TimeTree: Exploring Time Changing Hierarchies

Stuart K. Card*, Bongwon Suh*, Bryan A. Pendleton*, Jeffrey Heer**, John W. Bodnar***

*Palo Alto Research Center   **University of California at Berkeley   ***SAIC

ABSTRACT

Intelligence analysis often involves the task of gathering information about an organization. Knowledge about individuals in an organization and their relationships, often represented as a hierarchical organization chart, is crucial for understanding the organization. However, it is difficult for intelligence analysts to follow all individuals in an organization. Existing hierarchy visualizations have largely focused on the visualization of fixed structures and can not effectively depict the evolution of a hierarchy over time. We introduce TimeTree, a novel visualization tool designed to enable exploration of a changing hierarchy. TimeTree enables analysts to navigate the history of an organization, identify events associated with a specific entity (visualized on a TimeSlider), and explore an aggregate view of an individual's career path (a CareerTree). We demonstrate the utility of TimeTree by investigating a set of scenarios developed by an expert intelligence analyst. The scenarios are evaluated using a real dataset composed of eighteen thousand career events from more than eight thousand individuals. Insights gained from this analysis are presented.

CR Categories and Subject Descriptors: H.5.2 [Information Interfaces]: Graphical User Interface. I.3.6 [Methodology and Techniques]: Interaction Techniques.

Additional Keywords: TimeTree, DOI Tree, tree visualization, organizational chart, timeseries data, visual analytics

1   INTRODUCTION

Intelligence analysis often involves the task of gathering information about an organization. Knowledge about individuals in an organization and their relationships, often represented as a hierarchical organization chart, is crucial for understanding the organization. However, it is difficult for intelligence analysts to follow all individuals in an organization. Due to the difficulty of managing large time-varying hierarchies, only a few very experienced political analysts keep track of and identify the dynamics in very large organizations other than those of key individuals.

This problem is an instance of a more general problem in visual analytics. It has long been recognized that a hierarchical organization improves efficiency in various systems, as described in [26] and [28]. It is not surprising, therefore, that hierarchies occur often in natural and man-made systems and are important subjects for analysis. In addition, trees are relatively easy to

*E-mail: {card, suh, bpendlet}@parc.com
** jheer@cs.berkeley.edu
*** john.w.bodnar@saic.com
* Address: 3333 Coyote Hill Road Palo Alto, CA 94304

visualize and many methods have been developed for doing so [3][6]. But many hierarchies change over time. Not only does incumbency in a position change, but the structure of the hierarchy itself changes. In the case organizational analysis, these changes are all of interest [2].

In this paper we present new visual analytic techniques to help an analyst understand very large hierarchies that change through time. To do this, we create an instance of a *cross-product* visual analytic tool called *TimeTrees* (Figure 1) that stitches together two analytic dimensions, hierarchical structure and time, into one visualization space. This structure enables the analyst to answer questions and detect patterns of relationships in a complex large-scale hierarchy x time space that would be very difficult otherwise.
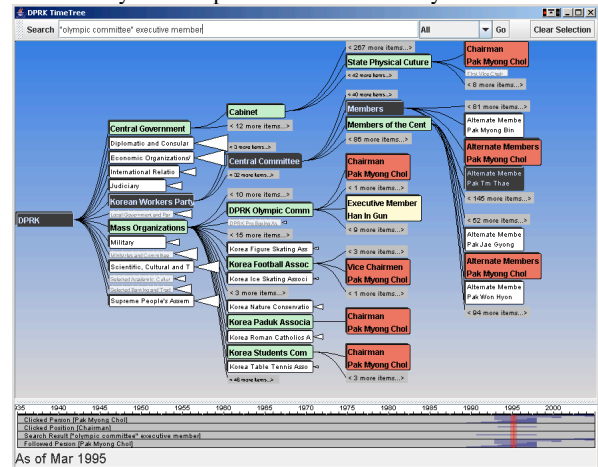

Figure 1: TimeTree allows users to interactively browse and search organizational hierarchies changing over time. (Data source: [7][23])[1]

2   TREES AND TIMELINES

Considerable work has been devoted to the layout of trees for information visualization. Trees are easier to lay out than general graphs, so they are often used when possible. Work on the use of trees for visualization is summarized in Di Battista et al. [6], Herman et al. [15], and Card et al. [3]. Early work assumed that trees were static, drawn with links and nodes, and often that the nodes were small points not containing significant information. Another approach has been to represent trees using containment, such as with a set of nested circles. The best known examples of containment trees are TreeMaps [17] and later variants. Containment trees have the advantage that they stay within a predetermined space, but there is little room for node content, especially for non-terminal nodes.

In order to accommodate larger trees and more content in the nodes, dynamic "focus+context" versions of trees were developed. Focus+context trees automatically choose portions of a tree to

---

[1] All organization and incumbency information shown in this paper is acquired from public sources [7][23].

show at any moment. Four main methods for determining how to show a focus+context tree have been proposed. Furnas [11][12] describes a class of fisheye techniques in which nodes are selected according to a degree of interest function, where the degree of interest of each node is determined by the sum of its intrinsic importance (such as distance from the root) and its distance from a focus node. Only those nodes above a certain threshold value are displayed. A second technique uses geometric distortion as in the hyperbolic tree [20]. Leung and Apperley [21] have addressed other geometric distortions. A third class of techniques is semantic zooming. As nodes grow beyond a certain size, their content changes, taking advantage of the greater space. Perlin and Fox [24] used this technique in their Pad system to expand a calendar. A final technique is to cluster nodes far from the point of focus. This technique was pioneered in SemNet [10]. Previously, we have developed Degree-of-Interest Trees [4] that combine all of these techniques. We subsequently introduced techniques to increase the scale of trees to hundreds of thousands of nodes and to lay them out more compactly [14]. This is the base on which we have built TimeTrees.

