Assignment 2

Author: Matthias Tilsner
Student-ID: 200882645
Course: COMP 6785

Submission: January 30, 2009

1 Chose two topics from the list you provided in Assignment 1. Find two recent peer-reviewed research papers for each topic, and provide a brief summary of the primary contributions of each paper

1.1 Topic 1: Visualizing Network Traffic for Administrators

1.1.1 Description

As distributed software systems become more and more popular, networks are increasingly used in order to deploy and provide applications. These networks require special attention by administrators assuring the ongoing operation of the network and maintaining a level of security enabling the clients to operate undisturbed and uncompromised. The visualization of ongoing network traffic can greatly improve this administration by visualizing life data thus simplifing the cognition of threats such as attacks or security holes. Other notable papers in this area are [11, 10, 12, 14, 4, 5, 28, 18].

1.1.2 Paper 1: [3]

This paper introduces the Visual Information Security Utility for Administration Live (VISUAL): a toll that visualizes patterns in the communication between external hosts and internal clients. By looking at these patterns, an administrator can easily acquire insights about network activity. The entire network traffic can hereby be monitored simultaneously, allowing the administrator both to get a good grasp of the overall situation. Applying Shneiderman's directive [23] to provide first an overview, then zooming in and filtering, and finally providing details, the administrator can then zoom in to acquire additional information such as ports, IP addresses, or the relative workload. This way, specific points of interest, such as increased communication with critical hosts, can be identified and examined.

1.1.3 Paper 2: [1]

This paper concentrates on detecting network intrusion by focusing on port activity. It suggests displaying the aggregated port activity in a stacked histogram to provide the administrator with an overview

of the current situation. By again applying Shneidermans's overview-zoom in and filter-details directive from [23], the administrator can afterwards zoom into those histograms to acquire more information about certain points of interest. The paper furthermore describes a number of intrusion attacks and demonstrates how the developed visualization can be utilized to detect them.

1.2 Topic 2: Visualizing search results

1.2.1 Description

Especially when retrieving information, search engines are required to order and sort the list of documents matching the query in order to equip the user with the means to easily find whatever he is looking for. Unfortunately, this list is commonly provided in a format that is hard to comprehend for the user. While programs and algorithms have no trouble comparing and evaluating a list of several thousand entries, human users are often overburdened. Especially when searching for words with different domains such as *Java*[1], common sorting algorithms often fail to provide the user with the means to grasp the results. Visualization can help the user to comprehend the result list and to understand the relationship of individual result items. This way, the user can identify the results that are relevent for him more easily. Other notable papers in this area are [26, 20, 6, 8, 25, 17, 2, 16, 19, 13, 24, 9, 15, 27, 21].

1.2.2 Paper 1: [22]

This paper compares textual, two-dimensional, and three-dimensional visualization for displaying search results. It discovers that finding a specific target is much easier in a textual presentation, while clustering and classification of documents benefit from 2D and 3D presentations. Furthermore, even though the three-dimensional application did have the worst performance, it also had the best learning rate, indicating that a trained user might be able to work with it very efficiently. In fact, test subject with greater computer skills responded more rapidly when working with the tree-dimensional presentation than when limited to two dimensions.

1.2.3 Paper 2: [29]

This paper concentrates on putting the results of a web search into relationship with one another. It provides a tool that retrieves the content of the first 50kB for each search result. By comparing the content of each result, similarities are discovered and the documents are put into relationship with one another. The result is a two-dimensional graph displaying the results as a map. Each result is presented by an icon with its size corresponding to the documents position in the result list. The icons are placed so that their distance corresponds to their similarity. Furthermore, similar documents are connected by threads. A table displays all details of the documents and by selecting an icon the downloaded content is displayed thus allowing a quick preview of the document. Especially the table was often used by the users, suggesting an increased emphasis on details. The lack of following Shneiderman's directive from above might have provided a larger acceptance.

[1] Java is both a programming language and an island

2 Briefly summarize what the "opponent process theory of colour" means for selecting color values for use in information visualization systems

The *opponent process theory of colour* postulates that the signals received by the three cones in the human eye are encoded into three different color channels: red-green, blue-yellow, and black-white. These channels are aligned orthogonal to one another, forming a octahedron. The HSB color space is based on this octahedron. A specific color is described by its **H**ue, representing the angle of spin around the black-white axis, its **S**aturation, meaning the distance from this axis, and by its **B**rightness, being the position on that axis.
The three channels result in in six different basic colors placed on their far ends: red, blue, green, yellow, white and black. Most cultures recognize these colors as the primary colors identifying and naming them before any other color as shown in [7]. Since those colors are cross-culturally identified as the most dominant and relevant colors, they should be used in information visualization for identifying distinct data categories. Due to their distance on the channels, these colors are easy to identify and to separate from one another. This way, the distinction between the different data categories is alleviated. When labeling continuous data, on the other hand, the different colors of the used spectrum must be perceptually ordered. This is necessary to ensure that all viewers can identify the order of the different colors correctly. By using a monotonous change on one or more of the color channels – for example from blue to yellow, from red to green, or from blue to green – this can be achieved.

References

[1] K. Abdullah, C. Lee, G. Conti, and J. Copeland. Visualizing network data for intrusion detection. In *Proceedings of the Sixth Annual IEEE SMC. Information Assurance Workshop*, pages 100–108, June 2005.

[2] O. Alonso and R. Baeza-Yates. Alternative implementation techniques for web text visualization. In *Proceedings of the First Latin American Web Congress*, pages 202–204, Nov. 2003.

