

Google+ Ripples: A Native Visualization of Information Flow

Fernanda Viégas
Google, Inc.
viegas@google.com

Martin Wattenberg
Google, Inc.
wattenberg@google.com

Jack Hebert
Google, Inc.
jhebert@google.com

Geoffrey Borggaard
Google, Inc.
geoffreyb@google.com

Alison Cichowlas
Google, Inc.
asc@google.com

Jonathan Feinberg
Google, Inc.
feinberg@google.com

Jon Orwant
Google, Inc.
orwant@google.com

Christopher R. Wren
Google, Inc.
cwren@google.com

ABSTRACT

G+ Ripples is a visualization of information flow that shows users how public posts are shared on Google+. Unlike other social network visualizations, Ripples exists as a “native” visualization: it is directly accessible from public posts on Google+. This unique position leads to both new constraints and new possibilities for design. We describe the visualization technique, which is a new mix of node-and-link and circular treemap metaphors. We then describe user reactions as well as some of the patterns of sharing that are made evident by the Ripples visualization.

Categories and Subject Descriptors

H5.2. Information interfaces and presentation: User Interfaces. Graphical user interfaces.

Keywords

Visualization, social networks, social data analysis.

1. INTRODUCTION

As long as people have been having conversations, news has spread. Opinions on movies, politics, and friends naturally travel as one person talks to another, who then talks to another. But the process has always been a bit of a mystery: If I tell you a piece of juicy gossip, whom will you tell? And whom will they tell, in turn?

With the advent of social networks that keep persistent histories of public conversations, this process is no longer completely opaque. It is possible to trace the paths of information, at least in public channels, as it flows from one person to the next. These paths are interesting from many points of view. For users of a social network, understanding how information travels can lend a better sense of the potential audience of a post; in turn, this understanding can lead to a feeling of control over their online world. For scientists the data is an invaluable tool for analyzing the mechanics of social networks.

There are two significant challenges in helping people understand information flow. First, online networks can generate large cascades of sharing, involving thousands of people. While it is possible to summarize some of this activity in a few data points or simple chart—say, a list of the people who have been quoted the most—any attempts to aggregate the data lose richness and detail. Thus a natural way to present chains of sharing is not with summary statistics, but with a visualization capable of showing a fuller story.

Copyright is held by the International World Wide Web Conference Committee (IW3C2). IW3C2 reserves the right to provide a hyperlink to the author's site if the Material is used in electronic media.
WWW 2013, May 13–17, 2013, Rio de Janeiro, Brazil.
ACM 978-1-4503-2035-1/13/05.

A second challenge, especially relevant for end users, is one of convenience. While it is possible for third-party sites to create visualizations of public social network data, such tools require a significant context switch. A user who wants to see how a post has spread needs to go to a completely different site. Unfortunately, on the web, requiring more than a click can be a significant barrier to adoption.

This paper introduces Google+ Ripples, a tool that addresses both of these challenges. It includes a visualization that can scale to show large information flows. And it does so natively, directly on the site: users of Google+ who find an interesting post can click a link to see, immediately, a picture of how it spread.

1.1 Google+ and sharing behavior

Google+ is an online social network, launched in 2011. Like networks such as Twitter and Facebook, Google+ allows users to create posts that consist of text and images, and may point to other media via URLs. The list of a user's posts is called their stream. Users can also share posts they read, *i.e.*, they create a post in their own stream referencing that post. A share can be as simple as a pointer to the original post, or it can include additional commentary. Users may comment on each other's posts as well, though the current version of Ripples does not include this data.

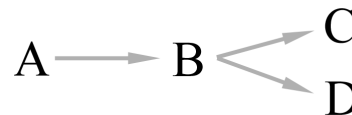


Figure 1. The structure of a simple reshare tree. A: original post; B: reshare of parent; C and D: reshares of B.

An interesting post may be shared and “reshared” in a chain or tree. In other words, Person A might write a post; Person B shares that post with some commentary; then People C and D reshare B's post with some additional commentary. Thus, the flow of dissemination of a post has a natural tree structure (Figure 1) where each post's “parent” is the one that it is resharing. In the example above, A would be the root, and C and D would be leaves.

The temporal aspect of this process is also of interest. Typically the distribution of post dates is bursty. In particular, we observed that for many posts, the vast majority of reshares occur soon after the initial post, with a few occurring much later. This observation echoes that of [10] as well as informal observations on other social networks. As we will see, this uneven distribution turns out to impose design constraints, since using time as a spatial dimension leads to inefficient, hard-to-read layouts.

Finally, it's important to note that not all posts are accessible to all users. A key feature of Google+ is that it enables fine-grained sharing, so that not every user has permission to see every post. It is also possible to share content publicly. As a result, activity on the site is a mix of public and non-public conversation. The Ripples system only displays public posts. While this provides just a partial view of the conversations on the site, it greatly simplifies the task of preserving user privacy.

1.2 Related Work

We discuss several types of related work. First, there is a long history of visualizations of conversations and online sharing. In our review of relevant work, we focus specifically on systems that show disaggregated data; displays of summary statistics alone are numerous but not directly of interest here. Second, there are many visualizations of online social networks that are not native to social platforms, but that depict network structure. Finally, since our system uses a data model that is a collection of trees, it draws on a long line of research into tree visualization. We will discuss this latter stream of work in the context of a detailed description of our visualization technique in Section 3.

There has been extensive work on ways to visualize conversations [5]. Much of this work has been based on "threads," which have a similar structure to the Ripples trees of reshares. There are practical differences, however. For example, a long email thread usually has multiple messages per person, but it is rare (though not impossible) for a user to appear more than once in a reshare tree. Furthermore, the largest Ripples reshare chains easily top 1,000 posts, whereas most email, usenet, or online forum threads are smaller.

One relevant system is Xiong's WebFan [27], which used a radial display of a conversation thread: an initial post in the center, with a subsequent replies branching out. The result was beautifully evocative of a conversation expanding and evolving, but at the same time suffered from crowding at the edges and had difficulty with labeling.

More recently, visualizations based on social network data have appeared. Many of these use data from Twitter: the service is broadly popular and has an open API; furthermore, the public nature of most Twitter posts ("tweets") makes it especially easy to work with. Twitter visualizations tend to fall into two categories. Some ask a user for their login information, and then show a customized view of their own conversations. Others are based on a stored set of public tweets.

