EuroVis Workshop on Visual Analytics (2014)
M. Pohl and J. Roberts (Editors)

A Visual Analytics Field Experiment to Evaluate Alternative Visualizations for Cyber Security Applications

F. Fischer[1], J. Davey[2], J. Fuchs[1], O. Thonnard[3], J. Kohlhammer[2,4] and D. A. Keim[1]

[1]University of Konstanz, Germany
[2]Fraunhofer IGD, Germany
[3]Symantec Research Labs, France
[4]Technische Universität Darmstadt, Germany

Abstract

The analysis and exploration of emerging threats in the Internet is important to better understand the behaviour of attackers and develop new methods to enhance cyber security. Fully automated algorithms alone are often not capable of providing actionable insights about the threat landscape. We therefore combine a multi-criteria clustering algorithm, tailor-made for the identification of such attack campaigns with three interactive visualizations, namely treemap representations, interactive node-link diagrams, and chord diagrams, to allow the analysts to visually explore and make sense of the resulting multi-dimensional clusters. To demonstrate the potential of the system, we share our lessons learned in conducting a field experiment with experts in a security response team and show how it helped them to gain new insights into various threat landscapes.

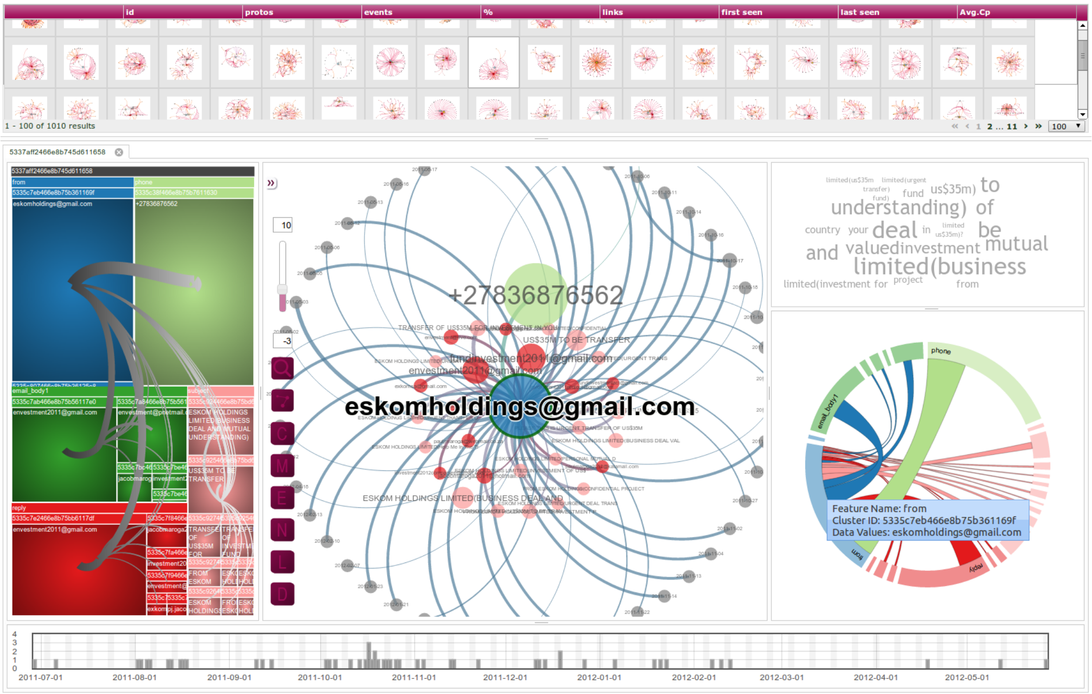
Categories and Subject Descriptors (according to ACM CCS): C.2.0 [Computer-Communication Networks]: General—Security and protection C.3.8 [Computer Graphics]: Application—H.5.2 [Information Interfaces and Presentation]: User Interfaces—

