Academic **N**etwork **Ex**plorer: Making Sense of Your Research through Interaction

Zipeng Liu

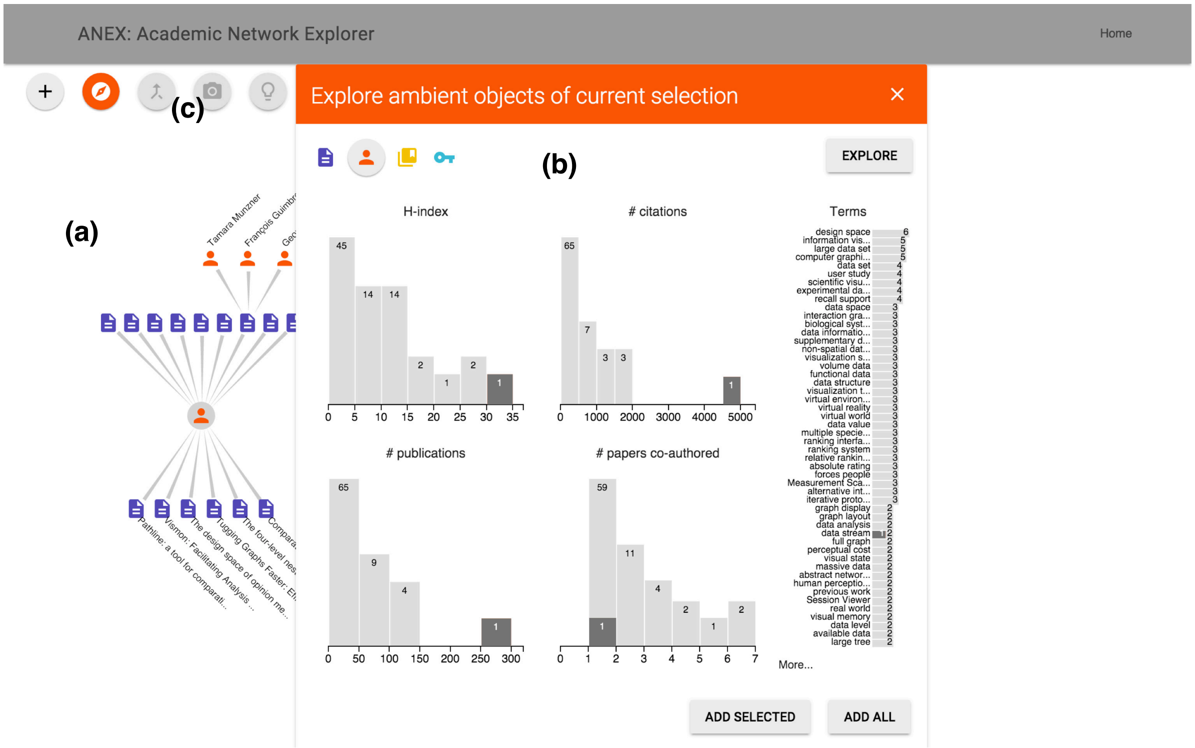


Fig. 1. The main interface of ANEX. It consists of (a) a infinite canvas holds the node-link diagram of all seeds and expanded objects; (b) an explore card (panel) showing the distributions of co-authors of a selected author on different attributes; (c) A tool bar contains handy function buttons: add seed, explore ambient objects, merge nodes, take snapshot of current system, show tips for next step. Four icons with four different colors represent the four types of objects throughout all the interfaces: paper, author, venue, term.

Abstract— Literature reading is a necessary but never easy task for researchers, especially for beginners of academia or in time limited circumstances. To make sense of literature, researchers need not only to read papers, but also to understand their context such as the authors who wrote them, where they are published and what are their impact on certain research areas. I propose a visualization system to explore and make sense of local academic network. The system, which I call Academic Network Explorer (ANEX), allows users browse and explore multiple objects like papers and authors around a seed, and also understand the relationships between these objects. ANEX is implemented with modern web based technology using a dataset of million of academic papers and authors, but ANEX is also suitable for other scholarly dataset with the same kind of information.

Index Terms—Information visualization, Network visualization, Multi-faceted data

1 Introduction

Making sense of literature requires paper reading, but usually it is impossible to read all papers in a researcher's interesting field, or even all related work of a paper he is working on. Experienced professors knows the history of his field, knows the majority of his community, knows what topics are hot spots and etc, which provide him with enough context to understand literature faster and deeper than young researchers. He has been there for many years, and information collected throughout the time constructs his wisdom in his mind.

As we are in the digital information century now, it is possible to create systems that can externalize this wisdom, which will greatly benefit students and young researchers who lack context of a field to quickly grasp the key knowledge they need. There have been already numerous effort in analyzing literature and bibligraphic data in the visualization, and what I propose here is another tool to explore an academic network.

