CiteWiz: A Tool for the Visualization of Scientific Citation Networks

Niklas Elmqvist **Philippas Tsigas**
{elm,tsigas}@cs.chalmers.se
Department of Computer Science and Engineering
Chalmers University of Technology and Göteborg University
412 96 Göteborg, Sweden

ABSTRACT

We present CiteWiz, an extensible framework for visualization of scientific citation networks. The system is based on a taxonomy of citation database usage for researchers, and provides a timeline visualization for overviews and an influence visualization for detailed views. The timeline displays the general chronology and importance of authors and articles in a citation database, whereas the influence visualization is implemented using the Growing Polygons technique, suitably modified to the context of browsing citation data. Using the latter technique, hierarchies of articles with potentially very long citation chains can be graphically represented. The visualization is augmented with mechanisms for parent-child visualization and suitable interaction techniques for interacting with the view hierarchy and the individual articles in the dataset. We also provide an interactive concept map for keywords and co-authorship using a basic force-directed graph layout scheme. A formal user study indicates that CiteWiz is significantly more efficient than traditional database interfaces for high-level analysis tasks relating to influence and overviews, and equally efficient for low-level tasks such as finding a paper and correlating bibliographical data.

Author Keywords

citation networks, bibliographic visualization, information visualization, causal relations

INTRODUCTION

One of the key tasks of scientific research is the study and management of existing work in a given field of inquiry. The specific nature of the tasks involved in this venture vary greatly depending on the situation and the role of the researcher; for a new student just entering a research area, the task is that of orientation within the existing work; for a reviewer, one of originality and correctness checking; for a conference organizer, one of chronological survey; and, finally, for an experienced scientist, one of staying abreast with new developments, confirming intuitions, and identifying current hot topics in his or her area of choice. Researchers spend a considerable portion of their time on these tasks, ample evidence that it is in everyone's best interest to streamline this process as much as possible, and that large time savings can likely be made.

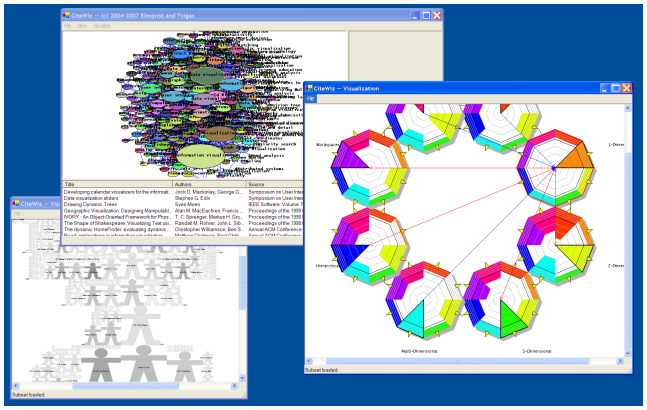

Figure 1. The CiteWiz prototype system running under Windows.

In this paper, we present CiteWiz, a tool for bibliographic visualization of the chronology and influences in networks of scientific articles. The tool contains three primary visualizations: a timeline visualization for overviews and navigation in a full citation database, an influence visualization for detailed views of a specific subset of the citation database, and an interactive concept map for exploring keywords and co-authorship in the database. Users would typically orient themselves in a citation database using the timeline and the concept map and then construct specialized subsets to study using the influence visualization. The tool was designed for use by researchers, scientists, and students alike, and its baseline features were established through extended discussions in a focus group consisting of such users. These discussions helped us formulate a taxonomy for citation database interaction that we believe is a valuable contribution to the area. Guided by this taxonomy and the focus group, we created a prototype implementation of the tool with a user interface that allows for normal browsing and filtering of the citation meta-data as well as building nested subsets of the dataset for visualization. We have conducted a

This is the author's version of a paper that appears in the *Information Visualization* journal, volume 6, issue 3, pages 215–232, 2007.

Correspondence: Niklas Elmqvist, INRIA/LRI, Bat 490, Université Paris-Sud, 91405 Orsay Cedex, France (Tel: +33 1 69 15 61 97).

Received: 14 August, 2006
Revised: 27 April, 2007
Revised: 22 August, 2007
Accepted 24 August, 2007

formal user study to assess the efficiency of the tool in comparison with standard web-based database interfaces. Our results indicate that CiteWiz is equally efficient as standard database interfaces for low-level analysis tasks such as finding papers and correlating authors, and significantly more efficient than standard databases for higher-level analysis tasks related to overviews and influences of bibliographical data.

Causality and influences both play a large role in tracing the history of ideas and trends in a scientific community, and these are core strengths of the Growing Polygons [12] technique. In order to allow us to make use of this technique, we show how to model citations in scientific articles using general causal relations, and we introduce the slightly relaxed concept of *influence* between articles in a citation database. We chose an article-centered approach (as opposed to an author-centered one) in our implementation, where the articles themselves are the active entities (represented by processes), and citations are the information-bearing messages between them. To allow the technique to cope with potentially huge datasets, we also improved its scalability in two different ways: we implemented multi-level process hierarchies for grouping sets of articles together, and we added a focus+context technique with variable time scale to handle long event histories. The visualization was accordingly supplemented with a number of interaction techniques to support these new features as well as techniques targeted specifically at citation visualization; these include collapsing and expanding the group hierarchy, navigating in the citation network by following backward and forward references, and getting details-on-demand of the complete bibliographical data for a specific paper.

The rest of this article is structured as follows: we begin by describing the state of the art in visualization of citation networks. We then give background information on citation networks and presents our taxonomy for its use. We present the CiteWiz platform, including the three visualizations. We then give a case study to highlight the CiteWiz workflow and how these visualizations complement each other. The following section describes the user study, and our results are presented in after that. We finish the paper with a discussion in and some conclusions.

RELATED WORK

The common model of viewing citation networks as directed graphs (see the next section) lends itself quite naturally to visualizing bibliographical data as simple node-link diagrams. However, node-link diagrams has two major weaknesses [17]: they scale poorly for dense networks, and (ii) require aggregation methods to reduce the density enough to be readable. In the context of citation networks, node-link representations show only local dependency information; it is easy to see direct citing and cited articles, but the user must traverse the graph in order to see dependencies more than one step away. The CiteWiz system presented in this article, on the other hand, provides the surrounding context by directly showing the dependencies of an article, yielding a much more straightforward way to see the chronology.

Modjeska et al. [25] propose a minimum set of functions necessary for effective bibliographic visualization: (i) display of complete bibliographic information, (ii) filtering by record fields, (iii) display of chronology and influence of articles, (iv) information views at different levels of detail, (v) multiple simultaneous views, and (vi) visualization of large search results. They also present the BIVTECI prototype system that partially implements this specification, but the visualization used in the tool is restricted to node-link diagrams with visualized attributes. CiteWiz also implements this minimum functionality, but instead employs the Growing Polygons causality visualization technique in order to handle larger search results and provide stronger chronology information.

The RefViz [29] system is a commercial visualization tool for scientific citation networks. It is based on two main views, a "galaxy" network view showing clusters of papers based on their conceptual relations, and a concept occurrence "matrix" view showing the distribution of concepts being discussed in relation to each other. This is orthogonal from the CiteWiz approach, where the basis for the visualization are the actual references in each paper.

The Butterfly [24] system provides a 3D visualization frontend of the DIALOG science citation databases, using the notion of "organic user interfaces" to build an information landscape as the user explores the results of various queries. Individual articles are represented by a butterfly-shaped 3D object with references and citers on the left and right wings, respectively, and various graphical cues are provided to orient the user when browsing the citation network. Butterfly uses a node-link diagram for overview and context, however, and has no mechanism for showing the cumulative influences and chronology of articles.

The results from the InfoVis 2004 contest [13] are of special interest to this work, especially since we use the same citation database for our prototype implementation. Many excellent entries, developed concurrently with our work, were presented at the contest. Ke, Börner and Viswanath [21] use a graph visualizer to show the relationship between the major papers in the database, scaling node size proportionally to the number of citations just like entities in the CiteWiz timeline visualization. Delest et al. [9] use a related approach. Keim et al. [22] employ InterRings [35] to visualize co-authorship, a technique somewhat similar to using Growing Polygons for co-authorship studies.

