare-Category Detection. NIPS 2004.
- J. He, and J. Carbonell. Nearest-Neighbor-Based Active Learning for Rare Category Detection. NIPS 2007.
- Lamba, H., & Akoglu, L. (2019). Learning On-the-Job to Re-rank Anomalies from Top-1 Feedback.
- D. Zhou, K. Wang, N. Cao, J. He. Rare Category Detection on Time-Evolving Graphs, IEEE International Conference on Data Mining (ICDM-2015), November 2015.
- V. Pavan, and W.-K. Wong. "Category detection using hierarchical mean shift." ACM SIGKDD 2009.
- J. He, and J. Carbonell. Co-Selection of Features and Instances for Unsupervised Rare Category Analysis. SDM 2010.
- J. Sun, H. Qu, D. Chakrabarti, and C. Faloutsos. Neighborhood formation and anomaly detection in bipartite graphs. ICDM 2005.
- H. Tong, C. Lin: Non-Negative Residual Matrix Factorization with Application to Graph Anomaly Detection. SDM 2011.
- Emaad A. Manzoor, Sadegh M. Milajerdi, Leman Akoglu: Fast Memory-efficient Anomaly Detection in Streaming Heterogeneous Graphs. KDD 2016.
- Dawei Zhou, Jingrui He, K. Selcuk Candan, Hasan Davulcu. MUVIR: Multi-View Rare Category Detection. IJCAI 2015.

References

Part II

- J. He, H. Tong, and J. Carbonell. Rare Category Characterization. ICDM 2010.
- P. Yang, J. He, J.Y. Pan. Learning Complex Rare Categories with Dual Heterogeneity. SDM 2015.
- E. Keogh, J. Lin and A. Fu (2005). HOT SAX: Efficiently Finding the Most Unusual Time Series Subsequence. In Proc. of the 5th IEEE International Conference on Data Mining (ICDM 2005), pp. 226 - 233., Houston, Texas, Nov 27-30, 2005.
- Chin-Chia Michael Yeh, Yan Zhu, Liudmila Ulanova, Nurjahan Begum, Yifei Ding, Hoang Anh Dau, Diego Furtado Silva, Abdullah Mueen, Eamonn Keogh (2016). Matrix Profile I: All Pairs Similarity Joins for Time Series: A Unifying View that Includes Motifs, Discords and Shapelets. IEEE ICDM 2016.
- Y. Bu, O. T.-W. Leung, A. W.-C. Fu, E. J. Keogh, J. Pei, and S. Meshkin, "WAT: Finding Top-K Discords in Time Series Database," in Proc. of the 7th SIAM Intl. Conf. on Data Mining (SDM), 2007, pp. 449–454.
- X.-y. Chen and Y.-y. Zhan, "Multi-scale Anomaly Detection Algorithm based on Infrequent Pattern of Time Series," Journal of Computational and Applied Mathematics, vol. 214, no. 1, pp. 227–237, Apr 2008.
- V. Chandola, V. Mithal, and V. Kumar, "A Comparative Evaluation of Anomaly Detection Techniques for Sequence Data," in Proc. of the 2008 8th IEEE Intl. Conf. on Data Mining (ICDM), 2008, pp. 743–748.
- S. Budalakoti, A. N. Srivastava, and M. E. Otey, "Anomaly Detection and Diagnosis Algorithms for Discrete Symbol Sequences with Applications to Airline Safety," IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications, vol. 39, no. 1, pp. 101–113, Jan 2009.
- Dawei Zhou, Jingrui He, Yu Cao, Jae-sun Seo. Bi-level Rare Temporal Pattern Detection. ICDM 2016.

References

Part II

- F. A. Gonz´alez and D. Dasgupta, “Anomaly Detection Using Real-Valued Negative Selection,” Genetic Programming and Evolvable Machines, vol. 4, no. 4, pp. 383–403, Dec 2003.
- D. Dasgupta and F. Nino, “A Comparison of Negative and Positive Selection Algorithms in Novel Pattern Detection,” in Proc. of the 2000 IEEE Intl. Conf. on Systems, Man, and Cybernetics, vol. 1, 2000, pp. 125–130.
- D. Endler, “Intrusion Detection Applying Machine Learning to Solaris Audit Data,” in Proc. of the 14th Annual Computer Security Applications Conf. (ACSAC), 1998, pp. 268–279.
- A. K. Gosh, J. Wanken, and F. Charron, “Detecting Anomalous and Unknown Intrusions Against Programs,” in Proc. of 18 IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 25, NO. 1, JANUARY 2014 the 14th Annual Computer Security Applications Conf. (ACSAC), 1998, pp. 259–267.
- A. Ghosh, A. Schwartzbard, and M. Schatz, “Learning Program Behavior Profiles for Intrusion Detection,” in Proc. of the 1st USENIX Workshop on Intrusion Detection and Network Monitoring, 1999, pp. 51–62.
- A. K. Ghosh and A. Schwartzbard, “A Study in using Neural Networks for Anomaly and Misuse Detection,” in Proc. of the 8th Conf. on USENIX Security Symposium (SSYM), 1999, pp. 12–23.
- B. Gao, H.-Y. Ma, and Y.-H. Yang, “HMMs (Hidden Markov Models) based on Anomaly Intrusion Detection Method,” in Proc. of the 2002 Intl. Conf. on Machine Learning and Cybernetics, vol. 1, 2002, pp. 381–385.
- A. W. Williams, S. M. Pertet, and P. Narasimhan. Tiresias: Black-box Failure Prediction in Distributed Systems. IPDPS 2007.

References

Part II

- L. Akoglu, M. McGlohon, C. Faloutsos. OddBall: Spotting Anomalies in Weighted Graphs. PAKDD, 2010.
- D. Zhou, S. Zhang, M. Y. Yildirim, S. Alcorn, H. Tong, H. Davulcu, J. He: A Local Algorithm for Structure-Preserving Graph Cut. KDD 2017.
- Yin, Hao, Austin R. Benson, Jure Leskovec, and David F. "Local higher-order graph clustering." KDD 2017.
- S. Zhang, Dawei Zhou, M. Y. Yildirim, S. Alcorn, J. He, H. Davulcu, H. Tong. HiDDen: Hierarchical Dense Subgraph Detection with Application to Financial Fraud Detection. SDM 2017.

Part III

- Huang, Chen, et al. Learning deep representation for imbalanced classification. ICCV 2016.
- D. Zhou, J. He, H. Yang, W. Fan. SPARC: Self-Paced Network Representation for Few-Shot Rare Category Characterization. KDD 2018.
- Wu, J., He, J., & Liu, Y. (2018). ImVerde: Vertex-Diminished Random Walk for Learning Network Representation from Imbalanced Data. arXiv preprint arXiv:1804.09222.
- HX. H. Dang, B. Micenkov ́a, I. Assent, and R. T. Ng. Local outlier detection with interpretation. ECML-PKDD 2013.
- N. Liu, D. Shin, and X. Hu. Contextual Outlier Interpretation. IJCAI 2018.
- Macha, Meghanath, and Leman Akoglu. "Explaining anomalies in groups with characterizing subspace rules." DMKD 2018.