Examples of such visualizations are too numerous to review comprehensively here, but we describe some key aspects of such work. Many use node-and-link diagrams to show aggregate patterns of retweeting and conversation, rather than the diffusion of a particular piece of information. An example closer to Ripples comes from SocialFlow [16]. Their technology tracks a specific subject, and appears to show a conversational network (not necessarily a tree) whose geometry is defined via a force-directed layout.

Project Cascade [11] is an installation created by the New York Times for internal analysis of the spread of the newspaper's stories on Twitter. The system runs off of a five-screen video wall controlled by an iPhone, and offers multiple perspectives on information flow. The most detailed of these views uses a 3D polar layout, in which points represent tweets, distance from the origin represents time since the first tweet with a story, and height

above the xy plane represents the depth of a retweet chain. The resulting visualizations are beautiful and dramatic, though the use of 3D, and time as a spatial variable, can bring challenges with labels.

All of the systems described above are third-party tools. They exist separately from the social networks they visualize, and a user needs to break context to go from a social network post to its visualization. In a few cases, these third-party tools can be "plugged into" an existing system. For instance, a number of Facebook widgets visualize users' social network; e.g., the Facebook Friends Wheel shows a local view of a user's friends, linking people who know each other [7]. In addition, some blogs include plugins that graph statistics on the blog itself, such as traffic over time [18].

One precedent for what we call *native* visualizations is YouTube Insights [22]. Next to many videos on YouTube there is an icon that, when clicked, displays a dashboard with a graph of viewership over time, a map of viewers' locations, and other data. This immediate, seamless transition from media object to visualization of its history is one of the inspirations for Ripples.

2. THE RIPPLES SYSTEM

Ripples, launched in October 2011, is a feature inside of the Google+ network. That initial version, which we call Ripples V1, was available in the US only and allowed users to track the sharing of an individual Google+ post. A second version, Ripples V2, was launched in March 2012. The V2 edition is internationally available and allows users to track the spread of all posts that contain a specific URL.

2.1 Accessing Ripples

Next to every post on Google+ is a drop-down menu offering a set of relevant user actions. On public posts that have been reshared at least once, this menu includes a "View Ripples" option, which takes the user directly to the Ripples page. The entire page, including the full visualization, usually loads in less than a second, leading to a seamless experience for the user.

2.2 The Ripples page

The Ripples page (Figure 2) is made of several modules that provide information about a given post. At the top is a snippet of the post itself (or the entire post, if it is short), along with instructions. In the center is the main visualization. We describe the visualization in greater detail in Section 3; briefly, Figure 2 shows a collection of posts that mention a particular viral video. Each solid-colored circle represents a tree of reshares. For example, the largest (lavender) circle shows the cascade of reshares of a post originally written by user Chris Pirillo, mentioning the video. To its left is a green circle with the label "Laughing Squid," which shows the tree of reshares of an initial post by user Laughing Squid. Within each circle, arrows and smaller circles show the structure of resharing: e.g., the thick arrow from user "Chris Pirillo" pointing to a circle labelled "Josh Armour" indicates that user Josh Armour reshared Chris Pirillo's post. The smaller circles within the Josh Armour circle indicate further reshares. Note that the visualization scales to relatively large numbers of nodes: Figure 2 depicts a forest of more than 4,500 shares. A final feature of the visualization, introduced in summer of 2012, is that it updates in near real time: within a few seconds of a new post, the visualization will show that post, with a brief "expanding ripples" animation to call attention to the new item.

