

A Visual Analytics System for Tracing Suspicious Entities in the Fisheries Domain using Knowledge Graphs

2023 IEEE VAST Challenge MC1: Outstanding Comprehensive Submission

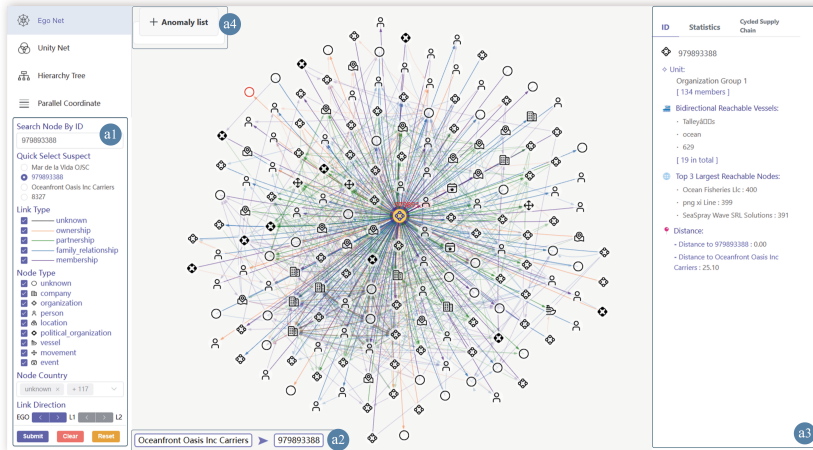
Jingfu Wu, Diyun Lu, Lei Chen, Ningyi Peng, Fan Yang, Xinyu Tang, Yuxin Ma



Southern University of Science and Technology, China

The VAST Challenge 2023 Mini-Challenge 1 provides a knowledge graph generated using Natural Language Processing (NLP) tools extracting entities and relationships from text articles, requiring the participants to figure out entities involved in illegal, unreported, and unregulated fishing (IUU) and extract their patterns.

Our system, FishHook, consists of 4 views embedded with rich interactions and a self-built math model: the Ego Net view, Unity Net view, Hierarchy Tree view and the Parallel Coordinates view. This integrated solution not only exhibits robustness and scalability but also presents a powerful tool for understanding complex knowledge graph



A. Ego Net

The force-directed graph shows the ego-net of the neighbors within the second order of the selected node. The selected node is colored blue, and suspected entries that suggest further exploration are colored red. Node types are encoded as icons and edge types are represented by colors. Single-click on any node highlights it yellow and the corresponding message card will be shown on (a3). Double-click a node to switch it to the center of the ego net, and immediately update the node ID to the history breadcrumbs (a2) for future stepping back.

a1. Filters

Directly search and submit any targeted node's ID to switch the center of the Ego Net view. Nodes' types and countries, edges' types together with in and out degree filtering are enabled for simplifying observation.

a2. History Breadcrumbs

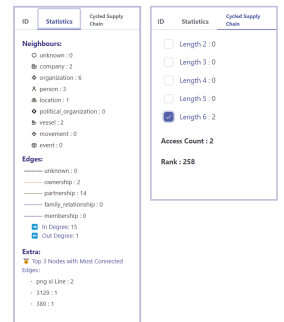
It records the nodes every time we switch the central point of the ego-net to form a path of explored entities.

a3. Message Card

ID Panel: Connected vessels, the neighbor node with the highest edge count, and the distance between this node and the anomalies identified by the LINE algorithm in the feature space.

Statistics Panel: Numerical information about the neighbors and edges can be found here, which corresponds to the content of the Parallel Coordinates view.

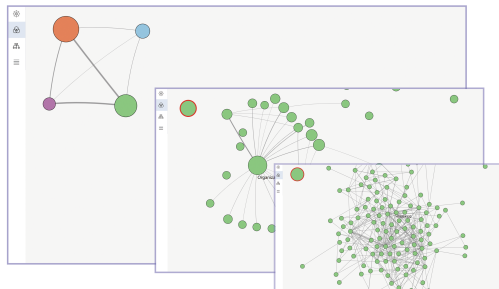
Cycles Panel: Regarding the anomaly cycle of supply chains in potential IUU behaviors, the number of cycles passing through the selected node and the anomalies of the length 2 to 6 are displayed on the information card. Check the box in the card, the related cycles will be highlighted in the ego net.



B. Unity Net

Extracting representative relationships of each type of node, nodes are merged into clusters through the connected relation. The size of the unity node represents its number of members.

Double-click each cluster to see further detailed connections, and double-click on individual nodes seamlessly switches to an ego-net view centered around the selected node.

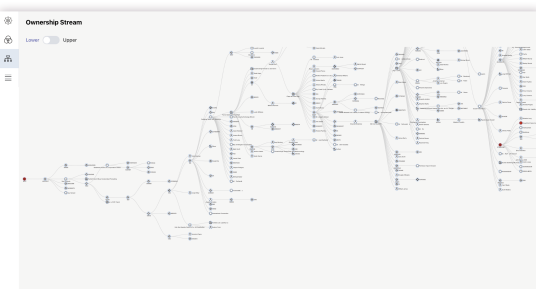


C. Hierarchy Tree

By traversing through the in and out ownership connections, the graph has been expanded into a tree structure.

Click on a node to expand or collapse its subtree. The clicked node is highlighted in blue, while suspicious nodes are in red.

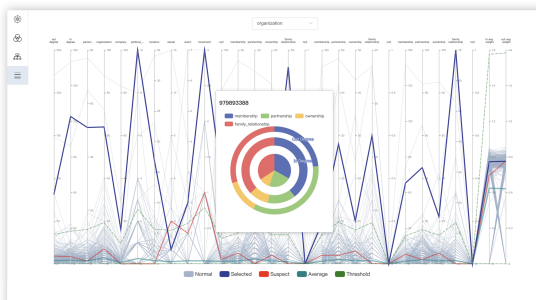
A cycled ownership stream including node 979893388 is detected.



D. Parallel Coordinates

Each line represents a feature vector of an entity's statistical and structural factors of its first and second-order neighbors, where hovering triggers the corresponding piechart.

Nodes of the same type are placed together for comparison. The selected node is colored blue, together with a pair of green threshold lines of 3-sigma for outliers identification.



a4. Anomaly List

Anomaly list shows nodes with high anomaly score (>60) under the mathematical model we construct considering both the graph structure of nodes and semantic information within the context of IUU.

$$S = R_n - \hat{R}_n + 0.1 \times F_{12} + P_{12} - 0.5 \times (\text{dist}_{(n, \text{node } 979893388)}) + \text{dist}_{(\text{node } 979893388, \text{Oceanfront})} + C \quad (1)$$

$$R_n = \frac{V_{\text{max}}}{P_{\text{max}}} \quad (2)$$

Switch to the *show in net*, nodes in this list will be marked red, whose darkness is positive correlated with its anomaly score. This list allows adding new nodes by right-clicking on the node and removal by clicking the delete button.

