Centrality Resilience in Complex Networks

Fariba Afrin Irany[1], Turja Kundu[1], Soumya Sarkar[2]

Animesh Mukherjee[3], Sanjukta Bhowmick[1]

University of North Texas[1], Microsoft India[2], IIT Kharagpur[3]

What Is Centrality Resilience?

Traditional Resilience:

- Focuses on **connectivity**
- Key question: “Is the network still connected after an attack?”

Centrality Resilience (Our Focus):

- Focuses on **path-based influence**
- Key question: “Do the top central nodes retain their rank?”

Why it matters:

- Critical in systems where **latency**, **flow**, or **influence** matters
- E.g., communication, transportation, biological, or social systems

Centrality Disruption vs Traditional Attacks

Path-Based Centralities:

- **Betweenness**: Number of shortest paths passing through a node
- **Closeness**: Inverse of the average shortest-path distance to all other nodes

Traditional vs Centrality Attacks

	Traditional Attack	**Centrality Attack**
Goal	Disconnect the network	Disrupt node rankings
Method	Remove nodes or edges	Remove edges
Visibility	Easily detectable	Often stealthy
Outcome	Structural damage	Functional degradation

Real-World Impacts

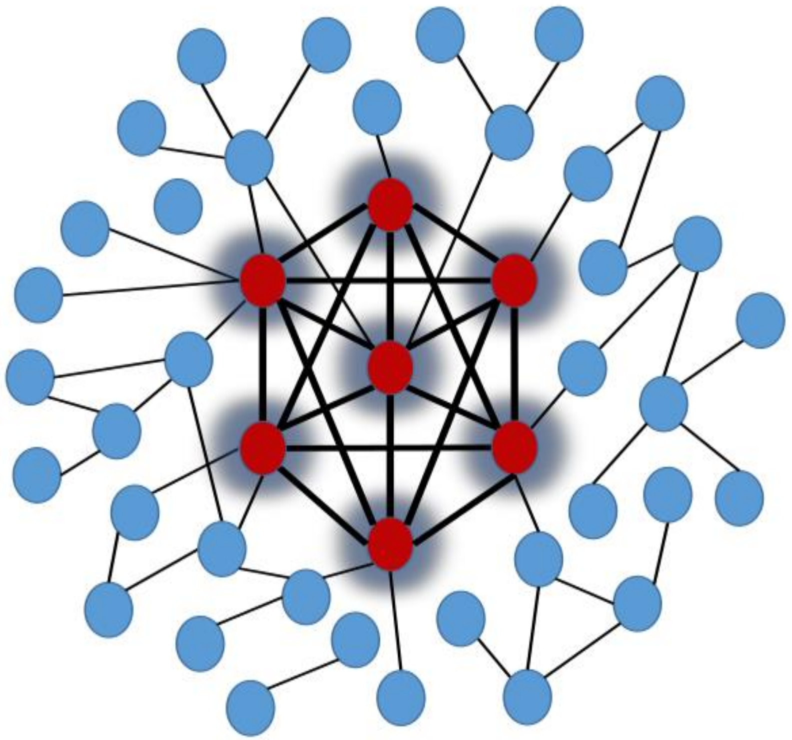
- Increased delays, misrouting, bottlenecks
- Network appears intact—but performs poorly

Applications: Cybersecurity • Load balancing • Influence control

Our focus is on **subtle yet powerful attacks** that degrade functionality without being obvious

Understanding Rich Clubs

A **rich club** is an assortative subgraph where all nodes exhibit **high values** of a centrality property—specifically, **betweenness** or **closeness centrality**.



Core Properties:

- Formed by nodes with **top-ranked centrality**
- Located in **dense regions**, often the **innermost cores**
- Exhibit strong **mutual connectivity** (assortative structure)
- Enable **efficient communication** across the network