I design the Academic Network Explorer, coined as ANEX, to enable a user to understand the context of certain objects he is interested in. Object in this report refers to four kinds of academic-related concept: paper, author, venue, term. I invent a rectangle-like layout algo-

- *Zipeng Liu is with University of British Columbia. E-mail: zipeng@cs.ubc.ca.*

This report is submitted in fulfillment of CPSC 547 Information Visualization, Computer Science, University of British Columbia, Fall 2015.

rithm to position the explored objects in order to present a clear path of history exploration and the current system state. In the following sections, I will first introduce related studies to ANEX, and then go through the data I use, the task ANEX supports. I will articulate the concrete design and the underlying thoughts of it, and show you how a typical usage scenario of ANEX. At last, I will discuss the limitations and drawbacks of ANEX, the future work and conclude it.

2 Related Work

This project is related to a few previous studies and systems spanning from different areas in visualization.

2.1 Literature Analysis

Starting from decades ago, before the "birth" of visualization, researchers have studied how to display paper citation data to augment writing reports on history of science [9]. Later when digital collection of papers became available, Chaomei Chen, interested in research trends, detected "research fronts" using burst detection and betweeness analysis, and then visulized them in a clustered node-link graph [6]. John Stasko's group from Georgia Tech visualized papers published in IEEE Information Visualization Conference from 1995 to 2012 [18]. Their CiteVis system could show citations and references of papers in details, and rankings by citations as well. They got some interesting findings and patterns in visualization publications. They also tried different methods such as CiteMatrix, CiteList. In the InfoVis 2004 contest, participants were required to come up with visualizations to analyze a dataset with eight years' of InfoVis publications. Researchers in PNNL used the tool IN-SPIRE [1] they built earlier for exploring large corpus to analyze the InfoVis literature [22]. They formed clusters of documents in their "Galaxy" view and showed research trends in "Theme" View where trends stood out as mountains on a plain. Lee et. al designed PaperLens [12] dedicated for this contest to present hot topics through years, sorted list of papers and authors and supported cross reference of papers and authors. SurVis [5] worked on carefullly surveyed literature collection in order to disseminate literature. Users such as survey authors could structure their references with the powerful selector interaction. There are also literature browsers like Treevis.net [17], Timeviz.net [3] to disseminate visualization literature. Unlike many of these previous studies, ANEX takes a different angle and focuses on visualizing the network indicated by citation, co-authorship and other relationships. Nees Jan van Eck et al. developed a software tool to construct and visualize bibligraphic network [21], but generally they followed the overview to details mantra, which contradicted with mine. The same method went with CiteWiz by Elmqvist et. al [8]. The exploration in ANEX only concerns about local networks.

2.2 Faceted Data

Scholarly dataset is also regarded as typical faceted data, which people have come up with creative ways to explore. Preliminary research on faceted data focused on browsing and searching [24, 16]. Marian Dork et. al proposed faceted information space [7], which used pivot interaction to enable strolling in the space. Their succinct design and slick transition inspired ANEX a lot. Jian Zhao et. al built PivotSlice [25] to easily browse multiple facets and find correlations between them. They treated facets as sets and supported expressive set manipulations through rich interaction techniques. PivotSlice was better in understanding overview of facets and relations in between, but lack of details in individual objects, which ANEX provides. Also, Keshif [23] took facets as sets to support filtering and comparison in a highly interactive , versatile document browser.

2.3 Network and Set

As we think my dataset as network data, there is work related from network visualization. Detangler [15] addressed cohesion problems of multiplex network by constructing a substrate and catalyst layer and enabling "leapfrog" between them. Kieffer et. al studied how human produced elegant orthogonal networks by hand and dervied guidelines to automate such process [11]. The layout algorithm we come up with was greatly inspired by them.

As we pointed out in the data abstraction (pretending that this is the final report :)), part of the data could be also visualized as sets [4]. We also leverage techniques from sets such as union and intersection operation to select our expanding interests (central objects), multiple linked views to visualize distributions of attributes [23], and rankings of attributes [13]. These techniques help users understand what are the ambient objects and filter exactly what they want.

3 Data

The dataset I use includes paper information, paper citation, author information and author collaboration. There are about two million papers, eight million citations, one million authors and four million collaborations. For papers, there are attributes such as title, authors, affiliations, publication venue, year, abstraction and references; for author, there is name, affiliations, count of published papers, extracted key terms and scholarly indexes like H-index. This dataset is provided by ArnetMiner [19], which scraped and collected publications and authors in computer science for years. It can be downloaded from https://aminer.org/billboard/AMinerNetwork.

It is a typical network data induced by multiple relations like citation, co-authorship. It can also be regarded as set data since there are natural belongingness like authors are in affiliations and papers are published in venues, and also artificial ones like authors are described by some key terms. Besides, as many previous work pointed out, this is a faceted data because objects have many different aspects: papers have a few associated attributes.

The data was scraped from several digital libraries which might use different representations. Inconsistency and deficiency was detected during preliminary data processing and validation. The format of venue and affiliation name is extremely diverse. Different entities (authors) turned out to be the same and should be merged. It misses important entities like there is no term "visualization". I will point out more specific findings on data deficiency in the case study section. Despite of poor data quality, this dataset does support illustrating the exploration pattern and design of ANEX.