Likewise, considerable work has been devoted to the visualization of timelines. Plaisant et al.'s Lifelines [25] created timelines for patient histories. Kumar et al. [19] analyzed information that could be presented in timelines and created a general framework, ITER, for modelling and presenting them. Timelines often have the problem that their extent is too long to fit on a display while still showing important detail. To address this problem, Spence [29] introduced the Bifocal Lens, a focus+context technique that semantically compresses items in the periphery. Mackinlay et al. [22] introduced the Perspective Wall, another focus+context visualization which used perspective to compress the periphery. Timelines can be articulated to show more complex relationships. Richter et al. [27] used stacked overview + detail timelines to orient the analyst. Jensen [16] created multiple timelines that could be stacked and linked to show relationships among events. Ahmed et al. [1] and Ke et al. [18] visualized the history of InfoVis publications and characterized the research field's evolution over time.

Very little work, however, has gone into combining both hierarchy and timelines. One system that has attempted this is the "History Wired" project at the Smithsonian Institute (http://historywired.si.edu/). This project uses a static TreeMap of 435 objects linked to a static timeline. The visualization allows simple relationships between an object and its time to be noted. Erten et al. [9] created a 2D and 3D animation framework to visualize the evolution of various types of graphs. Chen and Morris [5] visualized the evolution of co-citation networks and reported that tools for network evolution should maintain the cohesiveness of critical paths.

3 TIME CHANGING HIERARCHIES

Most analysts work with limited knowledge of an organization and its history. This is due partially to the difficulty of collecting the data, but also to the difficulty of representing, manipulating, and making sense of it. Information about organizations can provide valuable insights ranging from questions such as "Who are the members of the governing council?" to "Who is 'Mr. Big' for energy policy?" to "Who is a rising star and potentially the next head of the ministry?" As the number of people and positions increases, the data and relationships involved becomes overwhelming and the analysis becomes difficult. That is why this type of analysis is performed on a routine basis by only a very few experienced political analysts.

To develop our solution and test its scale in the study we shall describe, we encoded changes in the entire political hierarchy of a small country from about 1947 until 2004 as given by public sources [7][23]. This comprised approximately 8,000 people and 18,000 incumbencies (Person P is in a position O from time T_{start}, to T_{end}) derived from public sources. This scale is large enough to handle a real problem and beyond the scale that an unaided analyst could normally do.

3.1 Questions for a Visual Analytic Tool to Answer

We start with some task analysis. What questions would an analyst want our visual analytic tool to help answer about organizational hierarchy? (Of course this is just a subset of the analysts' interest.) One way of thinking about this question is to think about understanding the political structure of the country in the conceptual terms of situation awareness [8]. Situation awareness is about knowing generally what is going on in the whole of a situation. There are three levels: Perception of the basic entities involved (Level 1), comprehension of how these go together (Level 2), and the ability to extrapolate to anticipate the near future (Level 3). We therefore defined a set of questions at each of these levels in collaboration with a senior analyst. Table 1 describes task details in each category.

Table 1: Visual Analytic Tasks. The tasks are presented as questions about a large organization and divided into three levels of situation awareness.

LEVEL 1. DIRECT FACTS
1.1 Who had position O at time T?
1.2 Who were all the people that served in position O?
1.3 What positions did P have at Time T?
1.4 Did P have both O1 and O2? When?
1.5 What positions did P ever have in his career?
LEVEL 2. COMPREHENSION
2.1 Build a résumé and address the questions that résumés answer:
2.1.1 What can I tell about P from his employment/position history?
2.1.2 What *key positions* did P hold?
2.1.3 Who were the associates of P in position O?
2.2 Do a quick check on organizational roles:
2.2.1 Who is Mr. Big X?
2.2.2 Who is the Logistics Guy, Operations Guy, R&D Guy, etc?
2.2.3 Who are the people who have had lots of positions and might be the organizational "memory"?
2.3 Do a quick check on organizational function:
2.3.1 When did P1 and P2 work together?
2.3.2 Who were all working at O1 when event X occurred and might have been involved?
LEVEL 3. EXTRAPOLATION
3.1 Who is a rising power?
3.2 Who are in the inner circle of Mr. Big X?
3.3 Who are the technical experts that make the project go?
3.4 Who are the projects coordinators with links to other organizations?
3.5 What projects are advancing?

The purpose of a visual analytic tool is to aid an expert by providing cues that can trigger action, whether that action is an interpretation of a situation or the initiation of events. The purpose of the TimeTree is to aid the analyst in answering as many of these questions as possible as part of a larger interpretive project. Since it isn't the only method the analyst has and some of the questions are difficult, we do not expect it to necessarily be able to help with all, still the table defines the need inherent in the task. In support of these tasks, we now discuss the TimeTree's data and visual design.

4 THE TIMETREE VISUAL ANALYTIC TOOL DESIGN

TimeTree is an interactive visual analytic tool designed to support analysts to browse and search a large hierarchy changing over time. There are at least two challenges in designing such a tool. One is to build a tool that is scalable enough to handle large data sets while remaining interactive. The other is to design a visualization that is cognitively tractable: allowing the analyst to explore complex associations such as those that can be derived from hierarchy x time relationships. We address these challenges through the design of the basic data model and the design of the visual interaction.