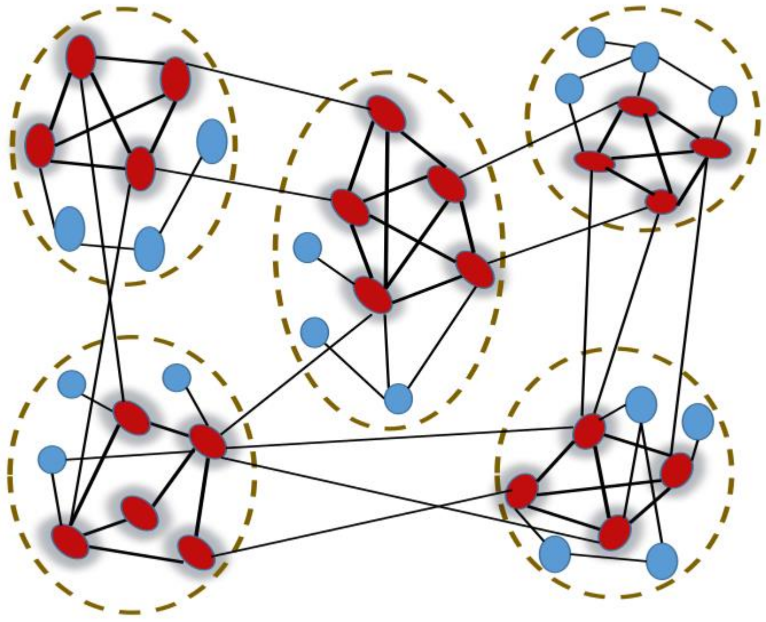
Beyond the Core – Scattered Rich Clubs

Problem: Not all networks have centralized rich clubs.

Scattered Rich Clubs:

A Scattered Rich Club is a collection of **disjoint, dense subgraphs** such that:

- Each subgraph contains at least **one high-centrality node**
- **Every node** in these subgraphs has a **centrality value above a fixed threshold** (based on betweenness or closeness centrality)



Key Characteristics:

- Subgraphs may appear in the **core, middle, or periphery**
- Reflect **decentralized regions of influence**

Key Steps and Contributions

- Generalize concept of Rich club to **Scattered Rich Clubs**
- Define a quantitative metric: **Degree of Scatteredness**
- Develop scalable Scattered Rich Clubs detection method via **Snowball sampling**
- Propose **edge-based attack models** leveraging Scattered Rich Clubs
- Validate results on 10 real-world datasets

Identifying Scattered Rich Clubs

Steps to Identify Clusters forming Scattered Rich Clubs

- **Seed Selection**:
  - Combine top-20 betweenness & closeness centrality nodes
- **Initial Clustering**:
  - Each node in seeds forms a cluster with its neighbors
- **Cluster Merging** :
  - Merge any pair of Cluster if their Jaccard Index > 0.1
- **Iterate Until Stable**:
  - Repeat merging until no more merges are possible

Final Output:

- Disjoint clusters = rich clubs
- Single cluster → centralized rich club
- multiple cluster → scattered rich clubs

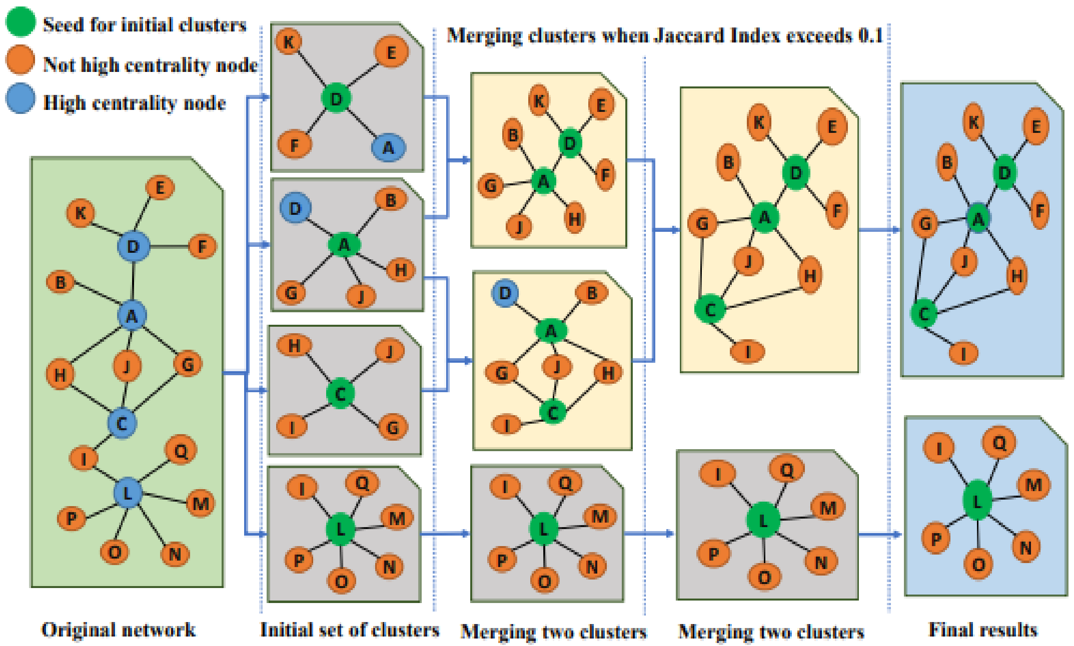


Fig. 2: Step-by-step illustration of the clustering algorithm.