4 Task Abstraction

To understand the problem I design for, I will first think of what are the typical and important task that target users do, and then generalize them using idioms in visualization to facilitate designing a solution.

4.1 Domain Specific Tasks

In this project, I focus on making sense of a local region, which users are interested in, of this giant academic network from multiple angles. Also, we facilitate finding related work of certain topics and provide overview of them. Notice that we are not interested in the main content of the published papers, which requires text related techniques.

ANEX should be able to answer the following questions. The list is not exhaustive, but indicates typical and basic functions. By integrating several tasks, users should be able to gain a mental image of their areas of interests and have a sense of general direction of what to read next.

- Making sense of a paper
  - What is this paper about? What are the related topics?
  - What are the references of this paper? Who cited it after its publication?
  - How important is this paper in terms of related topics
  - What are the similar papers?
- Making sense of a researcher
  - What does he work on?
  - What is he famous for? What's his contribution in the field?
  - Who are his co-authors? How do they do in the field?

  - Who are the most related researchers?
  - How does his work evolve through time?

- Making sense of a venue
  - What is this venue about?
  - What is its academic impact?
  - Who published papers on it?
  - What are the most influential work?

- Making sense of a few related topics
  - What are some related or similar key terms?
  - What are the some influential (must-read) papers?
  - Who are working on it? Who are the most influential authors?
  - How does hot research questions evolve?
  - What are the first-tier publication venues?

In general, all of these questions could be generalized as exploration of ambient objects around a centered object whose relations might be given by citation, authorship, collaboration, belongingness.

4.2 Visualization Task

According to the user needs, I abstract them to tasks in "visualization language" presented in the textbook [14].

- Locate (target known, location unknown) a seed object by text searching to start
- Explore ambient objects around a central object including presenting distributions of attributes of ambient objects and discovering trends and outliers
- Derive a subset of ambient objects of a center
- Present the path of exploration process

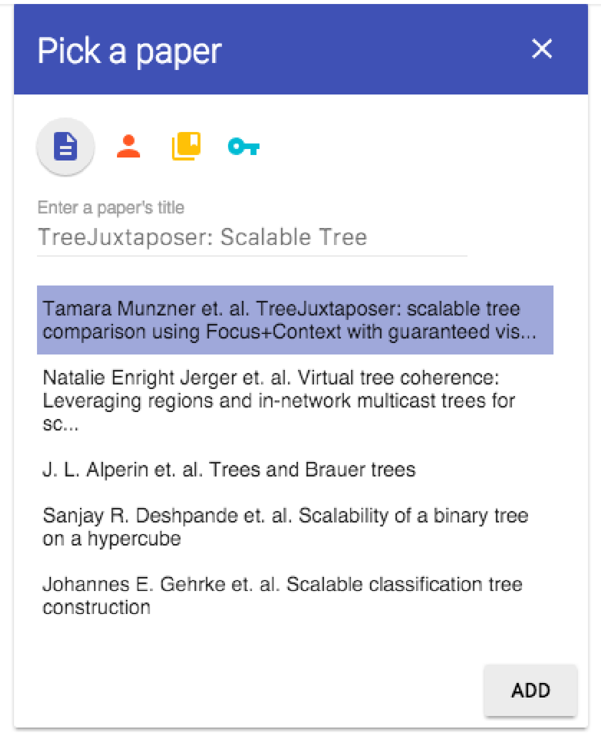
5 Design

I will start with the general design principle of ANEX, and then walk you through each part of the interface in this section, and articulate the design considerations behind it.

5.1 Design Rationale

The nature of this problem is exploring objects locally, and hence we are not interested in providing users with overview of the whole dataset, but instead, enabling expansion from a seed through rich interaction. ANEX follows the general guideline "from detail to overview via selections and aggregations" coined as DOSA by Stef van den Elzen et. al [20], except that the overview in our case is more of a local context.

The task abstraction highlights the networking feature (objects and relationships). ANEX's exploration process is adding up pieces by pieces starting from a seed through flexible and interactive expansion, which supports specifying different objects and selecting interesting subsets given attribute distributions. As a side effect of the expansion, users can grasp the features of attributes.



Enter a paper's title

TreeJuxtaposer: Scalable Tree

Tamara Munzner et. al. TreeJuxtaposer: scalable tree comparison using Focus+Context with guaranteed vis...

Natalie Enright Jerger et. al. Virtual tree coherence: Leveraging regions and in-network multicast trees for sc...