A number of citation-based visualization systems similar to CiteWiz have been proposed previously; VxInsight [8, 2] is a general knowledge management system where relations between articles (i.e. citations and keywords) are used to map the data objects to a 3D terrain that is rearranged using a force-directed layout scheme. CiteSpace [6] (recently updated to its second version) provide support for the full work process for studying a scientific community, including operations such as filtering, time slicing, pruning, merging, and visual inspection.

CiteWiz and the above-mentioned systems are all article-focused tools in that they emphasize the visualization of

articles and their interdependencies. A number of group-focused techniques have also been proposed, where the emphasis lies on representing the groupings and structure of a scientific domain through metrics such as relevance, bibliographic coupling [23], and co-citation [32]. Work in this area is numerous but peripheral to the system described in this paper; examples include [3, 4, 5, 7, 19].

Timelines, like the Newton's Shoulders visualization in CiteWiz, have been widely used for applications like personal histories [28], time-space diagrams for distributed systems [33, 1], and scientific publication analysis such as for research fronts [26] and historiographs [16]. Our approach provides an interactive and linked view that integrates well with the other visualizations in the CiteWiz system.

CITATION NETWORKS

Citation networks consist of bibliographical entries representing scientific publications, each being a tuple of attributes such as title, authors, source, date, abstract, keywords, etc. In addition, each entry has a number of references to other entries representing the citations found in the article. Thus, citation networks can be seen as directed graphs where each node represents an article, out edges represent cited papers (i.e. the dependencies of the current paper), and in edges represent citing papers. A citation graph is generally not acyclic since articles may mutually cite each other; this is sometimes the case when an author (or a team of authors) publishes two or more related articles to the same conference or journal issue.

From the publications data, it is possible to derive a host of related concepts, including authors, conferences, and the citations themselves. These may all be of varying interest depending on the current analysis task the user is trying to perform. In this paper, we present a specialized task taxonomy for citation database; see Plaisant et al. [27] for a general taxonomy of graph visualization tasks.

Traditional bibliographical databases generally provide means for searching, sorting, and filtering the citation data in various ways (examples include IEEE Xplore[1], the ACM Digital Library [10], and CiteSeer [18]). These database interfaces serve as suitable reference implementations when assessing new visualizations for citation networks.

Formative Evaluation

In order to identify the best way to analyze and visualize bibliographic data, we organized a formative user evaluation using a focus group of six active researchers from our department prior to starting the design and implementation of our platform. Our intention with this session was to derive the high-level concepts and tasks involved with the use of bibliographical data, including various situations when researchers make use of such databases.

The authors acted as moderators during the focus group session, first giving a five-minute introduction to citation networks and then posing a number of open-ended questions for discussion. The participants were then encouraged to discuss the questions in the group while the moderators took notes and drew a conceptual map of the problem domain (under the directions of the participants) on a whiteboard. The entire session lasted approximately one hour.

Discussions and results from the focus group session served as the foundation for our taxonomy of citation database interaction. The concept of roles—and the tasks and subtasks associated with each role—turned out to be a key component in most of the participants' thinking on the subject. This taxonomy, presented in the following section, has proven useful when discussing bibliographic visualization and the analysis tasks involved in this activity, but may have a bias towards a researcher's point of view; we plan to involve other users of citation databases (e.g. librarians) in future updates of the taxonomy.

Taxonomy of Citation Database Interaction

A researcher may assume any of a number of different roles when interacting with a citation database, and we have thus chosen to base our taxonomy on the concept of *user roles* and the goals and tasks associated with these. Clearly, a user has different *goals* to achieve depending his or her current role, and these govern which *tasks* need to be done. Using this taxonomy, we can make decisions about which user roles and goals we want a tool to support, and accordingly which tasks we must implement.

In the taxonomy below, the terms *group* and *subgroup* refer to any (potentially hierarchical) clustering of articles (and subgroups) according to some criteria, such as shared keywords, author, source, etc. An *event* is defined as any scientific community activity, such as a journal issue, a conference, a workshop, etc. Furthermore, we have categorized the user tasks depending on the focus of the task; making a distinction between (i) article- (P), (ii) event- (E), (iii) author- (A), and (iv) group-focused (G) user tasks is useful when discussing the nature of a visualization tool.

Table 1 presents the roles we have identified, including a short description of each role. Table 2 gives a listing of the individual goals of each role, as well as the tasks involved with completing that particular goal. Finally, Table 3 shows the different tasks, including their focus category. Note that these tasks operate on the current working group and not necessarily the entire database; for instance, task T3 should be interpreted as "find the most influential paper in the current group of papers".

Citations as Causal Relations

A causal ordering is a general relation that relates two *events* where one is the cause of the other. We can interpret citations in scientific articles as causal orderings in at least two different ways: either with authors as the active entities (processes) and their papers as events, or with papers as the active entities and a single event marking the paper's publication for each entity. For both cases, we represent citations by causal relations between the events. In this work, we have chosen the latter approach for the simple reason that the former causes problem with the visualization when authors

[1] http://ieeexplore.ieee.org/

Role	Description
Novice	A researcher that is new to a specific field; can either be a new student or an experienced researcher moving to a new area.
Expert	An experienced researcher with intimate knowledge of a field.
Reviewer	A researcher tasked with peer-reviewing a new paper, potentially from a field he or she has only passing knowledge of.
Organizer	A researcher responsible for organizing, editing, and/or steering an event (such as a conference or journal).
Evaluator	A person, such as a recruiter, tasked with evaluating the work of a specific researcher.

Table 1. User roles in citation database usage.

Role	Goal	Tasks
Novice	Orientation in a new area	T2-3, 5-6
	Find open problems	T4
Expert	Verify intuition	T1
	Stay updated	T1
	Find papers quickly	T1
Reviewer	Check originality	T2-3, 5
	Check correctness	T2-3
	Check references	T2, 5
Organizer	Identify hot topics	T4-6
	View event chronology	T7
	View event collaborations	T8
Evaluator	View author career	T7
	Assess author work	T2-3, 5

Table 2. Goals for each user role.

Task	Description	Focus
T1	Find a particular paper or author	A/P
T2	Find related papers	P
T3	Find the most influential paper or author	A/P
T4	Find hot topics (at a specific time)	G
T5	Partition an area into subareas	G
T6	Study overall citation network	A/P
T7	Study chronology	A/E/G/P
T8	Study collaboration	A/E/G/P

Table 3. Tasks for citation database interaction (A/E/G/P = author/event/group/paper).

combine to work together on a paper; thus, our visualization is fundamentally article-focused instead of author-focused.

Seeing that a citation in a scientific article can be modeled by a causal relation is straightforward: a citation implies that (a) the authors have read and somehow been influenced by the cited paper (and thus, indirectly, that the cited paper existed before the citing paper), and that (b) the citing paper has a dependency to the cited paper. Admittedly, mutual citations cannot be represented and must be either removed entirely or broken arbitrarily. However, in practice these occur seldomly, so this is a minor point. In this paper, we will use the term *influence*, which is a relaxed interpretation of causality in this context: if a paper A cites a paper B, the authors of A have been influenced (in some undefined way) by paper B, and this is reflected in the paper (put shortly, A has been influenced by B).

THE CITEWIZ PLATFORM

The CiteWiz system is a modularized bibliographic visualization platform based on a central citation dataset and a number of subsets that can be used as input for the available visualization techniques. The three primary visualizations in CiteWiz include a timeline visualization and an interactive concept for overview, and an influence visualization for detail views; the implementation of the former is called a Newton's Shoulders diagram, and the latter is an adaptation of the Growing Polygons causality visualization method. The interactive concept map visualization allows for exploring keywords and co-authorship in the citation database. Based on the taxonomy described in earlier, we developed the tool to be primarily article-focused, meaning that we emphasize the visualization of articles and their interdependencies, but sufficient provisions exist for author-, group-, and event-focused user tasks as well.

The system is a complete software platform for citation network visualization providing a suite of different visualization techniques for studying this kind of data. Figure 1 shows a picture of the prototype implementation running under Windows with all of the visualizations active. The system provides a basic dataflow model built around the central database and hierarchical subsets users define from the database. This allows for each view to be linked, causing selection and brushing in one view to immediately be updated in other views. The database viewer (shown as a table in the center of the image) ties together the visualizations into a coherent whole. The dataflow model makes the framework flexible enough to easily support adding new visualization techniques without code changes to the system core.

Users typically employ the timeline visualization and the concept map to orient themselves in a dataset and then study subsets of the dataset using the Growing Polygons method. Section presents a case study describing this work process.