Measuring Scatteredness

Let

H_i = Number of high-centrality nodes in cluster i

K = Number of Clusters

$$\text{Degree of Scatteredness} = \frac{\mathbf{1}}{\boldsymbol{K}}\sum_{\boldsymbol{i}}^{\boldsymbol{K}}\frac{\boldsymbol{H_i}}{\boldsymbol{i}}$$

Interpretation:

Degree of Scatteredness = 1 → all centrality nodes in one cluster

Degree of Scatteredness = 0 → maximum scattering of centrality nodes

Scatteredness in Real Networks

- Scatteredness on ten real-world networks from different domains:

Degree of scatteredness of Networks and high-centrality nodes prediction						
Dataset	**High-centrality nodes**	**Number of Clusters**	**Distribution of High Centrality Nodes**	**Degree of Scatteredness**	**Precision**	**Recall**
dmela	25	25	25(1)	0.152	0.07	0.88
euroroad	33	31	29(1), 2(2)	0.167	0.09	0.55
HepPh	28	21	17(1), 2(2), 1(4), 1(3)	0.293	0.04	0.96
CondMat	28	21	19(1), 1(2), 1(7)	0.362	0.14	0.79
as20000102	24	15	12(1), 2(5), 1(2)	0.402	0.70	0.79
caida	25	12	9(1), 1(11), 1(3), 1(2)	0.577	0.39	0.96
HepTh	26	10	6(1), 2(3), 1(2), 1(12)	0.609	0.15	0.77
email-univ	26	10	7(1), 2(2), 1(15)	0.683	0.15	0.85
AstroPh	31	9	6(1), 2(2), 1(21)	0.763	0.15	0.74
grid-fission-yeast	33	6	4(1), 1(26), 1(3)	0.862	0.30	0.45

Table 2: Degree of scatteredness and high-centrality nodes prediction via sampling. Multiplicity of clusters is shown as, K(M) = K clusters with M high-centrality nodes.

The Computational Challenge

Exact centrality computation for initial seeds:

- Time Complexity: **O(VE)** – infeasible for large graphs

Real-world constraints:

- Incomplete access
- Time/compute limits

Solution:

→ Snowball Sampling (Low time complexity)

- Rich clubs behave like **expanders** — reach many nodes quickly
- Snowball Sampling exploits this property to efficiently approximate high centrality nodes

Rich Club Detection via Snowball Sampling

- **Finding high centrality nodes by Snowball Sampling**
  - Start from high-degree, high-clustering seed nodes
  - Expand snowball sampling to cover 10% of nodes
  - Extract **core-periphery structure** of sampled subgraph
  - Label nodes in **top 2 cores** as high-centrality nodes
  - Repeat until convergence or max 40 runs
- **Output**
  - **Sampled subgraphs ≈ predicted rich clubs**
  - **Inner-core nodes ≈ high-centrality vertices**
- **Time complexity:**
  - O(T · E(S)); T is the maximum iterations. (edges in sampled subgraph **S**)