J. L. Alperin et. al. Trees and Brauer trees

Sanjay R. Deshpande et. al. Scalability of a binary tree on a hypercube

Johannes E. Gehrke et. al. Scalable classification tree construction

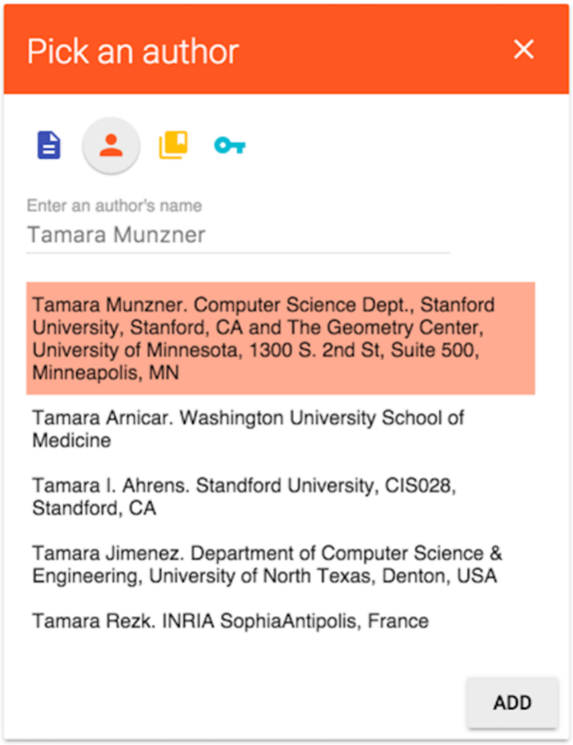
Fig. 2. Illustration of a seed card. A user is looking for a paper "TreeJuxtaposer". Color indigo is dedicated only for object type paper.

5.2 Seed Pick-up

In ANEX, a user needs to have something specific to start with, for example he knows a paper's title, at least part of it. A seed card, illustrated in Fig. 2 and Fig. 3, is the view for searching and adding a seed. I refer panel as card in ANEX to align with the concept in the material design by Google. The four icons in the card in four different colors represent four different types of objects paper, author, venue and term from left to right respectively. Each time the user changes the seed object type by clicking another icon, the color of elements on the card changes accordingly including header, background for selected object, searching indicator. I put a non-trivial effort on color consistency to make sure that the user is aware of what kind of objects are he interact with currently in anytime just by seeing the color. After selecting the type of seed, the user enters part of the words of his interest. As he is typing, ANEX automatically searches the database and returns the five most relevant objects, which he could select one seed from. The current selecting seed is highlighted with a slightly tinted color. In order not to overwhelm the server by sending too many queries while users are actively typing in, I debounce the trigger of query, which is a common technique used in text searching between client and server. Once the seed is selected, he could click on the add button to add the seed node to the main canvas for exploration.

5.3 Ambient Objects Exploration

To explore the ambient objects around a center, the user must specify the center first by clicking the node, which is shown in grey background. Fig. 1, Fig. 5 and Fig. 6 show the interface for exploring papers, co-authors and venues around the central author Tamara. Because I have more information on papers and authors, and venues and terms are actually derived from the formers' attributes, multiple linked histograms are shown for papers and authors, while a simple list of names of objects for venues and terms. Attribute values are binned according to data type, which is quantitative or ordinal, and range of values. For ordinal values such as venue, author and term, values are counted distinctively; for quantitative data such as number of citations, H-index, values are aggregated by bins, which are generated by divid-



Pick an author

Enter an author's name

Tamara Munzner

Tamara Munzner. Computer Science Dept., Stanford University, Stanford, CA and The Geometry Center, University of Minnesota, 1300 S. 2nd St, Suite 500, Minneapolis, MN

Tamara Arnicar. Washington University School of Medicine

Tamara I. Ahrens. Stanford University, CIS028, Stanford, CA

Tamara Jimenez. Department of Computer Science & Engineering, University of North Texas, Denton, USA

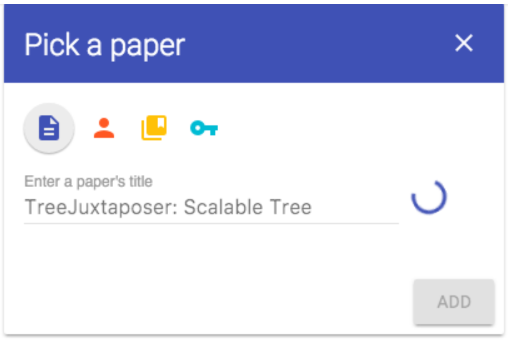
Tamara Rezk. INRIA SophiaAntipolis, France

Fig. 3. Illustration of a seed card. A user is looking for an author "Tamara Munzner". Color orange is dedicated only for object type author.

ing the value range evenly. Log scale on the height of bars are used in order to avoid too much divergence of the heights caused by diverging values, and the number of objects of each bar is shown on the top to amend visual imprecision caused by the log scale. Charts are linked with highlight: the user can hover on a bar to see how the object that fall into current bar are distributed in other histograms in grey; he can also click on the bars to select a subset of objects, which will be highlighted in the corresponding color. The user can choose between adding all ambient objects of the selected type and adding only the selected objects to the canvas.

5.4 Rectangular Node-link Graph

All the objects added to the canvas are shown in a node-link graph, as illustrated in Fig. 7. Node shapes are determined by the object types, and the same as icons in the seed card and explore card. Same goes with the color. Tapered edges are drawn between parent nodes and its children to illustrate the path of exploration. I specifically choose tapered edges over arrow edges or other type because study showed that it is the most evident one to indicate direction [10].

