VisAlert: From Idea to Product

S. Foresti and J. Agutter

Abstract Visalert is a visualization system designed to increase the monitoring and correlation capabilities of computer network analysts engaged in intrusion detection and prevention. VisAlert facilitates and promotes situational awareness in complex network environments by providing the user with a holistic view of network security to aid in the detection of sophisticated and malicious activities, and ability to zoom in-out information of interest. The system provides a mechanism to access data from multiple databases, and to correlate *who*, *what*, *when* and *where*. This chapter describes the design process that enabled the team to go from the conception of rough visual sketches to the implementation and deployment of a finished software. In addition, the chapter describes the issues that the interdisciplinary team had to address to carry the project from idea to product.

1 Introduction

This chapter describes the interactive development process to design a visualization system for computer network security. We want to focus on the design phase of the visualization and show how initial concepts were developed with lessons learned from these concepts. In addition, we wish to illuminate the design choices along the development path that ultimately led to a successful visualization paradigm. First, we will discuss the different people and roles of the interdisciplinary team. Second, we will describe the different design sketches that were developed. Third, we will describe the transformation of these pen–paper based sketches into a refined computer visualization scheme. Finally, we will briefly discuss the move from final static prototype to implementation.

S. Foresti and J. Agutter
University of Utah and Intellivis, Inc. e-mail: stefano@intellivis.com, jim@intellivis.com

1.1 The Project and Team

The project was started by the CROMDI team at the University of Utah. We were awarded a grant by the Intelligence Community (ARDA-DTO-IARPA) to research novel visualizations that would aid network analysts in detecting cyber-anomalies, with particular interest for stealthy attacks that are diluted in time, and very hard to detect.

The Center for the Representation of Multi-Dimensional Information is a Utah State Center of Excellence that performs science, R&D and commercialization of user centered interactive displays. Our team uses an interdisciplinary methodology that integrates cognitive psychology, visual design, computational and visualization methods, and knowledge from domain experts.

CROMDI had previously developed novel displays from idea to commercialization, in medicine. Prior to this project the team had no specific experience in network security: this enabled the team to approach the project by "thinking outside of the box" and come up with a novel visualization method.

1.2 The VisAlert Metaphor

VisAlert enables correlation of heterogeneous network data, leveraging the fact that all events possess what we term the "W^3" premise: *When, Where*, and *What* attributes:

- *When* refers to the point in time when the event happened.
- *Where* refers to the network node, e.g., an IP address, to which the event pertains.
- *What* refers to some indication of the type of the event, e.g., $\$\log = \text{snort},\ \text{gid} = 1,\ \text{sid} = 103\$$.

The visual layout, as shown in Fig. 1 maps the *Where* of an alert into the center of the circle. This is represented via a topology map of the network under scrutiny. The *What* of an alert instance is mapped to the different sections of the outside circular element. The *When* of an alert instance is mapped to the radial sections of the circle moving from most recent (closest to the topology map) to the past as it radiates outward. Alert instances are visualized as lines from the alert type on the outer ring, to the node location in the inner circle. The *When* space is divided into history periods, that show the number of alert instances that occurred in them, while the alert instance lines are only shown for a selected history period only.

Additional visual indicators encode information to increase the situational awareness of the user:

- The icon size increases when nodes experience several alerts. The assumption is that a node that is experiencing multiple unique alerts has a higher probability of malicious activity than one experiencing only one alert. The size will make it stand out and focus the user's attention so he or she can take action.

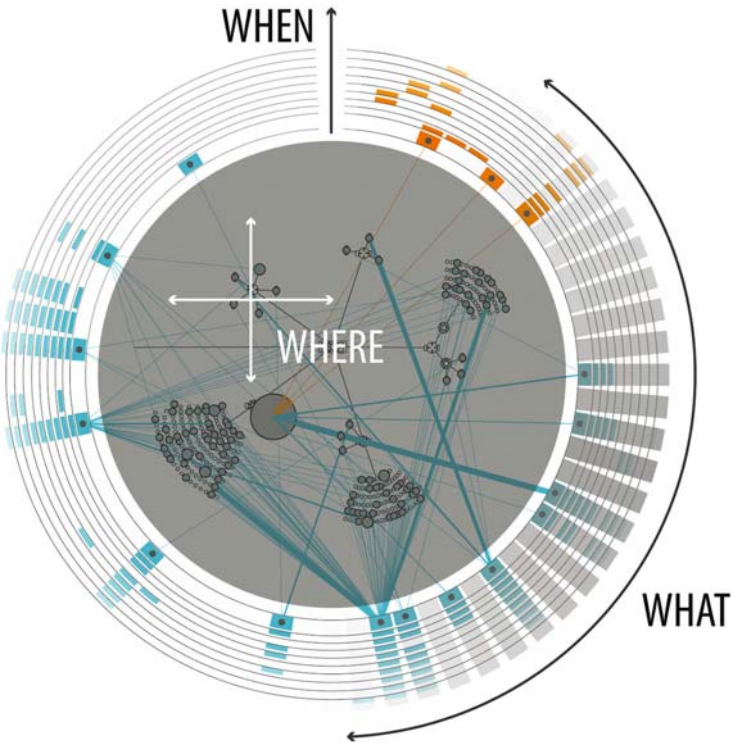

Fig. 1 An event is represented by a line connecting an alert type (*What*) at time (*When*) to a resource (*Where*). VisAlert here exhibits multiple alerts and relevant visual indicators to analysts: alert type via color coding, larger node size with larger number of different alert types, and larger beam size for persistence of the same problem

- The alert beams encode persistence of a particular problem. If a large number of the same alerts are triggered on a particular node over an interval of time the line changes in thickness to show how many alerts. In this manner, continual or recurring problems become evident very quickly, enabling the user to take swift action.
- Color is used to determine user selected ranges and severity levels.

