VAST CHALLENGE 2023
ILLEGAL, UNREPORTED AND UNREGULATED FISHING BUSTERS

Deng Zhan, Liu Jiaqi, Wilson Tan Xin Xian
SMU School of Computing and Information System

Introduction

FishEye International, a non-profit focused on countering illegal, unreported, and unregulated (IUU) fishing. FishEye has transformed the database into a knowledge graph which includes information about companies, owners, workers, and financial status. FishEye is aiming to use this graph to identify anomalies that could indicate a company is involved in IUU. The visual analytics tool which is to identify illegal fishing is motivated by the need to protect marine resources, conserve ecosystems and sustain fisheries.

Challenge Selection

Mini-Challenge 3: Detect Anomalies

The goal is to use various visual analytics to help FishEye better understand fishing business anomalies, as well as identify companies that may be engaged in illegal fishing.

1. Use visual analytics to identify anomalies in the business groups present in the knowledge graph.
2. Develop a visual analytics process to find similar businesses and group them. This analysis should focus on a business's most important features and present those features clearly to the user.
3. Measure similarity of businesses that you group in the previous question. Express confidence in your groupings visually.
4. Based on your visualizations, provide evidence for or against the case that anomalous companies are involved in illegal fishing.

Methodology

The app involved the usage of jsonlite, tidygraph, ggraph, visNetwork, graphlayouts, ggforce, skimr, tidytext, tidyverse, DT, ggiraph, tm, topicmodels, plotly, shiny, shinythemes, bslib, ggplot2, shinycssloaders, topicdoc packages.

Results and Insights

1. Anomalies Detection

Approach 1: Identifying Companies with Abnormal Structure.

In this approach, we aim to detect companies that might be involved in IUU fishing through their ownership and contact patterns. Companies with an abnormal or disproportionate number of beneficial owners or company contacts relative to their declared revenue are considered suspicious. Such companies may raise red flags for potential involvement in IUU fishing.

After analyzing the data, Aqua Aura SE Marine Life has emerged as the most suspicious company. It is associated with a total of 33 beneficial owners and 8 company contacts, indicating a significant deviation from the norm.
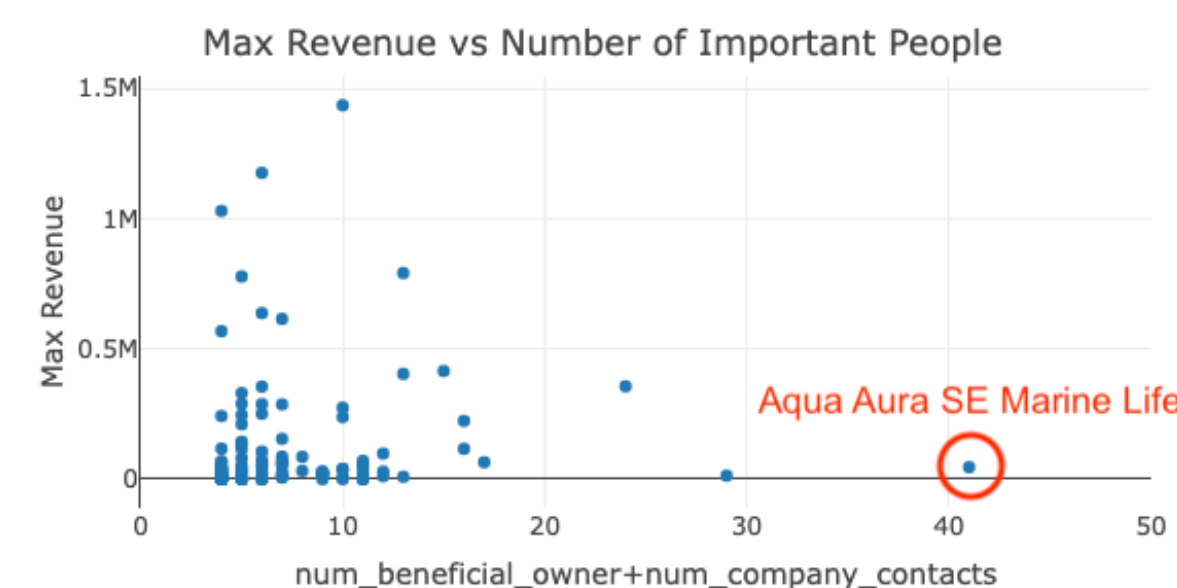
Approach 2: Identifying Companies Registered in Multiple Countries

In this approach, we focus on identifying companies that are registered in more than three countries. The rationale behind this is that companies operating in multiple countries might have complex ownership structures, which could make it easier for them to engage in IUU fishing activities across various jurisdictions.