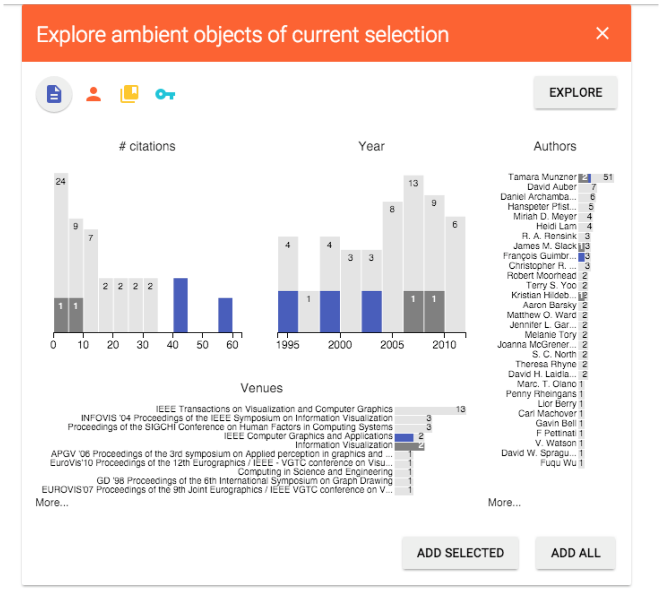


Pick a paper

Enter a paper's title

TreeJuxtaposer: Scalable Tree

Fig. 4. Illustration of a seed card while a user is searching for a paper.



Explore ambient objects of current selection
×
EXPLORE
# citations
24
9
7
2
2
2
2
1
1
0
10
20
30
40
50
60
Year
4
1
4
3
3
8
13
9
6
1
1
1995
2000
2005
2010
Authors
Tamara Munzner 2 51
David Auber 7
Daniel Archamba... 6
Hanspeter Pfist... 5
Miriah D. Meyer 4
Heidi Lam 4
R. A. Rensink 3
James M. Slack 1 3
François Guimbr... 3
Christopher R. ... 3
Robert Moorhead 2
Terry S. Yoo 2
Kristian Hildeb... 1 2
Aaron Barsky 2
Matthew O. Ward 2
Jennifer L. Gar... 2
Melanie Tory 2
Joanna McGrener... 2
S. C. North 2
Theresa Rhyne 2
David H. Laidla... 2
Marc. T. Olano 1
Penny Rheingans 1
Lior Berry 1
Carl Machover 1
Gavin Bell 1
F Pettinati 1
V. Watson 1
David W. Spragu... 1
Fuqu Wu 1
More...
Venues
IEEE Transactions on Visualization and Computer Graphics 13
INFOVIS '04 Proceedings of the IEEE Symposium on Information Visualization 3
Proceedings of the SIGCHI Conference on Human Factors in Computing Systems 3
IEEE Computer Graphics and Applications 2
Information Visualization 2
APGV '06 Proceedings of the 3rd symposium on Applied perception in graphics and ... 1
EuroVis'10 Proceedings of the 12th Eurographics / IEEE - VGTC conference on Visu... 1
Computing in Science and Engineering 1
GD '98 Proceedings of the 6th International Symposium on Graph Drawing 1
EUROVIS'07 Proceedings of the 9th Joint Eurographics / IEEE VGTC conference on V... 1
More...
ADD SELECTED
ADD ALL

Fig. 5. Illustration of an explore card where the distributions of papers on a few attributes that Tamara wrote are shown. Current hovering papers are in grey and current selected ones in indigo, which is the color for paper in ANEX.

Given that there is only four types of objects, I assume that there will be only a trivial number of expansion on each parent. To maintain the stability of the graph so that the user does not loose track of what have been explored, I designed a "rectangular-like" layout algorithm specifically for ANEX. The nodes already added would never be repositioned, and newly added nodes are positioned around them, so the user would have a feeling of growing a tree along his exploration process.

Here is how I calculate the positions of children, as shown in Fig. 8. The underlying rationale is to position them along a border the the rectangle around the center and also give them reasonable amount of room. I will use the north border as an example. Denote n as the number of children, d as the distance between the leftmost child and rightmost child, r as the distance between parent and leftmost child or rightmost child, and α as the included angle. d is given by a function:

