

Controlled Visual Search Engine with Data Integration and Exploration

K.BalaRaju, K.Raju, K.Vidya

Abstract— In the previous years there have been lot of changes made in various online social networking services and developing the dramatic popularity of these networking services, in which millions of members publicly articulate mutual “friendship” relations. Guided by ethnographic research of these online communities, we have designed and implemented a visualization system for playful end-user exploration and navigation of large-scale online social networks. Our design builds upon familiar node-link network layouts to contribute techniques for exploring connectivity in large graph structures, supporting visual search and analysis for making people understand easily by automatically identifying and visualizing community structures. Both public installation and controlled studies of the system provide evidence of the system’s usability, capacity for facilitating discovery, and potential for fun and engaged social activity.

Index Terms— social networks, visualization, graphs, community, exploration, play.

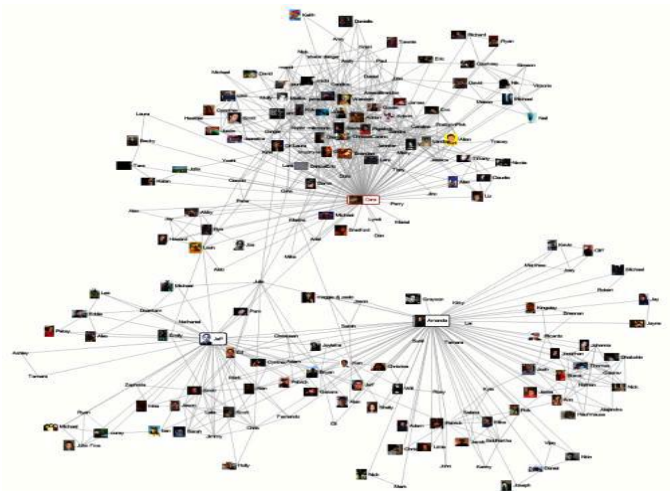
I INTRODUCTION

The development and evolution of internet has given rise to many forms of online sociality, including e-mail. Usenet, instant messaging, blogging, and online dating services. In 2003, another form of online community acquired stunning popularity: online social networking services. In addition to descriptive personal profiles, members of such communities publicly articulate mutual “friendship” links with other members, creating a browseable network of social relations. Although pre-dated by other services such as sixdegrees.com in 1997, social networking services emerged as a veritable phenomenon with the dramatic rise of Friendster.com, which rapidly amassed millions of users beginning in the spring of 2003. Friendster’s success was closely followed by a number of other popular services, including Tribe.net and Orkut.com, quickly earning these services the acronym “YASNS” or “Yet Another Social Networking Service”.

Though the users of these systems have constructed massive graph structures of social connectivity, typical web interfaces to these social networks remain relatively impoverished, showing only the network connections of single individuals in a linear list on a web page. Articulated connections between one’s own “friends” in these systems are obscured, and can be unearthed constructs, such as shared patterns of interests and community structures can be even harder to discern. This has problematic implementation for members’ ability to efficiently explore their online community and gauge both the scale and individuals to which their self-reported personal information is exposed.

In this paper we present the Vizster, a visualization system for exploring such online social networks. Vizster builds on ethnographic research of online social networking services and previous work in social network visualization to provide a system by which members of such online communities can explore their articulated social network in the playful manner that they desire. Our system is motivated by both social and

technological concerns. On the social side, we are interested in better facilitating the discovery of people, connections, and community to promote increased awareness of community structure and information exposure, while maintaining a fun and socially engaged online space. From the perspective of information visualization, we are interested in crafting techniques for the situated exploration and visual analysis of large



graph structures by a specific end-user population.

Figure 1: Vizster depicting three intersecting social networks

Our design builds upon familiar node-link network layouts to provide novel techniques for exploring inter-connectivity in large graph structures, supporting visual search and analysis of network and profile data, customizing spring-embedded layouts, and automatically identifying and visualizing community structures. We have used the resulting system to visualize a 1.5 million member crawl of the popular Friendster service, collected in winter 2003. We first summarize related work in social network visualization and share

relevant findings from an ethnographic investigation of the Friendster service. We then describe the design of the Vizster visualization in detail, followed by observations from a public installation and informal studies of the system's usage.