1. Introduction and Related Work

The behavioural analysis of attackers in the Internet is a challenging, but highly relevant field of research. It is important to understand their modus operandi to mitigate attacks and develop new methods to protect network infrastructures, customers, and to identify fraud. However, threat actors may belong to various organizations that operate in different ways making it hard to differentiate them based on common behaviour. Fully automated data mining algorithms can help to address this challenge, but when used alone they are often not capable of providing actionable insights, because human analysts can hardly understand the results generated by these algorithms. In this work we leverage the multi-criteria clustering algorithm TRIAGE [TMD10], which was designed to support threat intelligence and attack investigation tasks. We integrate three interactive visualizations developed within the VIS-SENSE [VS13] project to facilitate the interpretation and sense-making of the resulting multi-dimensional clusters. The applicability of the algorithmic approach has been previously shown in different uses cases on various datasets [ITC*13, TBO*12, TD11, CLT*10].

In the field of visual analytics, evaluation is quite challenging [vW13]. On the one hand, real-world scenarios often have no ground truth, and on the other hand, only experts can identify and validate insights. User studies in the lab are not an appropriate or realistic approach to judge the usefulness of visual analytics applications, which require in-depth domain knowledge. Shneiderman and Plaisant propose the use of “Multi-dimensional In-depth Long-term Case studies (MILCs)” [SP06], which is a promising long-term evaluation approach. However, this is hard to achieve in practice due to the lack of financial support and willingness of experts to participate in such studies. Shiravi et al. [SSG12] provide an extensive overview of visualization systems for network security and conclude that “only a couple have performed usability studies”. It is also stated, that “one of the reasons that security visualization systems, despite their great potential, are not often incorporated [...] is the result of failing to address the focal points of user experience”. We tried to address this issue and gathered feedback about the user experience for the three visualization techniques. Within VIS-SENSE, we had the chance to conduct a two-day *field experiment* [Car08] with security experts from an

DOI: 10.2312/eurova.20141144



Feature Name: from
Cluster ID: 5335c7eb466e8b75b361169f
Data Values: eskomholdings@gmail.com

Figure 1: *After feature selection and analysis, the shown visualization display can be used to explore the MDC clusters. The small-multiple view at the top can be used to select MDCs. The* Treemap View (TV), *the* Graph View (GV), *and the* Chord View (CV) *show the respective MDC of a well-known scam campaign impersonating the company "Eskom Holdings" [ITC*13].*

operational response team while observing them, how they worked with their own data using our visual analytics application deployed as prototype system on their premises. The three main contributions of our work are: (i) The adaptation of several well-known visualizations to enhance the interactive analysis of threat landscapes. (ii) The implementation of a web-based system to visually explore and make sense of the complex results of the TRIAGE algorithm. (iii) Sharing the results and lessons learned of conducting the field experiment with domain experts.

2. Data Analytics for Threat Intelligence

The TRIAGE algorithm uses a combination of graph-based analysis and data aggregation methods as used in multi-criteria decision analysis [TMD10]. The system can generally be applied to various security-related datasets consisting of individual events, for the purpose of identifying groups of related events that might have a common root cause, *e.g.*, series of cyber attacks sourced by the same attackers or *threat group*. In a spam e-mail dataset, for example, each message represents one event with different features (e.g., sender address, recipient, subject), denoted as F_k, with $k = (1, ..., n)$. For each feature an undirected edge-weighted graph $G_k(V_k, E_k, w_k)$ is created, where the vertices V_k represent the message features, and the edges E_k weighted by the function w_k reflect similarities among messages [TD11]. Afterwards the different weighted graphs G_k are combined using an aggregation function. The resulting multi-dimensional clusters (MDCs) represent groups of events correlated by a number of features, where the combination of correlated features may vary within the same cluster, depending on the data fusion model. In a spam dataset such MDCs are likely to reflect individual spam campaigns containing messages having similar characteristics, and hence a common root cause. A visual analytics system can help the analyst: (i) during feature selection to decide which features to include, (ii) during the parametrization of the aggregation function to incorporate experience and domain knowledge, and (iii) during cluster interpretation to understand the structure of campaigns, formulate hypothesis and attain insights.