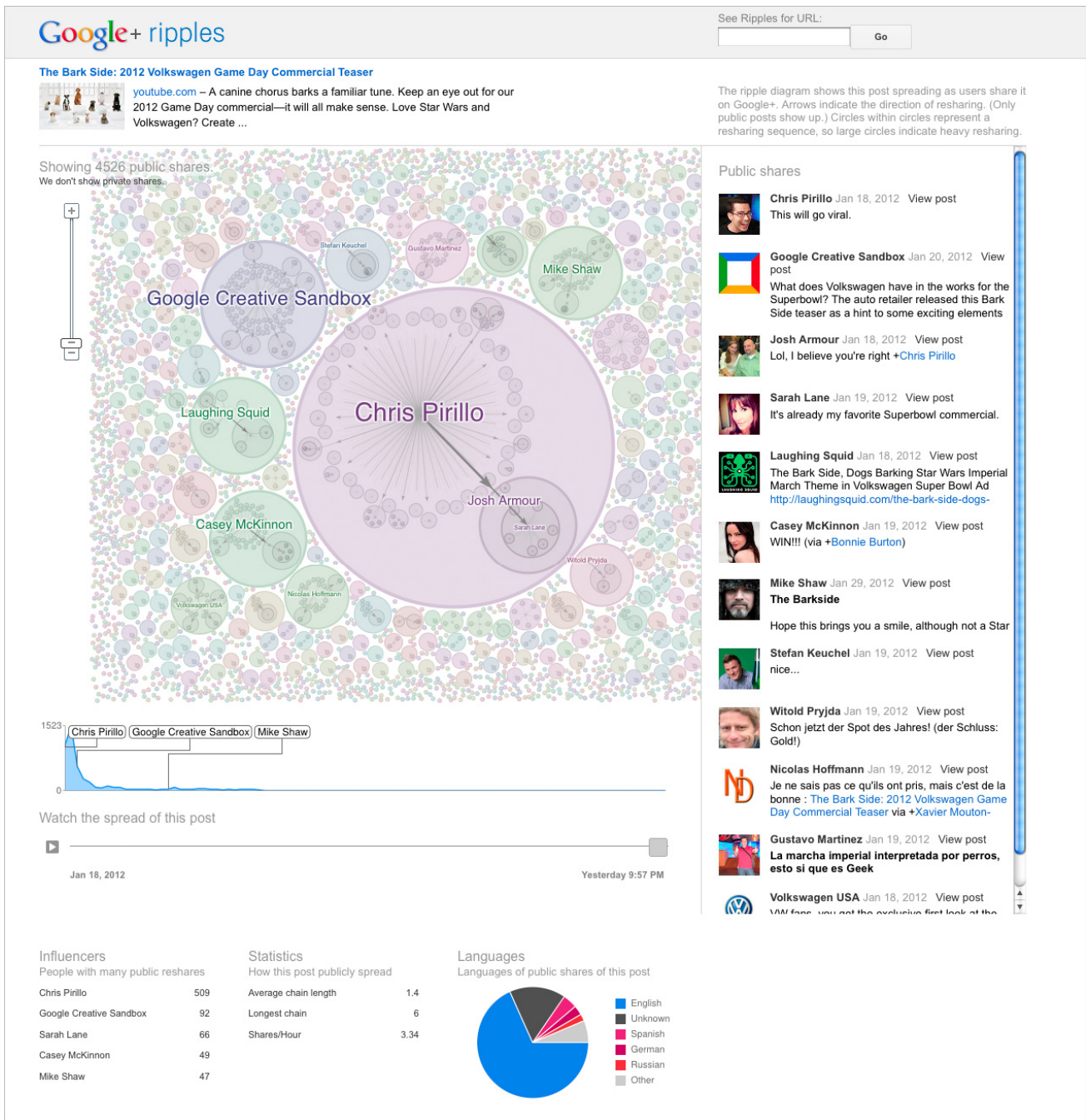


Figure 2. Full Ripples page. At the top of the page is a snippet of the post itself along with instructions. In the center is the main visualization. Below the visualization is a timeline of shares over time and a small table of statistics about the share graph. A sidebar shows the contents of posts made by the most prominent users in view.

Below the visualization is a graph showing the volume of shares over time. As described below, the graph is linked to the visualization, letting users “play back” the sharing history. Underneath the graph is a small table of statistics about the share tree. We include a list of the top “influencers”—that is, the set of people who were individually reshared the most. (For the purposes of this graph, we simply define “influence” as the total number of direct reshares.) The points in time when these people made posts are labeled in the graph as well.

Obviously there are other ways to define influence, and early iterations of the system contained more than one list, corresponding to a variety of metrics. For example, another natural measure of “influence” is the total number of downstream shares, rather than direct shares. In informal user tests, however, we discovered that the use of multiple metrics was often confusing. People often expected a list of influencers ordered by direct reshares, so in the end we used that measure alone.

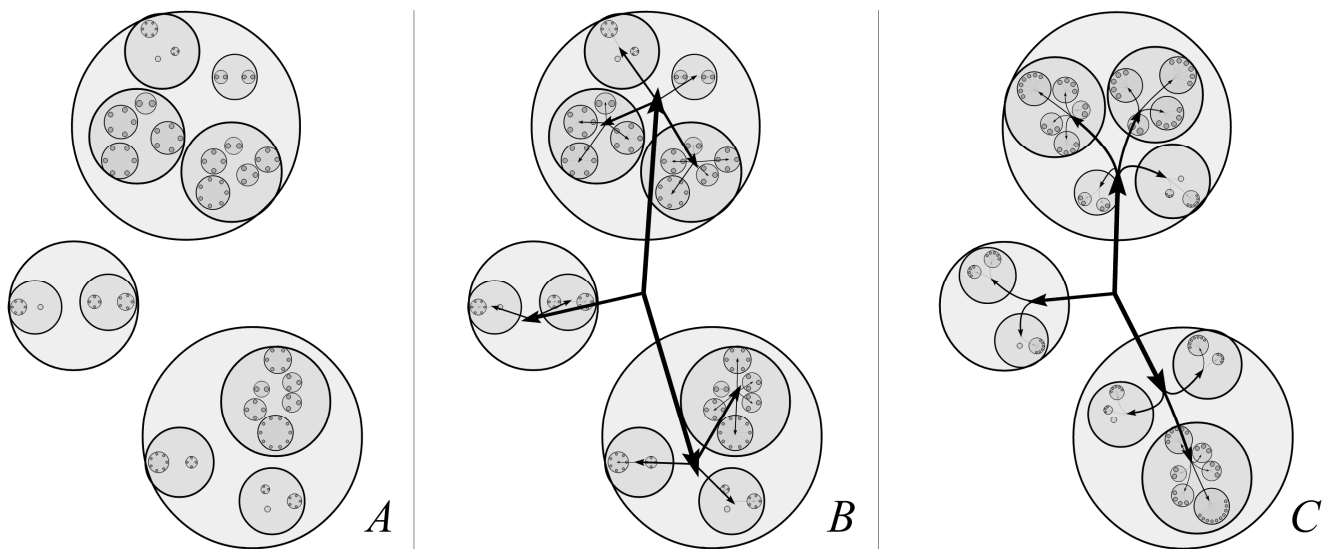


Figure 3. Graphical treatment of directionality. A: no arrows. B: Angular arrows. C: curved arrows along with rotation of circle contents so that the centroid lies opposite the point of entry of an arrow to maximize the appearance of flow.

A second panel provides summary statistics about the share graph itself. We display the average chain length, the length of the deepest chain, and the “shares per hour” for the life of the post. As discussed in section 4, the first two figures seem to reveal information about the type of post and how it was perceived. The “shares per hour” figure—which generally decays over time, after an initial burst—can be useful in identifying “viral” posts. Finally, a third panel shows a pie chart of the distribution of languages of posts. Due to the significant international membership of Google+, this pie chart is helpful in understanding the readership of a given post.

3. VISUALIZATION TECHNIQUE

The centerpiece of the Ripples system is a visualization of share trees. Before going into the details of the visualization, we describe some of the data and design constraints and what we learned from early iterations.

3.1 Design Process

For each public post on Google+, we have access to all of its downstream public reshares, along with a tree structure on those reshares. We also know the timestamp and author of each reshare. In Ripples V1, this constituted the entire data set. Ripples V2, however, displays a forest of trees, showing all posts that mention a given URL (Figure 4). In theory we had more data available to us, including the full text of each reshare, but we could not include all this information in the visualization due to latency and browser memory constraints.

Our first designs were organized temporally. In these views, dots represented posts, the x -axis corresponded to time, and the y -axis corresponded to a depth-first list of tree node. Unfortunately, this approach had multiple problems. Most items see significant sharing early on, and then very little later (e.g. Figure 2 timeline). As a result, basing the x positions of items on time created a crowded view. In a few cases there were multiple “waves” of posts, in which structure could be seen, but these were the exception and not the rule.