II RELATED WORK

Visualization of social networks has a rich history, particularly within the social sciences, where node-link depictions of social relations have been employed as an analytical tool since at least the 1930s: Linton Freeman documents the history of social network visualization within sociological research, providing examples of the ways in which spatial position, color, size, and shape can all be used to encode information[8]. For example, networks can be arranged on a map to represent the geographic distribution of a population. Alternatively, algorithmically generated layouts have useful spatial properties: a force-directed layout can be quite effective for spatially grouping connected communities, while a radial layout intuitively portrays network distances from a central actor. Color, size, and shape have been used to encode both topological and non-topological properties such as centrality, categorization, and gender.

In recent years, such approaches have been effectively used in the analysis of domains such as e-mail communication [7], early online social networks [2], and co-authorship networks in scientific publications [15]. There are a number of systems for generating such social network visualizations and performing statistical analyses for the purpose of sociological research, including UCInet [19], JUNG [12], and GUESS [4]. In addition to sociological research, there have been numerous social visualization projects intended for end-users rather than outside researchers.

Though often utilizing techniques seen in analytical domains, these systems present users with a sense of social place [10] different from more detached analytical environments. Nardi et al.'s ContactMap [14] uses spatial grouping and color to redundantly code for various community groups within a visualization of a user's e-mail contacts. TouchGraph [18] uses a force-directed layout to present a network visualization of users of the Live Journal online community, allowing personal networks to be expanded or contracted by user interaction. BuddyZoo [5] analyses users' instant messaging (IM) buddy lists to present a static network visualization of their IM contacts. Boyd and Potter's social network fragments[9] visualizes personal e-mail archives in a zoomable network view and includes temporal filtering to visualize contacts at various time periods. They found their system to be particularly well suited as a memory aid for past contexts and contacts. Despite this wealth of social network visualization, we believe there is still a need for new designs and techniques, especially as articulated social networks became increasingly common in web services for signifying various types of relationships. Though

limited network views have proven manageable and useful, means for better exploring extended network contexts and providing useful cues for exploration are possible. Visualization of profile attributes unique to online social networks is needed, and techniques for incorporating analytical tools within the simplified domain of end-user visualization may prove useful. In crafting Vizster, we aspired to create a visualization system that builds upon the success of existing work, introducing an increased potential for exploration while maintaining a sense of playful social place. To do so, we turned to ethnographic investigations of the Friendster community to guide our design.

III DESCRIPTION

In text retrieval, full text search refers to techniques for searching a single computer-stored document or a collection in a full-text based. Full text search is distinguished from searches based on metadata or on parts of the original text represented in databases. In a full text search, the search engine examines all of the words in every stored document as it tries to match search words supplied by the user. Full text searching techniques became common in online bibliographic databases in the 1990s. Many websites and application programs provide full text search capabilities. Some web search engines such as Alta Vista employ full text search techniques while others index only a portion of the WebPages examined by its indexing systems. However, when the number of documents to search is potentially large or the quantity of search queries to perform is substantial, the problem of full text search is often divided in to two tasks: indexing and searching.

The indexing search will scan the text of all documents and build a list of search terms, often called an index, but correctly named a concordance. In search stage, when performing a specific query, only the index is referenced rather than the text of original documents. The extended version of search engines are Visual search engines which shows the graphical representation of link establishment with hierarchy levels. Visualization is done on the basis keywords relevant data, etc. The basic idea behind this visual search tools may vary with different sites, each and every search tool maintains their own ideas to run their site functionality. Most of the tools allow users to explore the search results in more depth with the representation of similar search items. Friendster was designed to be an online friendship site with profiles, demographic and interest driven search, and a private messaging system. What made Friendster unique was its articulated social networking component and testimonial feature. When users joined the site, they were required to develop a profile that asked for demographic and a description of tastes. While most users surfed the network by clicking on the friends of others profiles, there was also a search tool available,

more closely “tag-alone” with their friends, reinforcing their limited connections. Meanwhile, nodes with higher connectivity still group into communities, but with a reduced density to increase readability.