For more detail about the visualization scheme refer to Foresti et al. (2006).

2 Related Work

2.1 Visualization of Network Security

The growing community of VizSec indicates the relevance and number of problems that need to be tackled in network security, where visualization is seen as a potential

solution or aid. Many of these techniques have been demonstrated to be effective at allowing users to see malicious activities such as worm or DoS attacks (Teoh et al., 2002).

One general theme that can be found in the previous work of visualization of security is that it was driven by the identification of specific problems (Teoh et al., 2003) or the monitoring of specific data types (Cox et al., 1996; Estrin et al., 2000). Much research has been done on visual techniques to improve visual pattern matching (D'Amico and Larkin, 2001), and to improve recognition of known events or levels of threat (Polla et al., 1998).

Traditional representations and network alert reporting techniques tend to use a single sensor – single indicator (SSSI) display paradigm. Each sensor has a unique way of representing its information (indicator) and does not depend on information gathered by other sensors. The benefit of such an approach lies in the separation of the various sensors. Each sensor's indicator can thus be optimized for the particular data produced by its sensor, and the user can pick and choose which sensors to use in an analysis. Furthermore, the failure of one sensor does not impact the capability of the rest of the system.

Consequently, the separation between the various sensors is also the weakness of this representation technique. Because each indicator is isolated, the user must observe, condense, and integrate information generated by the independent sensors across the entire enterprise. This process of sequential, piecewise data gathering makes it difficult to develop a coherent, real-time understanding of the interrelationship between the information being displayed.

2.2 Design

The discipline of design has collected a comprehensive knowledge base of the nature, methods, and value of basic 2D and 3D design and their relationship to human collective and individual psychology and behavior. This knowledge base consists of basic principles (e.g., scale, shape, rhythm, color, structure), elements (e.g., line, figures, objects, space) and organizational rules (e.g., hierarchy, layering, symmetry) (Arnheim, 1977; Bogdan, 2002; Wong, 1972, 1977).

Our team determined that information visualization tools are more effective for decision making if developed with an iterative design process that permits simultaneous attention to multiple perspectives, skills and knowledge-bases. We also found that the *design process* allowed for a spontaneous and natural way of socially engaging a wide range of disciplines and individuals working in a very difficult problem. This is in line with existing knowledge that the *design studio model* in general and the design process in particular are a successful working laboratory and methodology for addressing open-ended, fuzzy, and multivariable problems (Cross, 1982; Rowe, 1987).

2.3 Inter-Disciplinary Collaboration

Collaborative success is ultimately grounded in the careful structuring of a team's group dynamics, which are based on clear roles, respect, trust, values, shared goals, and a common language (Friedman, 1997). Accommodating different methods, techniques, positions, interests, standards, languages, perspectives, knowledge, expectations of people from different disciplines takes considerable time and effort, as one has to overcome prejudices each field has of the others (Kraut et al., 1988).

Our team determined that the production of technology that meets users' needs requires the involvement of several roles and perspectives, and that trust among different disciplines is both essential to success, and becomes stronger with the demonstrated value of the work.

3 Technical Approach

3.1 The Team Dynamics

The CROMDI team utilizes an iterative interdisciplinary process to design, a built-in evaluation process to verify its design output, and a business approach to meet customer needs. The lifecycle of the project included several dozens of people at different levels of involvement and periods in time.

We can identify the following teams or roles, each one addressing the problem from their specialty but in direct collaboration with others according to needs. The Design Team establishes the overall rhythm of the process, and interacts with all other teams at different times in a modality similar to that of the traditional design studio. Following is a description of the roles and their tasks involved in the design process:

1. The *Client* is the actual organization or user asking/supporting the development of a new data representation solution to a particular information problem.
2. The *Application Team* takes the role of the specialist and works as middle-person between the Client and the Design and Psychology Teams (3 and 4). This team "translates" the client needs and requirements into programmatic needs. This team also works as critic and adviser to our research group at large during the design process. Application teams have been in Medicine, Finance, and Network Security.
3. The *Design Team* is in charge of developing the data representation scheme following a collaborative design process and using special principles and techniques discussed elsewhere (Bermudez et al., 2000). During the initial phases, the design team works very closely with (2) and (4). As the schemes become more final, the design team begins to have direct contacts with (1) and with (5).
4. The *Psychology Team* extracts the mental model experts in the field of application use to make sense and act upon the data. During the initial phase, (4) works

in close relationship with (2) to study (1). In later phases, this team is heavily involved in the evaluation of the representation schemes developed by (3) and implemented by (5).