4.1 The TimeTree Information Model

An organization chart is usually composed of two types of information: 1) organization hierarchy (e.g. vice-presidents work for CEOs) and 2) individuals' incumbency information in the organization (e.g. person P is in position O from time T_{start} to time T_{end}). Figure 2 explains how to convert the information into the visual representation of the TimeTree.

Organizational Hierarchy (A-1): Positions in an organization can be easily modeled into supervisor-supervisee relationships as *<Position O_{super}, Position O_{sub}>*, which denote that an individual in a position O_{super} supervises individuals in O_{sub}. The whole organization hierarchy can be built by aggregating the supervisor-supervisee relationships in an organization (B in Figure 2).

We recognize that this static organization model creates trade-offs for succinctly representing the evolution of the organizational hierarchy. Such hierarchies change over time due to changes such as reorganizations, creating of new divisions, the closing of an office, and so on. In our information model, however, instead of capturing such high-level organizational changes, we instead use only individual incumbencies as a basis, because: 1) incumbency information is sufficient to completely represent organizational changes 2) we found that acquiring this type of information is difficult (reorganizations are often not well documented), and 3) most importantly, we found that static organization structure with visually penalized nodes makes it easier for users to track entities (see section 4.2).

Incumbency Information (A-2): of individuals in an organization can be encoded as a four dimensional tuple as:

$$<Person\ P,\ Position\ O,\ Time\ T_{start},\ Time\ T_{end}>$$

which denotes that a person P is in a position O from time T_{start} to time T_{end}. In this model, we maintain time bounds for each incumbency record. Accurate time values for each record are crucial to providing accurate inferences, but time ranges are subject to ambiguity and availability. Thus, we choose the time bounds to be the based on the most accurate available information. If accurate information is unavailable, we use the closest approximation, which is often implied from the publishing timeline of the information source.

Organization Evolution (C): By combining the Organizational Hierarchy and the Incumbency Information, we create a complete conceptual model capable of generating several cross-cutting views of the state of an organization. The following views can be induced from this model and a set of constraints specific to the view being generated, as described below:

Personal Event Strip (D-1): The personal event strip, D-1 in Figure 2, is a summarization view of all career positions that an individual assumed during his/her career. Information about where and how long a chosen individual assumed a position throughout his/her career is acquired through a filter defined as:

$$Personal\ Events\ (P_0) =$$
$$\{< O_i,\ Tstart_i,\ Tend_i> |\ if\ <P_0,\ O_i,\ Tstart_i,\ Tend_i>\ exists\ \}$$

Each filtered tuple, denoting a career incident, is overlaid as a thin line on the time line strip as in D-1 of Figure 2 and Figure 7. Multiple lines would be superimposed on the personal event strip if an individual assumed multiple positions in his/her career. When the chosen individual held multiple positions at once, the lines are overlapped accordingly. Implementation details will be discussed in a later section.

Positional Event Strip (D-2): Similarly, all events associated with a position can be overlaid on the timeline as shown in Figure 2 D-2. Given a position O_0, the position event overlay can be defined as:
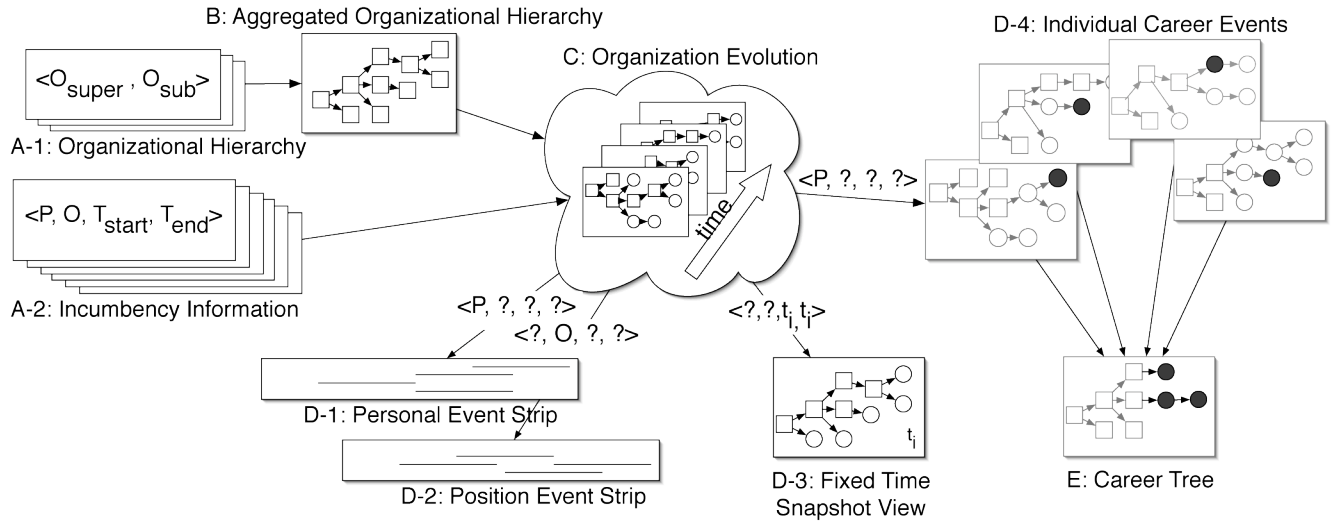

Figure 2: TimeTree Information Model. Squares denote positions and circles denote incumbents. Organization evolution is derived along the timeline by mapping a set of Incumbency Information (A-2) onto Aggregated Organization Hierarchy (B; generated by aggregating <O_{super},O_{sub}>). Four different types of filtering techniques (D-1 through D-4) are used to derive a set of information that drives user interaction in TimeTree.