[3] R. Ball, G. A. Fink, and C. North. Home-centric visualization of network traffic for security administration. In *Proceedings of the 2004 ACM Workshop on Visualization and Data Mining for Computer Security*, pages 55–64, New York, NY, USA, 2004. ACM.

[4] R. Becker, S. Eick, E. Miller, and A. Wilks. Dynamic graphics for network visualization. In *Proceedings of the First IEEE Conference on Visualization*, pages 93–96, 467, Oct 1990.

[5] R. Becker, S. Eick, and A. Wilks. Visualizing network data. *IEEE Transactions on Visualization and Computer Graphics*, 1(1):16–28, Mar 1995.

[6] C. Benjamin, W. Woon, and K. Wong. A graphical and convenient tool for document comparison and visualization. pages 362–367, May 2008.

[7] B. Berlin and P. Kay. *Basic Color Terms: Their Universality and Evolution.* University of California Press, 1969.

[8] K. Einsfeld, S. Agne, M. Deller, A. Ebert, B. Klein, and C. Reuschling. Dynamic visualization and navigation of semantic virtual environments, July 2006.

[9] O. Hoeber. *A Study on Interactive Visualization for Web Information Retrieval.* PhD thesis, Department of Computer Science, University of Regina, Regina, SK, Canada, April 2007.

[10] M. Iliofotou, P. Pappu, M. Faloutsos, M. Mitzenmacher, S. Singh, and G. Varghese. Network monitoring using traffic dispersion graphs (tdgs). In *Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement*, pages 315–320, New York, NY, USA, 2007. ACM.

[11] D. Keim, F. Mansmann, J. Schneidewind, and T. Schreck. Monitoring network traffic with radial traffic analyzer, 31 2006-Nov. 2 2006.

[12] A. Komlodi, J. R. Goodall, and W. G. Lutters. An information visualization framework for intrusion detection. In *CHI '04 extended abstracts on Human Factors in Computing Systems*, page 1743, New York, NY, USA, 2004. ACM.

[13] M. Konchady, R. D'Amore, and G. Valley. A web based visualization for documents. In *Proceedings of the 1998 workshop on New Paradigms in Information Visualization and Manipulation*, pages 13–19, New York, NY, USA, 1998. ACM.

[14] E. L. Malécot, M. Kohara, Y. Hori, and K. Sakurai. Interactively combining 2d and 3d visualization for network traffic monitoring. In *Proceedings of the 3rd International Workshop on Visualization for Computer Security*, pages 123–127, New York, NY, USA, 2006. ACM.

[15] S. Mukherjea and Y. Hara. Visualizing world-wide web search engine results. In *Proceedings of the 1999 IEEE International Conference on Information Visualization*, pages 400–405, 1999.

[16] S. Mukherjea, K. Hirata, and Y. Hara. Visualizing the results of multimedia web search engines. In *Proceedings of the 1996 IEEE Symposium on Information Visualization*, pages 64–65, 122, Oct 1996.

[17] L. T. Nowell, R. K. France, D. Hix, L. S. Heath, and E. A. Fox. Visualizing search results: some alternatives to query-document similarity. In *Proceedings of the 19th annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 67–75, New York, NY, USA, 1996. ACM.

[18] A. Oline and D. Reiners. Exploring three-dimensional visualization for intrusion detection. In *Proceedings of the 2005 IEEE Workshop on Visualization for Computer Security*, pages 113–120, Oct. 2005.

[19] F. Paulovich, R. Pinho, C. Botha, A. Heijs, and R. Minghim. Pex-web: Content-based visualization of web search results. In *Proceedings of the 12th International Conference on Information Visualisation*, pages 208–214, July 2008.

[20] A. Rauber and H. Bina. "'andreas, rauber'? conference pages are over there, german documents on the lower left...": an "old-fashioned" approach to web search results visualization. In *Proceedings*

of the 11th International Workshop on Database and Expert Systems Applications, pages 615–619, 2000.

[21] D. Roussinov. Internet search using adaptive visualization. In *CHI '99 extended abstracts on Human Factors in Computing Systems*, pages 69–70, New York, NY, USA, 1999. ACM.

[22] M. M. Sebrechts, J. V. Cugini, S. J. Laskowski, J. Vasilakis, and M. S. Miller. Visualization of search results: a comparative evaluation of text, 2d, and 3d interfaces. In *Proceedings of the 22nd annual international ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 3–10, New York, NY, USA, 1999. ACM.

[23] B. Shneiderman. The eyes have it: a task by data type taxonomy for information visualizations. pages 336–343, Sep 1996.

[24] M. Tvarozek and M. Bielikova. Personalized view-based search and visualization as a means for deep/semantic web data access. In *Proceeding of the 17th International Conference on World Wide Web*, pages 1023–1024, New York, NY, USA, 2008. ACM.

[25] E. Weippl. Visualizing content based relations in texts. In *Proceedings of the Second Australasian User Interface Conference*, pages 34–41, 2001.

[26] M. Weiss-Lijn, J. McDonnell, and L. James. Visualising document content with metadata to facilitate goal-directed search. In *Proceedings of the Fifth International Conference on Information Visualisation*, pages 71–76, 2001.

[27] P. C. Wong. Adding a visualization feature to web search engines: It's time, Nov.-Dec. 2008.

[28] K. Yoshida, Y. Shomura, and Y. Watanabe. Visualizing network status. In *Proceedings of the 2007 International Conference on Machine Learning and Cybernetics*, volume 4, pages 2094–2099, Aug. 2007.

[29] C. M. Zaina and M. C. C. Baranauskas. Revealing relationships in search engine results. In *Proceedings of the 2005 Latin American Conference on Human-Computer Interaction*, pages 120–127, New York, NY, USA, 2005. ACM.