Datasets and Subsets

CiteWiz has a central citation dataset that is used for all queries and visualizations. The dataset is represented by a data table containing publications, represented by rows in the table. Each row has fields for attributes such as title, authors, source, keywords, abstract, etc. Entries also have a list

of references to other rows in the dataset, such as other papers cited in a paper. The dataset is loaded from disk using a simple XML-based file format for citation meta-data that was designed for the InfoVis 2004 contest [13]. Any dataset using the same format will be compatible with the CiteWiz system.

Users can browse, filter, sort, and search the dataset in the CiteWiz application. In addition, users can also build *nested subsets* of the dataset for visualization; these are hierarchical subsets of the central dataset. This makes it possible to build complex structures of nested groups according to some criteria relevant to the user; for instance, when studying a dataset containing citation data for a specific conference over a period of time, one might create groups for each conference year, and the papers could then be arranged in subgroups representing the different sessions for each conference. Other groupings are possible and depend on the user's goals. For instance, when performing author-focused tasks, it might be useful to create groups for each author in the dataset and add their papers, allowing for easy study of author chronology and collaboration.

Subsets can be saved and loaded to disk using another straightforward XML format; each view file is associated with a specific dataset file, and uses the internal identifiers to refer to bibliographical entries in the dataset.

Timeline Visualization

The overview visualization of the CiteWiz tool is informally referred to as a Newton's Shoulders diagram[2]. This visualization creates a timeline of either articles or authors in the central CiteWiz citation database, displaying each entity as an icon on the timeline according to their publication date (or the date of their first publication, in the case of authors). The user task we want to support is overview, specifically the tasks T3, T7 and T8 in our taxonomy.

The surface area of each icon is scaled proportionally to the amount of citations the article or author has received (rounded up so that the icon conforms to a uniform grid). The timeline is split up into suitable time units (years or months), and each time segment gets assigned space on the timeline equal to the size of the largest entity in the segment. The icons representing the entities for each time segment are then laid out using a greedy algorithm that places the entities in descending size within the allocated space on the timeline, always trying to minimize the distance to the centerpoint of the diagram. An example of such a Newton's Shoulders diagram can be seen in Figure 2 depicting a modest-sized citation database of some 1,000 authors. Our implementation allows the user to zoom and pan continuously in the visualization.

Furthermore, we can orient the timeline vertically and use human figures for the entity icons, giving the impression of people standing on the shoulders of others. This is exactly the metaphor we had in mind when designing the visualization, and matches the intuition of the work of a researcher resting on the work of those who came before her. The diagram now tells us the relative chronology of researchers in a specific field, and instantly shows the most influential authors and their relationships.

These diagrams can be modified to show additional dimensions by applying color to the entity icons. The choice of metric to display this way can be chosen arbitrarily; one ad-hoc metric for authors could be *citation density*, which we define as the total number of citations for an author divided by the total number of publications written by the author (i.e. a kind of "average paper quality" metric). Another, slightly more complex, metric would involve weighing citations for an author or article by their age so that recently cited articles or authors get a stronger and more visible color than older ones, signifying that this article or author is involved in a "hot topic".

In the example in Figure 2, the large and black figures at the bottom stand out—these correspond to authors of the paper "Visualizing the Non-Visual: Spatial Analysis and Interaction with Information from Text Documents" [34] that was published at the first IEEE Information Visualization conference in 1995, and has the highest citation count within the conference. Other prolific authors include Christopher Ahlberg (appears in 1995), John Stasko (1995, duplicated in 2000), Jock Mackinlay (1997), Stuart Card (1997), and George Robertson (1997). Note that the latter three would be dominant if the whole InfoVis 2004 contest database had been used in the figure (as opposed to just the InfoVis conferences from 1995 to 2002).

Interactive Concept Map

To further support orientation in a citation database, the CiteWiz platform also includes an interactive concept map visualization based on a basic force-directed graph layout scheme [11]. The purpose of this visualization is to give users an overview of the keywords and co-authorship in the database, but additional concept graphs can also be visualized, including author influences and article authorship (i.e. connect articles with at least one shared co-author). The user task we are addressing here is again overview, specifically the tasks T3, T4, and T8 in our taxonomy.

The concept map visualization is implemented as a simple spring-based system where each concept is mapped to a node and associations between concepts are undirected edges. Unlike full-fledged force-directed layout schemes such as Kamada-Kawai [20] and Fruchterman-Rheingold [14], where a lot of emphasis lies in reaching equilibrium, we use a very simple physically-based spring simulation to achieve an animated view of the concepts moving around. This allows the user to directly manipulate and rearrange the concept map as it is stabilizing.

For purposes of physical simulation, nodes have an associated weight (which is used to derive the node size), and edges are modelled as linear springs with a specific spring constant and a normal length. We employ Hooke's Law to derive the force exerted on two nodes connected by a spring, and sum these up to form the resulting force. Furthermore, we also connect each node to the center of the canvas using

[2]So named after Sir Isaac Newton's famous quote in a letter to Robert Hooke in 1676, "If I have seen further, it is by standing on the shoulders of giants."

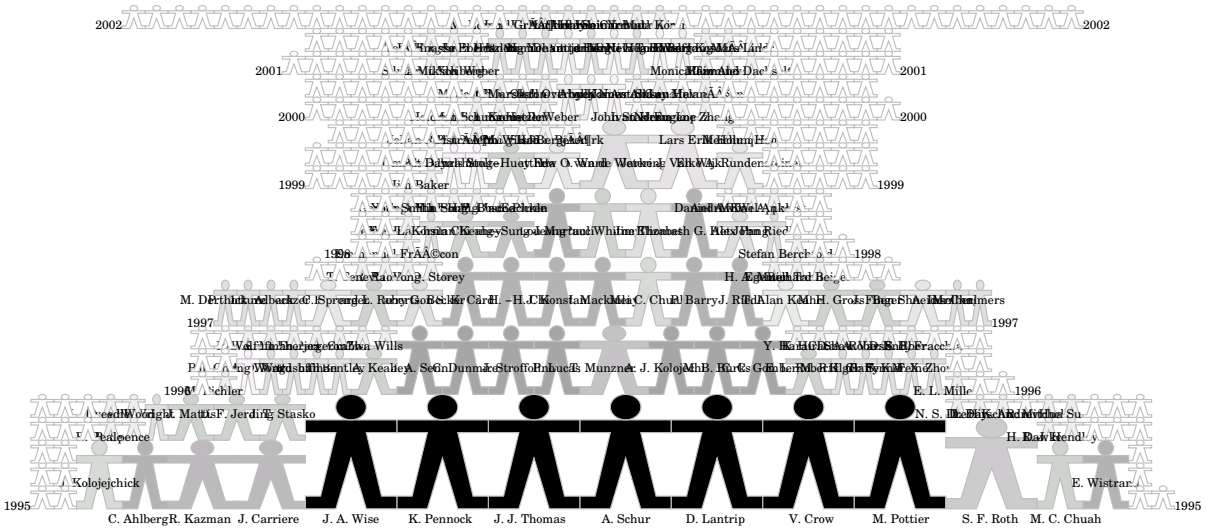

Figure 2. Timeline visualization of authors of the IEEE InfoVis conference. Time runs vertically from 1995 (bottom) to 2002 (top), and the size of each glyph represents the number of citations for each author.

an invisible "leash" to force nodes to gravitate towards the center of the visualization. Both of these are modulated by the node mass. Dampers are associated with each node to help the system reach a stable state. Finally, we use a gravity law to generate repelling forces between nodes to avoid them overlapping each other.

In our implementation, we perform real-time numerical simulation of the physical system to show the result. Users can interact with individual nodes by clicking and dragging on them, causing a spring to be temporarily connected between the mouse pointer and the node. The user can also control the dampening, repelling, spring and leash constants for the visualization.

We construct various kinds of concept maps depending on the user task. For keyword maps, we traverse the database and add all keywords as nodes, using their individual frequency as mass. Links are created between any two keywords that appear in the same article. See Figure 4 for a keyword concept map of the IV04 contest dataset. We can plainly see the main themes of the citation database, i.e. information visualization, data visualization, visualization, and so on. A slider allows the user to put a threshold on displaying labels, for instance only for keywords which occur 10 times or more in the dataset.