3. Making Sense of Data Clusters Using Visualizations

The deployed visual analytics application contains visual dashboards, charts, tables for feature selection, and cluster visualizations to cover the whole analysis workflow. The focus of this study is the usage of several state-of-the-art visualizations that help the analysts to explore multi-dimensional clusters during the final cluster-interpretation phase. In the following, we focus on the usage of three well-known techniques as seen in Figure 1 that can be used to explore individual MDCs: The *Treemap View (TV)*, the *Graph View (GV)*, and the *Chord View (CV)*.

The visualization modules were built on top of our Visual Analytics Suite for Cyber Security (VACS), which is a web-based research framework providing visualizations and a secure REST interface to remote datasets and algorithms of multiple project partners. This modular architecture helps to interdisciplinarily develop visual analytics applications that enables us to work on sensitive datasets and novel algorithms, while preserving the rights of the property owners.

The *Treemap View* in Figure 1 provides an overview of the features, mapped to colour, and their value occurrences. Each coloured rectangle on the upper level represents a feature, containing further rectangles representing cluster prototypes. The more frequently a value, the bigger the corresponding rectangle in the squarified treemap [BHW00]. Interaction enables the user to zoom in and reveal splines to show the event co-occurrences of values in entities. Treemap representations with splines are also used in related security applications [FMK*08] for the exploration of network traffic, while treemaps alone are commonly used to provide overviews for file systems forensics and malware analysis [HPPT08, THGF09].

The *Graph View* shows the relationships between feature values, which is widely used in various security applications [Mar08]. Each node represents a value occurring in the cluster, whereas an edge indicates the co-occurrence of a pair of values in an event of the dataset. The node sizes are mapped to the number of events and the thickness of the edges is determined by the number of co-occurrences. The graph is highly interactive and provides, zooming, panning, re-positioning nodes, and the modification of edge thickness, label size, and node size. To handle large datasets, a sampling can be applied and the layout is calculated on the server-side using Graphviz [EGK*02].

The interactive circular *Chord View* enables the exploration of all relations between the different feature clusters composing the MDC. The circle segments on the edge of the view represent the values, their colour is determined by their feature. Interactive highlighting shows which feature clusters have co-occurring events. When the users selects a feature cluster, the shown chords encode the number of co-occurring events to the other feature clusters. The implementation is based on [BOH11] using an approach similar to Circos [KSB*09], which is widely used to analyse complex datasets.

4. Field Experiment with Security Response Experts

The two-day field experiment was conducted in November 2013 [VS13] and carried out on the premises of Symantec Security Response in Dublin, Ireland and involved six participants with a solid background of cyber security threat analysis. The study was focused on collecting a qualitative assessment of the visual analytics system and to evaluate the user experience of the interactive visualizations. It consisted of three phases. First, a general introduction to the goals and results of the VIS-SENSE project were given, followed by an interactive demonstration using data known by the project partners. Then, the main part of the field experiment consisted in a hands-on session, in which the participants used the system for analysing their own data. The demonstration was designed to show how a typical exploratory session would be carried out. It illustrated the use of the main functionalities, such as the overview, search, and visualization features and showed how to find, confirm, and explain interesting patterns. The main task was to "explore clusters to understand the reasons why these entities have been grouped together" including questions like: Which customers are targeted? What are the strongest correlative features and characteristics of a campaign? What are the most significant coalitions of features that are linking entities?

