Visual Analytics of Cell Phone Data using MobiVis and OntoVis

Carlos D. Correa Tarik Crnovrsanin Christopher Muelder Zeqian Shen Ryan Armstrong
James Shearer Kwan-Liu Ma *

Visualization and Interface Design Innovation (VIDI) Group
University of California, Davis

ABSTRACT

MobiVis is a visual analytics tools to aid in the process of processing and understanding complex relational data, such as social networks. At the core of these tools is the ability to filter complex networks structurally and semantically, which helps us discover clusters and patterns in the organization of social networks. Semantic filtering is obtained via an ontology graph, based on another visual analytics tool, called OntoVis. In this summary, we describe how these tools where used to analyze one of the mini-challenges of the 2008 VAST challenge.

Keywords: Visual Analytics, Heterogeneous Graph Visualization

Index Terms: I.3.6 [Methodology and Techniques]: Interaction Techniques—

1 OVERVIEW OF THE TOOLS

Our visual analytics tools is a combination of two tools into a single visualization. MobiVis is a visual analytics system designed for exploration and discovery of mobile data [1]. To aid in understanding complex relationships in these networks, MobiVis allows users to apply semantic and temporal filtering, via the interactive timechart and an ontology graph (OntoVis).

The ontology graph, shown in Fig. 1 on the top right corner, is a high-level graph that represents the different entity types and their relationships. These entities are obtained automatically (e.g., from each attribute in the data), or user-defined depending on the task at hand. In our case, we derive entities such as people, cell tower and calls. The ontology graph then enables the user to apply semantic filters on the data, and define association rules between disparate entities. This capability is an extension of the semantic filtering capabilities of another of our visual analytics tools, OntoVis [2].

Since a cell phone networks can be characterized as spatio-temporal social networks, providing the temporal information is important. We use 2D time charts to define different time scales and colors denoting a certain attribute. In the example shown in Fig. 1 (bottom), the timechart shows daily and hourly activity of a given person of interest. It shows, for example, repeating patterns or anomalies in the behavior. In this example, we notice a change in behavior after a day of complete "silence". The interactive timechart allows the analyst to observe the social network in any given interval.

2 VISUAL ANALYTICS PROCESS

2.1 Overview via Structural Abstraction

MobiVis allows the analyst to obtain a quick glance of the entire data set, both spatially and temporally. Unfortunately, complex networks become quickly cluttered and it is difficult to find interesting

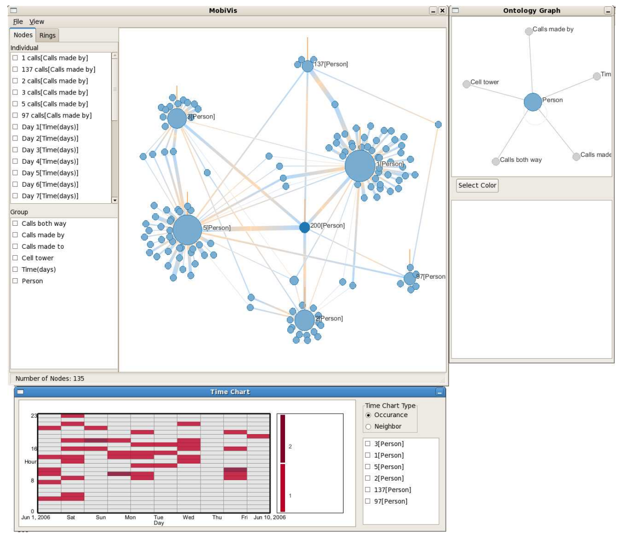

Figure 1: Overview of MobiVis. The central view provides a force-directed layout of cell phones, where a link indicates communication between them. The upper right corner shows the ontology graph, which allows the user to highlight entity types of interest, such as people, calls and cell towers. The bottom view shows the interactive time chart, where color boxes denote time periods of activity.

patterns. To obtain structural abstractions, we use OntoVis, which allows us to consider connectivity and node degree to condense the network. Fig. 1 shows an structural abstraction after considering only second degree of separation from a given node of interest. This capability proves to be very useful when analysts have leads to follow.

2.2 Discovery via Semantic Associations

One of the capabilities of MobiVis is semantic filtering. This is done via an ontology graph, which is a high-level graph that encodes the different entity types and relationships among them in the data. For example, mobile phone data can be characterized by person, calls, cell towers, duration and time, among other possible criteria. MobiVis allows us to select different association rules and display the social actors that satisfy those rules. Fig. 2 shows several association rules for a subset of the VAST challenge data set. These rules specify the calls made to two particular cell phones (identifiers 200 and 300), shown in orange, and the association of people (blue nodes) to the cell towers they used when making a call (green nodes). With this semantic abstraction, we can now have an idea of the geographical sparsity of the social network of a person of interest, which otherwise could not be obtained by looking at each call in isolation.

2.3 Temporal Filtering and Details on Demand

One of the advantages of MobiVis is the ability to look at large heterogeneous cell phone data via simple abstractions. However,

*correac@cs.ucdavis.edu, turokhunter@gmail.com, cwmuelder@ucdavis.edu, zqshen@ucdavis.edu, rnarmstrong@ucdavis.edu, shearer@cs.ucdavis.edu, ma@cs.ucdavis.edu