B. Basic Interaction:

Basic interaction is done with simple mouse operations. Clicking a node causes the corresponding member profile to appear in the profile panel to the right of the network view. Dragging a node moves it around in the space: upon release it is again subject to the ongoing force-directed layout. Double-clicking a node causes it to either expand or contract, depending on its current state. Upon expansion the egocentric network for that member is added to the visualization. The expanded node assumes a fixed position, subject to repositioning by the user. Double-clicking an expanded node causes that member’s egocentric network to disappear and the node again becomes subject to the layout.

This mechanism allows users to selectively grow or hide aspects of the network, increasing the scope of the visualization while maintaining the context of already expanded members. The force-directed layout adjusts to these changes, with friends shared between expanded members moving to an intermediate position.

C. Exploration: Connectivity Highlighting:

To aid network understanding, Vizster highlights nodes based on connectivity in the larger network context. When the mouse hovers over a node, it causes that person, that person’s friends, and visible friends-of-friends to highlight. A graded color scale moving from red-orange to orange to yellow- redundantly coding through both intensity and hue- is used to visualize network distance. All other members in the graph are correspondingly desaturated, with images reverting to grayscale, to provide a figure-ground separation between highlighted and non-highlighted nodes (Figure 3).

Moving the mouse off the node causes the highlighting to disappear, but with a half-second hysteresis applied to prevent annoying visual “bouncing” as consecutive nodes are visited by the mouse pointer. Clicking and holding a node for 1 second will make the current highlighting persistent, after which clicking the background or another node will resume the normal highlighting interaction. Though readily understandable, egocentric views suffer from a lack of greater network context. To provide cues to the greater network topology, connectivity is computed using the complete backing graph rather than the visualized graph. As a result, people connected to selected person through intermediate friends in the larger network highlight even if those intermediate nodes are not currently visible (Figure 4).

These connectivity queries, however, do not pass through the expanded nodes, as they are obviously intermediate to currently visible nodes, and the goal is to help provide cues to currently obscured connections.

D. Exploration: Linkage Views:

As expanding a member’s egocentric network can significantly change the display, in response to user feedback we have also incorporated linkage views that enable the user to visualize intermediate nodes in a lightweight manner. A user can click a node to select it, mouse over another node of interest, and then tap the space bar to visualize all connections between the selected node and the current mouse target. This causes any intermediate nodes not in the currently expanded networks to also be displayed, providing a complete view of the two-hop connectivity between the nodes (Figure 5).

E. Navigation:

The network visualization employs a mix of both manual and automated panning and zooming for navigating the space. Panning is performed by dragging the background of the display with the left mouse button down. The display automatically pans when a new node is expanded, centering on the newly expanded network.

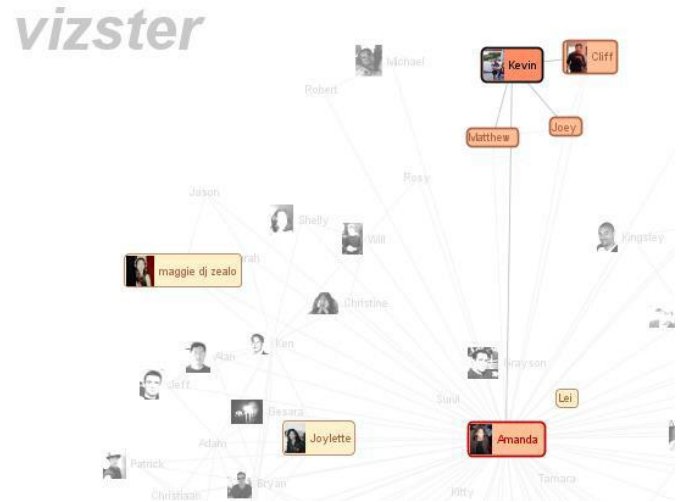


Figure 3: Connectivity Highlighting. The highlighted person “Kevin” is connected to “maggie dj zealo”, “Joylette”, and “Lei” (all shown in yellow) by friends not currently visible