5. The *Computing Team* addresses algorithms, software development, and distributed computing implementation. This team is particularly involved during the final phases of a project. At that time, it works closely with (3) and receives advice from (2) and (4). Its initial input consists of providing prescriptive and "budgetary" advice regarding actual computer implementations of the data display. In areas of application relating to computing (e.g., networking monitoring), this team may also become (2).
6. The *Administration Team* is focused in supporting the day-to-day operation of our research group as well as seeking new areas of work, recruiting consultants and collaborators, managing intellectual property, etc.
7. *Consultants*: external reviewers enter the process at critical times to evaluate the ongoing results and provide an unbiased review of our team's design effort.

3.2 The Design Process

The interactive work of all these people and disciplines occurs within an over reaching ideology of design as a function of human needs and behavior and the interaction between operator and display. As a result, the design process follows the concept of a "hermeneutic circle" (Snodgrass and Coyne, 1990). This concept is an iterative process of implementing a design, learning and understanding from discussion and feedback from the targeted users, and subsequent design refinement. First, the problem and the metaphors for the information that will be displayed are defined. Next, an iterative process via "dialogical exchanges" is used to gain additional insight into the design. New interpretations are discovered and the design is refined. The design development process contains a second feedback loop of iterative evaluation for design usability and intuitiveness.

Each design refinement is evaluated using a testing protocol. The results of these evaluations are methodically analyzed to elucidate design changes while minimizing designer bias. This process also minimizes alterations to the requirements and the design, late in the display designs lifecycle, when changes are more costly (e.g., a change during the design phase is less costly than a change after the display has been deployed). This methodology is successful because the design is evaluated and redesigned during each phase of development, with the intent that the majority of design changes occur in the early stages of development.

The development methodology of the VisAlert display system followed an iterative design process that allowed for many possible design solutions to be explored. The team first developed numerous sketches. These sketches were then refined and mad into conceptual computer based display concepts. These were then iteratively refined and developed until a final solution was achieved. Following the completion of the design phase, an iterative refinement period was conducted with end users. This ensured that the final solution would fit with the uses needs.

3.3 Sketches

Figure 2 shows an early sketch representing the idea to collapse many variables into one as a primary indicator. The position in the *y* dimension shows the amount of the deviation from normal both for the primary indicator, and the variables combined to create the primary indicator. This concept helped the team think about nesting of variables and the hierarchical representation of network data to create primary indicators from a series of disparate variables.

The sketch in Fig. 3 explored a metaphoric representation idea, based upon a large group of environmental formations such as a field of flowers. When flowers bloom they can be seen standing out of the rest of the group because they are visually unique and distinct. This follows general gestalt principles found in design. The design concept further explores different types of unique flower types to encode additional information.

The sketch in Fig. 4 explored time evolution of variables, which are represented by the *y* dimension's height. In addition, items could be placed on a series of quadrant grids that are subdivided to represent problem severity and problem relevance. Glyphs could represent different graphic primitives to indicate type of data or variable presented.

The sketch in Fig. 5 examined the idea of the "inside" protected by a firewall surrounding the enclave from the Internet "outside". The multiple paths through the firewall were representative of unauthorized breaches in the firewall and the