For co-authorship maps, on the other hand, we add authors as nodes, using their number of citations as mass, and connect each pair of authors with a link for every article they have co-authored. See Figure 5 for a co-authorship map for the same dataset. Again, this visualization allows us to quickly pinpoint the main contributors to the citation database and the clusters they form.

Influence Visualization

After having studied the general shape of a citation database using the Newton's Shoulder visualization, users are able to build subsets tailored to answer their specific queries about

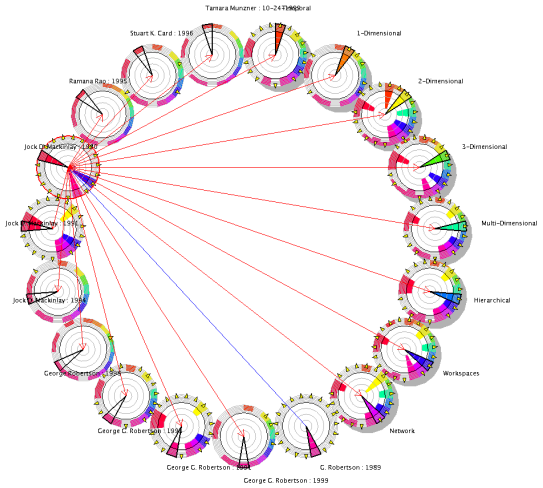

Figure 3. George Robertson's impact on the previously-defined research areas in information visualization. The currently selected paper ("Rapid controlled movement through a virtual 3D workspace") has citations (red arrows) in all of the research areas.

the dataset. These subsets form the input for the Growing Polygons [12] method for visualization of general causal relations, suitably modified to be able to handle citation networks and the scalability issues associated with these. We believe the focus on influence and causality visualization in the Growing Polygons technique makes it very well suited to visual exploration of citation networks. The technique uses a combination of 2D shapes, color, and animation to graphically represent a system of n active processes as n-sided so-called *process polygons* showing the influences affecting each process. As time progresses, the process polygons grow from zero to full size (i.e. from the center and out), and the sectors of each polygon fill in as messages are received by other processes, signifying the information transfer.

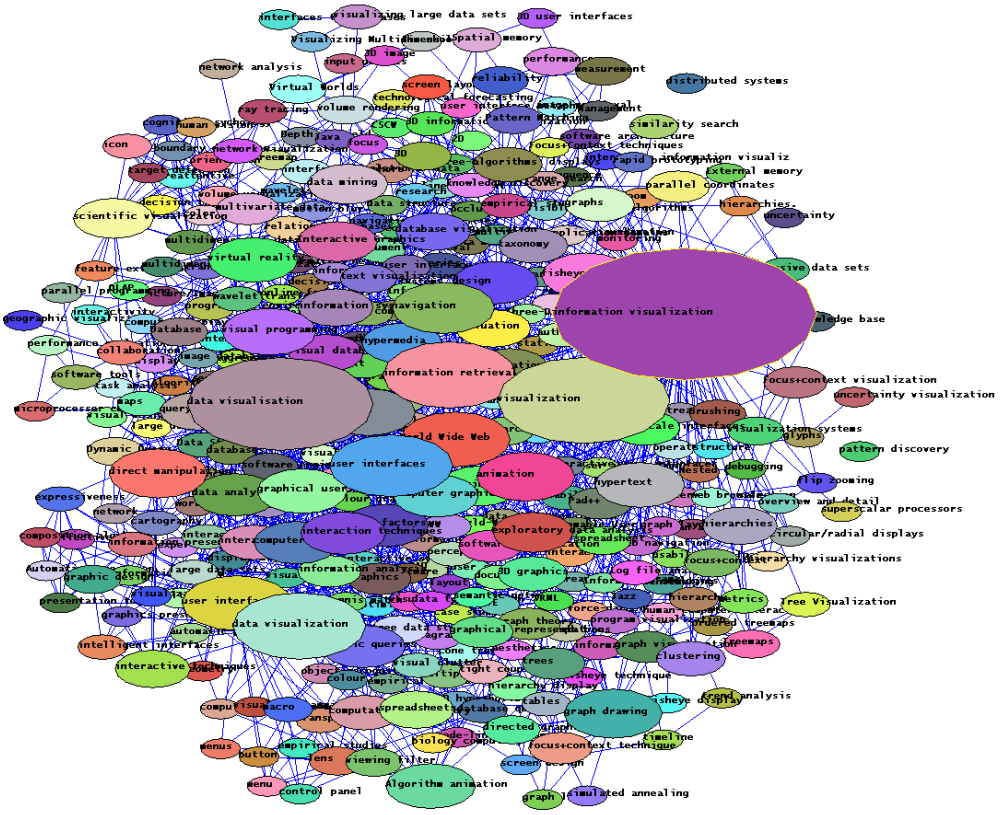

Figure 4. Concept map for the keywords in the InfoVis dataset. Node size represents the number of occurrences of a specific term, and links between nodes represent keywords that occur together.

Figure 6 shows an example system of five processes labelled $p0$ to $p4$ and colored black, green, blue, cyan, and red, respectively. Each of the five polygons represent a specific process and have been segmented into sectors showing dependencies to other processes. Time goes from the center and out, the light grey rings indicating discrete time units. Each process has its own sector outlined by a black triangle with its color showing when in time it was active, and filled-in areas of the other sectors marking influences to others. In this example, we can for instance see that $p0$ seems to have a dependency (a causal relation) to all other processes since all of its sectors are filled in, whereas $p2$ only has a reference to one other process, the red $p4$.

In our adaptation of the original technique, articles form the processes in the visualization (thus represented by *article polygons*), and citations are messages from a source (cited) article to a destination (citing) article. Thus, articles are assigned unique colors using a simple color scale and citations are drawn as arrows showing influences from one process (article) to another. This mimics the information transfer implicit when authors reference another paper. Even if articles are more or less static once published, this article-focused approach gives us a way to easily see the influences and chronology of a set of articles, including global transitivity information for each article. The user tasks in the taxonomy we address are primarily T2, T3, T4, and T6.

In order to make effective use of the Growing Polygons method in this context, we addressed two scalability issues in relation to (i) long execution times, and (ii) large quantities of visualized articles. For the former issue concerning time scalability, the problem lies in that visualizing a large citation network may result in very long chains of causality, and the visualization will then run out of space for displaying individual time segments. For the latter case, the quantity scalability issue comes from the fact that visualizing a sufficiently large amount of articles means that each individual article gets assigned a very small polygon sector and it will thus be difficult to distinguish between neighboring sectors. Both of these issues can be partially addressed through zooming mechanisms, but this instead results in loss of overview.

Our solution for these concerns in the modified, more scalable version of the Growing Polygons method is two-fold:

1. We introduce a focus+context [15] technique based on adjustable *linear time windows* that lets the user concentrate on certain areas of the execution while still retaining the context of the surrounding history (i.e. the focus view and the overview are integrated in the same visual space, as opposed to overview+detail techniques where the views are spatially or temporally separate). For exampe, Figures 7 show these time windows where the outer part of the time has been compressed and only the inner three time units are shown at full detail.