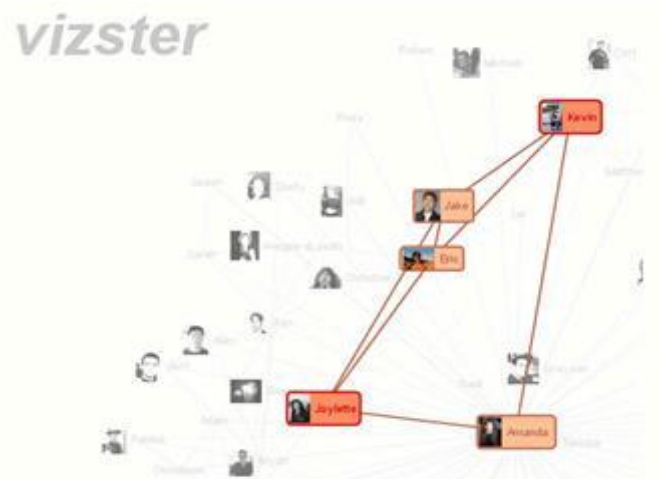


Figure 4: Linkage Views. Clicking on “Kevin” and then tapping the space bar while hovering over “Joylette” (or vice versa) causes all intermediate nodes to be visualized

Manual zooming is performed by holding down the right mouse button and moving the mouse up or down. Simply clicking the right mouse button causes the display to automatically pan and zoom such that the entire visualized network fits within the display. The rendering components update to draw higher resolution photos when zoomed-in to double the normal scale. In some cases it is desirable to see a particular region while still maintaining the context of a zoomed-out network.

To support this, we use geometric distortion within the context of the force-directed layout to provide a focus+context display (Figure 5). When the display is zoomed-out such that node text is no longer legible, users can hold down the left mouse button on a node for a second to “inflate” the currently highlighted subset of nodes. The nodes balloon to become more readily visible and their mass in the force-directed layout is correspondingly increased to cause the nodes to push further apart and prevent occlusion.

F. Search:

Vizster also supports keyword search of the visualized network. As items become visible, their profile attributes are added to an in-memory search index. A search box allows users to type in search queries. Search hits are indicated by “auras” around matching nodes (Figure 6), with non-matching nodes again deemphasized through desaturation. The search auras are presented in a dark purplish color, intended to leverage blue-yellow color opposition to balance against the hues of the highlighting. This contrast was inspired by the graceful balance of Mark Rothko’s painting No. 14. As users mouse over the network with search auras present, highlighting is performed as usual, desaturating the text and images of non-highlighted search hits. The search auras themselves remain visible, however, to balance competing search and connectivity queries. Search auras for highlighted nodes also grow slightly larger to maintain visibility.



Figure 5: Focus+Context views by inflating highlighted nodes.

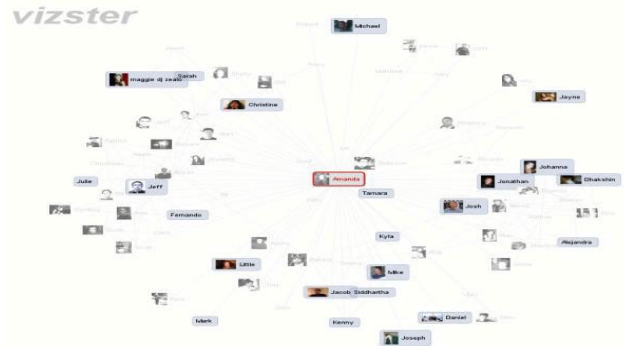


Figure 6: Visualized search results for the query “student”.

Searches are executed by clicking text in the profile panel or typing in the search box. Occupation lifelong student
Search>> student

Additionally, users can issue search queries from the profile panel in a direct manipulation fashion. As the mouse pointer moves over specific words, the word will highlight if there are at least 2 matching search hits for that term. Clicking the word will then execute a search, with the text of the query appearing in the editable search box. Double clicking a word will initiate a search for the surrounding phrase. Manually selecting text by mouse drag similarly executes a search for the selected text.