Position Events (O_0) =

{<P_i, $Tstart_i$, $Tend_i$> | if <P_i, O_0, $Tstart_i$, $Tend_i$> exists }

In this strip, all career events associated with a specific position are overlaid in a way similar to the personal event strip.

Fixed Time Snapshot View (D-3): A snapshot of an organization chart is a fixed view of the organizational structure at a specific time point. The snapshot view is built by collecting the career events of each person who was incumbent at the given point of time. As shown in D-3 of Figure 2, the incumbency information can be queried from a database using the following conditions.

TimeTree(T_0) = {<P_i, O_i> | $Tstart_i$ <= T_0 and T_0 <= $Tend_i$ }

This view is built by composing the selected incumbent information with the static (aggregated) organization structure.

CareerTree View (E): This view is similar to the Personal Event Strip in a way that it tries to summarize an individual's career path. However, this view projects all positions of his/her career path on the top of the aggregated organization structure.

CareerTree(P_0) =

{ <P_0, O_i, $Tstart_i$, $Tend_i$> | if < P_0, O_i, $Tstart_i$, $Tend_i$ > exists }

D-4 and E in Figure 2 explains the conceptual process of how to build a CareerTree for an individual. From a series of organization charts shown in Figure 2 C, a set of organization charts on which a target individual P_0 ever appears are filtered out as in Figure 2 D-4. A CareerTree of an individual P_0 is built by aggregating the filtered organization charts for the individual, building the view in Figure 2 E. This single CareerTree contains all positions that an individual has ever held.

4.2   The TimeTree Visual Design

Visually, the TimeTree interface is composed of three major parts; a DOI (Degree-Of-Interest) Tree hierarchy visualization, a search interface, and a TimeSlider Timeline.

The DOITree visualization is a central part of TimeTree and is used to display a snapshot of an organization chart. The DOITree used in TimeTree takes advantages of techniques introduced in Card and Nation [4] and Heer and Card [14]. A basic rule of thumb is to assign a higher DOI value to important nodes and give a lower DOI value for less imperative information. We expanded the DOI model in [14] and added a new set of visual schema (Table 2) to support visual analytic tasks.

In TimeTree, four types of focus are available, to which TimeTree assigns the highest DOI values. They are: 1) nodes matching the *Follow* condition (an explicit way of specifying high DOI; it will be explained in the next section), 2) nodes matching the search condition, 3) the most recently clicked node, and 4) the tree root. Based on these foci, TimeTree computes DOI values of each node in the tree by propagating DOI values, as described in [4], calculating a DOI value that approximates the semantic distance of each node from the foci. The DOI value assigned to each node is then used for laying out the tree and determining the visual representation of the node. For these representations, we defined a new set of interest-inducing primitives, described in Table 2.

Table 2: TimeTree visual feedback schema

VISUAL EFFECT	TYPE OF NODE
Aggregated	Low DOI
Ghosted	No content subtrees
Red Color	Followed entities
Yellow Color	Occupied search results
Pale Green	Parents of high DOI nodes
Deep Green	Recently clicked nodes
Darkened Color	Node under the cursor and parents of that node
Triangular aggregate sub-tree	Sub-trees composed of low DOI nodes
<## more items>	Groupings of low-DOI sibling nodes

The visual schema (in Table 2 and illustrated in Figure 3) is designed to facilitate sophisticated interactions by providing visual feedback to the analyst, allowing them to understand the significance of a node based on its visual representation.
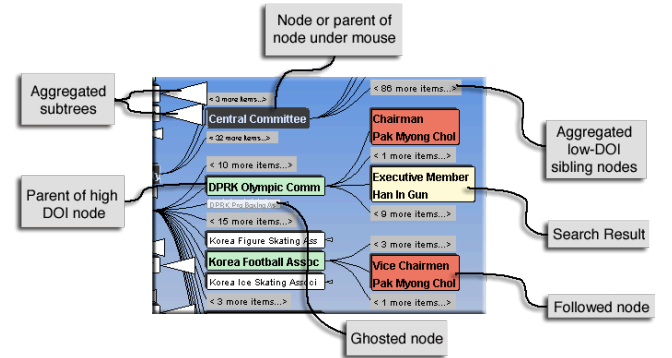

Figure 3: DOITree primitives and visual representation examples. A tree node visualized based on its DOI value as well as associated entities. (Data source: [7][23])

As with previous implementations of the DOITree [14], when there are too many nodes to fit into the limited screen space, we elide nodes of lower DOI values until the tree fits on the screen. Nodes that have been elided are represented via aggregate nodes. If an entire sub-tree is aggregated, a triangular representation of the aggregated nodes is added to the parent node's rendering. When multiple siblings are to be aggregated due to low DOI, they are grouped together into a single node representing the set (Aggregated sibling nodes in Figure 3).

DOITree dynamically responds to analysts' interest. Whenever a user clicks a node, specifies a search condition, or adds an entity into the Follow Set, DOITree immediately recalculates DOI values of all tree nodes and re-performs the layout, producing an updated view of the organization chart.

