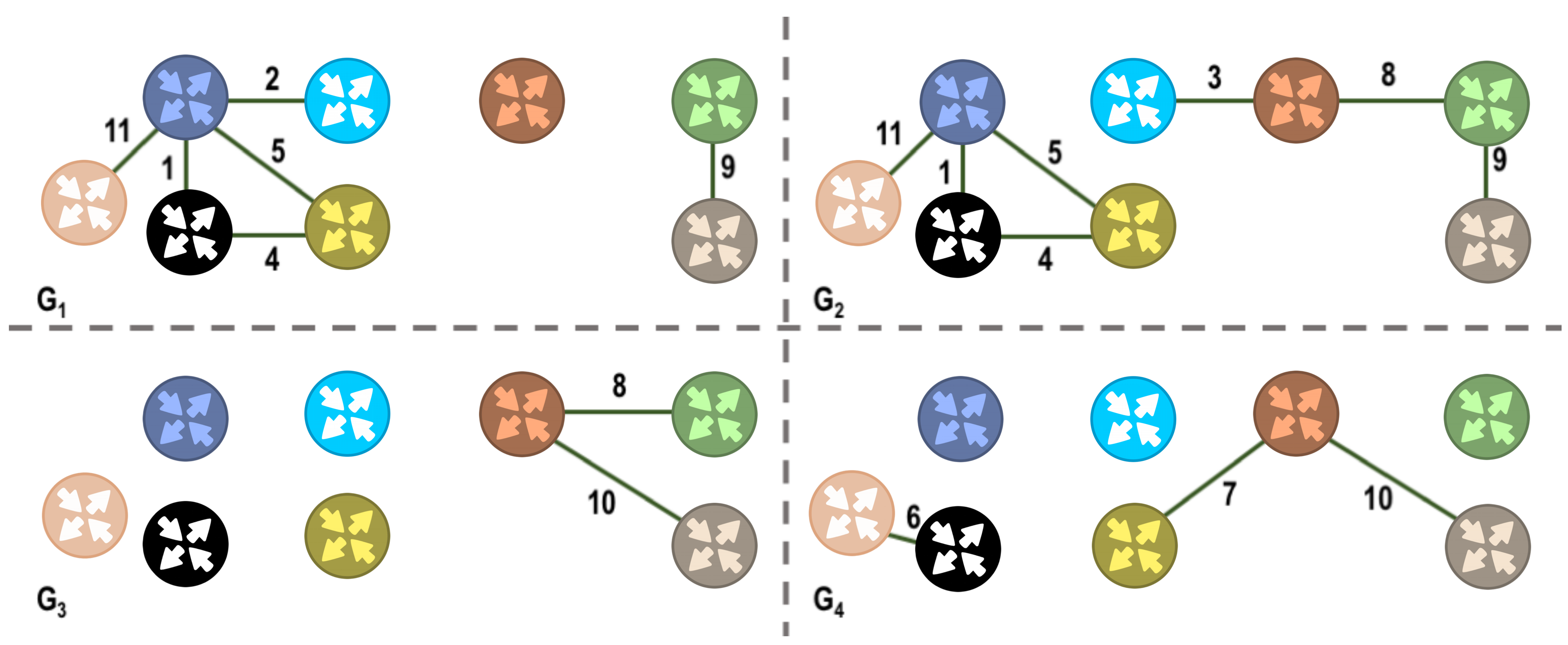
ExCoDE: a Tool for Discovering and Visualizing Regions of Correlation in Dynamic Networks

UNIVERSITY OF TRENTO

Giulia Preti[1], Polina Rozenshtein[2], Aristides Gionis[2], Yannis Velegrakis[3]

[1]University of Trento [2]Aalto University [3]Utrecht University

The BGP Routing Topology



A **fault** at a router induces changes in all the **outgoing routes**.

Changes at a group of routes at the **same time** may be caused by a **common root cause**.

Routes 1, 5, 9, 11 are **correlated**!

Routes 1, 5, 11 are **closed together**. They may have been affected by the **same** router **failure**!

We can **isolate** the root causes of the issues!

Finding Dense Correlated Subgraphs

INPUT: a dynamic network, a **correlation** threshold τ, a **density** threshold σ, and a **similarity** threshold ε

GOAL: find all the maximal **τ-correlated σ-dense** subgraphs with **ε-limited overlap**

In static networks, the density of a subgraph is the **average node degree**

In dynamic networks, the average node degree **changes over time**!