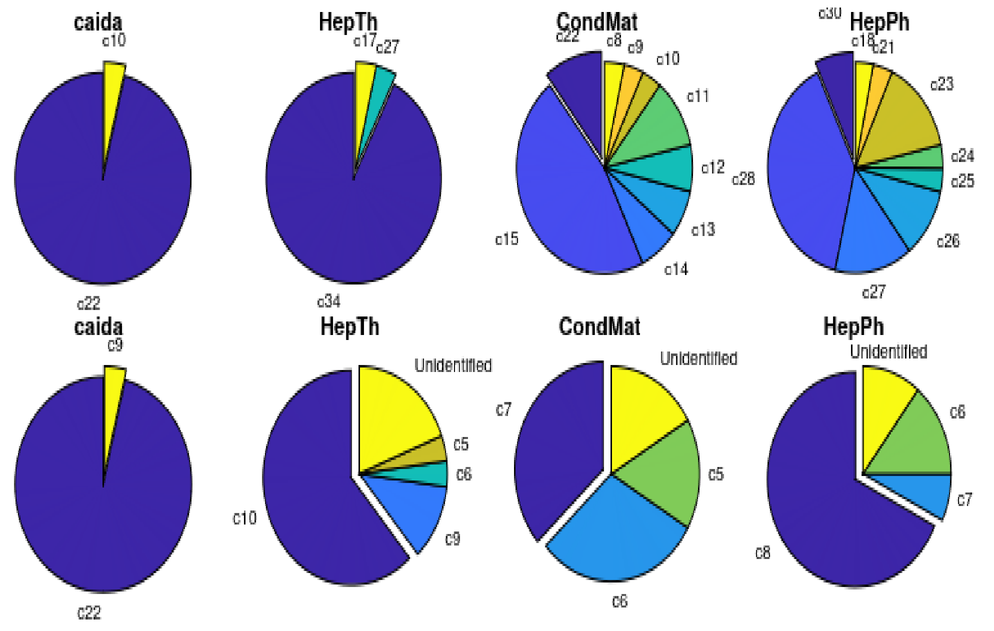


Fig. 3: Distribution of high-centrality nodes across network cores. Top: original networks. Bottom: networks sampled using the snowball in a single run with 10% of nodes.

High-Centrality Nodes Prediction Result

- High Centrality nodes prediction results on ten real-world networks from different domains:

Degree of scatteredness of Networks and high-centrality nodes prediction						
Dataset	**High-centrality nodes**	**Number of Clusters**	**Distribution of High Centrality Nodes**	**Degree of Scatteredness**	**Precision**	**Recall**
dmela	25	25	25(1)	0.152	0.07	0.88
euroroad	33	31	29(1), 2(2)	0.167	0.09	0.55
HepPh	28	21	17(1), 2(2), 1(4), 1(3)	0.293	0.04	0.96
CondMat	28	21	19(1), 1(2), 1(7)	0.362	0.14	0.79
as20000102	24	15	12(1), 2(5), 1(2)	0.402	0.70	0.79
caida	25	12	9(1), 1(11), 1(3), 1(2)	0.577	0.39	0.96
HepTh	26	10	6(1), 2(3), 1(2), 1(12)	0.609	0.15	0.77
email-univ	26	10	7(1), 2(2), 1(15)	0.683	0.15	0.85
AstroPh	31	9	6(1), 2(2), 1(21)	0.763	0.15	0.74
grid-fission-yeast	33	6	4(1), 1(26), 1(3)	0.862	0.30	0.45

Table 2: Degree of scatteredness and high-centrality nodes prediction via sampling. Multiplicity of clusters is shown as, K(M) = K clusters with M high-centrality nodes.

Comparative Evaluation – Attack Performance

Attack Strategy Overview

- Use snowball-sampled subgraphs to identify rich clubs
- Focus on nodes in **innermost and second innermost cores**
- Select **edges for removal** where:
    - **Both endpoints** have high core number
    - (If needed) At least **one endpoint** meets the condition

Attack Budget

- Remove **2%, 4%, 6%, 8%** of total edges
- Stop when budget is met or no qualifying edges remain

Time complexity: O(E(S)) (edges in sampled subgraph **S**)

Comparative Evaluation – Attack Performance

Baseline Methods Compared:

- Shell-Max
  - Removes edges with high k-core value
- Edge Betweenness Attack
  - Removes edges with highest edge betweenness centrality

Evaluation Metric:

Compare Top-20 central nodes **before vs. after** attack

Use **Jaccard Index**:

$$J = \frac{|A \cap B|}{|A \cup B|}$$

Lower J → More centrality disruption

Results – Closeness Centrality Disruption

Effectiveness:

- Rich Club based attack outperforms Shell-Max
- Comparable to Betweenness-based attacks

However, Time Complexity:

Method	Time Complexity
Betweenness Attack	O(VE)
Rich Club Attack (Ours)	O(E(S))