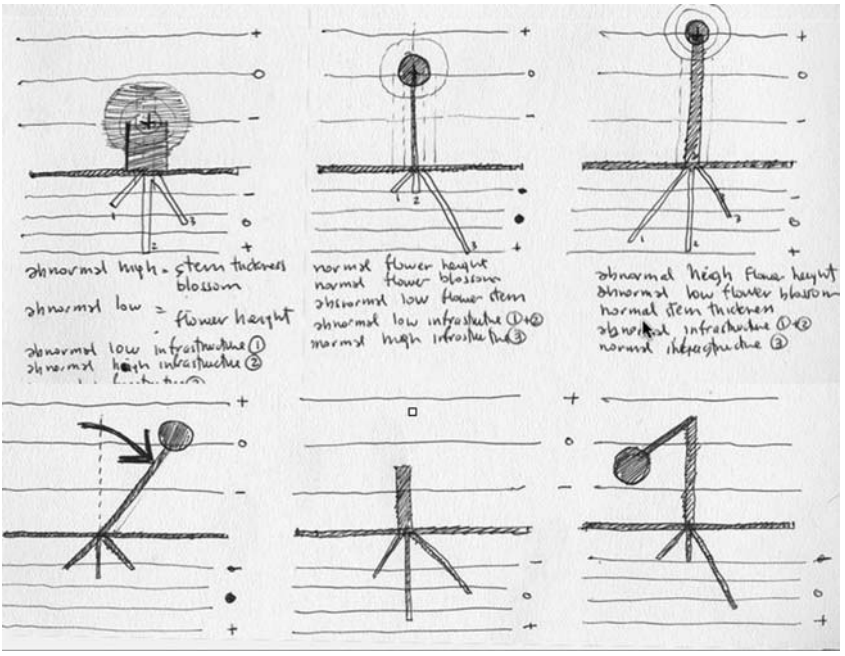

Fig. 2 Stem sketch

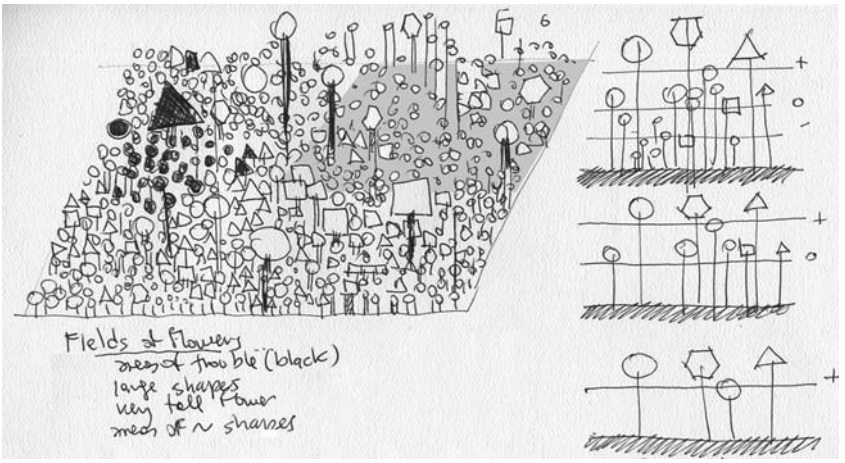

Fig. 3 "Field of Flowers" sketch

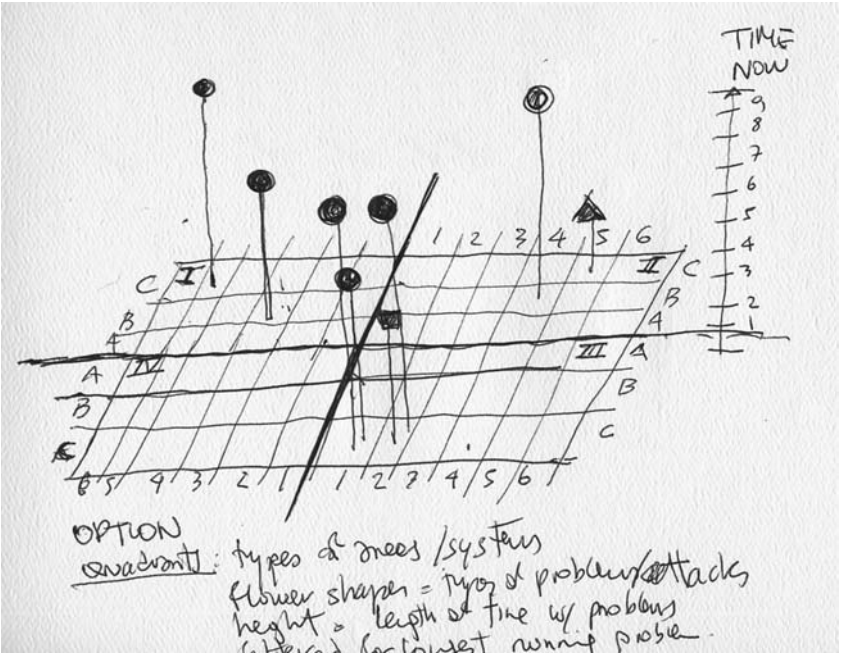

Fig. 4 Quadrants sketch

compromised machines. This sketch was the beginning of a network typology or the representation of "*where*" to provide users with a context of network events.

The concept in Fig. 6 explored the idea of a firewall through a literal representation of walls and increased levels of security as you move from outside to inside. It was thought that the most valuable assets would be placed in the center section, and

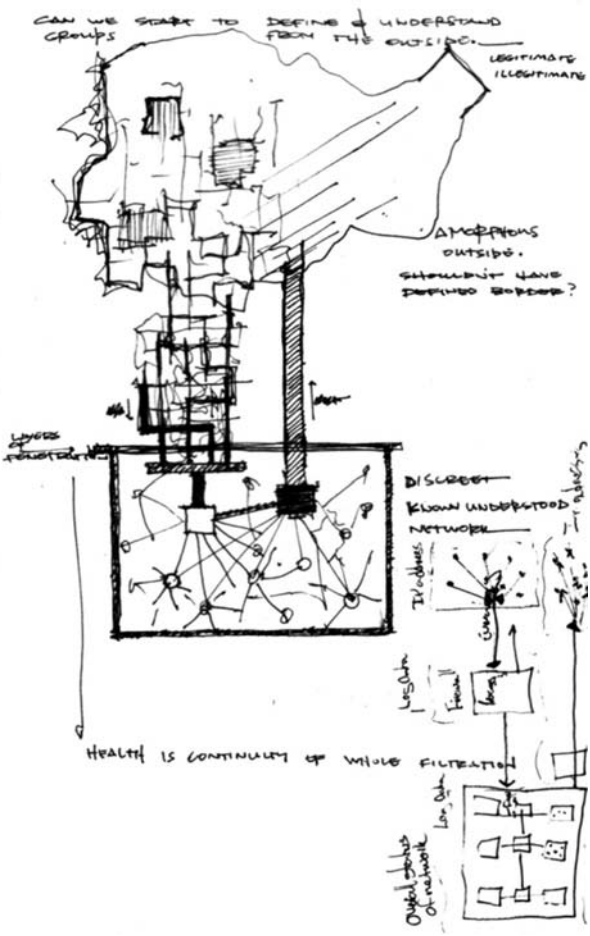

Fig. 5 In-and-out sketch

granted only selected access. Unauthorized breaches in the firewall could be shown as breaks in the wall.

3.4 Refined Conceptual Ideas

After the sketches had been discussed and revised, we then proceeded to move some of the more promising ideas into a more refined and computer-based representation. During this step we examined issues of scaling, usage of color, interfaces, and data handling.