Density of a subgraph H in a **dynamic** network:

ρ^k_m is the **minimum** density across the snapshots where H is **active**

ρ^k_a is the **average** density across the snapshots where H is **active**

H is **active** at time t, if **at least k%** of its edges **exists in t**

H is **σ-dense** if $\rho^k(H) \geq \sigma$

Correlation of a subgraph H in a **dynamic** network:

$c_m(H)$ is the **minimum Pearson** correlation between edges in H

H is **τ-correlated** if $c_m(H) \geq \tau$

Similarity of two subgraphs H and J:

s(H, J) is the **Jaccard** similarity between the edge sets

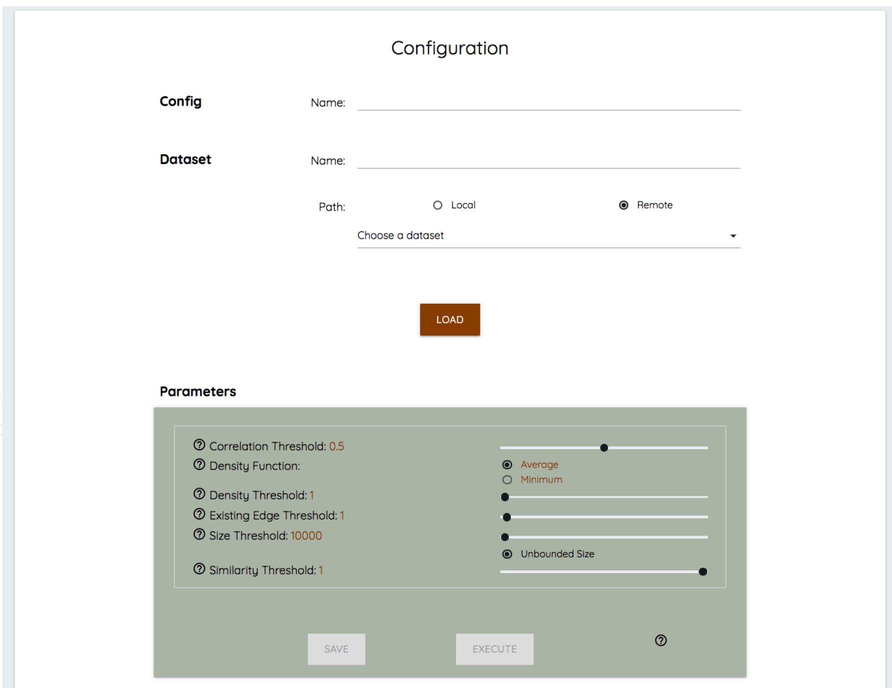
A set of subgraphs have **ε-limited overlap** if all $s(H, J) \leq \epsilon$

HOW?

1. Enumerate **maximal groups** of correlated edges
   a. Build **auxiliary graph** of correlated edges
   b. Find **maximal cliques**
2. Extract **dense subgroups** from each clique
   a. Find **connected components**
   b. For each component, iteratively remove the lowest-degree node **until** the group **becomes dense** or empty

Demo Scenario

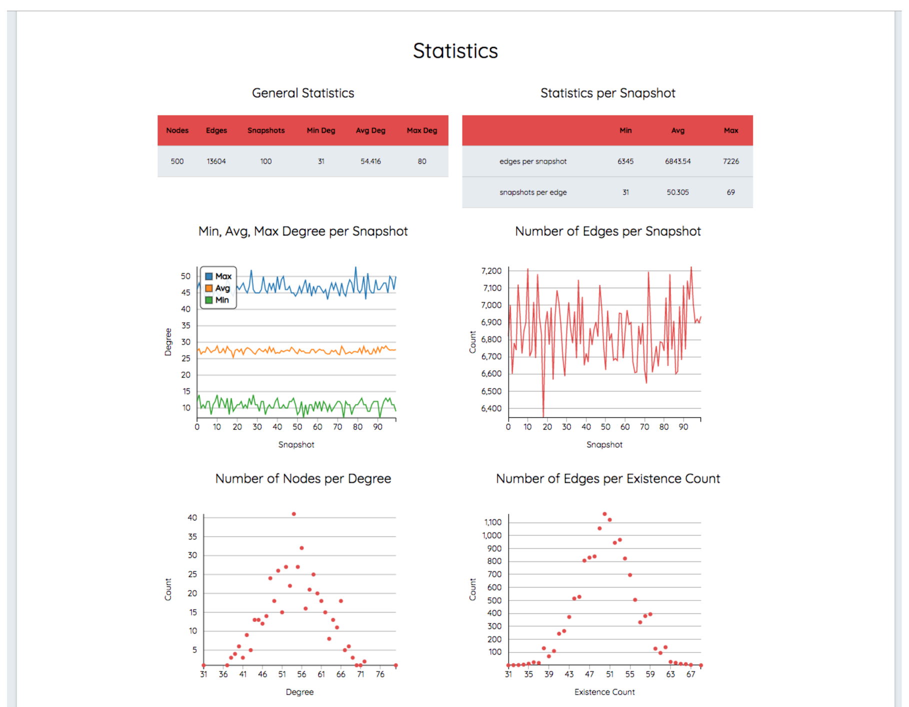
Loading the Network



PARAMETERS
- Correlation between edges
- How the density is computed
- Density for a group of edges
- How many edges to be active
- Upperbound on the group size
- Upperbound on the similarity