G. Attribute Visualization using “X-Ray” mode:

In addition to generating search queries, the profile panel also includes controls to further visualize the individual profile attributes of members. By clicking on a check box next to an attribute within the profile panel. Vizster will enter an “X-ray” mode visualizing attribute values such as age, number of friends, gender, relationship status or time since last login. The backgrounds of the display turns back to clearly indicate the modal state, member photos are removed, and node color is used to visualize the currently selected attribute. Additionally, a color legend appears in the upper-right corner of the display. In this state, all the other operations discussed still function as normal. Highlighting and search still desaturate non-matching items, though node colors no longer change under highlighting.

Figure 7 shows Vizster in X-ray mode visualizing the number of friends each person has in the service. Figure 8 shows the use of X-ray mode to visualize each member’s gender. Figure 9 displays the results of including a mouse-over a highlight, and search results for the query “student”. Community groupings are shown, but consistent with the X-ray metaphors are reduced to simple outlines.

H. Visualizing Community Structure:

Although the force-directed layout does an admirable job of clustering the network into communities based on linkage, Vizster also supports explicit visualization of community structures. Inferred community groupings

are visibly represented as “blobs” surrounding their community members, taking advantage of low spatial

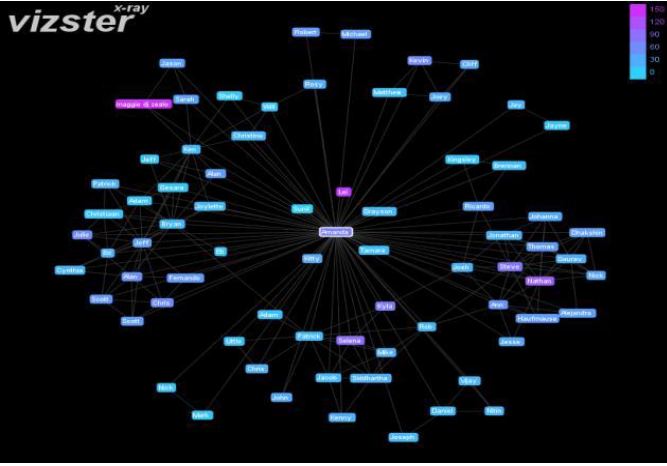


Figure 7: X-Ray mode visualizing the number of friends

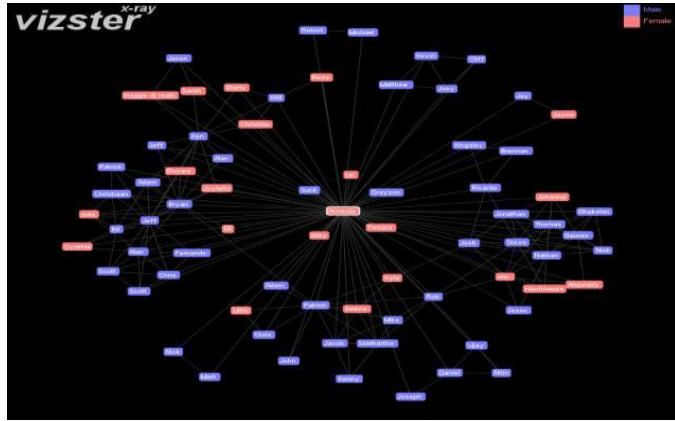


Figure 8: X-Ray mode visualizing the gender frequencies to make community structures apparent (Figure 10). These blobs are computed by taking the coordinates of consistent members, calculating the convex hull, and then using quadratic splines to provide curved borders. As members within a community are moved about, the blob boundaries adjust accordingly for a dynamic, playful character. The blobs are translucently rendered, with colors determined by hue sampling.

This semi-transparency allows the blobs to act as Venn diagrams if members are assigned to multiple communities. To further spring forces to support the spatial separation of communities. As manual creation of large communities can prove somewhat tedious, Vizster includes tools for the automatic determination of community structure. For this task, we have used Newman’s community identification algorithm [16], which identified group structures based solely on link analysis. The algorithm employs hierarchical agglomerative clustering, first placing each node in its own community, then greedily merging groups based upon a metric that attempts to maximize within cluster linkage while minimizing between cluster linkage.

