

# 基于最大连通图的犯罪团伙发现

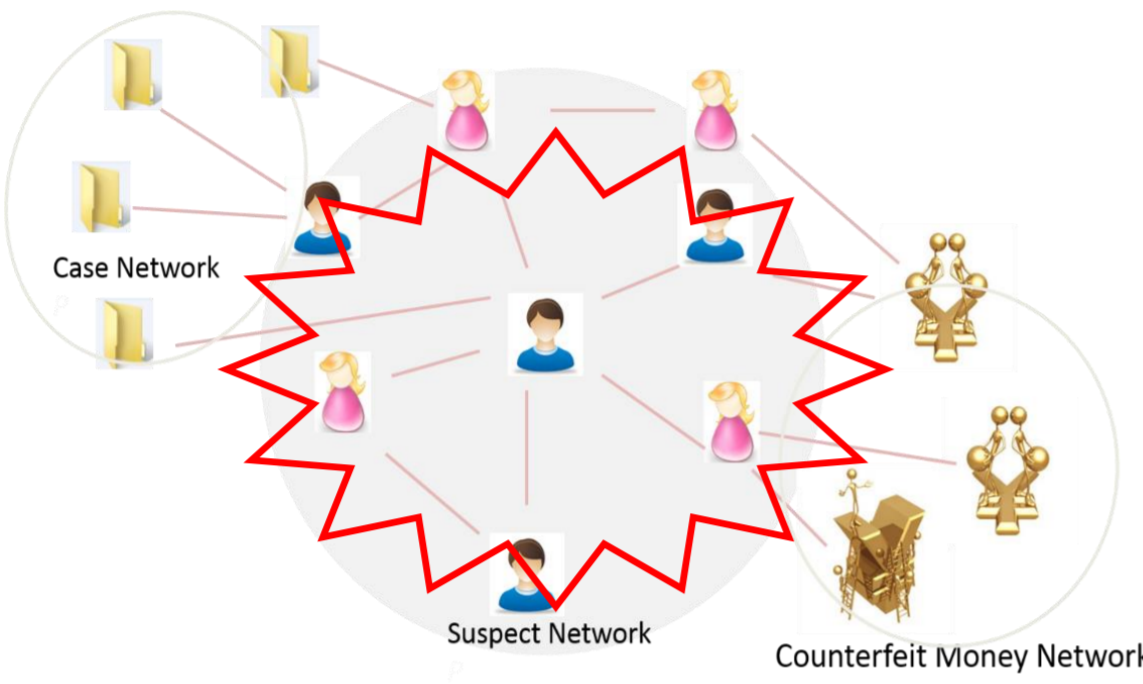
## Identification of Criminal Gang Based on Maximum Connectivity Graphs