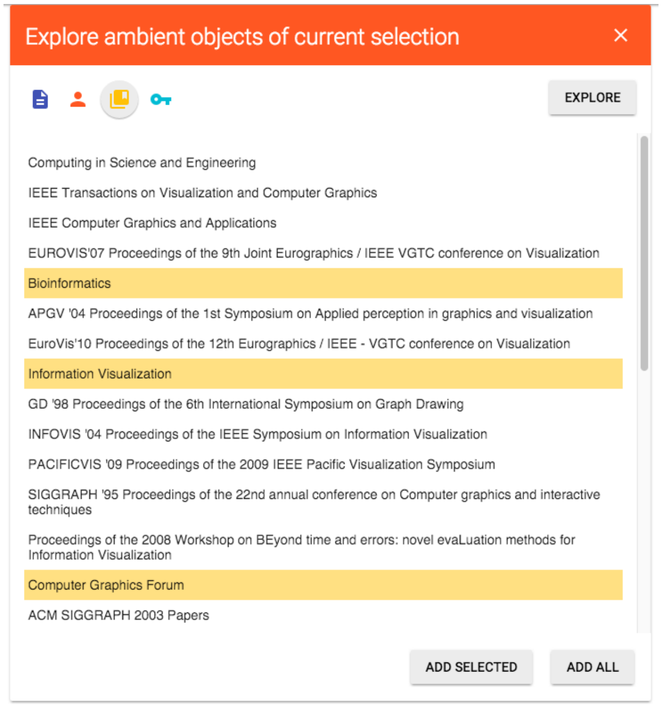
$$D(n) = (A(n-1)^{-1} + B)n, n \geq 2 \qquad (1)$$

where $A+B$ and B are the maximum and minimum distance between two consecutive children. Once d is given, I assign a boundary on r: $min(r) \leq r \leq max(r)$, and then α is the average of the maximum α achieved by $min(r)$ and the minimum α achieved by $max(r)$. In the current implementation, I use a configuration of $A = 40, B = 20, min(r) = 50, max(r) = 400$. After d, r and α are set, I translate them into the offset coordinates from the parent and position each child accordingly.

Labels are shown for the most "outer" levels of nodes to avoid occlusion. More specifically, labels of a parent and also its direct siblings, which are added in the same batch and located along the same border of a rectangle, are hidden once the children are "grown" from the parent. Labels of children are shown by default. Due to the unique property of this rectangular layout, labels are only allowed to be read along four directions as illustrated in Fig. 7 for both readability and aesthetics. Long labels are automatically chopped to avoid crowded text.

5.5 Styling

I employ material design [2] as the major styling guidelines for ANEX to create an unified user experience. Icons are all from Google de-



Explore ambient objects of current selection ×

EXPLORE

Computing in Science and Engineering

IEEE Transactions on Visualization and Computer Graphics

IEEE Computer Graphics and Applications

EUROVIS'07 Proceedings of the 9th Joint Eurographics / IEEE VGTC conference on Visualization

Bioinformatics

APGV '04 Proceedings of the 1st Symposium on Applied perception in graphics and visualization

EuroVis'10 Proceedings of the 12th Eurographics / IEEE - VGTC conference on Visualization

Information Visualization

GD '98 Proceedings of the 6th International Symposium on Graph Drawing

INFOVIS '04 Proceedings of the IEEE Symposium on Information Visualization

PACIFICVIS '09 Proceedings of the 2009 IEEE Pacific Visualization Symposium

SIGGRAPH '95 Proceedings of the 22nd annual conference on Computer graphics and interactive techniques

Proceedings of the 2008 Workshop on BEyond time and errors: novel evaLuation methods for Information Visualization

Computer Graphics Forum

ACM SIGGRAPH 2003 Papers

ADD SELECTED

ADD ALL

Fig. 6. Illustration of an explore card which shows a list of venues that Tamara had her papers published in. Selected venues are highlighted with ember background.

sign, and colors also from its color palette. Color usage is consistent throughout all interfaces in ANEX to provide awareness of current interested object type.

6 Implementation

The system is implemented in a typical Browser-server (BS) architecture given that there is no way to load gigabytes of data into the browser all at once and that user only query a small amount of data in each step.

I pre-processed all the raw data in plain text documents and put it into a MongoDB database with Python. During the pre-processing, I created indexes and look-up tables in a Redis cache so that I can process multiple chunks of data in parallel for efficiency, and also stored the indexes in MongoDB to cross reference related objects. Then I wrote a NodeJS server using Express to set up RESTFul APIs to search, retrieve and explore documents.

On the forntend, I picked ReactJS as the major UI rendering library and Reflux as the application architecture. Although I could have chosen something that I am familiar with, I want to try React+Reflux for the sake of learning since React is a rather revolutionary way of thinking which is more like functional programming than mainstream framework like AngularJS. I also used D3 for creating the histograms and bar charts in the explore card though it was quite awkward to integrate D3 with React. For the material design, I used the Material Design Lite UI library and Material Icons from Google.

7 Case Study

Here I will illustrate how to use ANEX by walk you through a typical scenario and I will also reveal some of the problems in the dataset.

As a graduate student who knows a little about visualization , I am interested in Professor Tamara Munzner. First I initiate a seed card and select the author icon. I enter her name on the text field and after a few seconds, the system shows me a list of related authors with their first affiliation recorded in this data, as illustrated in Fig. 3. I select and add this seed to the canvas.