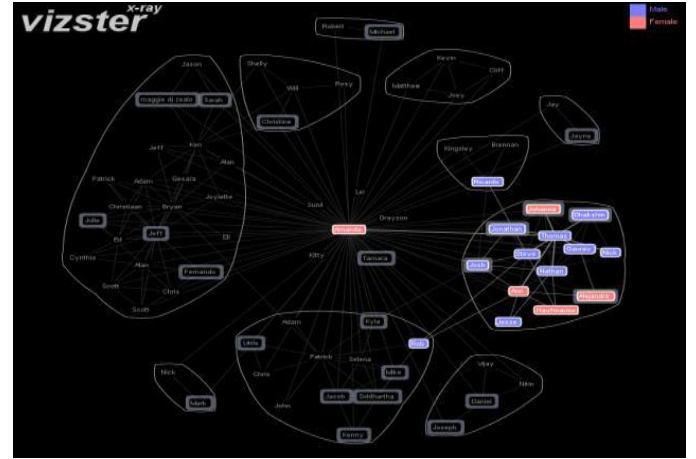


Figure 9: X-Ray mode visualizing the genders, search hits, mouse-over highlight and community structures

Users can instigate automated community analysis by clicking the community analysis button at the bottom-left of the display. The community structure algorithm is then run on the adjacency matrix of the currently visible network. Currently expanded nodes are omitted from the input matrix to avoid unduly collapsing communities the goal is to show community relations surrounding the expanded nodes. The results of the community analysis are then visualized using communities identified at the optimal value of the clustering metric.

Our design hides the complexity of the algorithm from users, letting them explore these data mining results through simple widgets and visual analysis. Like most data mining methods, the community analysis is not perfect, as it may identify communities at higher or lower granularities than those desired by the user or make assignments otherwise perceived as problematic. To help combat this, a community slider is provided to explore the various states of the clustering (Figures 10, 11). Moving the slider to the far left reverts the display to the initial state of the clustering, while moving the slider progressively to the right reveals each Friends 73 Gender Female merge performed by the algorithm. Thus the slider allows the user to interactively explore clustering states by moving through progressive slices of the computed cluster tree. In practice, we have observed users employ the slider until the communities “look right” to them. The current community visualization persists when additional nodes are expanded or contracted; when the visible network changes, an update button is added to the control panel, allowing users to re-run the community analysis.

The current algorithm assigns members to unique communities. In future work it may be useful to generalize the underlying algorithm to support soft or “fuzzy” clustering techniques permitting multiple memberships in communities. We also plan to introduce manual editing so that users may further adjust communities as desired.

I. Summary:

In summary, the Vizster design constitutes a visual environment for the exploration and analysis of online social networks, including both topological and profile data. The scale of displayed information and layout were chosen to support observed behaviour and capabilities, and allow users to expand visualized networks while maintaining landmarks. Interacting highlighting is used to explore friendship relations and unearth “hidden” connections in the large network structure. Panning, zooming and distortion techniques are provided to help users navigate visualized networks. Interactive search and attribute visualization (“X-ray” mode) enable visual exploration of member profile data. Finally, visual community analysis is provided to help users construct and explore higher-level structures of their online communities.

V IMPLEMENTATION NOTES

Vizster was written in Java using the prefuse visualization toolkit [11]. The toolkit’s filters, configurable physics stimulation, image loading and catching, renderers, and interactive controls greatly facilitated the implementation of our design. We also wrote extensions for database connectivity and perform connectivity highlighting and linkage views. To support generalized keyword search, we also integrated the Lucene search engine into the prefuse framework. Network and profile data for the visualization were collected using a custom web crawler and stored in a backing MySQL database. To support highlighting connectivity queries and faster expansion response times, 2nd order networks are loaded from the database upon expansion of nodes, though only the 1st order networks are immediately visualized. The only caveat is the increased memory requirements of the application.