We use animation to help analysts to maintain an understanding of a tree when parts of it are being modified. The organization chart changes over time and the DOITree has to be updated accordingly. Whenever it is possible, all visual changes such as location or color of nodes are animated. We also animate transitions involving changes in an entity's position. This is especially useful because it provides continuity, so that the analyst can reuse his or her spatial memory of the organizational layout.
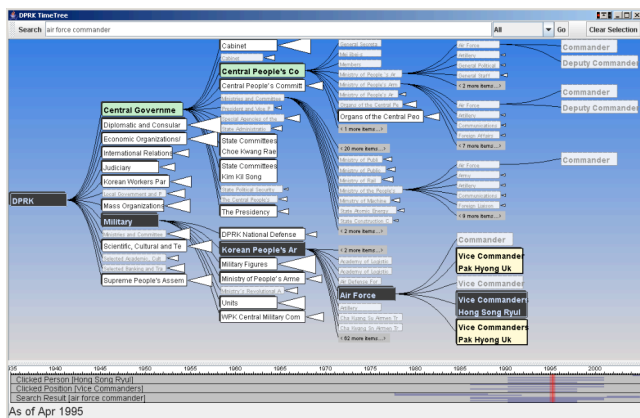
Figure 4: Search in TimeTree. When a node matches with the given search condition, it is assigned a high DOI value. If occupied, the node will expand further to show its occupant and become highlighted in yellow. Otherwise, the node will only be made visible, but not highlighted or expanded. (Data source: [7][23])

For finding rather than discovering information, TimeTree supports keyword search against attributes of Person or Position. When a node matches with the specified search query, it is highlighted with a pale yellow color as shown in Figure 3 and Figure 4. In addition, the "Search Result" strip in the TimeSlider (described below) is updated with career event information for all matching entities.

For instance, an analyst was trying to uncover information about the query "air force commander". The TimeTree shows results, as in Figure 4, allowing the analyst to continue exploring, leaving the search results as additional foci available to guide discovery.

The TimeSlider is much like a standard slider widget. While it serves the function of selecting a point in time to display in the DOITree, its value comes from the overlaid timeline. The TimeSlider is made up of several information strips, each representing a career event or events. As shown in Figure 7, individual career events are rendered on the information strip. When there are too many career events to fit in the available space, we use alpha compositing to overlap them together to give an analyst a sense of the density of events.

There are three default strips: the last clicked person (a Personal Event Strip), the last clicked position (a Position Event Strip), and the search results. As explained in section 4.1, we can create a Personal Event Strip and a Position Event Strip from available information.

From these views, the analyst can easily make rapid discoveries. For instance, (1) the last clicked individual in Figure 4 holds several positions as of April 1995, and (2) his career events span all the way to the most current data. It is also possible to discover the fact that (3) several others worked with him in the same capacity. This information is continuously available as the analyst explores the tree.

4.3 Tracking an Entity over Time

To facilitate tracking entities in an organization chart, TimeTree allows an analyst to set any person or position into the Follow Set. The Follow Set in TimeTree is defined by a set of entities in which the analyst has explicitly specified a persistent interest. When navigating, entities in the Follow Set are always displayed, if they have any incumbencies at the current time. Once an item is in the Follow Set, matching nodes are given increased saliency by increased size and distinct coloring.

In the TimeTree, we support two different types of following, by individual, or by position. When an individual is added into the Follow Set, the tree displays all incumbencies of that individual as shown in Figure 5. If a position is added to the Follow Set, the tree will display all officials who occupy that position, or an empty node, for positions with no current occupants. An analyst can add an entity to the Follow Set by double-clicking (to select the person) or by choosing from a context menu.

Since TimeTree uses a DOI Tree to display trees, the Follow Set is implemented as an additional set of high degree-of-interest nodes, artificially injecting high DOI values nodes that match entities in the Follow Set.
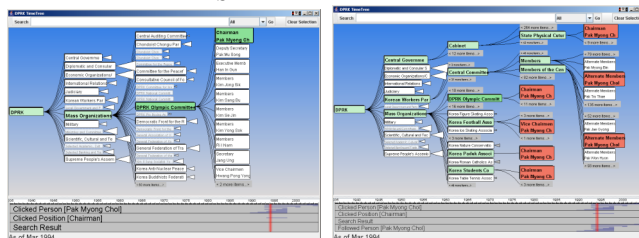

Figure 5: Left: A TimeTree with no entities in the Follow Set and one focused node. Right: The analyst has chosen to add the focus node seen in the left figure to the Follow Set. This causes all other positions of the chosen individual to be revealed. (Data source: [7][23])

As shown in the right panel of Figure 5, to further allow discovery, a new strip is added into the TimeSlider representing each entity in the Follow Set. Because the Follow Set acts as a persistent attribute, this timeline strip enables the analyst to quickly discover relationships between entities over time. There is a permanent strip representing the recently clicked node and position. Thus, it is easy to compare timelines of followed entities with the timeline of the most recently clicked node.
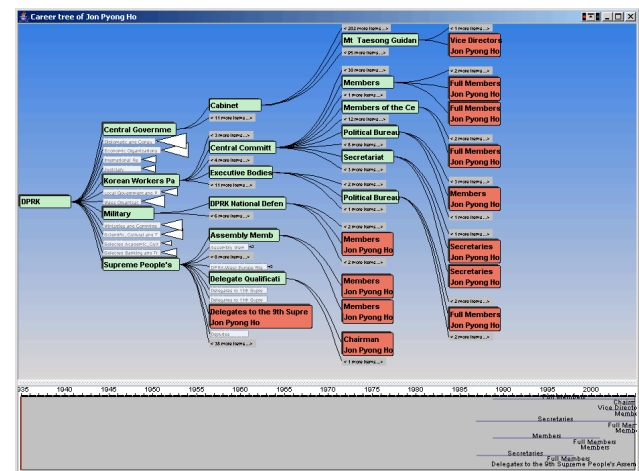