Hui Li, Jixin Zou, Xubin Liu, Dengfeng Wu, Kegang Wei, Peijun Li, Shuren Xu

Contact : lihui2012@otcaix.iscas.ac.cn

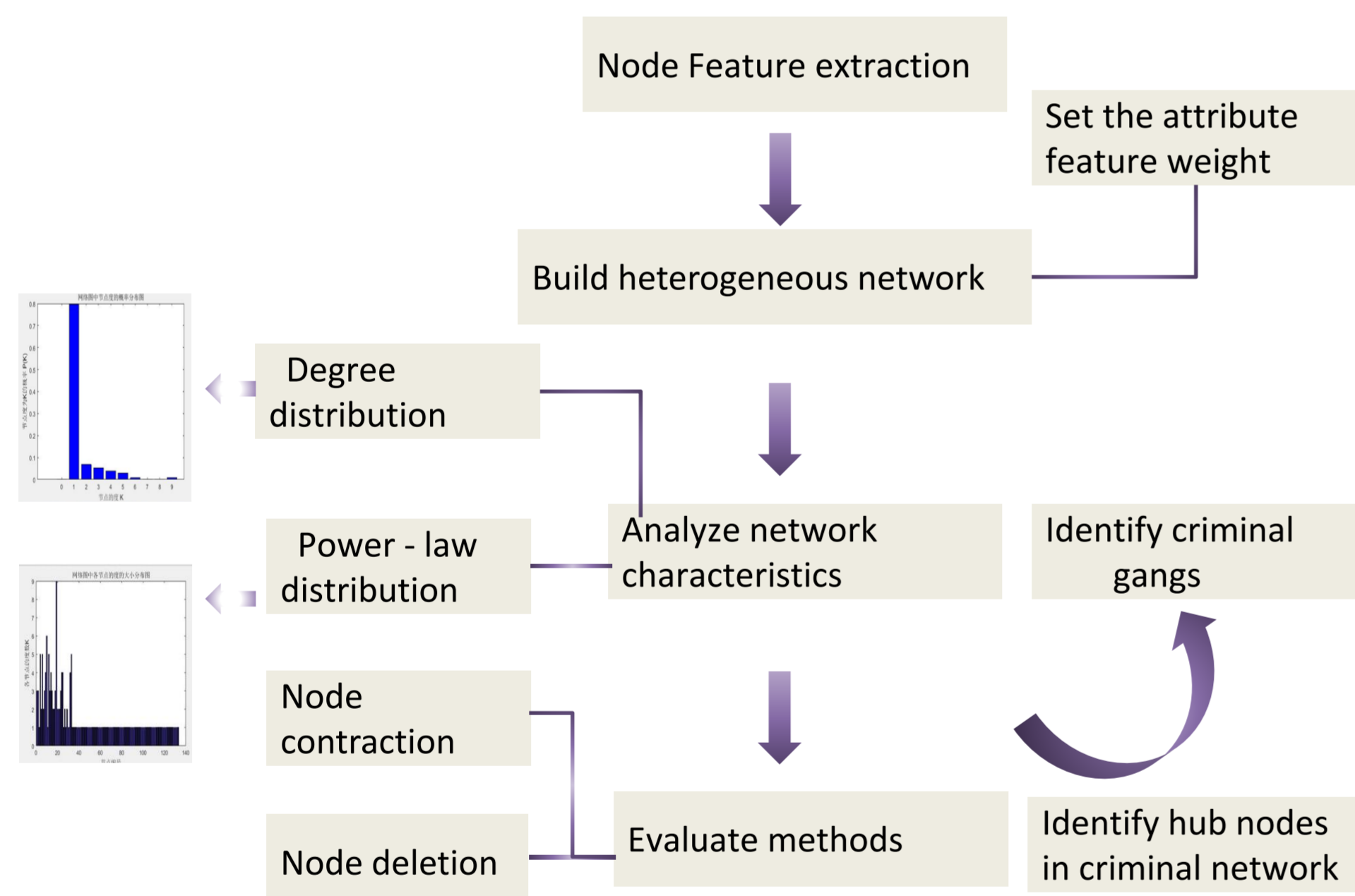
### Motivation

- With the development of big data and data mining techniques, network analysis can reveal entities, interdependencies and vulnerabilities of nodes in criminal network.
- Acts of counterfeit crime are considered one of the greatest threats to national security and defense.



### Methods

- Identify important nodes based on degree distribution considering topological features.
- Build heterogeneous network and analyze criminal network characteristics.
- Build the largest graph in criminal network considering attributes features.
- Identify criminal gangs from the network.



### Challenge

- The node type of criminal network is different. The criminal network have various nodes, such as case, suspect, and money. At the same time, each node has different attributes and interdependencies.
- With regards to structural and dynamic properties, we use data mining techniques to identify important nodes in heterogeneous network.
- Build the network to express the criminal gang to facilitate greater understanding of complex networks.

### Dataset

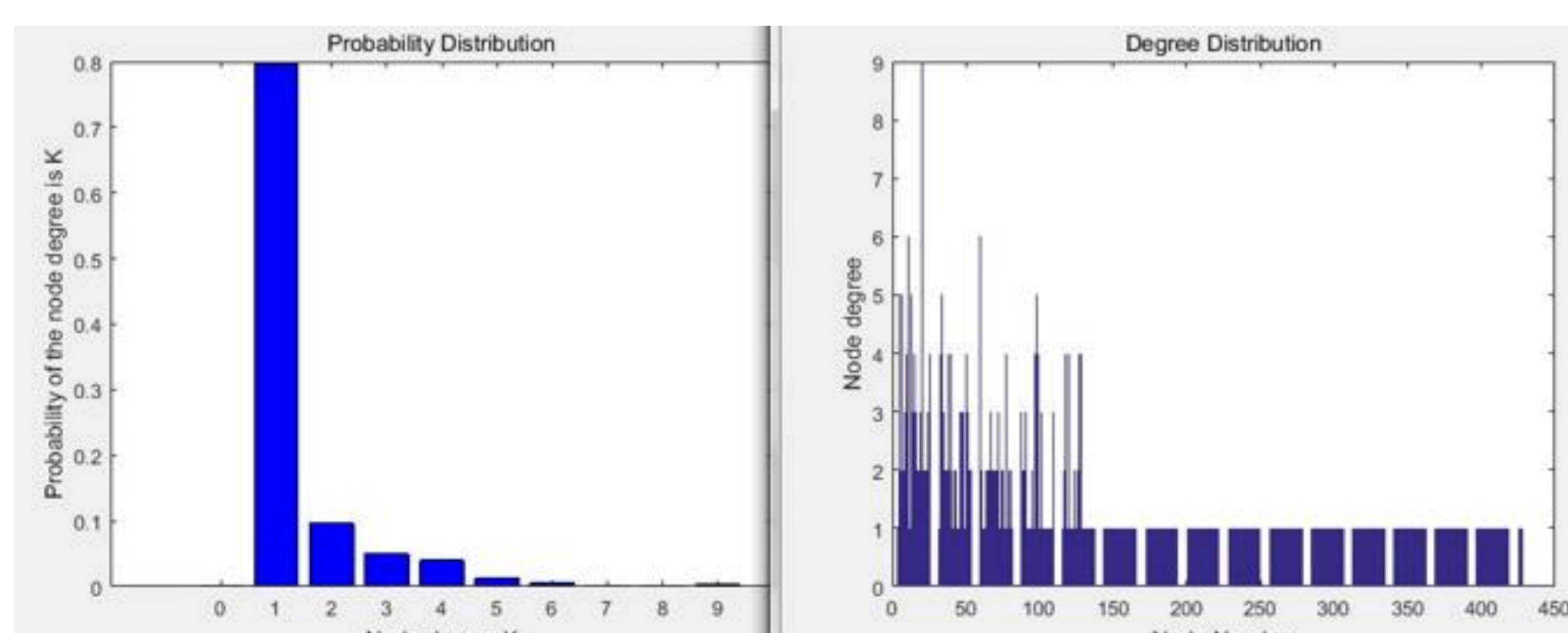
Based on the crime incident data provided by The Institute of Forensic Science, we summarized some of the important features of the dataset. The list is as follows :

