Birds of a feather: A method for detecting suspicious clusters of companies in the UK

Keywords: firm ownership networks, shell companies, community detection, machine learning, network analysis

Extended Abstract

The international financial system facilitates the construction of complex networks of company ownership that are often exploited for malicious intent [5]. The industry of offshore service providers help hide, obscure, and launder trillions of dollars each year while leaving little trace. Although this issue has long been recognised, governments and the consortiums designed to fight corruption acknowledge that we are no closer to dismantling corruption networks as these networks adapt and innovate faster than policymakers can make policy. Recent implementations of databases on company directors, beneficial ownership, and their physical location were intended to reduce the effectiveness of shell companies as a vehicle for obscuring information [3, 2]. Nevertheless, shell companies continue as the preferred instrument for sheltering profits from oversight. We assert that databases on companies are currently under-exploited and propose a method for detecting clusters of suspicious companies. We argue that company ownership structures adhere to patterns of specialists and generalists, where some clusters of companies under shared ownership self-organize based on their domain expertise ('specialists'), while others diversify the industries they participate in by owning companies in many areas ('generalists') [1]. Though there exist legitimate reasons for clusters of companies to be diversified, we find evidence that a handful of generalist clusters may actually be specialists at hiding wealth through the formation of shell companies (Fig. 1). We generate a network of company ownership in the United Kingdom and use data from the International Consortium of Investigative Journalism to train a machine learning algorithm to detect suspicious companies. Our findings help re-conceptualize the role of firms who use domain expertise to circumvent government efforts at revealing the real owners of companies and offers a methodology that can adapt to evolving strategies of obscuring beneficial ownership.

Specifically, we rely on a sample of 500,000 companies and their officers in the UK's Companies House registrar. Each company is required to list their directors and address, and they must comply with other legal requirements reported in the registrar. We collect this information and match companies that appear in the Panama and Pandora paper leaks [4]. These data are used to construct a network of company ownership within the UK. Data on leaked companies compliance, registration, and position within the network (characterised through node level metrics, and community membership) are used to train a random forest model to identify companies resembling the shell companies that appear in the leaked documents. The model is applied to the full sample of companies and used to generate a "suspiciousness score" for each company, indicating the extent to which it resembles those companies named in the leaks. Including information on community membership not only reduces the number of false positive suspicious companies identified by the machine learning algorithm but additionally facilitates a focus on communities containing a large concentration of potential shell companies. We argue that although clusters of nodes with shared ownership within these communities may look like generalist or diversified investments, they may actually be specialists in the creation of companies that can be used for illicit purposes. Our approach provides law enforcement and policymakers with a tool to identify groups of companies that merit attention, facilitating a more efficient use of enforcement resources.

References

[1] Fricke, D. and Roukny, T. (2020). Generalists and specialists in the credit market. *Journal of Banking and Finance*, 112.

[2] Luna, D. K., Palshikar, G. K., Apte, M., and Bhattacharya, A. (2018). Finding shell company accounts using anomaly detection. In *Proceedings of the ACM India Joint International Conference on Data Science and Management of Data*, pages 167–174.

[3] Rocha-Salazar, J.-d.-J., Segovia-Vargas, M.-J., and Camacho-Miñano, M.-d.-M. (2022). Detection of shell companies in financial institutions using dynamic social network. *Expert Systems with Applications*, 207:117981.

[4] Tiwari, M., Ferrill, J., and Mehrotra, V. (2022). Using graph database platforms to fight money laundering: advocating large scale adoption. *Journal of Money Laundering Control*, ahead-of-print(ahead-of-print).

[5] Tiwari, M., Gepp, A., and Kumar, K. (2020). A review of money laundering literature: the state of research in key areas. *Pacific Accounting Review*.

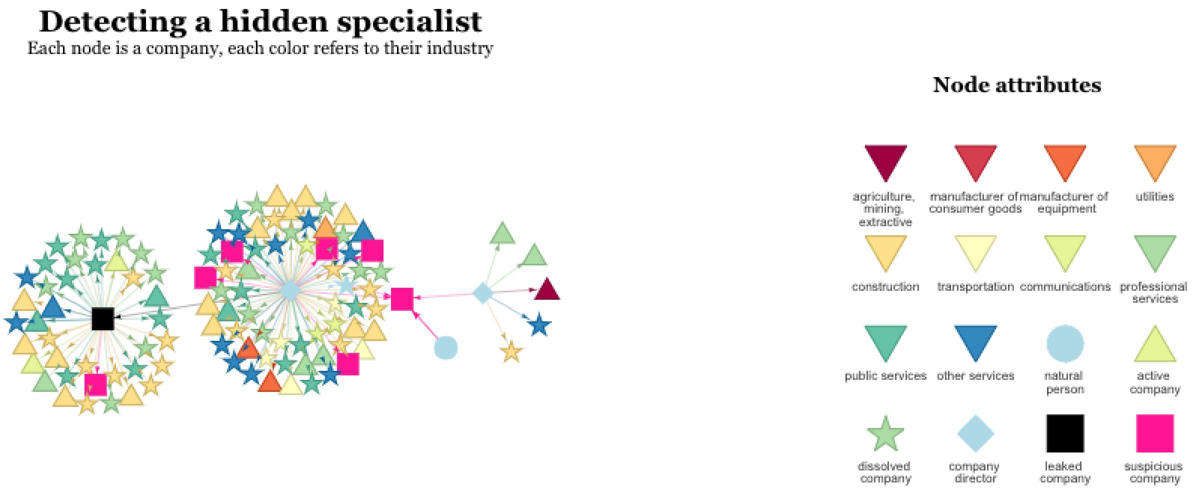

Figure 1: A "generalist" community of companies owned by the same individual. The black square is a company reported in a leak, the pink squares are those identified by the algorithm as suspicious.