Figure 6: Career Tree of an individual. This visualization provides a compact overview of the career path of the selected individual. All career events of the individual are expanded and highlighted. (Data source: [7][23])

4.4 CareerTree

We developed a new visualization technique that summarizes the career path of an individual. While the snapshot view provides a view of the organization at a fixed point in time, allowing temporal exploration, the Career Tree overlays the entire career path of an individual over the static structure of the organization tree. This view, therefore, essentially removes time from the

dynamic hierarchy x time interactive visualization. This facilitates determining what connections and impact an individual might have had on various parts of the organization's hierarchy.

In this view, organizational subtrees not involved in the individual's career path are collapsed and displayed with diminished visual saliency. This provides a concise summary of the individual's career.

While the combined tree view doesn't contain temporal markup, an enlarged TimeSlider strip provides chronological display of the individual's positions. As the analyst explores position nodes in the CareerTree, matching events on the TimeSlider are highlighted accordingly, and vice versa.

5 USE OF THE TIME AGAINST INTELLIGENCE TASKS

In this section we examine two usage scenarios in detail. These scenarios were chosen to highlight the usability of the TimeTree, as well as exhibit solutions to the visual analytic tasks identified in section 2. Each action which answers one of the analytic task problems is annotated in brackets with the item number of the task or tasks in Table 1. All examples in the following sections are from public sources: [7][23].

5.1 Scenario One - Browsing An Organization

This task introduces a number of useful techniques for discovering facts using the TimeTree.

As an analyst browses the tree, the analyst clicks on nodes of interest and TimeTree reacts to the clicks by expanding the clicked nodes. Suppose an analyst discovered an individual serving as the chair of the Olympic committee [Task 1.1] in the mid 90s, as in the left image of Figure 5. When the node is selected, the TimeSlider Person view is updated as in Figure 7.

Figure 7: TimeSlider overlays associated career events as blue lines. (Source: [7][23])

Each line in this view represents a career event of the selected person [Task 1.3]. In this scenario, during the mid-90s, there were many career events for this individual. It is also easy to immediately notice who else was serving with him on that committee [Task 2.1.3].

Figure 8 TimeTree visualizes the evolution of an organization. TimeTree helps an intelligence analyst to easily identify a target individual's career changes. Left: Displaying a snapshot from the mid 1990's. Middle: The same snapshot, as of 2000. Right: As of 2004. When the analyst changes the time, the transition to the new state is smoothly animated from the preceding one. (Data source: [7][23])

The analyst double-clicks the node to add the individual into the Follow Set. This reveals all instances of this individual in the currently selected time, and adds a new timeline to the TimeSlider specific to the selected individual as in the right image of Figure 5. By examining the highlighted nodes, it is revealed that this individual was not only serving on the Olympic Committee, but also other athletic associations. He had political posts in the

cabinet and the Central Committee as well. [Task 2.1.1 and Task 2.1.2]

To explore how the tree changes over time, the analyst drags the selector in the TimeSlider to a new date. Moving the view to the year 2000, as seen in the middle panel of Figure 8, it is immediately obvious that the individual's career lost its focus on athletics, with just the political posts remaining. Changing the view to the latest data (the right panel of Figure 8) reveals that this individual assumed a new role as a Vice Minister in one of the cabinet ministries. [Task 1.5]

5.2 Scenario Two - Evaluating Intelligence Report

Suppose an analyst receives a report that needs verification. The report details information reported by a defector, which he claims to have received from another individual. First, an analyst searches for each individual. The TimeTree displays available information based on the individual's name [Task 1.3]. An analyst then double-clicks an instance of the individual in the search results to add them to the global "follow" set. This allows the analyst to track each one in any version of the tree as shown in Figure 5. [Task 2.1.1] Since the report being investigated focuses on information given to the defected official prior to defecting, we first note that our tree consistently shows that official having vanished from the organization around the time of their defection. We see this in the TimeSlider at the bottom of Figure 9.

To understand the relationship between the two individuals, we drag the TimeSlider to the last time the defector had positions in the organization. It is immediately apparent that these two individuals served together, actually, in two different organizations, as highlighted in Figure 9. [Task 2.3.1 partially]
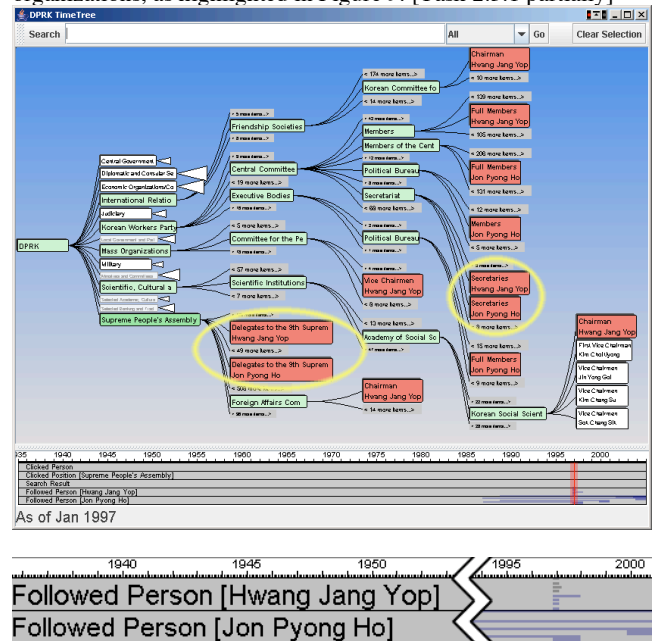