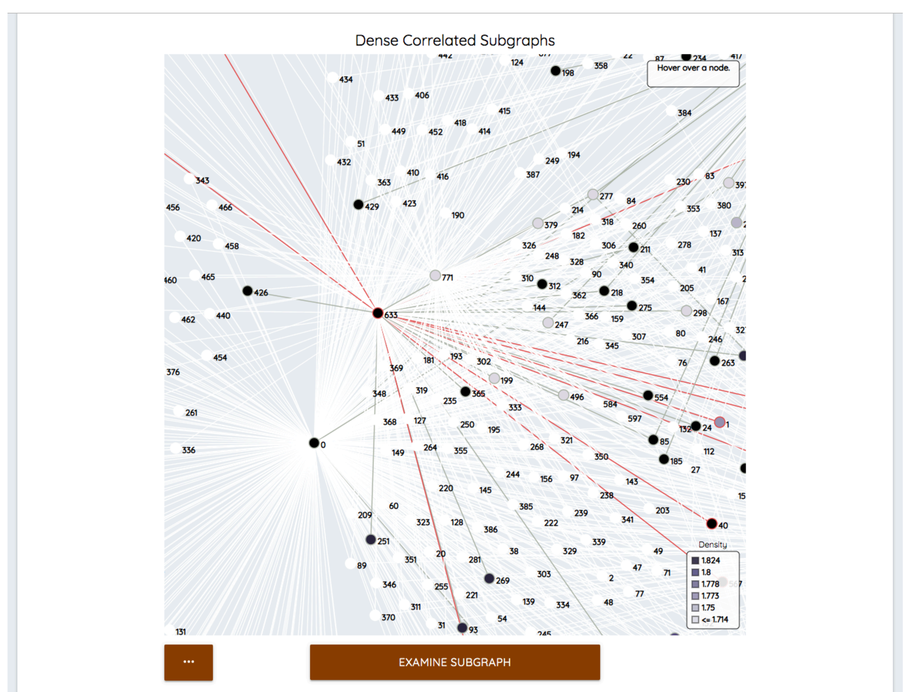
BGP routing topology network
Routing tables from Aug 29 to Aug 31 (Hurricane Katrina)

Network Statistics



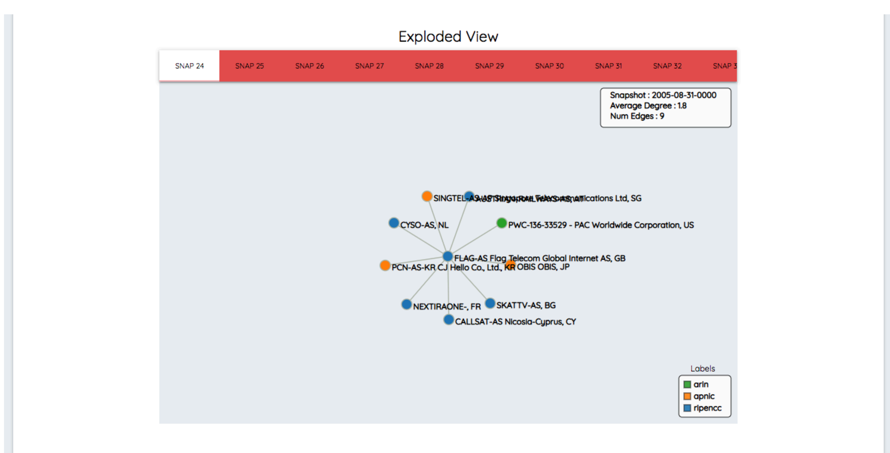
Minimum, average, and maximum node degree in each snapshot
Number of edges per snapshot
Node degree distribution in the union graph
Edge existence distribution

Viewing the Dense Subgraphs



Union graph with dense correlated groups highlighted
Routing paths that changed similarly over time and topologically close
Instability regions originated from the failure of the same router

Subgraph Exploration



Analyze a particular dense subgraph
The network analyst can see
- countries affected by the disaster
- issues caused by the same failure