VI USAGE OBSERVATION

To evaluate and further guide the design of our visualization, we observed usage in two environments: a public installation at a large party and an informal laboratory setting. While we were certainly interested in gauging Vizster’s utility and usability, we were also interested in larger patterns of discovery-finding unknown people, connections communities of relevance- and in people’s social and affective reactions to the visualization. Through experimentation, most party-goers and study participants quickly made sense of the connectivity highlighting. Often times, a lone node would highlight far away from the current locus of interaction, eliciting such as “why did that light up?” This would then trigger exploration, with users further expanding the network to discover the intermediary friends that were not yet visible.

Users would play with the slider to get a feel for

how the feature worked, and then settle with a configuration that “looked right” to them. Nearly everyone then began retelling stories about the various communities and their members. Prototypical comments included “Look, it found all the people from Harvard”. In general, people tended to impart a higher level of sophistication to the community analysis than is implemented: some remain convinced that it compares members’ profile data. Though they often parameterized the analysis using the slider, users seemed to discount their own input into this semi-automatic process.

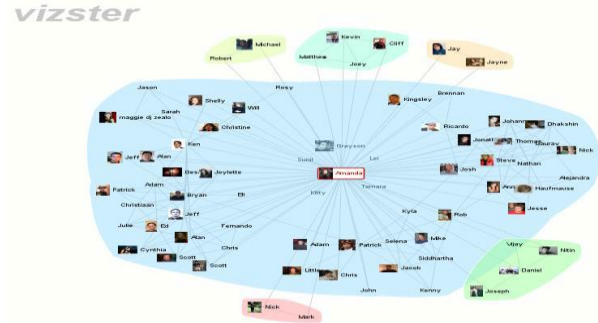


Figure 10: Community structure visualization using algorithmically determined optimum

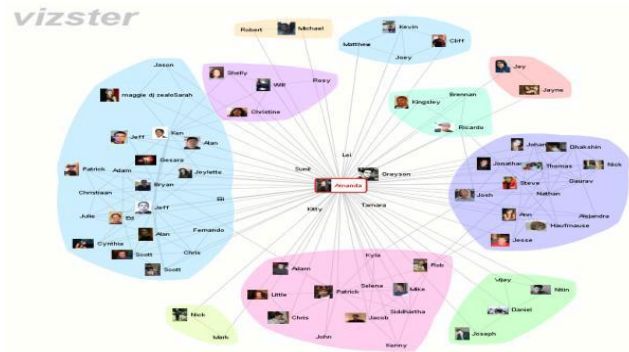


Figure 11: Community structure visualization after the community slider has been dragged to the right.

Users easily discovered the X-ray mode, often commenting on the results of their ad-hoc analysis. It proved particularly powerful when coupled with searches, as users naturally specified conjunctive queries by running search queries on top of the X-ray state. As we expected, however, users in both contexts showed a strong preference for imagery for navigating the networks. The lab context also provided valuable feedback on various features. In particular, linkage views were introduced in response to early user requests for a means of quickly unearthing connections without expanding another network. However, in later studies we found that both linkage views and the ability to inflate nodes were not discovered through experimentation, though were appreciated after consulting the reference. One participant suggested adding a mechanism to opportunistically suggest features to users- an approach

with a successful track record in video games.

In all, we found that the visualization was used by users to both explore and play with their networks-expanding the network to quite large depth, performing visual analyses, and exploring community structures while simultaneously engaging in social narratives. Although more formal comparatives and longitudinal studies would be required to more fully assess the utility of the individual features introduced, the engaged yet playful exploration and discovery we observed helps validate a number of our design decisions.

VII CONCLUSION

In this paper we have described the design of Vizster, a system for end-user visual exploration of online social networks. Grounded in an ethnographic investigation of the Friendster social networking service, the resulting Visiter system is an experiment in applying the fruits of information visualization within an end-user context more characteristic of play than analysis. As such, Vizster contributes a number of network analysis techniques designed to both reveal information and maintain a playful social environment.

These features include connectivity highlighting and linkage views for users to explore greater network contexts, X-ray mode and profile search for the visual analysis of member profile data and community analysis for investigation higher-level community structures. Observations of usage found that users find these features both useful and enjoyable for exploring their personal online networks. In the future work, we are interested in exploring the extension of these techniques to more traditional analytical domains, where more rigorous evaluation should be applied to evaluate their benefits and trade-offs. As forms of online sociality continue to spread and diversity, we also anticipate the need for further design approaches balancing utility and experience, enabling information visualization technologies to take on an increasingly important role in both mediating and making sense of our shared social landscape.