Building upon the idea of the firewall, we developed a multi-dimensional icon that could map 3 variables (Fig. 7). These information bricks or blocks could be then

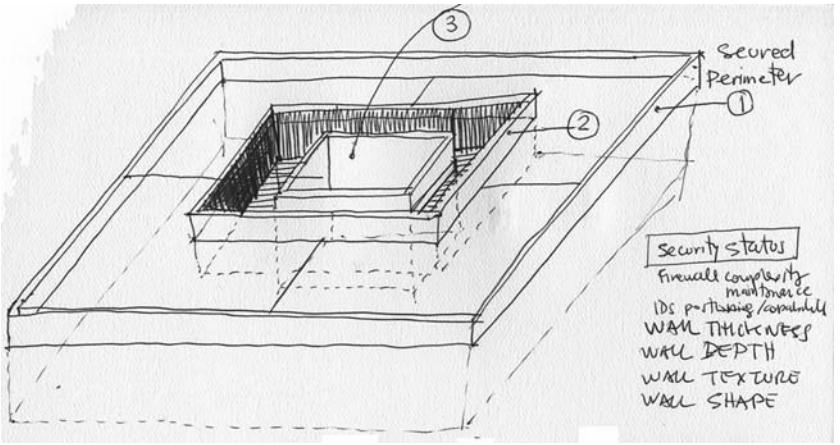

Fig. 6 Firewall sketch

arranged together to create an information structure. The blocks could be grouped together to show information that would be functionally related, such as information about all servers grouped together. The objective was to aid network administrators to quickly identify problems in specific portions of their network.

These icons could then be grouped together scaled to show more information. The images in Figs. 8 and 9 show how patterns of particular problems would emerge, and then allow the network operator to drill down by zooming in. This concept also explored what the interface might look like and what additional tools would be needed, such as filtering and zooming capabilities.

The next evolution in Fig. 10 was to develop the idea of the firewall icon as a larger representation and make it a placeholder for network topology and organization, thus combining two of the previous ideas. This combined the *where* and *what* of the display, even if limited to representing only firewall data through a metaphor.

The next step (Fig. 11) was to develop the sketch into a refined image to see what a complete display might look like. In this step we included different information around the four sides of the display such as firewall logs and alerts. In addition, the network topology was further developed. The concept of correlation between network alerts made its first appearance, as evidenced by the red connector lines to a particular machine, which indicates that a particular machine had experienced three different types of alerts.

A further refinement, as shown in Fig. 12, introduced the idea of time represented as radiating rings from the center. This enabled to include the *when* of an event, and sparked the idea of the representation of *what, where*, and *when* of network events. The last refinement was to create a radial representation so we could include many different types of alerts that could be functionally grouped.

The final iteration of the design was to include a network topology in the center and refine the look and feel of the concept, resulting in the VisAlert concept as described in the introduction.

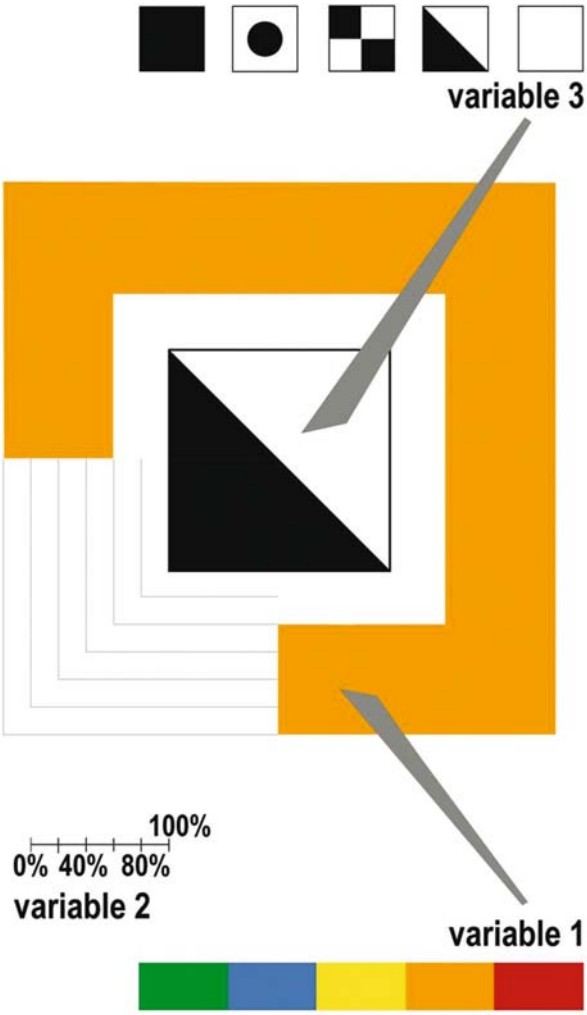

Fig. 7 Firewall icon

3.5 Implementation

The VisAlert prototype technology was implemented in C++ and first deployed at the Air Force Research Lab (AFRL) in Rome, New York. During the testing phase we collected comments and suggested features. VisAlert generated very positive response: users specifically noted its effectiveness, simplicity, and flexibility. They stated that it provided increased situational awareness to detect, diagnose and respond to network events and anomalies.