We tried several variations on x -axis positioning to avoid the visual overplotting caused by bursts of activity. In one, we set the x position to the order of posting, and in the other we set the x -position of a node to its depth in the sharing tree. While these made the tree structure clearer for small trees, the visualization remained impossible to read for graphs with large branching factors. The basic problem is that there aren’t enough pixels vertically to fit thousands of posts.

To solve this problem, we decided we needed a more space-efficient layout. We tried both a standard pivot-by-middle treemap layout [17] and a jigsaw treemap [25]. The irregular arrangement of a jigsaw treemap was immediately deemed unattractive and hard to read. A standard rectangular treemap gave a decent, efficient arrangement of leaf nodes, but finding a natural location for higher-level nodes—usually the most important ones in this context—was problematic.

Most crucially, both methods suffered from a fatal aesthetic drawback. Neither “looked” like flow of information. Understanding the diagrams felt like decoding a complicated message. Connecting shares to their parents with arrows just resulted in an unreadable mess, as the connecting lines collided with other shares beneath them. Given that a native visualization aimed at end users needs to be readily comprehensible, we decided we needed a technique that would allow familiar metaphors to indicate directionality of flow.

3.2 Circle Packing

Our design challenge was to find a visualization technique that could combine the best aspects of these variations. In particular, we wanted a visualization method that met three criteria:

1. Space-efficient layout of leaf nodes, as in a treemap
2. A natural sense of hierarchy, where parents were more prominent than their children
3. Space to overlay arrows from parents to children, to show directionality of flow

To achieve this combination, we used a variant of a circular treemap.

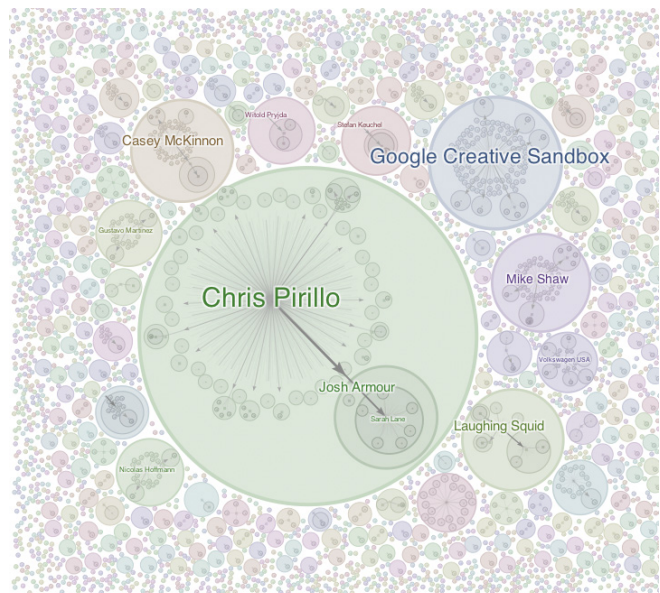
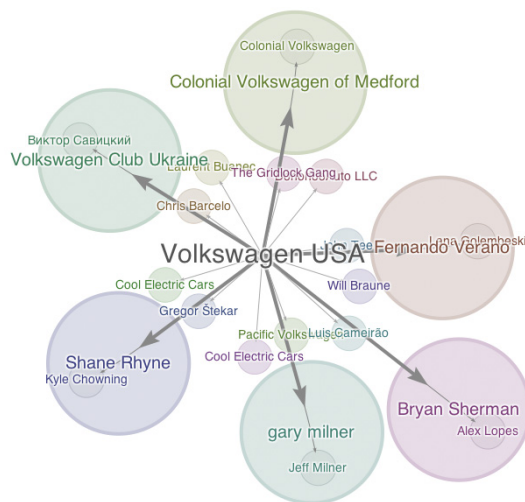


Figure 4: Visualizing the spread of “The Bark Side,” a Volkswagen teaser video for the 2012 Super Bowl. Left: Ripples V1, showing the spread of the original Volkswagen post (i.e., a single tree). Right: Ripples V2, showing all shares of the video (the entire forest). The arrow shows Volkswagen’s tree. It is clear that the video spread organically through many channels.

3.2.1 Previous work

Circle-packing visualizations and recursive circular visualizations of trees have been discussed in many contexts. Our technique can be viewed as a hybrid version of previous techniques, which we describe now.

Several systems include visualizations that represent a data set by circles of varying sizes. For example, the “bubble charts” of Many Eyes [20] used this technique. The layout algorithm involved a greedy randomized method with larger items placed toward the center of the screen. More recently, the d3 library [3] provides an interesting layout algorithm that puts circles in a kind of snaking spiral, producing a very different look from the placement algorithm described below.

Hierarchical circle-packing visualizations, also called circular treemaps, have also been proposed [2][6][24][26]. This visualization technique has not become popular, perhaps because the packing is not as space-efficient as rectangular treemaps: circles do not fit snugly together and so leave a significant amount of whitespace. Ironically, in our context this is as much a feature as a bug, since it provides room to overlay connections without creating a visual mess.

Finally, some graph visualization toolkits offer a “balloon layout” for trees [8], in which nodes are arranged in circles around their parents. Balloons offer a natural sense of hierarchy, but not the space efficiency of a treemap-style layout, and for trees with a high branching factor become unreadable.

3.3 The Ripples visualization: a “balloon treemap”

The goal of the Ripples visualization is to combine the relative space-efficiency of a circular treemap with the structural clarity of a balloon graph layout. To do so, we make a set of subtle but important changes to the basic geometry along with careful attention to rendering details. We call this hybrid a *balloon treemap*.

3.3.1 Heterogeneous circle packing geometry

As mentioned before, there are many different ways to pack circles. Existing visualizations typically pick one layout algorithm and use it recursively. After some experimentation, however, we decided that different levels in the share tree hierarchy look best with different layout algorithms.

The layout at the top level depends on whether Ripples is showing the chain of shares for a single post (a network that is a tree) or all posts and shares that mention a particular URL (a forest)—see Figure 4. When placing circles corresponding to the individual trees in a forest, the most natural choice was a layout in which big circles were generally toward the center and small circles could be placed in between large ones if necessary.