Figure 9: Compare two individuals' careers on one TimeTree. Upper: Both Hwang and Jon are highlighted. Lower: TimeSlider showing Hwang's career ending abruptly in 1997. (Data source: [7][23])

To further investigate whether the information could have come from the specified individual, we open a CareerTree view, as in Figure 6. Since the CareerTree is a cross-cutting view of the organization focusing on a single individual, it's easy to quickly determine whether the individual served in posts that would have given him access to the information in the intelligence

report.[Task 2.1.1 and Task 2.1.2] In this case, the individual had a position working in the National Defence Committee. With that information, the analyst concludes that the report is consistent.

6 DISCUSSION

This visual analytic tool takes on a task that would be difficult to handle without computational support. It allows questions to be asked about an entire country worth of political positions incumbencies over more than half of a century. Currently, a senior political analyst would know key figures, but would not have access to the texture of relationships that can be extracted from the TimeTree, both because of the scale of the amount of data and the possible complex interactions between hierarchy and time.

Our philosophy was to develop this visual analytic tool collaboratively and iteratively with an analyst. The tool as described in this paper has been through two iterations so far. We had a senior analyst rate the tool on the ability to help answer the questions in Table 1. The result is represented in Table 3. We assigned a mark of ++ to tasks the tool addressed directly, + to the tasks for which the tool was useful, but several steps were required, and 0 for tasks the tool didn't address. This gives us a rough assessment and helps direct us for a next round of development.

Table 3 TimeTree task ratings

LEVEL 1. DIRECT FACTS	
1.1 Who had position O at time T?	++
1.2 Who were all the people that served in position O?	++
1.3 What positions did P have at Time T?	++
1.4 Did P have both O1 and O2? When?	++
1.5 What positions did P ever have in his career?	++
LEVEL 2. COMPREHENSION	
2.1 Build a résumé and address the questions that résumés answer: 2.1.1 What can I tell about P from his employment/position history?	+
2.1.2 What *key positions* did P hold?	+
2.1.3 Who were the associates of P in position O?	+
2.2 Do a quick check on organizational roles: 2.2.1 Who is Mr. Big X?	+
2.2.2 Who is the Logistics Guy, Operations Guy, R&D Guy, etc?	0
2.2.3 Who are the people who have had lots of positions and might be the organizational "memory"?	+
2.3 Do a quick check on organizational function: 2.3.1 When did P1 and P2 work together?	+
2.3.2 Who were all working at O1 when event X occurred and might have been involved?	++
LEVEL 3. EXTRAPOLATION	
3.1 Who is a rising power?	+
3.2 Who are in the inner circle of Mr. Big X?	+
3.3 Who are the technical experts that make the project go?	0
3.4 Who are the projects coordinators with links to other organizations?	0
3.5 What projects are advancing?	0

The overall pattern of these ratings is clear (Table 4). Currently, the TimeTree does relatively well at Level 1 tasks (retrieval of basic facts), is helpful in many Level 2 tasks (comprehension of how facts go together), and less developed for Level 3 tasks (ability to extrapolate the near future). We are, in fact, working to create ways in which more subtle relationships can be discovered from the TimeTree.

Table 4: Distribution of ratings.

	0	+	++
Level 1	--	--	100%
Level 2	12%	75%	12%
Level 3	60%	40%	--

The TimeTree responds to recommendations in the National Visualization and Analytics Center roadmap report [30]. In particular, recommendation 2.4 recommended

Identify and focus on core conceptual schemas and create visually based components that support the analytical reasoning tasks associated with these schemas.

This study goes after the core conceptual schemas of organizational hierarchy and time. Hierarchies and time occur frequently in many analytic tasks. We have therefore thought they were promising areas to develop. But more than just focusing on single dimensions, there is an opportunity for visual analytics in focusing on cross product visualizations that take two areas and look at their possible combinations together. Due to the complexity of these combinations, associated tasks are even more difficult for un-aided humans, and hence are fertile areas in which visual analytic tools might make a substantial difference.

Computationally, TimeTrees are able to handle the scale in this range in which unaided humans have difficulties. The TimeTree is built off the base of prefuse [13]. Previous versions of DOI Trees have shown that the trees scale well to over 600,000 nodes. The greater computational burden of the TimeTrees does reduce the scaling somewhat, but the TimeTrees fairly easily handle the current database of 18,000 incumbencies and should be adequate for even larger databases.

TimeTree visual analytic tools would be possible for many applications involving hierarchy and time beyond intelligence analysis. Of course they could be used directly for analyzing corporations, but they could also be used for analyzing websites, for file systems, and for many other uses given the ubiquity of hierarchies.