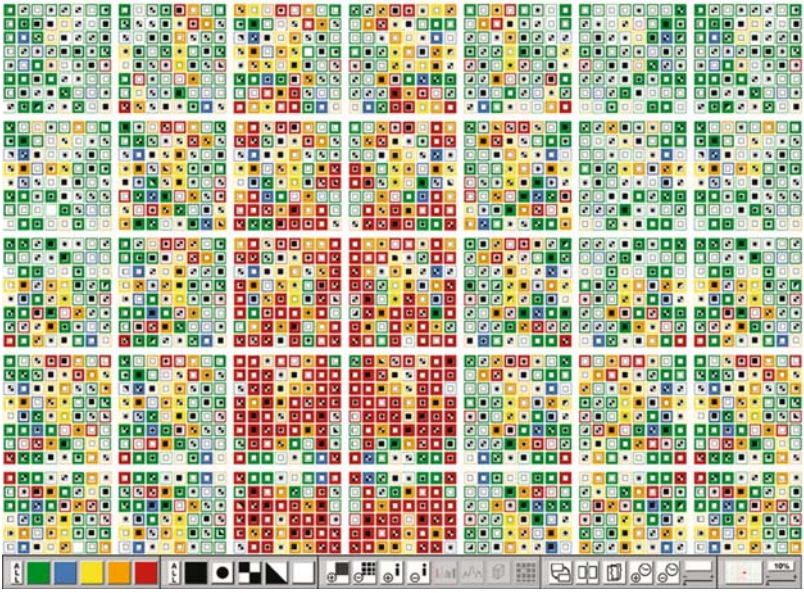

Fig. 8 Composite firewall icons – $O(2^{12})$

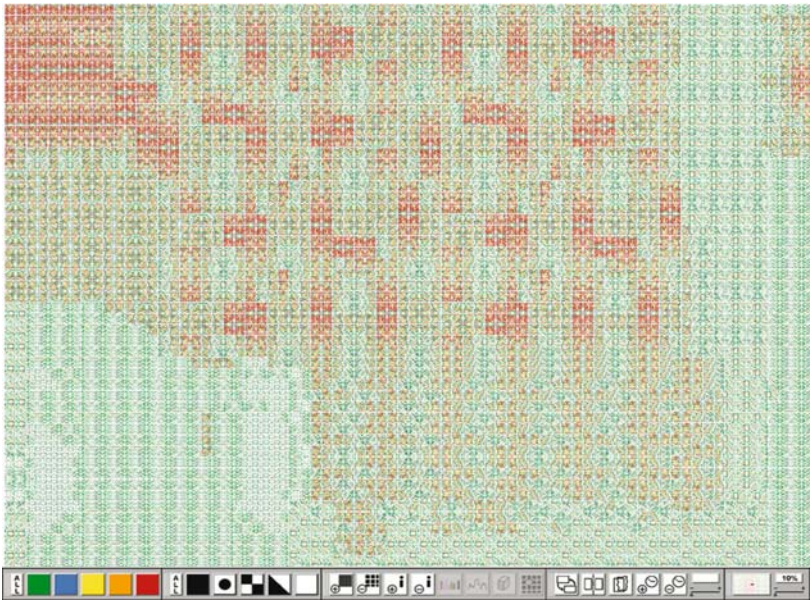

Fig. 9 Composite firewall icons – $O(2^{24})$

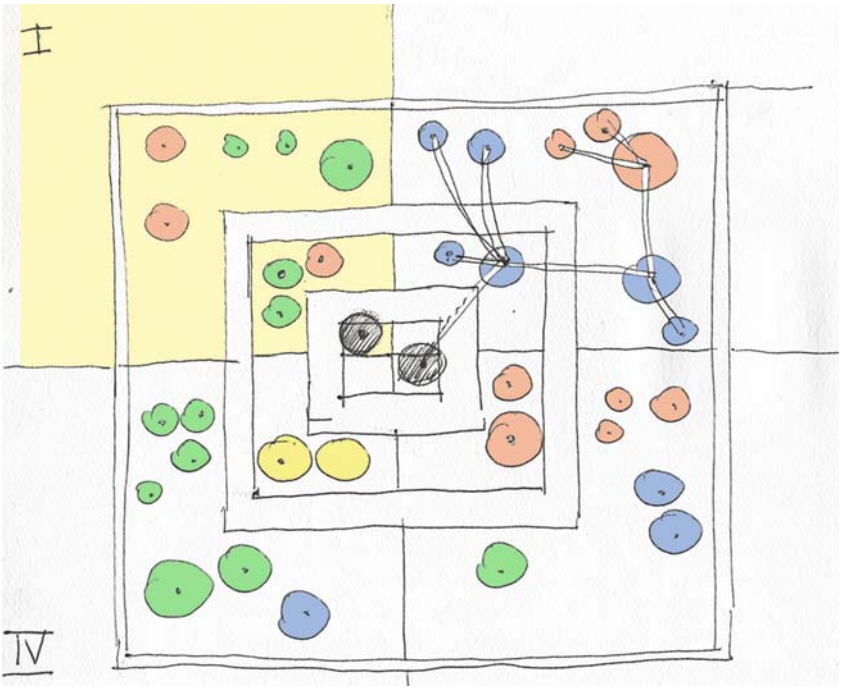

Fig. 10 Firewall topology sketch

However, the analysts wanted more deployment flexibility and a technology that would be operating system and database agnostic. We then completely rewrote the software in Java and integrated Hibernate database management software to accomplish the task (Fig. 13). Currently the system is being evaluated by several commercial network security software vendors.