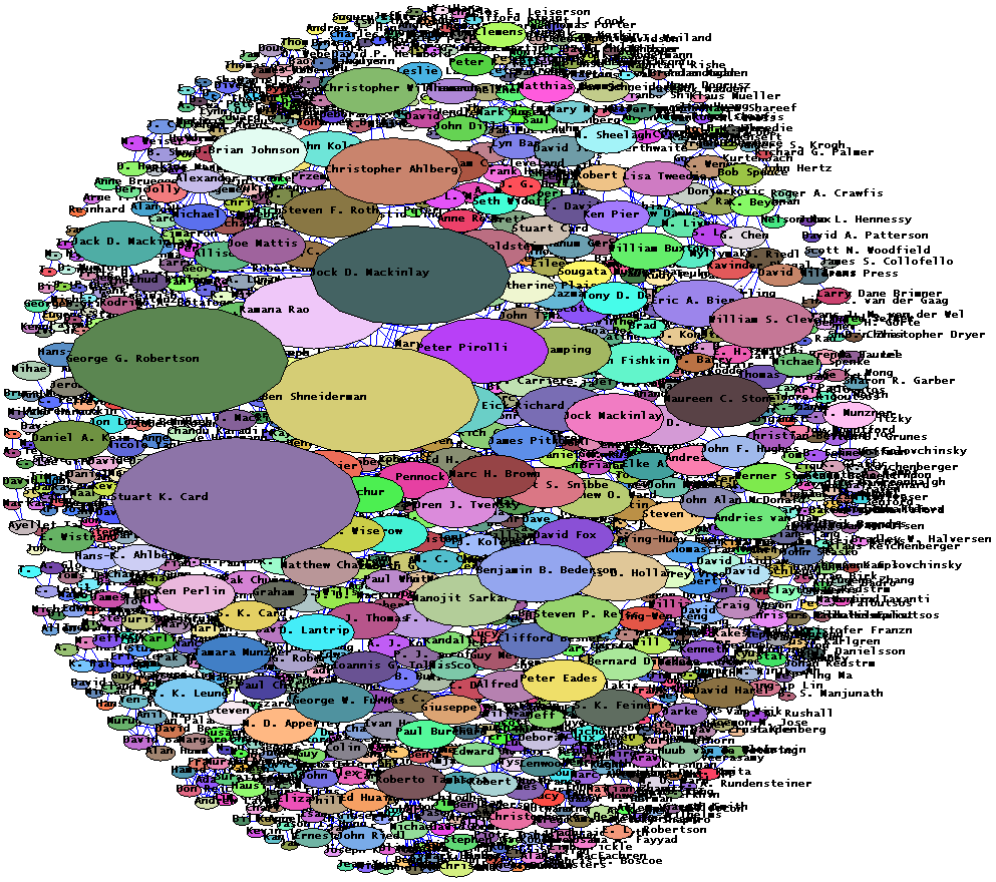

Figure 5. Concept map for co-authorship in the InfoVis dataset. Node size represents the number of citations for each author, and the strength of the springs binding two nodes is proportional to the amount of co-authored papers.

2. Secondly, we address the quantity concern by modifying the Growing Polygon technique to handle *nested subsets* instead of flat article lists (this was our incentive for the distinction between datasets and subsets in the design of CiteWiz). Note that the purpose of the influence visualization is not for overview of huge datasets, but for detailed studies of customized subsets (on the order of a hundred documents).

For example, consider the example in Figure 7. Here, we have grouped together the papers of a number of important authors in information visualization into hierarchical subsets. The colors in each polygon representing a specific author shows which other authors this author has referenced and when. The outlined "pie slice" for each polygon shows the time span during which the author has been active and publishing.

Hierarchical Subsets

In order to allow the Growing Polygons technique to handle a large quantity of articles, we modify the visualization to be able to render hierarchical groups of articles instead of single articles. These correspond directly to the subsets of the central dataset built by the users. The view hierarchy is visualized by treating an article group as a normal article, except that the group will have the cumulative influences of all of its children. We derive these influences by a simple postorder traversal of the hierarchy, building the influence timelines of the internal nodes from the bottom up (i.e. starting with the articles in the leaves of the tree). The currently visible nodes are rendered as normal article polygons, with the single exception that groups (i.e. non-leaves) have a drop shadow to signify that the polygon represents more than one article.

Figure 8 shows an example of this hierarchical grouping of subsets; here, we have created three nodes representing the prolific authors Stuart Card (yellow), Jock Mackinlay (purple), and George Robertson (cyan), and populated each node with their individual papers. The visualization shows that we have expanded each of these group nodes to expose the papers; a circle chord with the appropriate color is drawn on the background to indicate group membership for the nodes. The thick yellow lines in the figure connect instances of the same entity, allowing the user in this case to see all of the articles these three researched co-authored together.

Interaction Techniques

In our modified version of the Growing Polygons technique, we provide two simple interaction techniques for browsing and exploring the article hierarcy: users can either click directly in the visualization to expand and collapse article groups (using the left and right mouse buttons, respectively),

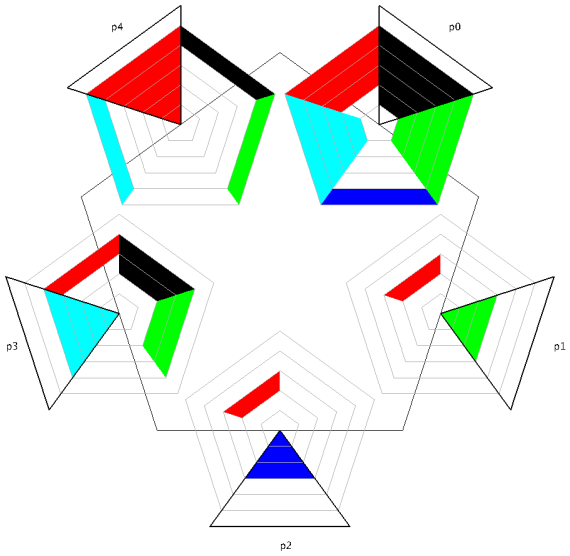

Figure 6. Example Growing Polygons influence visualization with 5 actors/processes.

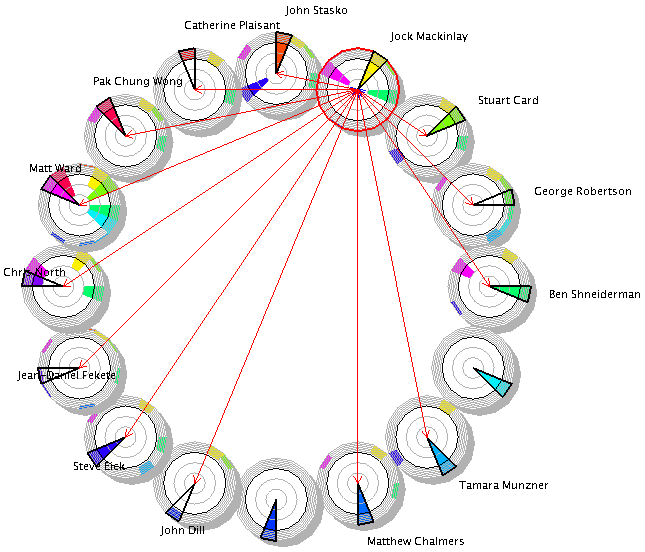

Figure 7. Visualization of prominent information visualization authors using the Growing Polygons method.

or they can use a separate tree navigation window to study the structure of the hierarchy. The same tree window can also be used to search for the full or partial name of a specific article, and the tree will be expanded to the level of the article to show the search result.

In addition to these techniques, we also provide an overview map window with a color legend and clickable fields for quickly jumping to a specific article polygon.

Details-On-Demand

As suggested by both Shneiderman [31] and Modjeska et al. [25], bibliographic visualization tools need to provide a mechanism to show the complete bibliographical data of an article. In CiteWiz, this is handled by a detail window that gives the full meta-data of the currently selected article. Users can easily navigate through the references of articles, as well as moving back and forth in the article history of the window.

To make citations evident at a quick glance, we add unobtrusive yellow arrows to the perimeter of each node to indicate an outward dependency (i.e. a citation) to another node. We also augment the currently selected node in the visualization with blue arrows pointing from the cited nodes, and with red arrows pointing to citing nodes—again for convenience reasons. Figure 9 shows an influence visualization of the publications dataset organized into research areas where both of these features are visible: all areas have citations to all of the other areas (as indicated by the yellow arrows for each sector of all the nodes), and references to and from the currently selected node ("1-Dimensional") are shown as blue and red arrows.

Implementation

The CiteWiz tool is implemented as a C++ application running under the Windows operating system. It uses standard OpenGL for efficient 2D rendering, and the Windows Forms library for the graphical user interface components.

CASE STUDY: IDENTIFYING INFLUENTIAL AUTHORS IN INFOVIS

Consider an analyst interested in identifying influential authors in the IEEE InfoVis 2004 contest dataset using the CiteWiz tool. The contest dataset consists of bibliographical data for all of the papers presented at the InfoVis conferences from 1995 to 2002, as well as any articles cited by these conferences. The analysts launches her analysis by loading the XML file containing this dataset into the CiteWiz tool.

One useful starting point for someone who is new to a research area would be to look at the CiteWiz interactive concept map for the keywords. Figure 4 quickly gives our analyst a sense of the contents of the articles in the data. Not surprisingly, variations of the term "visualization" seems to be the single most commonly occurring concept in the dataset.