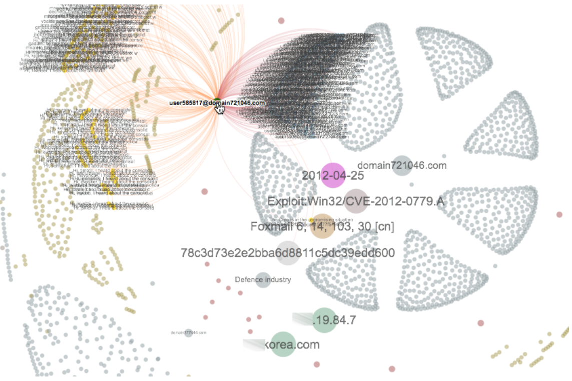

user585817@domain721046.com
domain721046.com
2012-04-25
Exploit:Win32/CVE-2012-0779.A
Foxmail 6, 14, 103, 30 [cn]
78c3d73e2e2bba6d8811c5dc39edd600
Defence industry
.19.84.7
korea.com

Figure 2: *Example of an MDC found during the field study, attributed to a notable espionage campaign.*

4.1. Hands-On Session

In the hands-on session, the participants had approximately two hours to analyse their data with our visual analytics system. There were four users (three analysts and one designer) actively using the TRIAGE application on three laptops. Their dataset consisted of 44 features and approximately 100 000 entities and had not previously been analysed with TRIAGE and was completely unknown to the project partners. Due to a lack of experience with the dataset, the parameters for the clustering algorithms were chosen based on what had worked well previously with other similar datasets. Based on density and cardinality 10 features were selected as input to the clustering algorithm. In spite of these challenges, we were able to find clusters suitable for exploratory analysis and hypothesis formulation and validation. An example is illustrated in Figure 2, which represents a notable cyber espionage campaign that affected two large defence industries in April 2012, and was attributed to the Elderwood gang [Sym12]. To acquaint the analysts with the software, a series of simple, predefined interactive tasks, and general questions were given to the participants. However, the analysts were able to freely use the software to explore their dataset. The project partners observed the participants passively, but were available on request to answer questions and provide guidance to the participants. After the hands-on session an informal discussion was conducted and feedback of the participants was recorded. At the end of the experiment, a summary was presented to all participants and further interested parties.

4.2. Results of the Field Study

During the introductory presentations the participants posed detailed questions about the techniques, hardware, and software. They appeared to see the applicability of the visualizations to their own work. In particular, the participants wanted to know more about the potential for the integration of visualization components into other environments. They stated that their goal in the field experiment was to try out the components to see whether they could achieve the tasks, they previously had to do manually, faster with VIS-SENSE technologies. We found this encouraging and it explained the overall high degree of interaction during the meeting. The users began working on the tasks set for them during the hands-on session and had little trouble achieving basic tasks. In some instances, a few words or a sentence from the observers was required to guide the participants, but no deeper explanations were necessary. The participants diverged from the structured tasks frequently to engage in more exploratory activities, returning occasionally to the tasks. In this way they were able to *test out* each of the interactive features of the interface. Some participants started targeted searches for specific phenomena in the data, copied attributes from the TRIAGE application and compared them manually with other datasets and internal systems. One participant began a deeper exploration of a cluster in GV, repositioning nodes and conducting a closer examination of connections. The cluster showed a cyber-criminal campaign. The participant was able to identify and characterize distinct phases of the campaign and used the visualization to explain the modus operandi of the attacker to another participant. Other similar spontaneous discussions between users about their findings occurred.

The participants had some difficulty acquainting themselves with the UI due to missing UI features and lack of UI documentation. For example, they applied filters in tables and expected similar filtered views of the data in the visualizations. However, this feature was not yet implemented and the lack of linking in the displayed data led to some confusion. They also complaint about the lack of meta-data integration. In general, GV was perceived as the most useful of the three alternative visualizations. Indeed, GV was used most intensively by the participants. To avoid the influence of layout and positional preference and to force the analyst to focus on each visualization individually, each visualization was shown in full-screen and not as integrated display as seen in Figure 1. The other two were tried out initially, but not pursued much subsequently. Participants generally preferred TV to CV while the latter was criticised as lacking usefulness for their workflow. However, it still may be useful for short overviews of relations in very large datasets. A participant commented that their most common workflow is of an investigative nature; drilling down into the data and exploring details. Thus, visualizations focused on providing an overview without possibilities for deeper interactive exploration are not very useful for them. In addition, GV was the most interactive of the three views. Overall, it was concluded, that GV was best suited for their need of detailed structural exploration for medium sized MDCs, TV provided an helpful and compact overview, while the least preferred CV mostly focused on exploration of relations between clusters within a MDC.