At lower levels (or the top level, when showing only a single tree of shares) we discovered that the large-circles-in-center layout suffered from a bad problem with labels: it essentially guaranteed that the label for a node would collide with label for its children. To avoid this problem, for deeper levels of the tree we used the opposite layout algorithm, one in which a “hole” of whitespace was placed at the center, then small circles around that, with the largest ones on the outside.

Our algorithm in both cases is greedy placement of circles, either largest first or smallest first, searching along a spiral [19]. To speed collision detection we use a *kd-tree* structure [12]. As a special case, for large sets of items that are all the same size, we use a golden-ratio spiral to layout instantly, without needing to perform any collision detection calculations, as in [9] and [14]. The resulting algorithm can handle real-world trees of several thousand nodes, all in less than a second.

We also included some special-case layouts for sets consisting of fewer than 10 circles, all of equal size. These simple, special cases were common at the lowest level of hierarchy. In these cases, we simply laid circles out around a 180-degree angle. The reason we used a 180-degree angle rather than the full 360 degrees relates to the directionality considerations below.

3.3.2 Clarifying Tree Topology

In a circular treemap, the tree topology can be hard to discern. To make a more readable display, we combined the circular layout with the directed arrows typically seen in a traditional graph layout. This is a natural idea, and has been explored in the context of rectangular treemaps by Nguyen and Huang [15]. Our design has a few new elements, however. First, we made sure to use smooth curves, rather than straight lines, to connect items; this helped provide a feel of “flow” from bigger circles to smaller ones. In addition, this technique helped clarify directionality: with polygonal paths we found that direction was ambiguous, so that users would need to carefully read arrows.

Another small but helpful touch was to rotate circle arrangements so they pointed “outward”. In other words, we found the centroid of the circles, and rotated the arrangement so it was opposite from the direction of “flow” into the circle. Especially in the case of small sets of items, which were arranged in a half-circle, this resulted in a pleasing, clear sense of directionality. See Figures 3b and 3c, which show example trees with and without these design optimizations.

To help distinguish the different top-level trees (or top-level branches of a single-tree forest) we used a palette of gentle pastels. Each tree/subtree was assigned a unique, mostly transparent color; circles in that tree are then filled with this color. The effect is to convey tree depth with overall darkness (as the mostly transparent colors overlap) and to visually separate neighboring circles. We also tried alternative color schemes—including encoding additional data, such as showing a share’s timestamp via color. Unfortunately, it was surprisingly hard to read a graph colored by post time, as the scattered colors ended up looking like a jumble.

3.4 Rendering, Interaction and Animation

The visualization is implemented in JavaScript, and all rendering is performed using the HTML5 Canvas element. For large data sets (more than 1,000 items) rendering in full detail becomes too

slow for smooth animations. To avoid choppy motion, during animations we use a level-of-detail calculation to draw only the largest, most prominent visual features. After animations finish, we then draw the full visualization into an offscreen image, and perform a smooth crossfade to draw that image to the screen.

We provide multiple ways of interacting with the main visualization. The primary interface for seeing more details is to zoom and pan. We use essentially the same interface as Google Maps, partly because it seemed like it made sense to users, and partly for consistency between Google properties. However, we also allow users to zoom directly into the “sphere of influence” of a user by clicking on their name or circle.

As the user moves the viewpoint, a sidebar shows the contents of posts made by the most prominent users in view. This is a standard linked view, but an interesting twist arose in the internationalization process. In Ripples V2, the system became available in 60 languages. As a by-product, the internationalization team added logic so that on right-to-left languages (e.g., Hebrew, Arabic) the sidebar switched sides. This is a good example of the kind of flexibility required by native visualizations.

As a user moves the mouse over circles, a “hovercard” appears, displaying post details (Figure 5). The hovercard also includes a “Follow” button, which allows users to add new people/entities to their network. The fact that users can take advantage of the visualization to affect change in their social network is a benefit of a native visualization.

Finally, we wish to show the evolution of the tree over time. Although often share trees grow in a single burst, sometimes there are more interesting patterns. To provide support for exploration of temporal growth, we added an interactive slider to the time graph below the main visualization. Moving the slider, or clicking a “play” button, provides an animated view of the graph, filtering to show only the posts made before or on a given date.

As with zooming and panning, drawing the full graph is too slow to allow smooth motion. In the time animation, however, we can’t resort to drawing a low-resolution view—often the appearance of large groups of graphically small leaf nodes is an important aspect of the evolution of a share tree. To provide at least the illusion of interactive behaviour, we used a small trick. As the slider moves forward, from time t_1 to time t_2 , we simply draw the shares that appeared between t_1 and t_2 . This allows quick animation during forward movement of time, with especially fast drawing when the increment $t_1 - t_2$ is small. When moving the slider backward in time, we redraw the entire graph for that time. This is a slow process, but luckily, it is also uncommon. Thus we achieve smooth motion in the cases when the “play” button is hit and when users move the slider slowly in the forward direction.

A second type of animation occurs when a new post appears and Ripples needs to update the view. In this case, the rendering of the new tree or forest is done in an offscreen image, which is then smoothly blended on top of the existing image. The shifts of layout caused by the insertion of a new node helps call attention to the change, but for large numbers of shares the new node isn’t always obvious. To address this problem, we added a short animation of expanding circles (like water ripples) around the new node.

4. PATTERNS OF INFORMATION FLOW

We believe the Ripples visualization brings to light some interesting patterns of social sharing. Both for their intrinsic interest, and because they illustrate how the visualization can illuminate this area of research, we describe them in more detail below.



Figure 5: Hovercard showing details of a user’s post and a “Follow” button, which may be used to follow the highlighted user directly from the visualization.

4.1 Broad but shallow patterns

A common feature of social networks is the presence of “celebrities,” users who have many followers, and whose posts can be highly influential. These celebrities can of course be Hollywood stars, but can also be well known in a particular field, or represent prominent businesses or media outlets. We often observed large share trees when celebrities posted. Yet these trees were notably shallow, with few long reshare chains. Figure 6 (right) shows an example, where Selena Gomez (the well-known teen actress and singer) posted a Valentine’s Day message. Many followers shared it, but the spreading rapidly faded out.