Now, to get an understanding of the major players and seminal works in the dataset, the analysis starts looking at co-authorship. Figure 5 shows the CiteWiz co-authorship concept map for this dataset. While there certainly are a lot of authors who have contributed to the research area over the years, a few of them stand out: notably Stuart Card, George Robertson, Jock Mackinlay, and Ben Shneiderman. The analyst can now select these four entities and add them as categories in the blank working subset that CiteWiz maintains for selecting parts of the whole dataset. The subset is simply a tree viewer where the nodes are bibliographic entries (i.e. papers) that can be optionally ordered into categories.

Satisfied with this insight, our analyst moves on to study the chronological aspects of the dataset using the interactive timeline. Having selected the above four actors in the co-authorship visualization, CiteWiz will ensure that these actors are highlighted in the timeline to allow for easy cor-

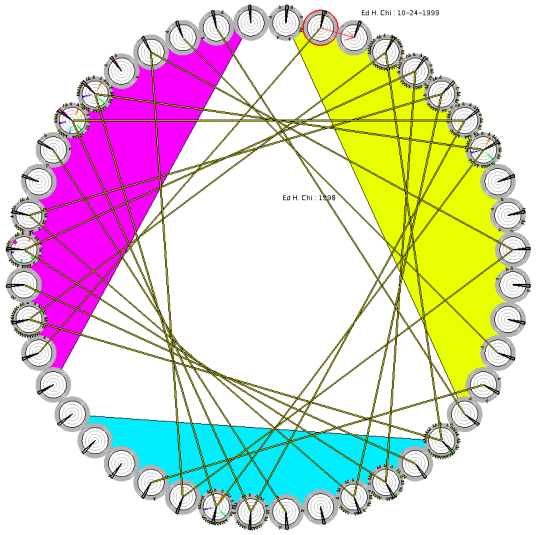

Figure 8. Influence visualization of the publications of Stuart Card (yellow), Jock Mackinlay (purple), and George Robertson (cyan). The thick yellow lines indicate co-authorship, i.e. papers that appear more than once in the visualization, and the small yellow arrows indicate citations from other papers.

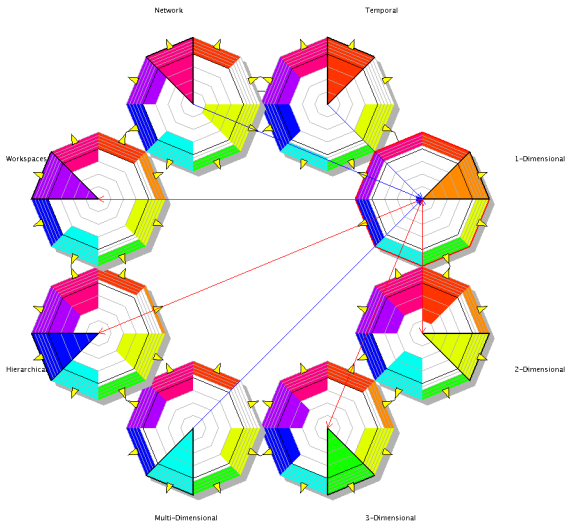

Figure 9. Papers from the InfoVis 2004 contest grouped into user-defined hierarchical subsets representing research areas and visualized using the influence visualization.

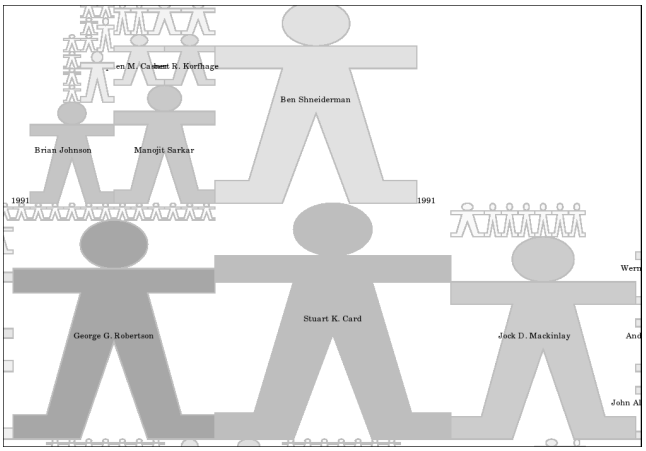

Figure 10. Close-up of a part of the Newton's Shoulders diagram for the InfoVis 2005 contest dataset showing the four most prolific actors.

relation. Figure 10 shows a close-up of the interactive timeline centered on the four actors. This particular visualization shows that Ben Shneiderman's work in the dataset is first published in 1991, whereas Card, Robertson, and Mackinlay's contributions appeared the year before. Given this information, our analyst decides to study the latter three authors in more detail to see how they relate to each other.

As it happens, our analyst quickly discovers that the trio Card, Robertson, and Mackinlay has a very interesting collaboration pattern when she goes back to the main CiteWiz window to manage the working subset. Using filters on the main dataset, the analyst populates each of the three categories corresponding to each author with the papers that author has worked on. Then she creates a Growing Polygons influence visualization for the whole working subset.

Figure 8 is the resulting visualization. Each small circle represents one paper, grouped into one of three sectors representing each author (yellow for Stuart Card, purple for Jock Mackinlay, and cyan for George Robertson). The thick yellow identity lines in the visualization show that many of the papers exist in more than one sector, i.e. they are co-authored by two or all of the three actors under study. Selecting individual articles calls up detailed bibliographic information in a separate window, allowing the analyst to get more details on demand. After some additional work, our analyst notes that most of the joint papers in the working subset came about when all of the three authors worked together at Xerox PARC in the early 1990s (including one of the most-cited HCI articles "Cone Trees: Animated 3D Visualizations of Hierarchical Information" [30]).

USER STUDY

The purpose of the CiteWiz citation visualizer is to provide researchers with additional tools for analyzing citation data beyond the standard low-level features available in most traditional database interfaces. We selected the IEEE Xplore database as a good baseline database web interface for comparison. Our hypotheses were the following:

- CiteWiz will perform as well as IEEE Xplore for finding papers and correlating bibliographical data, and
- CiteWiz will perform significantly better for higher-level tasks.

Subjects

The main target audience for the CiteWiz tool are active researchers, which dramatically limits our pool of available subjects. In total, 10 unpaid test subjects, 9 of which were male, participated in this study. All subjects were researchers at our department, but were carefully screened to have no previous knowledge of the IEEE InfoVis community and the InfoVis 2004 contest citation database. However, all

subjects had considerable previous experience in the use of general citation database interfaces. Ages ranged from 25 to 40. All subjects had normal or corrected-to-normal vision.

Equipment

The study was run on an Intel Pentium III 1 GHz desktop computer with 512 MB of memory and a 19 inch color display. The machine was equipped with a NVidia Geforce 3 graphics card and a 19-inch monitor with the display resolution set to 1280×1024.

Task

We selected three different tasks related to citation database interaction from our taxonomy presented earlier in this paper: T1, T3, and T8 (see Table 3). We designed the scenarios for T1 and T8 to consist of low-level analysis tasks such as searching, filtering, and correlating basic bibliographical data (finding the title of a paper given different search criteria—such as one of the authors, a specific conference, or a term in the abstract—and studying different instances of author collaboration). The scenario for T3, on the other hand, required a higher-level analysis of influence and structure of the citation network to find the most influential paper in the dataset.

Experimental Conditions

We designed the test to be a between-subjects comparative study of a traditional database interface versus our CiteWiz citation visualizer. We selected the IEEE Xplore web-based database interface as a suitable representative of traditional database interfaces. IEEE Xplore is widely used among scientists all over the world to access the bibliographical data and fulltexts of IEEE publications and supports all standard search and filtering features. Thus, the independent variable was INTERFACE, with two levels, "CiteWiz" and "Xplore".

Dataset

Given the use of the IEEE Xplore interface as the baseline condition, we selected all of the papers of the IEEE InfoVis conferences from 1995 to 2002 as our test database (175 articles). Albeit a small dataset, this was a necessary delimitation for us to be able to use the same database for both tools. In order to remove all distractions, we were able to design our own search interface to the IEEE Xplore database (essentially a cleaner version of the standard IEEE Xplore Basic Search), allowing us to constrain searches to the InfoVis conference and provide a browseable list of the InfoVis proceedings sorted by year. The CiteWiz XML-based database, on the other hand, was adapted from a subset of the InfoVis 2004 contest database [13].