Name	Values	Important attributes of Nodes
Number of case nodes	1058	Case_id; Case_Location; Case_happent_time;
Number of suspect nodes	2090	Suspect_name; Suspect_id; Nationality; ID number; Home_adress;
Number of money nodes	500	Fmid;Currency;Category; Denomination; Number;
Affiliation of case and suspect	13487	Case_id; Suspect_id;
Affiliation of case and currency	16298	Case_id; Fmid;

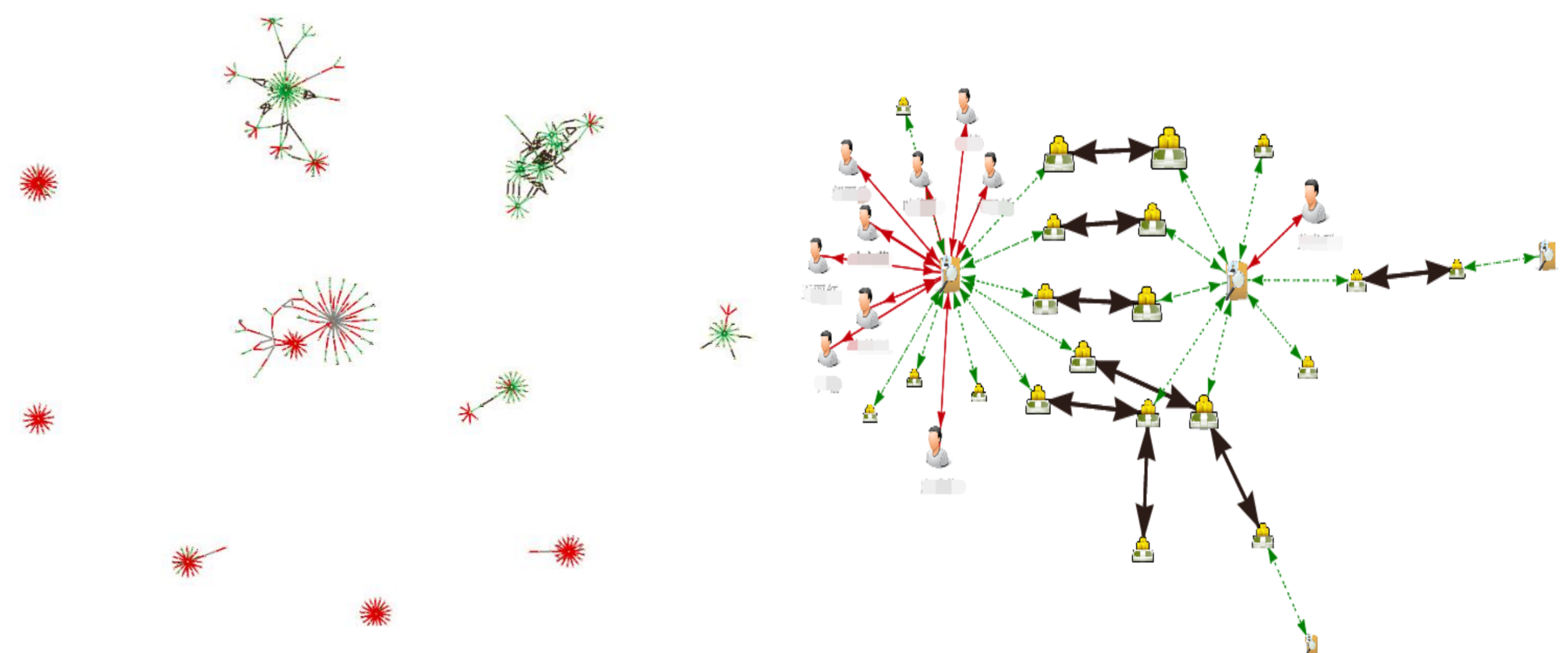
### Results & Evaluation

With such a complex real-world landscape, We regarded network analysis on the interdependencies and interconnectivity between nodes, which enrich our understanding of the underlying network structure and behavior to support solution design and decision making.

From the graph, we found that the degree distribution of criminal network is a power-rate distribution, which means a small part of nodes are large and important. We just focused on these central nodes. Respectively, we used cross-validation in our experiment under different network sizes.



a) Criminal network characteristics



b) Top 10 important nodes in Criminal network

c) A subgraph of a node (待发表论文)