We observed this pattern many times, and not just with popular-culture celebrities. Similar effects occur with several technology pundits. An interesting area for future research would be to investigate the significance of these shallow sharing patterns. Are they due to the fact that almost everyone interested in the original post will likely read it at the source? Or perhaps fans overestimate the actual interest of the post when they reshare it?

4.2 Distributed shallow posting

Some URLs seemed to show highly distributed, shallow sharing. That is, there were many initiators of reshare trees, none of which dominate the graph. For example, the image in Figure 6, left, shows the reshare graph for a YouTube video (a viral hit in which a man shoots a gun at his daughter’s laptop in a bid to teach her online etiquette). The origin of such evenly distributed sharing would be an interesting subject for future research. One possibility is that no “celebrity” wanted to endorse the video’s message. Another possibility is that a celebrity shared the video on another social network or in the mainstream media, and that this pattern reflects the same dynamics as the celebrity pattern of 4.1, but with the celebrity invisible.

4.3 Heterogeneous sharing

Of course, not all posts rely on celebrities to spread. Figure 7, left, shows the Ripples for a classic “viral” YouTube video (“Dollar Shave Club – Our Blades are F**king Great”). The first image shows the full forest of shares, which includes more intricate reshare patterns in bigger circles. Zooming in on the circle indicated by the red arrow

(Figure 7, right) reveals further intricate sharing, culminating in a chain that is eight people deep.

This network combines multiple types of sharing. In addition to deep chains, the forest contains many small, shallow share trees. There are also a few people who were very widely reshared; several of these are celebrities within the Google+ network. One might speculate that this heterogeneous collection indicates content that has an unusually broad appeal. Indeed, the Dollar Shave Club video received over three million views, and its rapid spread became the subject of articles in Time Magazine, BusinessWeek, and the Los Angeles Times. An interesting question for future research would be to see if a high level of heterogeneity, detected early after an initial post, could predict future “viral” spread of the post.

Another example of heterogeneous sharing comes from the Ripple of a petition to the White House requiring free access to taxpayer-funded research. The petition URL received around 500 public shares on Google+ (Figure 8), including posts from a number of influential people. As a result the Ripple contains several broad and deep sharing circles. For instance, the Open Science Federation was a strong influencer, as was Tim O’Reilly, a well-known publisher. Separately, writers and scholars—American, Danish and Australian—started their own individual threads. The Ripples visualization shows the diversity of these communities.

An interesting temporal pattern, possibly a reflection of this diversity, is there is no one overwhelming “spike” in the graph of shares over time. Whereas celebrity posts tend to see an initial spike and rapidly fizzle out (e.g. Figure 6, right), here the activity shows multiple peaks over time.

5. REACTIONS TO RIPPLES

A natural way to gauge the efficacy of Ripples was to study reactions in its native environment, the online world. In particular, we wished to investigate two general questions. First, what is the range of usage of Ripples? That is, can we begin to catalogue different ways—common or not—that users have engaged with the visualization in the real world? Second, is there any particular set of people that consistently derives utility from Ripples?

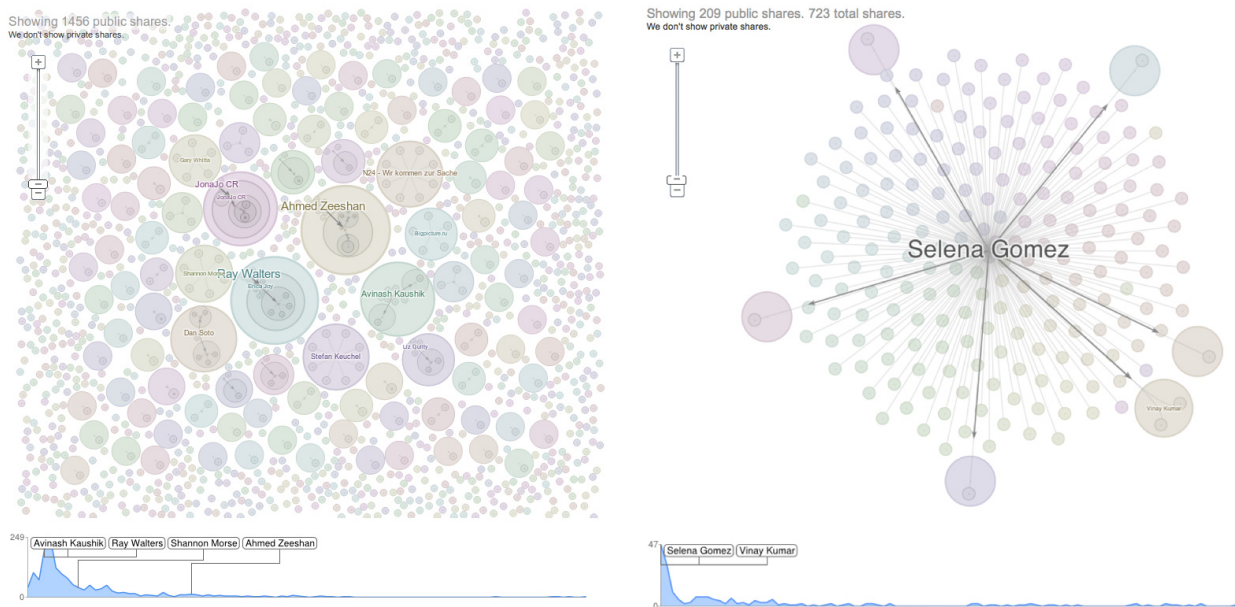


Figure 6. Left: Broad “forest” with deep sharing trees. Right: Celebrity pattern, where sharing action is broad and shallow.