A participant commented that the system did open many new possibilities for data exploration and representation. The system was perceived as very useful to speed up analysis tasks. However, the participants provided many constructive suggestions for improvement, in particular for further enhancing user interactions and data analytics capabilities.

5. Conclusions and Future Work

We presented a web-based visual analytics application to analyse multi-dimensional clusters to support the TRIAGE algorithm and enhance attack investigation tasks associated with it. We conducted a field experiment to gather qualitative feedback from domain experts specifically on the usage of three widely used visualizations. Furthermore, we identified primary tasks for alternative visual representations on the basis of the feedback of the analysts. The detailed feedback can be summarized in three areas, which will guide our future work and research directions: (i) Parametrization for the clustering of unknown datasets proved to be challenging, which further strengthened the importance of visual feature and parameter selection to make justified decisions. (ii) The feedback showed the importance of highly interactive visualizations. Slight improvements (e.g., filtering, highlighting multiple elements) in the visualizations can lead to considerable changes in user experience and it may have a strong impact on the usability to solve real-world problems. (iii) Inconsistent design decisions easily cause confusion. In collaboratively developed software, inconsistencies in design are common but should be avoided. The research prototypes discussed in this paper have since been integrated into Symantec's internal research framework to analyse security datasets and are actively used for other activities.

Acknowledgments

We thank Symantec Security Response in Dublin for their great participation despite their full schedule and the reviewers for their valuable comments. The research leading to these results has received funding from the European Commission's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 257495, "Visual Analytic Representation of Large Datasets for Enhancing Network Security" (VIS-SENSE).

References

[BHW00] BRULS M., HUIZING K., WIJK J.: Squarified treemaps. In *Data Visualization 2000*, Leeuw W., Liere R., (Eds.), Eurographics. Springer Vienna, 2000, pp. 33–42. doi:10.1007/978-3-7091-6783-0_4. 3

[BOH11] BOSTOCK M., OGIEVETSKY V., HEER J.: D3 data-driven documents. *IEEE Transactions on Visualization and Computer Graphics 17*, 12 (Dec. 2011), 2301–2309. doi:10.1109/TVCG.2011.185. 3

[Car08] CARPENDALE S.: Evaluating information visualizations. In *Information Visualization*, Kerren A., Stasko J., Fekete J.-D., North C., (Eds.), vol. 4950 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2008, pp. 19–45. doi:10.1007/978-3-540-70956-5_2. 1

[CLT*10] COVA M., LEITA C., THONNARD O., KEROMYTIS A., DACIER M.: An analysis of rogue av campaigns. In *Recent Advances in Intrusion Detection*, Jha S., Sommer R., Kreibich C., (Eds.), vol. 6307 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2010, pp. 442–463. doi:10.1007/978-3-642-15512-3_23. 1

[EGK*02] ELLSON J., GANSNER E., KOUTSOFIOS L., NORTH S., WOODHULL G.: Graphviz— open source graph drawing tools. In *Graph Drawing*, Mutzel P., Jünger M., Leipert S., (Eds.), vol. 2265 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2002, pp. 483–484. doi:10.1007/3-540-45848-4_57. 3

[FMK*08] FISCHER F., MANSMANN F., KEIM D. A., PIETZKO S., WALDVOGEL M.: Large-scale network monitoring for visual analysis of attacks. In *Proceedings of the 5th International Workshop on Visualization for Computer Security* (Berlin, Heidelberg, 2008), VizSec '08, Springer-Verlag, pp. 111–118. doi:10.1007/978-3-540-85933-8_11. 3