VIII. ACKNOWLEDGEMENT

We would to thank the anonymous referee for helpful comments.

IX. REFERENCES

- [1]. Boyd, d. Friendster and publicly Articulated Social Networks. CHI 2004, Vienna, Austria, 1279-1282.
- [2]. Adamic, A., O.Buyukkokten, E. Adar. A social Network Caught in the web. First Monday, 8 (6), 2003.
- [3]. Barnes, J., P. Hut. A Hierarchical O (N log N) Force Calculation Algorithm. Nature, 324,4 December 1986.
- [4]. Adar. E, GUESS. The Graph Exploration System. <http://www.hpl.hp.com/research/idl/projects/graphs/index.html>

dex.html

- [5]. BuddyZoo. <http://buddyzoo.com>
- [6]. Donath, J., d. boyd. Public displays of connection. BT Technology Journal, 22 (4), October 2004, 71-82.
- [7]. Fisher, D., P. Dourish Social and temporal structures in everyday collaboration. CHI2004, Vienna, Austria, 551-558.
- [8]. Ghoniem, M., J.-D. Fekete, P. Castagliola. A Comparison of the Readability of Graphs Using Node-Link and Matrix-Based Representations. InfoVis 2004, Austin, TX, 17-24.
- [9]. VisualThesaurus. <http://www.visualthesaurus.com>
- [10]. Heer, J., S.K. Card, J.A. Landay. Prefuse. A Toolkit for Interactive Information Visualization CHI 2005, Portland, OR, 421-430.
- [11]. JUNG: Java Universal Network/Graph Framework. <http://jung.sf.net/>
- [12]. Lucene Search Engine. <http://lucene.apache.org>
- [13]. Nardi, B., S. Whittaker, E. Isaacs, M. Creech, J. Johnson, J. Hainsworth. ContactMap: Integrating Communication and Information Through Visualizing Personal Social Networks. Communications of the ACM, April 2002.
- [14]. Newman, M.E.J. Co-authorship networks and patterns of scientific collaboration. Proc. Natl.Acad.Sciences, 101, 2004, 5200-5205.
- [15]. Newman, M.E.J. Fast Algorithm for detecting community structure in networks. Physical Review E, 69, 2004, 066133.
- [16]. Shneiderman, B. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. 1996 IEEE Conference on Visual Languages, Boulder, CO, 336-34
- [17]. TouchGraph Live Journal Browser. TouchGraph. LLC. http://www.touchgraph.com/TG_LJ_Browser.html
- [18]. UCINET: Social Network Analysis Software <http://www.analytictech.com/ucinet.htm>
- [19]. Viegas, F., d. Boyd. D. Nguyen, J. Potter and J. Donath. Digital Artifacts for Remembering and Storytelling: Postl history and Social Network Fragments. HICSS-37, Big Island, HI, 2002.

Short Bio Data for the Author



Mr. K. Balaraju received his B.Tech in Computer science and Engineering from Christu Jyothi Institute Of Technology And Sciences, JNTU, Hyderabad and Pursuing M.Tech in Computer science (Software Engineering) from Aurora's Technological And Research Institute, JNTU, Hyderabad.

Email: balaraju.kothapalli@gmail.com



Mr. K. Raju received his M.Tech in Computer science (Software Engineering) from Aurora's Engineering College, JNTU, Hyderabad and working as Assistant Professor in the Department of Computer science and Engineering in Supraja Institute of Technology And Science, JNTU, Hyderabad.

Email: kothapally.raju87@gmail.com



Mrs. K. Vidya working as Associate Professor in the Department of Computer Science and Engineering, Aurora's Technological And Research Institute College with a teaching experience of 8 years. She has received her M.Tech in Computer science (Image Processing). Her areas of interest includes Mobile Computing, Information Security and Mobile Commerce.

Email: kammavidya@gmail.com