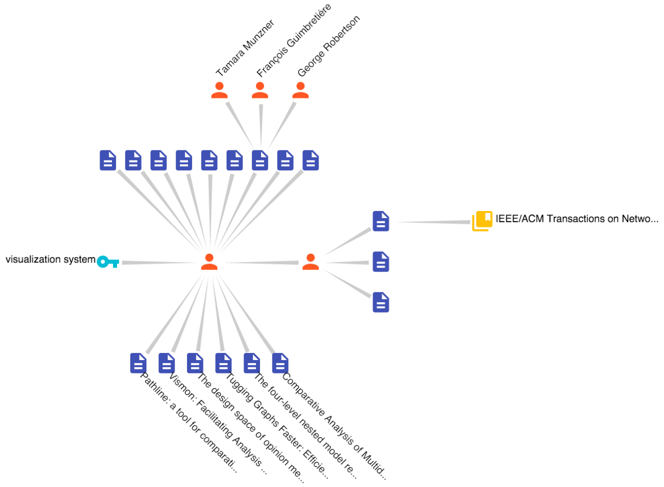

Tamara Munzner
François Guimbretiére
George Robertson
IEEE/ACM Transactions on Netwo...
visualization system
Pathline: a tool for comparati...
Vismon: Facilitating Analysis ...
The design space of opinion me...
Tugging Graphs Faster: Efficie...
The four-level nested model re...
Comparative Analysis of Multid...

Fig. 7. Illustration of the rectangular graph layout.

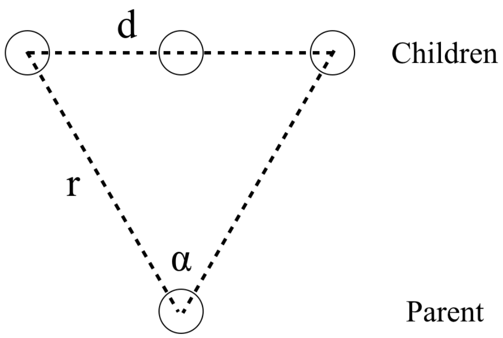


Fig. 8. Illustration of calculation of children positions.

Next I am going to explore her ambient objects. I open up an explore card on Tamara, select paper type, and the distributions of papers her wrote are shown (Fig. 5). Notice that there are a total of dozens of papers in this dataset, which is obviously not complete. According to Google Scholar, she published more than 120 papers. The same thing happens to the citation. Throughout this exploration, I find out the dataset is rather incomplete in the amount of papers, authors and citations. By hovering on the bars of the citation histogram, I can see how papers with different citations distributed across years, published venues and co-authors. I discover that the number of citations generally increase if the paper was published longer, which makes sense that earlier papers has a bigger probability to be cited. I pick and add the three highly-cited papers.

Then I look into her co-authors (Fig. 1) and pick up the most highly-cited researcher who also has the most publications. When I turn to the venues she had published in, things are pretty untidy, as you can see in Fig. 6. The names of venues are not consistent: sometimes with the year sometimes not, some with publisher and some not, etc. I guess this is mostly caused by automatic extraction algorithm of the Arnetminer, who provides the dataset, and it requires exhaustive human effort to correct it.

I can further explore the objects that I added around Tamara with the same actions for example who co-authored her most cited paper, who is the most cited co-author and what field he is in.

8 Discussion and Future Work

Because of this special exploration pattern of a local network, ANEX will not provide any grand big pictures of the entire dataset. Actually I change one of the guideline proposed by Stef [20] from "from detail to overview via selections and aggregations (DOSA)" to "from detail to context via selections and aggregations (DCSA)". However, the effectiveness of this "DCSA" is not proven and remains unknown currently. Assume the data is clean and complete, the quality of the produced network of ANEX largely depends on the user interaction. To be more specific, it depends on the selected subset of ambient objects and the order of adding them. There is no guidance on how to conduct the exploration effectively and efficiently.

The scalability is a problem in the layout algorithm. The rectangular feature limits number of times that a user can expand from a parent to only four, but hopefully he would not want to exceed four because there is only four type of objects. However, this is not guaranteed since he could add different subsets of the same type of ambient objects. Also, if one side of the rectangle is too long, it might cross other side, which results in occlusion of both nodes and labels. There is a vague upper bound of the number of nodes on one side: probably around dozens, because there is not enough pixels on the screen. The layout algorithm can be improved to address this problem such as using regular polygons instead of rectangles.

The poor quality of the dataset is evident, but thinking in another angle, clean and complete data is seldom available, and usually we should face and deal with the horrible reality. One possible future work is to correct the objects during the exploration. If we have enough users, the dataset would become tidy and complete. There are many tools for data cleaning and wrangling, but maybe cleaning and wrangling along the exploration and visualization is another choice.

For the current implementation, the dragging of a node sometimes lags if there are hundreds of nodes on the screen because of the awkwardness of library integration. Due to limited time, some of the features are not implemented but only has an entry on the interface, for example the merge nodes and snapshots.

9 Conclusion

I design and implement the Academic Network Explorer to enable a user to understand the context of certain objects he is interested in. I invent a rectangle-like layout algorithm to position the explored objects in order to present a clear path of history exploration and the current system state.