Aqua Aura SE Marine Life is flagged again in this approach as it has operations in 9 countries.
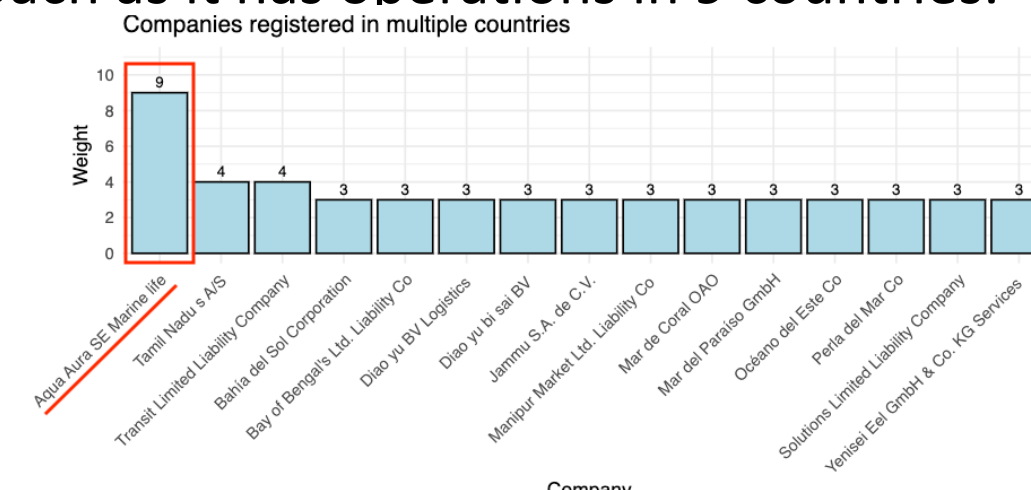
2. Similar Business Grouping

LDA topic modelling was used to group the nodes based on their product_services. Basic text analysis preprocessing was performed such as lowercase conversion, punctuation, numbers and stopwords removal. The data was then fitted into an LDA model of between 3 to 8 topics.

One limitation is that a substantial number of nodes do not have product service data available, so there is no meaningful groups assigned for them.
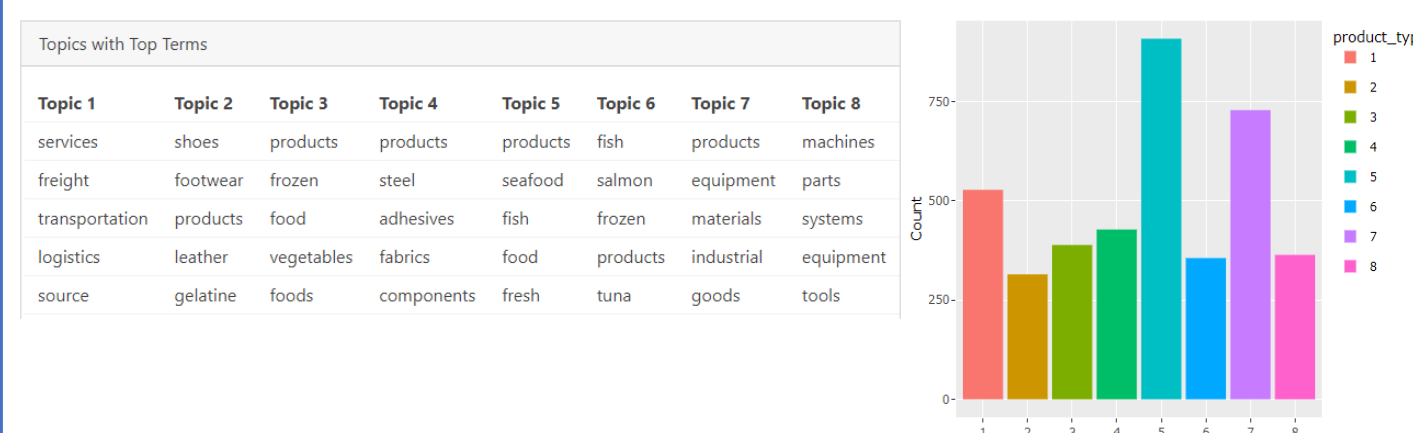
3. Similarity Comparison of Groupings