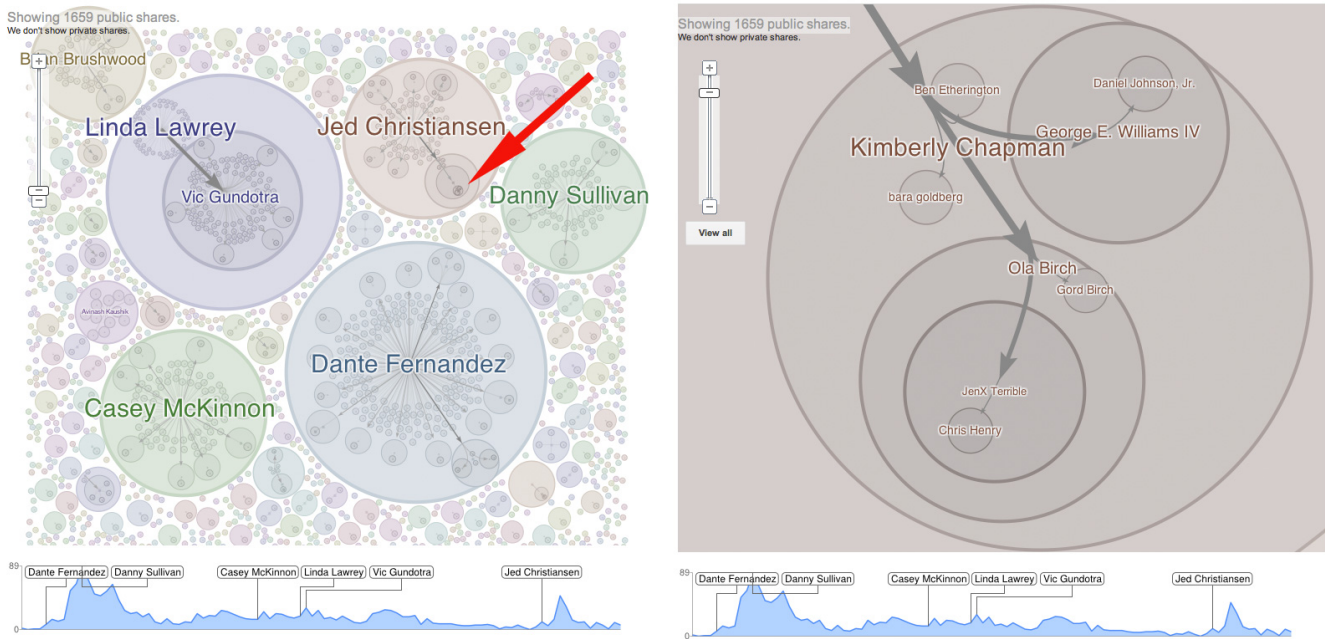


Figure 7. Left: complex share forest of a classic “viral” YouTube video (“Dollar Shave Club – Our Blades are Fking Great”). Right: zooming in on the circle indicated by the red arrow reveals a deep branching structure.**

5.1 Data collection

We collected two different samples of data on user reactions, one related to immediate post-launch commentary, and the other looking at later reactions. As one might expect, waves of reactions followed the initial launch of Ripples, as well as the introduction of new features. (Days after the V2 launch, for example, new tweets about Ripples were arriving at the rate of once every 10 seconds.) During the two post-launch periods we looked at as many Ripples-related blogs, articles, and social network posts as we could—sampling on an ad hoc basis—and noted the different ways people reacted to and used Ripples. We call this set of data the “Launch Sample.”

These immediate reactions provided useful information on the range of usage scenarios, but they are obviously susceptible to novelty effects and do not provide evidence that the visualization provided sustained utility. To investigate further, a year after the Ripples launch we collected a second sample by performing a Google search for “google plus ripples” and examining the first 100 resulting links. Of these, 36 led to web pages or social media posts that were written in English and provided detailed descriptions of Ripples usage (rather than, for example, simply a link or quote from another site). Within this subset, 17 were written more than a month after the launch of a new feature. We call this subset the “Year Sample.”

These online reactions provide an initial probe into how people interact with Ripples in real-world situations. There are, of course, many possible sources of error. To name a few, people who post online are likely to have unusually extreme views; in some situations, people may be more likely to post positive opinions than negative ones. (Then again, online users are traditionally not shy about criticizing products they do not like.) There may also be important use cases and opinions that do not make it online. Nonetheless, the broad themes we uncovered seem like a first step

toward understanding the effects of Ripples, and may be a useful guide for future studies of the system.

Usage Scenarios in the Launch Sample

As one would expect, after the initial launch of Ripples there was a wave of reaction on Google+ itself. Comments on Twitter, for instance, were generally positive or interested, ranging from *Google, I’m so glad we are friends. You make such awesome*

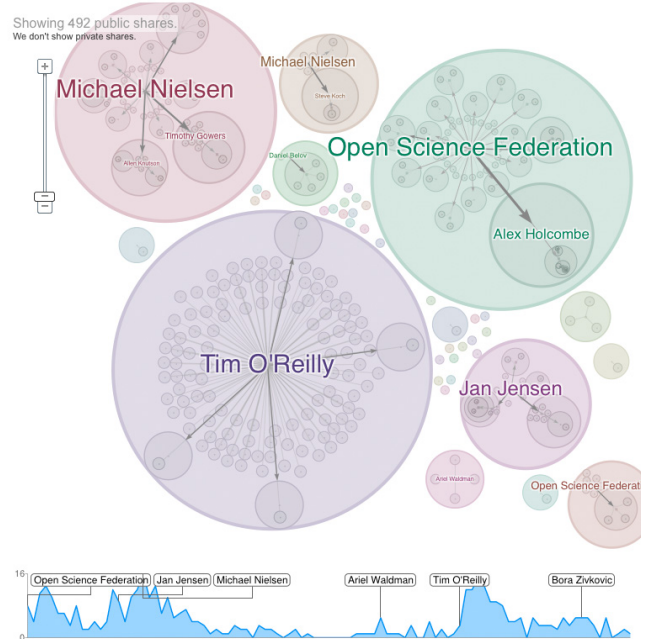


Figure 8. Ripple of a petition to the White House requiring free access to taxpayer-funded research.

"tools for data scientists" to "Google+ Ripples are slicker than snor". We saw people use the system itself to understand and play with the new feature. For example, several groups of people competed to form the largest share trees. Eventually one Ripple created purely for this purpose grew to more than 4,000 reshares. This playful activity, which as a by-product served to introduce many people to the tool, is reminiscent of some of the games played by users of the Many Eyes system [21].