I learned a lot from this course project from data wrangling to visualization designs to implementation. Especially I learned the "React" way of thinking in user interface rendering, which is more functional like than any other conventional frontend frameworks, and I understand the hardness to write non-trivial interactions on the current React library. Things become even harder if D3 or other libraries are integrated since their underlying rationale are fundamentally different.

Acknowledgments

I would like to thank Professor Tamara Munzner for her invaluable guidance on this project.

References

[1] IN-SPIRE. http://in-spire.pnnl.gov/.

[2] Material design guidelines. https://www.google.com/design/spec/material-design/introduction.html.

[3] W. Aigner, S. Miksch, H. Schumann, and C. Tominski. *Visualization of time-oriented data*. Springer Science & Business Media, 2011.

[4] B. Alsallakh, L. Micallef, W. Aigner, H. Hauser, S. Miksch, and P. Rodgers. Visualizing sets and set-typed data : State-of-the-art and future challenges. In *Eurographics Conference on Visualization (EuroVis)*.

[5] F. Beck, S. Koch, and D. Weiskopf. Visual analysis and dissemination of scientific literature collections with survis. *Visualization and Computer Graphics, IEEE Transactions on*, 22(1):180–189, Jan 2016.

[6] C. Chen. Citespace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *journal of the american society for information science and technology*, 57(3):359–377, 2006.

[7] M. Dork, N. H. Riche, G. Ramos, and S. Dumais. Pivotpaths: Strolling through faceted information spaces. *Visualization and Computer Graphics, IEEE Transactions on*, 18(12):2709–2718, 2012.

[8] N. Elmqvist and P. Tsigas. Citewiz: a tool for the visualization of scientific citation networks. *Information Visualization*, 6(3):215–232, 2007.

[9] E. Garfield, I. H. Sher, and R. J. Torpie. The use of citation data in writing the history of science. Technical report, Institute for Scientific Information, 1964.

[10] D. Holten and J. J. van Wijk. A user study on visualizing directed edges in graphs. In *Proc. the SIGCHI Conference on Human Factors in Computing Systems*, pages 2299–2308. ACM, 2009.

[11] S. Kieffer, T. Dwyer, K. Marriott, and M. Wybrow. Hola: Human-like orthogonal network layout. *Visualization and Computer Graphics, IEEE Transactions on*, 22(1):349–358, Jan 2016.

[12] B. Lee, M. Czerwinski, G. Robertson, and B. B. Bederson. Understanding eight years of infovis conferences using paperlens. 2004.

[13] A. Lex, N. Gehlenborg, H. Strobelt, R. Vuillemot, and H. Pfister. UpSet: Visualization of intersecting sets. *IEEE Transactions on Visualization and Computer Graphics (InfoVis '14)*, 20(12):1983–1992, 2014.

[14] T. Munzner. *Visualization Analysis and Design*. AK Peters Visualization Series. CRC Press, 2014.

[15] B. Renoust, G. Melançon, and T. Munzner. Detangler: Visual analytics for multiplex networks. *Comput. Graph. Forum*, 34(3):321–330, 2015.

[16] m. schraefel, M. Wilson, A. Russell, and D. A. Smith. mSpace: Improving information access to multimedia domains with multimodal exploratory search. *Communications of the ACM*, 49(4):47–49, 2006.

[17] H.-J. Schulz. Treevis.net: A tree visualization reference. *Computer Graphics and Applications, IEEE*, 31(6):11–15, Nov 2011.

[18] J. Stasko, J. Choo, Y. Han, M. Hu, H. Pileggi, R. Sadanaand, and C. D. Stolper. Citevis: Exploring conference paper citation data visually. *Posters of IEEE InfoVis*, 2013.

[19] J. Tang, J. Zhang, L. Yao, J. Li, L. Zhang, and Z. Su. Arnetminer: Extraction and mining of academic social networks. In *Proc. 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 990–998, 2008.

[20] S. van den Elzen and J. van Wijk. Multivariate network exploration and presentation: From detail to overview via selections and aggregations. *Visualization and Computer Graphics, IEEE Transactions on*, 20(12):2310–2319, Dec 2014.

[21] L. Waltman, N. J. van Eck, and E. C. Noyons. A unified approach to mapping and clustering of bibliometric networks. *Journal of Informetrics*, 4(4):629 – 635, 2010.

[22] P. C. Wong et al. IN-SPIRE Infovis 2004 contest entry. 2004.

[23] M. Yalcin, N. Elmqvist, and B. Bederson. Aggreset: Rich and scalable set exploration using visualizations of element aggregations. *Visualization and Computer Graphics, IEEE Transactions on*, 22(1):688–697, Jan 2016.

[24] K.-P. Yee, K. Swearingen, K. Li, and M. Hearst. Faceted metadata for image search and browsing. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 401–408. ACM, 2003.

[25] J. Zhao, C. Collins, F. Chevalier, and R. Balakrishnan. Interactive exploration of implicit and explicit relations in faceted datasets. *Visualization and Computer Graphics, IEEE Transactions on*, 19(12):2080–2089, 2013.