As businesses were grouped using topic modelling based on their product_services description into product types, the distribution of businesses' revenue and country can be examined to determine if businesses within the same grouping are similar or not.

Boxplot visualisation was used for revenue, as it is a continuous variable, while histogram visualisation was used for country counts.

As the design of the Shiny Application allows users to determine the number of business groups (between 3 to 8), conclusions may differ based on users' inputs. However, based on interacting with the application, our group has observed that different product types are clearly dominated by different countries. This gives us confidence that the topic modelling in part (2) does yield meaningful business groupings.

For revenue, it is not as useful to measure the similarity between business groups, as the distribution of revenue across different product types is relatively similar and the variance of revenue within each product group is quite wide.
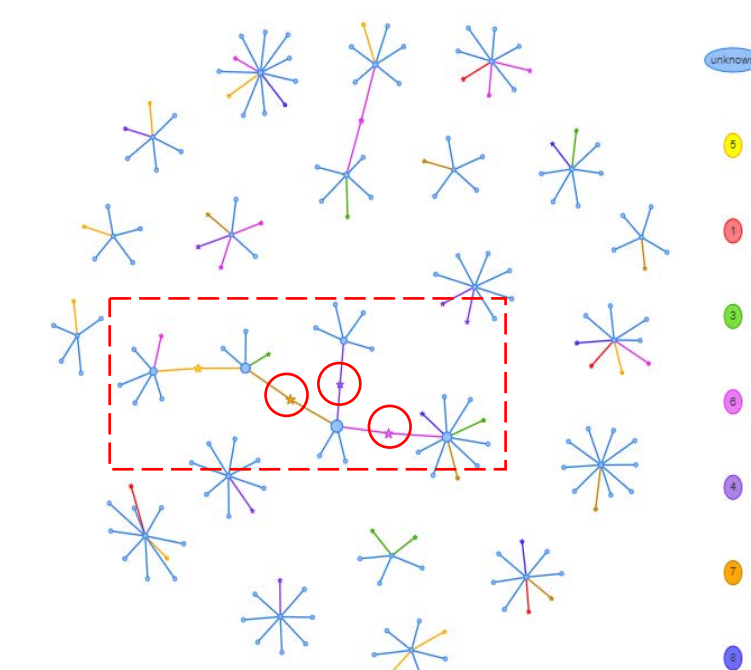
4. Anomalous Companies

Companies with an abnormal or disproportionate number of beneficial owners relative to their declared revenue are suspicious.

- *Aqua Aura SE Marine Life (The company has 33 beneficial owners and registered in 9 countries)*

On the other hand, a person who owns many companies might be involved in IUU as well. We plotted network focusing on the beneficial owners who have >= 5 companies. From the network graphs below, we can observe a big cluster includes several high betweenness companies. Beneficial owner "Jessica Brown" is the essential linkage in this cluster. Therefore her 1-hop neighbouring companies are likely to be involved in IUU fishing.

- *BlueTide GmbH & Co. KG*
- *Mar del Oeste*
- *West Fish GmbH Transport*

Conclusions

This Rshiny App allows users to have a closer look at the network and zoom in to anomalous structures and companies that might be involved in IUU fishing with full extent of interactivity and adjustable input. Currently some variables value are arbitrarily set such as the number of topics and company ownership threshold, and they might not be the best choice. Future work can be done to improve on the optimum value selection.

Data Used

https://vast-challenge.github.io/2023/MC3.html