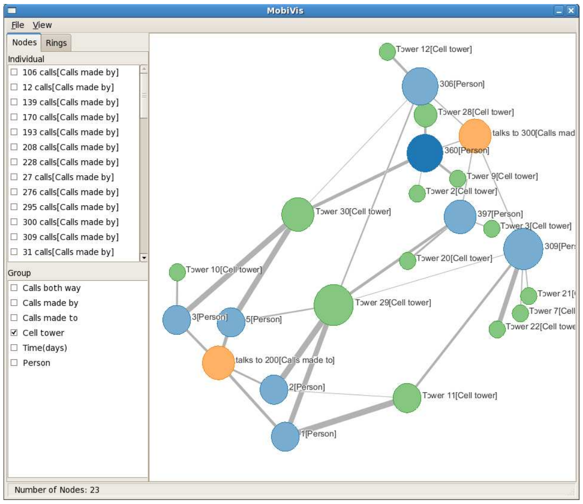

Figure 2: Example of semantic abstraction in MobiVis. Here, all people who called two cell phones of interest (200 and 300, shown as orange nodes), and their associated towers (green nodes). The difference in geographic sparsity of these two cell phones (believed to belong to the same person) is evident from this visualization.

abstractions usually summarize or ignore detailed data to be effective. Thus, detailed information about a particular call on or around a period in time may be difficult to visualize. One way to visualize particular call behavior is done with the interactive time chart, as shown in Fig. 1. Selecting a particular time period allows the analyst to filter the data temporally.

We can also improve the temporal filtering with *call graphs*, which represent calls and people in a two-dimensional plot. An example is shown in Fig. 3. Here, each person generates a time line, and links are created for every call made between any pair. We can see an almost uniform pattern for every single day. Call graphs help us visualize the call behavior of specific people (shown in different colors), and allows us to discover salient changes in their communication patterns. For example, it becomes evident that some actors stopped communicating after day 7. We also tried different layouts for the call graphs, such as radially (Fig. 4).

2.4 Cyclic Discovery

An important aspect of the analytics process is that, after interpreting the details found in the discovery stage, an analyst can form new hypotheses and make new inquiries. In this example, there is an interesting event around day 8, when the actors of interest stopped talking for the most part. Following this new lead requires new semantic and temporal abstractions. Once the analyst finds other actors of interest, he/she can visualize the detailed sequence of calls in the call graph and obtain new findings. For example, Fig. 4 shows what seems to be a similar pattern of calls in two different days, for two different sets of phones. With the new leads, a new structural abstraction can be created that reveals the overall relationship between these entities.

3 Lessons Learned and Conclusion

MobiVis and OntoVis are general tools for visualizing large heterogeneous networks, that can be tailored easily to a specific task. We found that the process of filtering data, both structurally and semantically, enables the analyst to discover patterns and relationships of interest that may be hidden when considering the data as a whole. We expected the discovery process for the VAST challenge to be a top-down approach, whereby we formulate hypothesis about a single person and progressively refine our findings. However, it turned

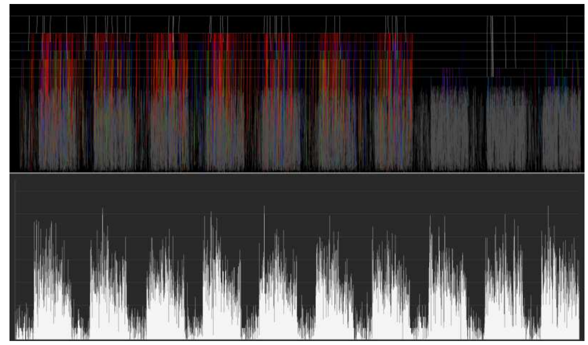

Figure 3: Time line visualization of calls. We enhance the capabilities of MobiVis to visualize call data in a different 2D space.

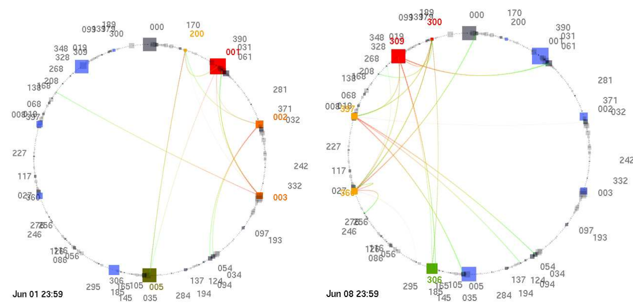

Figure 4: A different layout for the call graph, where nodes are placed in a circular layout. Links indicate calls (where direction is encoded as a color gradient from red to green). Although an overview is difficult to visualize, it provides detail information on periods of interest. Here, we show the calls among actors at two different stages. Similar patterns where found on nodes that where believed to be linked.

out to be the product of several repetitions, from an overview, using structural and semantic abstraction, down to the call distribution over time, using call graphs. This detailed information (after being filtered and abstracted) was useful for two purposes. It validates some of the findings done at the higher level, but also opens up questions for further inquiry. Repeating the cycle with new leads resulted in a more complete understanding of the entire data. We expect to make the transition from high-level abstractions, such as the ones provided by MobiVis, to detailed information (i.e. call graphs) more transparent. We believe that this will be critical for the analysts to provide correct interpretations of their findings.

MobiVis and OntoVis prove to be important tools for analyzing large heterogeneous relational data. The VAST 2008 challenge data set is a good example where brute-force approaches provide little useful information, and semantic and structural abstraction is critical. It also provides interesting changes over time that can be discovered with cyclic analysis and easy access to detail information, via interactive call graphs.

Acknowledgements

This research was supported by the National Science Foundation under contracts CCF 0808896 and CNS 0716691.

References

[1] Z. Shen and K.-L. Ma. Mobivis: A visualization system for exploring mobile data. In *Proceedings of IEEE Pacific Visualization Symposium*, pages 175–182, 2008.

[2] Z. Shen, K.-L. Ma, and T. Eliassi-Rad. Visual analysis of large heterogeneous social networks by semantic and structural abstractions. *IEEE Transactions on Visualization and Computer Graphics*, 12(6):1427–1439, 2006.