Procedure

Each session lasted approximately one hour. Participants were given a scripted introduction to the research problem and the CiteWiz project. This was followed by a short training period for the assigned interface (CiteWiz or IEEE Xplore) lasting between ten and fifteen minutes. The subject indicated when he or she was ready to proceed.

Participants were then given the three tasks (in paper form) and were asked to solve them using the available tool (CiteWiz or IEEE Xplore). For the CiteWiz tool, all visualizations were at the user's disposal. Completion time was capped at 15 minutes to avoid runaway tasks; subjects were given the option to abandon a troublesome task, in which case the completion time was set to the cap. Each participant was asked to fill out a questionnaire after having completed it.

RESULTS

The main findings of the user study confirmed both of our hypotheses: that (i) there is no significant difference in efficiency for CiteWiz and IEEE Xplore for simple tasks involving finding papers and collerating basic citation data, and that (ii) CiteWiz is significantly more efficient for a higher-level task involving the study of dependencies and influences of a set of articles.

Performance

The mean times of solving a full task set (i.e. all three tasks) using IEEE Xplore and CiteWiz were 20 minutes and 2 seconds (s.d. 158 seconds) and 8 minutes 5 seconds (s.d. 72 seconds), respectively. This was a statistically significant difference ($t(8) = 9.20, p < .001$).

For task T1, the mean completion times were 3 minutes 15 seconds (s.d. 61 seconds) for IEEE Xplore versus 3 minutes 20 seconds (s.d. 24 seconds) for CiteWiz, a nonsignificant difference ($t(8) = -.163, p < .875$). No user managed to solve task T3 within the 900 second time cap using IEEE Xplore (two subjects completed the task, three abandoned the task); the CiteWiz completion time was 2 minutes 34 seconds (s.d. 38 seconds). This was clearly a significant difference ($t(8) = 43.7, p < .001$). Finally, for task T8, the mean completion times were 1 minute 46 seconds (s.d. 100 seconds) versus 2 minutes 20 seconds (s.d. 25 seconds) for IEEE Xplore and CiteWiz, respectively. Again, like for T1, this was not a significant difference ($t(8) = -.511, p < .623$).

Our informal observations indicate that the participants used an approach that is very similar to the "information-seeking mantra" [31] of overview first, zoom and filter, and then details, and that the CiteWiz tool supported them in this process. The timeline and concept visualizations were used by the subjects to orient themselves in the dataset before constructing specialized database subsets and studying them in the influence visualization.

Correctness

No subjects using the IEEE Xplore managed to correctly solve task T3 (even when exceeding the time cap), while all subjects using CiteWiz correctly solved T3. All subjects registered correct answers on all other tasks.

Subjective Ratings

The ratings from the post-test questionnaire overall show encouraging results; see Table 4 for an overview. Note especially the responses to question Q2b, which show subjective ratings strongly in favor of CiteWiz over IEEE Xplore

	Question	Xplore	CiteWiz
Q1.	Ease-of-use	3.40 (1.67)	4.20 (1.10)
Q2.	Efficiency		
	(a) Find paper	4.40 (.55)	4.20 (.84)
	(b) Most influential	1.00 (.00)	4.60 (.55)
	(c) Collaboration	2.40 (.89)	3.20 (1.10)
Q3.	Enjoyability	2.60 (.89)	4.20 (.84)

Table 4. Mean (standard deviation) responses to 5-point Likert-scale questions.

(significant down to $p < .05$ using a Kruskal-Willis test). The difference for question Q3 on enjoyability was also significant ($p < .05$). Thus, users consistently perceived the CiteWiz tool as more enjoyable to use than IEEE Xplore.

For questions Q2a and Q2c, which constitute the low-level tasks of the study, there were no significant difference between the two interfaces (Kruskal-Willis test, $p = .73$ and $p = .22$, respectively). In other words, for these questions, users were more or less equally satisfied with IEEE Xplore and CiteWiz. The difference in perceived ease-of-use (Q1) was also not statistically significant ($p = .42$).

DISCUSSION

The results from our study shows initial evidence that a citation visualization tool can be beneficial for a sampling of the tasks discussed in the taxonomy presented in the beginning of this paper. The main findings were the following:

- Completion times for CiteWiz and IEEE Xplore were similar for low-level analytical tasks, but significantly faster for a more complex task;
- Correctness were again similar for low-level analytical tasks, while only participants using CiteWiz were able to solve the high-level task; and
- Subjective ratings show a clear preference for CiteWiz for high-level analytical tasks, but no particular preference of low-level ones.

In the following sections we will try to explain these findings and see how they generalize. We also discuss some limitations to CiteWiz and describe how we will improve the platform in the future.

Explaining the Results

Our expectations of the results of the user study was that CiteWiz and the IEEE Xplore tool would perform equally well at low-level tasks related to basic searching, sorting and correlation of bibliographical data, and that higher-level tasks involving assessing influences and structure of the citation network would yield a significantly higher efficiency for CiteWiz.

Overall, these expectations were fulfilled, but for the low-level author collaboration task, IEEE Xplore subjects were on average 34 seconds faster. We believe this is due to the fact that CiteWiz is essentially an article-centered tool, so finding author information requires an extra step of filtering, whereas the same task is performed in IEEE Xplore using a simple query. Future versions of the CiteWiz tool should probably maintain an author (and maybe even a conference) graph superimposed on top of the article graph.

The reason for these results has a simple explanation—CiteWiz was built to support a wider array of decision-making tasks than a standard database interface like IEEE Xplore. However, the sheer size of current citation databases and the high density of edges connecting bibliographic entities—such as papers, authors, and conferences—suggests that visualization in itself may be the best way to approach the problem.

Of course, it is certainly possible to improve standard database interfaces with better support for these higher-level analysis tasks, but this is not always practical; for instance, the IEEE publications database does not contain reference information (the ACM Digital Library does, however, but this database lacks the IEEE publications we used for the CiteWiz tool).

Generalizing the Results

In general, attaining full ecological validity is difficult when studying a new visualization tool such as CiteWiz for which there exists no real baseline comparison. Nevertheless, the purpose of this work was mainly to target the deficiencies of existing standard tools, and in this regard we succeeded.

The visual browsing and exploration features that CiteWiz provide are very hard to measure qualitatively in comparison to standard databases, but the test subjects expressed enthusiasm when exposed to this visualization and some were very eager to use the tool in their own area of research. From the formative evaluation as well as these comments from the test subjects, it seems clear that the general motivation for this paper is valid: managing and staying abreast with publications in a research area is time-consuming.

Since all subjects really should be active researchers for the experimental results to be valid, subject recruitment was difficult and we were only able to enlist 10 participants in the user study. Furthermore, the fact that we only had access to one dataset (the InfoVis 2004 one) forced us to perform a between-subjects study where half of the subjects used the CiteWiz tool and the other half IEEE Xplore. This means that the subject groups are rather small. However, the results do indicate significant improvements, and informal communication with the subjects reinforce these findings.

As a final note, the task set in the user study is limited, but this was a deliberate design decision due to the small feature set that the IEEE Xplore database provides. Choosing a more complex task set would give an unfair advantage to our tool, and would also punish the test subjects who used IEEE Xplore. We believe that the user study shows that there is room for improvement, and that the techniques presented in this paper are viable alternatives to traditional database interfaces.

Limitations to the CiteWiz Platform

The strength of the CiteWiz platform lies in showing influences and chronology in a citation database. All three of the visualizations that make up the framework emphasize citation structure and information flow. However, the tool lacks some of the more standard visualizations, such as co-authorship node-link diagrams, keyword burst analysis, and historiographs [16], that other tools like CiteSpace [6] provide.

Observations and informal interviews with study participants indicated that they thought the CiteWiz user interface was difficult to use, in particular the zoomable navigation controls for the visualizations. Many participants thought that the search interface for the tool was too simplistic and requested the ability to build more complex queries. Also, we noted that few participants took full advantage of the power of freely building working subsets of the citation database during the test sessions. We have been discussing the use of clustering for helping users build effective working subsets in the future.

Nevertheless, CiteWiz builds on the basic concept of what we believe a full-fledged citation visualization tool should be: an extensible dataflow framework with a citation database at its core to which any number of different visualizations can be plugged in.

Future Work

Results and observations from our user study brought to light a number of interesting avenues for future work. Possible extensions to CiteWiz includes the design of new bibliographic visualization techniques to provide alternate views of the dataset as well as the afore-mentioned clustering algorithms for automatic construction of nested subsets.