7 ACKNOWLEDGEMENT

REFERENCES

[1] Ahmed, A., Dwyer, T., Myrray, C., Song, L., and Wu, Y. X. (2004) WilmaScope GraphVisualisation. In proceedings of InfoVis '04.

[2] Bodnar, J. W. (2003). Warning analysis for the Information Age. Washington, D.C.: Joint Military Intelligence College.

[3] Card, S. K., Mackinlay, J. D., Shneiderman, B. (1999) Readings in Information Visualization: Using Vision to Think. San Francisco, CA, Morgan Kaufmann.

[4] Card, S. K. and Nation, D. (2002) Degree-of-Interest Trees: A Component of an Attention-Reactive User Interface. In proceedings of International Conference on Advanced Visual Interfaces, Trento, Italy.

[5] Chen, C., Morris, S. (2003) Visualizing evolving networks: Minimum spanning trees versus Pathfinder networks. In Proceedings

of the IEEE Symposium on Information Visualization (InfoVis'03). Oct 19-24, 2003. pp. 67-74

[6]   Di Battista, G., Eades, P., Tamassia, R., and Tollis, I. G. (1999) Graph Drawing: Algorithms for the Visualization of Graphs. Upper Saddle River, NJ: Prentice Hall.

[7]   Directory of officials of the Democratic People's Republic of Korea: Volume for 1992, 1988, and 1978. Washington, D.C.: Central Intelligence Agency, National Foreign Assessment Center.

[8]   Endsley, M. (2003). Designing for Situation Awareness. Boca Raton, FL: CRC.Press.

[9]   C. Erten, P. J. Harding, S. G. Kobourov, K. Wampler, and G. Yee. (2003) GraphAEL: Graph animations with evolving layouts. In 11th Symposium on Graph Drawing, pp. 98–110.

[10] Fairchild, K.M., Poltrock, S.E. and Furnas, G.W. SemNet: Three-dimensional graphic representation of large knowledge bases. In Cognitive Science and its Application for Human-Computer Interface. Erlbaum, Hillsdale, NJ., pp. 201-233.

[11] Furnas, G. W. (1986) Generalized Fisheye Views. In proceedings of SIGCHI '86 Conference.

[12] Furnas, G. W. (2001) The FISHEYE View: A New Look at Structured Files, In Readings in Information Visualization. Morgan Kaufmann.

[13] Heer, J., Card, S. K. and Landlay, J. (2005) Prefuse: A Toolkit for Interactive Information Visualization, In proceedings of SIGCHI 2005, ACM Press.

[14] Heer, J. and Card, S. K. (2004) DOITrees Revisited: Scalable, Space-Constrained Visualization of Hierarchical Data. In proceedings of International Conference on Advanced Visual Interfaces. Gallipoli, Italy.

[15] Herman, I., Melancon, G., and Marshall, M. S. (2000)

[16] Jensen, M. (2003) Visualizing Complex Semantic Timelines, NewsBlip Technical Report NBTR2003-001, http://newsblip.com

[17] Johnson, B. and Shneiderman, B. (1991). Space-filling approach to the visualization of hierarchical information structures, in Proceedings of IEEE Visualization '91, 284-291.

[18] Ke, W., Borner, K., and Viswanath, L. (2004) Analysis and Visualization of the IV 2004 Contest Dataset. In proceedings of InfoVis '04

[19] Vijay Kumar, Richard Furuta, Robert B. Allen (1998) Metadata Visualization for Digital Libraries: Interactive Timeline Editing and Review. ACM DL 1998, pp. 126-133

[20] J. Lamping and R. Rao, (1994) Laying out and Visualizing Large Trees Using a Hyperbolic Space. In proceedings of UIST'94, ACM Symposium on User Interface Software and Technology.

[21] Leung, Y., and Apperley, M. D. (1994): A Review and Taxonomy of Distortion-Oriented Presentation Techniques. ACM Transactions on Computer-Human Interaction, 1 (2), 126-160.

[22] J. D. Mackinlay, G. G. Robertson, and S. K. Card. The perspective wall: detail and context smoothly integrated. In Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through technology, pages 173--176, 1991.

[23] North Korea Directory (Chosen Minshu Shugi Jinmin Kyowakoku Soshikibetsu Jinmeibo). Volume for 1997, 2000, and 2004. Tokyo. Rajiopuresu (Radio Press).

[24] Perlin, K. and Fox, D. (1993) Pad: an alternative approach to the computer interface. SIGGRAPH 1993, pp. 57-64

[25] Plaisant, C., Milash, B., Rose, A., Widoff, S. and Shneiderman, B. (1996) LifeLines: Visualizing personal histories. In Proceedings of ACM CHI 96 Conference on Human Factors in Computing Systems, pp. 221-227.

[26] Resnikoff, H. L. (1989). The Illusion of Reality. New York: Springer-Verlag.

[27] Richter, H., Brotherton, J., Abowd, G. D., Truong, K. (1999) A Multi-Scale Timeline Slider for Stream Visualization and Control. GVU Center, Georgia Institute of Technology, Technical Report GIT-GVU-99-30.

[28] Simon, H. A. (1973). The Sciences of the Artificial. Cambridge, MA: MIT Press.

[29] Spence, R. and Apperley, M. D. (1982) "Database navigation: An office environment for the professional," Behaviour and Information Technology, vol. 1, pp. 43-54, 1982.

[30] Thomas, J. J. and Cook, K. A. (2005) Illuminating the Path: The Research and Development Agenda for Visual Analytics. IEEE CS Press.