4 Future Work

Ongoing and future work is in the following areas that are conducive to the complete user experience with her information space to make decisions.

- *Interaction*. The word "visualization" indicates how the data is presented to the user. The next step is to increase the ability of the user to fully interact with the information space, thus providing more effective ways to input data and to control which data is displayed.
- *Extensibility*. This refers to the ability of the visualization software to be extended and include new modules, and moreover to interact with other tools that users are familiar with. This includes the ability or users to encode and correlate their own alert algorithms.

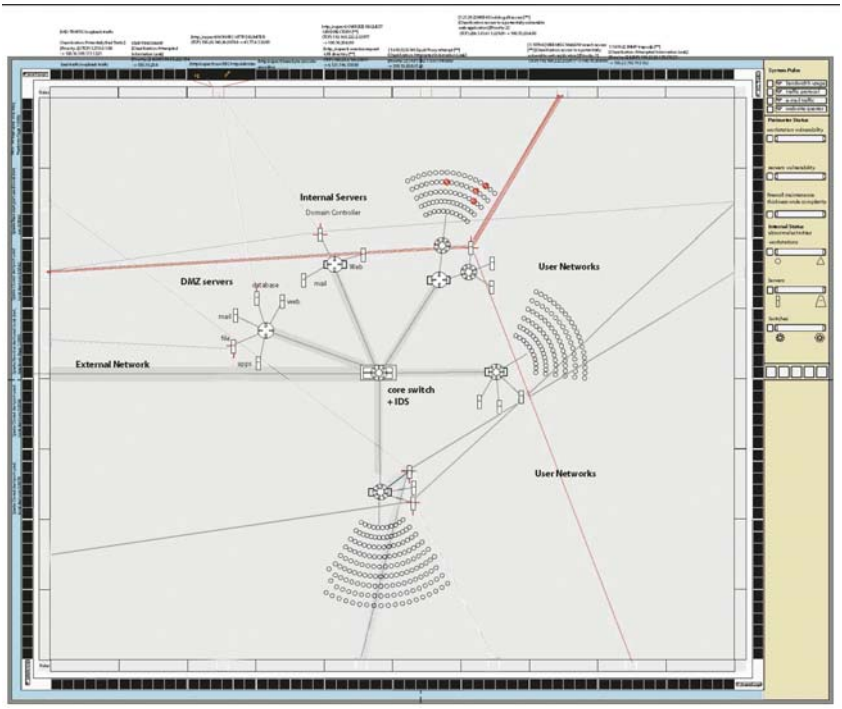

Fig. 11 Square visualization screen

- *Continuity*. This refers to ability to present multiple perspectives and level of details of the information space, and to smoothly transfer from view to view while maintaining the context and references.

More specifically to network security, future work includes the design of additional visualization structures that enable analysts to perform hypothesis testing of events and details, and reporting summaries to decision makers. The complete VisAlert system could then evolve in a visual continuum that would allow seamless transition from a holistic view of the system all the way to detail drill down.

5 Conclusions

The design process, starting from "thinking outside of the box", creating several concepts, refining ideas, and combining them is an organic and effective method to produce visual concepts that encode the relevant information and enable users to enhance the way they work, use their information space, and make decisions.

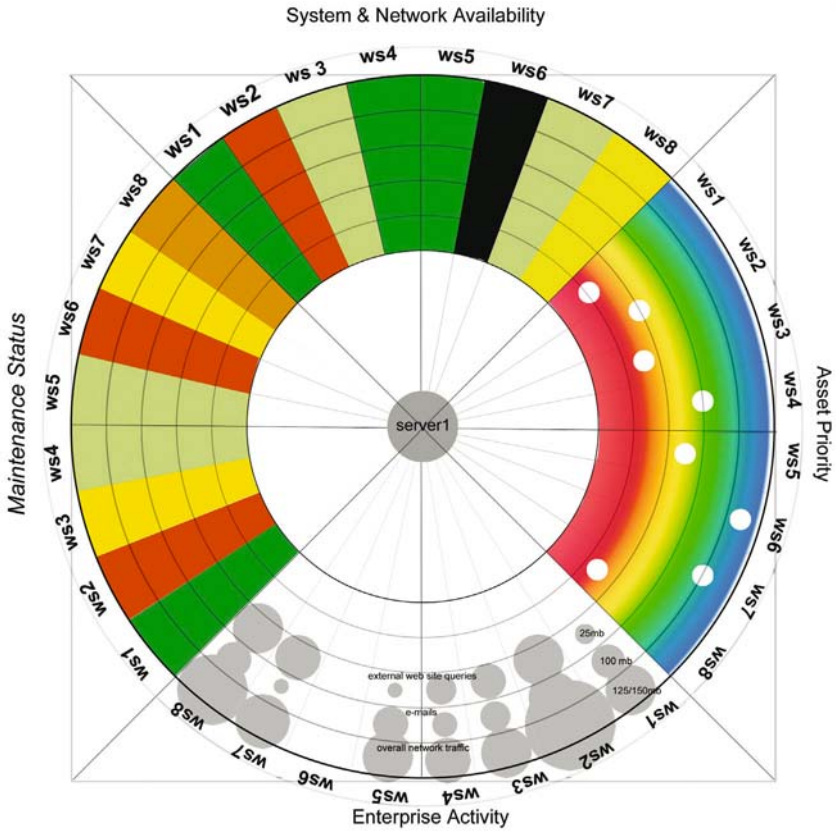

Fig. 12 Radial visualization

In order to optimize the chances that technology is actually used, the development requires interacting with users from the very beginning to address their specific problems, tasks, and mental models.