[HPPT08] HEITZMANN A., PALAZZI B., PAPAMANTHOU C., TAMASSIA R.: Effective visualization of file system access-control. In *Visualization for Computer Security*, Goodall J., Conti G., Ma K.-L., (Eds.), vol. 5210 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2008, pp. 18–25. doi:10.1007/978-3-540-85933-8_2. 3

[ITC*13] ISACENKOVA J., THONNARD O., COSTIN A., BALZAROTTI D., FRANCILLON A.: Inside the scam jungle: A closer look at 419 scam email operations. In *Security and Privacy Workshops (SPW), 2013 IEEE* (May 2013), pp. 143–150. doi:10.1109/SPW.2013.15. 1, 2

[KSB*09] KRZYWINSKI M. I., SCHEIN J. E., BIROL I., CONNORS J., GASCOYNE R., HORSMAN D., JONES S. J., MARRA M. A.: Circos: An information aesthetic for comparative genomics. *Genome Research* (2009). doi:10.1101/gr.092759.109. 3

[Mar08] MARTY R.: *Applied Security Visualization*, 1 ed. Addison-Wesley Professional, 2008. 3

[SP06] SHNEIDERMAN B., PLAISANT C.: Strategies for evaluating information visualization tools: Multi-dimensional in-depth long-term case studies. In *Proceedings of the 2006 AVI Workshop on BEyond Time and Errors: Novel Evaluation Methods for Information Visualization* (New York, NY, USA, 2006), BELIV '06, ACM, pp. 1–7. doi:10.1145/1168149.1168158. 1

[SSG12] SHIRAVI H., SHIRAVI A., GHORBANI A.: A survey of visualization systems for network security. *Visualization and Computer Graphics, IEEE Transactions on 18*, 8 (Aug 2012), 1313–1329. doi:10.1109/TVCG.2011.144. 1

[Sym12] SYMANTEC SECURITY RESPONSE: The Elderwood Project. http://www.symantec.com/connect/blogs/elderwood-project, Sep 2012. 3

[TBO*12] THONNARD O., BILGE L., O'GORMAN G., KIERNAN S., LEE M.: Industrial espionage and targeted attacks: Understanding the characteristics of an escalating threat. In *Research in Attacks, Intrusions, and Defenses*, Balzarotti D., Stolfo S., Cova M., (Eds.), vol. 7462 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2012, pp. 64–85. doi:10.1007/978-3-642-33338-5_4. 1

[TD11] THONNARD O., DACIER M.: A strategic analysis of spam botnets operations. In *Proceedings of the 8th Annual Collaboration, Electronic Messaging, Anti-Abuse and Spam Conference* (New York, NY, USA, 2011), CEAS '11, ACM, pp. 162–171. doi:10.1145/2030376.2030395. 1, 2

[THGF09] TRINIUS P., HOLZ T., GOBEL J., FREILING F.: Visual analysis of malware behavior using treemaps and thread graphs. In *Visualization for Cyber Security, 2009. VizSec 2009. 6th International Workshop on* (Oct 2009), pp. 33–38. doi:10.1109/VIZSEC.2009.5375540. 3

[TMD10] THONNARD O., MEES W., DACIER M.: On a multicriteria clustering approach for attack attribution. *SIGKDD Explor. Newsl. 12*, 1 (Nov. 2010), 11–20. doi:10.1145/1882471.1882474. 1, 2

[VS13] VIS-SENSE: Deliverable 6.3 - VIS-SENSE Framework Evaluation, 2013. URL: http://www.vis-sense.eu/. 1, 3

[vW13] VAN WIJK J.: Evaluation: A challenge for visual analytics. *Computer 46*, 7 (July 2013), 56–60. doi:10.1109/MC.2013.151. 1