Future design iterations of CiteWiz should be improved with better search and query functionality, include additional visualizations from the literature, and also support a wider range of non-visual analysis methods (such as centrality, co-authorship, and simple descriptive statistics). It would also be interesting to perform a longitudinal field study of the tool being used by a small number of researchers over a longer period of time.

CONCLUSIONS

We have described CiteWiz, a platform for bibliographic visualization. The platform includes a timeline visualization and an interactive concept map for overview, and a modified version of the Growing Polygons method for detailed studies. The tool is based on a taxonomy of the usage of citation databases. The timeline visualization, informally called a Newton's Shoulders diagram, constructs timelines of articles or authors showing the causality and citations in a citation database. The modifications to the Growing Polygons technique, on the other hand, were aimed primarily at adapting the method to citation networks, and included provisions for rendering hierarchies of articles rather than flat lists, and a focus+context technique with user-controlled time windows to more easily support long citation chains. Finally, we presented the formal user study we have conducted, a between-subjects comparative analysis of CiteWiz in relation to the standard IEEE Xplore web-based database interface. Our results confirm our intuition, that CiteWiz and IEEE Xplore perform equally well for low-level citation interaction tasks such as correlating bibliographical data, and that CiteWiz is significantly more efficient to use for higher-level tasks such as influence and citation structure assessment.

Acknowledgments

The authors would like to thank the colleagues at our department for their thoughts and feedback during the focus group session. We also thank Jean-Daniel Fekete, Georges Grinstein, and Catherine Plaisant for providing the InfoVis 2004 contest dataset that we used for testing.

Thank you to the anonymous reviewers for their helpful comments.

REFERENCES

1. Thomas Bemmerl and Peter Braum. Visualization of message passing parallel programs with the TOPSYS parallel programming environment. *Journal of Parallel and Distributed Computing*, 18(2):118–128, June 1993.
2. Kevin W. Boyack, Brian N. Wylie, and George S. Davidson. Domain visualization using VxInsight for science and technology management. *Journal of the American Society for Information Science and Technology*, 53(9):764–774, 2002.
3. Ulrik Brandes and Thomas Willhal. Visualization of bibliographic networks with a reshaped landscape metaphor. In *Proceedings of the Symposium on Data Visualisation 2002*, pages 159–164, 2002.
4. Matthew Chalmers and Paul Chitson. Bead: Explorations in information visualization. In *Proceedings of the ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 330–337, 1992.
5. Chaomei Chen. Visualising semantic spaces and author co-citation networks in digital libraries. *Information Processing and Management*, 35(3):401–420, 1999.
6. Chaomei 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. Chaomei Chen and Steven Morris. Visualizing evolving networks: Minimum spanning trees versus pathfinder networks. In *Proceedings of the IEEE Symposium on Information Visualization 2003*, pages 67–74, 2003.
8. George S. Davidson, Bruce Hendrickson, David K. Johnson, Charles E. Meyers, and Brian N. Wylie. Knowledge mining with VxInsight: Discovery through interaction. *Journal of Intelligent Information Systems*, 11(3):259–285, 1998.
9. Maylis Delest, Tamara Munzner, David Auber, and Jean-Philippe Domenger. Exploring InfoVis publication history with Tulip. InfoVis 2004 Contest.

10. Peter J. Denning. The ACM Digital Library goes live. *Communications of the ACM*, 40(7):28–29, July 1997.

11. Giuseppe DiBattista, Peter Eades, Roberto Tamassia, and Ioannis G. Tollis. *Graph Drawing: Algorithms for the Visualization of Graphs*. Prentice Hall, 1999.

12. Niklas Elmqvist and Philippas Tsigas. Causality visualization using animated growing polygons. In *Proceedings of the IEEE Symposium on Information Visualization 2003*, pages 189–196, 2003.

13. Jean-Daniel Fekete, Georges Grinstein, and Catherine Plaisant. IEEE InfoVis 2004 Contest. available at `www.cs.umd.edu/hcil/iv04contest`, 2004.

14. Thomas M. J. Fruchterman and Edward M. Reingold. Graph drawing by force-directed placement. *Software—Practice and Experience*, 21(11):1129–1164, November 1991.

15. George W. Furnas. Generalized fisheye views. In *Proceedings of the ACM CHI'86 Conference on Human Factors in Computer Systems*, pages 16–23, 1986.

16. Eugene Garfield. Historiographs, librarianship, and the history of science. *Toward a theory of librarianship*, pages 380–402, 1973.

17. Mohammad Ghoniem, Jean-Daniel Fekete, and Philippe Castagliola. On the readability of graphs using node-link and matrix-based representations: a controlled experiment and statistical analysis. *Information Visualization*, 4(2):114–135, 2005.

18. C. Lee Giles, Kurt Bollacker, and Steve Lawrence. CiteSeer: An automatic citation indexing system. In *Digital Libraries 98 - The Third ACM Conference on Digital Libraries*, pages 89–98, 1998.

19. Matthias Hemmje, Clemens Kunkel, and Alexander Willett. Lyberworld – A visualization user interface supporting fulltext retrieval. In *Proceedings of the ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 249–259, 1994.

20. Tomihisa Kamada and Satoru Kawai. An algorithm for drawing general undirected graphs. *Information Processing Letters*, 31(1):7–15, 12 April 1989.

21. Weimao Ke, Katy Borner, and Lalitha Viswanath. Major information visualization authors, papers and topics in the ACM library. InfoVis 2004 Contest.

22. Daniel A. Keim, Helmut Barro, Christian Panse, Jorn Scheidewind, and Mike Sips. Exploring and visualizing the history of InfoVis. InfoVis 2004 Contest.

23. Michael M. Kessler. Bibliographic coupling between scientific papers. *American Documentation*, 14(1):10–25, 1963.

24. Jock D. Mackinlay, Ramana Rao, and Stuart K. Card. An organic user interface for searching citation links. In *Proceedings of ACM CHI'95 Conference on Human Factors in Computing Systems*, pages 67–73, 1995.

25. David Modjeska, Vassilios Tzerpos, Petros Faloutsos, and Michalis Faloutsos. BIVTECI: A bibliographic visualization tool. In *Proceedings of the 1996 Conference of the Centre of Advanced Studies on Collaborative Research*, page 28, 1996.

26. Steven A. Morris, Gary G. Yen, Zheng Wu, and Benyam Asnake. Time line visualization of research fronts. *Journal of the American Society for Information Science and Technology*, 54(5):413–422, 2003.

27. Catherine Plaisant, Bongshin Lee, Cynthia Sims Parr, Jean-Daniel Fekete, and Nathalie Henry. Task taxonomy for graph visualization. In *Proceedings of BEyond time and errors: novel evaLuation methods for Information Visualization (BELIV'06)*, pages 82–86, 2006.

28. Catherine Plaisant, Brett Milash, Anne Rose, Seth Widoff, and Ben Shneiderman. Lifelines: Visualizing personal histories. In *Proceedings of ACM CHI'96 Conference on Human Factors in Computing Systems*, pages 221–227, 1996.

29. Thomson ResearchSoft. RefViz, 2006. *http://www.refviz.com*.

30. George G. Robertson, Jock D. Mackinlay, and Stuart K. Card. Cone trees: Animated 3D visualizations of hierarchical information. In *Proceedings of the ACM CHI'91 Conference on Human Factors in Computing Systems*, pages 189–194, 1991.

31. Ben Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings of the IEEE Symposium on Visual Languages*, pages 336–343, 1996.

32. Henry G. Small. Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science*, 24(4):265–269, 1973.

33. David Socha, Mary L. Bailey, and David Notkin. Voyeur: Graphical views of parallel programs. In *Proceedings of the ACM SIGPLAN/SIGOPS Workshop on Parallel and Distributed Debugging*, pages 206–215, 1989.

34. James A. Wise, James J. Thomas, Kelly Pennock, David Lantrip, Marc Pottier, Anne Schur, and Vern Crow. Visualizing the non-visual: Spatial analysis and interaction with information from text documents. In *Proceedings of the IEEE Symposium on Information Visualization*, pages 51–58, 1995.

35. Jing Yang, Matthew O. Ward, and Elke A. Rundensteiner. InterRing: An interactive tool for visually navigating and manipulating hierarchical structures. In *Proceedings of the IEEE Symposium on Information Visualization 2002*, pages 77–84, 2002.