A very important technical characteristic of VisAlert that enabled user acceptance was the ability to fuse in one view *any* and *all* the data of choice of the user, and to filter out the unwanted one.

User centered design can be addressed very effectively by involving an interdisciplinary team that builds trust and value in different roles and perspectives.

Acknowledgements As indicated in the description of the team and roles, there are numerous people and organization that have had a significant contribution in VisAlert's success. This includes the whole CROMDI team and other personnel at the University of Utah (with particular gratitude to Julio Bermudez and Shaun Moon), Utah State University (with particular gratitude to Robert Erbacher), the AFRL, Battelle, and Skaion Corporation.

This work was supported in part by a grant from the IC-ARDA, the Utah State Center of Excellence Program, and DARPA SBIR.

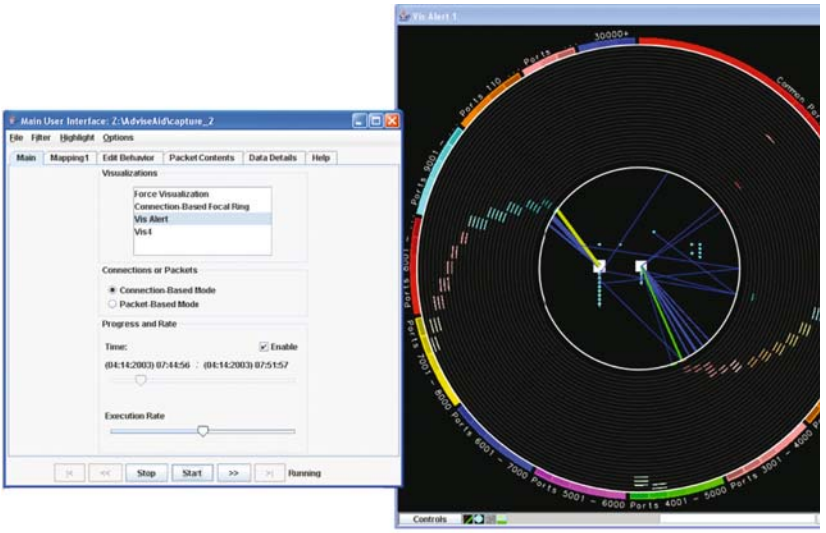

Fig. 13 Screen shot of the current VisAlert product

References

Arnheim R (1977) The Dynamics of Architectural Form. Berkeley: University of California Press.

Bermudez J, Agutter J, Westenskow D, Foresti S, et al. (2000) Data Representation Architecture, in M. Clayton and G. Vasquez de Velasco (eds): ACADIA 2000. Washington DC, pp. 91–102.

Bogdan C (2002) The Semiotic of Visual Languages. New York: Columbia University Press.

Cox K, Eick S, He T (1996) 3D geographic network displays, ACM Sigmod Record, vol. 25, no. 50.

Cross N (1982) Designerly Ways of Knowing, Design Studies vol. 3, no. 4, pp. 221–227.

D'Amico A, Larkin M (2001) Methods of visualizing temporal patterns in and mission impact of computer security breaches, Proceedings of the DARPA Information Survivability Conference and Exposition (DISCEX II'01), vol. 1, pp. 343–354.

Estrin D, Handley M, Heidermann J, et al. (2000) Network visualization with NAM, the VINT network animator, IEEE Computer, vol. 33, pp. 63–68.

Foresti S, Agutter J, Livnat Y, Moon S, et al. (2006) VisAlert: visual correlation of network alerts, IEEE Computer Graphics and Applications, vol. 26, no. 2, pp. 48–59.

Friedman B (1997) Human Values and the Design of Computer Technology, Center for the Study of Language and Information, Stanford, CA.

Kraut R, Galegher J, Egido C (1988) Tasks and relationships in scientific research collaborations, Human–Computer Interaction, vol. 3, pp. 31–58.

Polla D, McConnell J, et al. (1998) A framework for cooperative intrusion detection, Proceedings of the 21st National Information Systems Security Conference, pp. 361–373.

Rowe P (1987) Design Thinking. Cambridge: The MIT Press.

Snodgrass A, Coyne R (1990) Is designing hermeneutical? Technical Report, Sydney Australia.

Teoh S, Ma K, Wu S, et al. (2002) Case study: interactive visualization for internet security, Proceedings of the IEEE Conference on Visualization, pp. 505–508.

Teoh S, Ma K, Wu S, et al. (2003) Visual exploration process for the analysis of internet routing data, Proceedings of the IEEE Conference on Visualization, pp. 523–530.

Wong W (1972) Principles of 2-D Design. New York: Van Nostrand Reynolds.

Wong W (1977) Principles of 3-D Design. New York: Van Nostrand Reynolds.