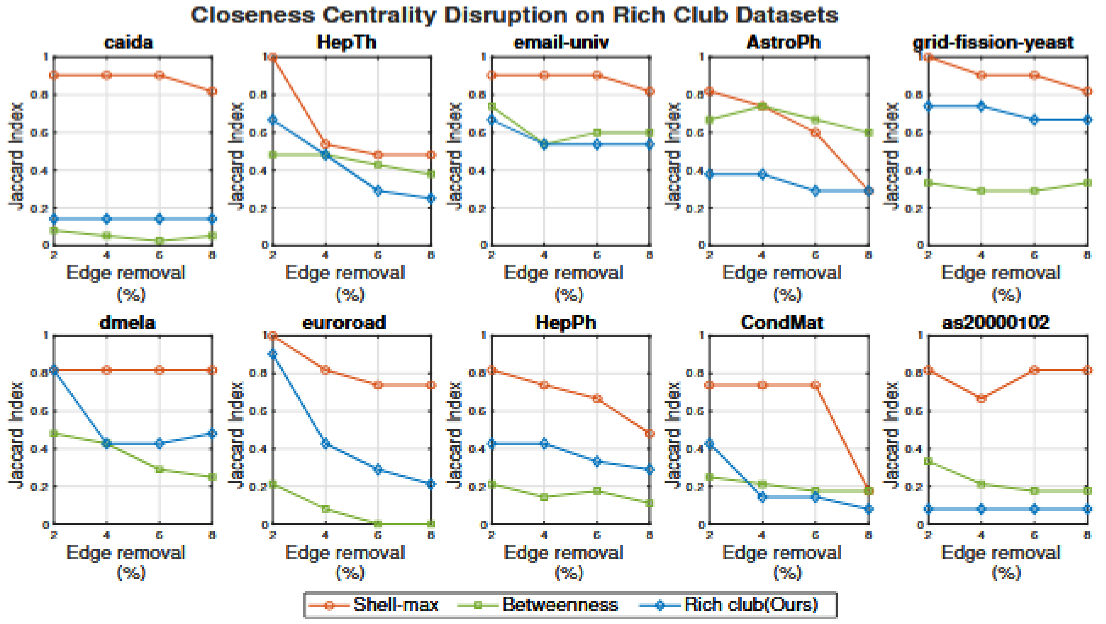


Fig. 4: Closeness centrality disruption under shell-max (red), betweenness (green), rich club (blue) attacks. Top: scatteredness > .5. Bottom: scatteredness < .5.

Rich Club attack balances **high impact** and **low cost**, making it a **scalable alternative** for large networks.

Results – Betweenness Centrality Disruption

Similar trends as with closeness

Rich Club Based method:

- High disruption than Shell-Max and comparable to Betweenness-based attacks
- Lower computational cost
- Effective across domains

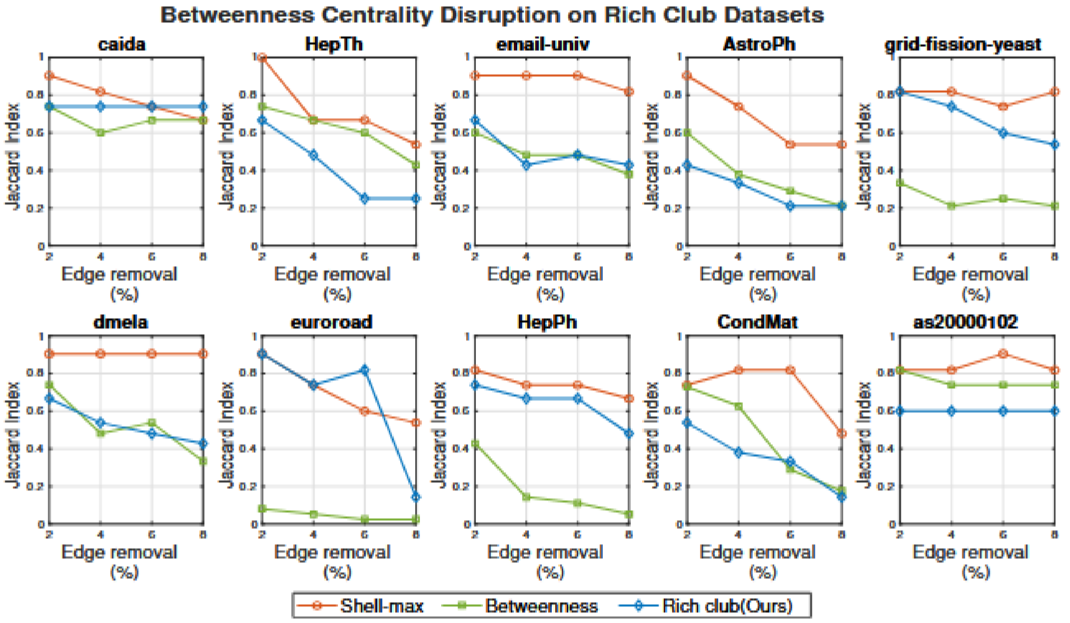


Fig. 5: Betweenness centrality disruption under shell-max, betweenness, rich club attacks. Top: scatteredness > .5. Bottom: scatteredness < .5.

Rich Club attack balances **high impact** and **low cost**, making it a **scalable alternative** for large networks.

Conclusion & Future Work

Key Insights

- High-centrality nodes are often scattered, not confined to a single core
- Scattered Rich Clubs provide a broader model of influence in networks
- Snowball sampling enables scalable detection of Scattered Rich Clubs
- Snowball sampling-based edge attacks cause greater disruption than traditional methods while being computationally efficient

Future Directions

- Extend analysis to dynamic and evolving networks
- Study the interaction of Scattered Rich Clubs with community structure