Users also speculated about the ultimate utility of Ripples. Here the reaction was mixed. A good example is the set of comments in a post from Guy Kawasaki, which asked, "Other than being interesting, are you finding Ripples actionable?" Some answers were negative ("No" and "they are neat to look at but I haven't found what they are useful for.") At the other end of the spectrum, one person said, "absolutely actionable," with another providing details: "definitely actionable... I went to a profile of someone of influence in my space, checked out a popular post (one that was public), checked the Rippled [sic] and I was able to see who shared and re-shared (a/k/a the "ripple") and circled up the "influential" people down the stream."

Social media reactions to the launch of Ripples were not confined to Google+. A significant amount of commentary came on Twitter, as well as on blogs. Again, we saw broadly positive reactions, such as "AMAZEBALLS" and "Beautiful. And it just works." Several tweets also gave opinions about potential usage, by end users seeking to understand their online world ("Now I get to be the social webthorologist and peek at the webs we weave") and by people who analyze social media professionally ("They've given brands a real-time analytic tool to see how users engage with their content.") A few users did express confusion with the interface, e.g.: "I'm not going to lie, it took me a while to get my head around Google+ Ripples"

In some cases, people used Ripples to find new individuals to follow. One user, whose post saw over five thousand shares, pointed out that "you can actually use Ripples to better your experience on Google+ by finding the 'like minded' users who will enhance your engagement." Another person used Ripples to "return favors" by thanking others for sharing their content. As one blogger put it: "Should you be so fortunate as to have one of your Google+ posts be reshared enough for it to generate a Ripple, pay attention to the list of people who have publicly shared it. Take a moment to click the "view post" link after each one and leave a brief "Thanks for re-sharing this!" comment..."

Overall, we saw three main themes in the Launch Sample of reactions. First, "coolness": Many posts described Ripples as cool or innovative. A second reaction—occasionally from the same people—was skepticism about the utility of the visualization. A third consistent theme, utility, usually came from people who identified as professionals in the field of social media marketing; they focused less on "coolness" and aesthetics, and much more on the system's direct practical impact on their work.

5.2 Social Media Analytics: A Common Use Case

Subsequent online reactions provided evidence that Ripples found serious usage in the community of people who analyze social media as part of their business practice. The Year Sample was overwhelmingly made up of posts relating to business uses of Ripples. Of the 17 posts that were made more than a month after launch, 15 described business use cases, all in a positive light. For example, several months after launch, a Guardian article called

Ripples "a stunning visual representation of who within your connections are the most influential, allowing you to reach out to those contacts with targeted messaging." [23]. A separate article [4] illustrated a complex reaction to the visualization's aesthetics: "The technology looks fancy, but it can be quite practical for those really looking to analyze their Google+ presence." Interestingly, in this article the "fanciness" of the aesthetics was perceived as a potential negative.

The online reactions cited above give a general sense of excitement over the potential for Ripples as an analysis tool, but how was it actually used in practice? While much analysis happens out of the public eye, social media marketers—perhaps not surprisingly—sometimes work in public as well. To give a sense of how Ripples is used, we describe one blog entry in detail, Tim Moore Online [13]. Moore, according to the blog, decided to "run a test" whose "purpose was simple: To compare the engagement levels by users on Google+ and Facebook."

Moore posted the same item to both Google+ and Facebook. He began by looking at engagement metrics, but then to get more detail on how his Google+ post spread, he looked at several views of a Ripples diagram. In each, he annotated his findings. Typical statements were "Notice Sandra Parlow, Arleen Boyd, Jed Kim and CathiBea Stevenson stand out as the key amplifiers..." and "Drilling deeper, you'll see Sandra Parlow as the key amplifier, with 30,108 who have her in circles."

The visualization of clear utility to this analyst; as he put it, "Using Google+ Ripples, was helpful in analyzing 'how' the public posts migrate through the channel". In addition, the images were a key part of his narrative explaining his results.

5.3 Summary of reactions

Based on the public reaction, it seems that there are two sets of users who express interest in Ripples. First, in the immediate post-launch periods, a highly visible set of people seemed to appreciate the visualization on the basis of aesthetics and novelty. A second more focused group is made of people who routinely analyze social media, typically for purposes of marketing and publicity. Within this group, we saw strong public enthusiasm, along with explanations (blog posts, YouTube videos, etc.) of how they used the system.

Encouragingly for the visualization technique itself, we did not uncover many comments complaining about understanding the diagrams. We saw discussion about the utility of the system, but few negative comments focused on the legibility of the diagrams themselves. Of course, this result could be biased by the self-selected nature of the people who wrote about Ripples; an interesting area for future work would be to verify the legibility of Ripples in a more controlled setting.

6. CONCLUSION AND FUTURE WORK

Google+ Ripples is a visualization of information flow on a social network. Unlike other systems with similar goals, Ripples is a native visualization—that is, it exists as part of the system it visualizes, so that users can access it immediately and seamlessly.

In addition to this unusual position in the user experience, Ripples illustrates a new visualization technique, a hybrid of a circular treemap and a traditional node-and-link tree diagram. This "balloon treemap" makes efficient use of space, while providing an easily understandable metaphor for the diffusion of information.

User reaction to the visualization has been generally positive, with a strong reaction on social media sites. We have seen users invest significant effort in the tool, making many dozens of videos showing it in use, and writing detailed blog entries describing how analysts can gain benefits from it.

We have also seen that the tool has uncovered interesting patterns of information flow. Indeed, one important area for future work would be to use this new visualization technique, likely in conjunction with statistical analysis, to create models for information flow. In particular, it would be good to add quantitative analysis to our informal observations of different patterns. It is possible that these patterns may show how share tree topology can help predict virality of a post, or help classify content.

There is also room for additional research on the visualization technique. A significant amount of work and design went into creating a visualization technique that would scale to large share meanwhile, smaller share trees were sufficiently legible.

Nonetheless, for relatively small trees (say, under 20 shares) the visualization doesn't seem optimal. Indeed, for these situations some of our early designs, which showed very simple traditional tree structures, probably would have been clearer and more pleasing. One could imagine setting a threshold size, below which the simpler visualization is shown, but this could be confusing to users. It would be interesting to develop techniques that would "scale down" smoothly—that is, that would allow the visualization to vary seamlessly between the circular treemap view at large scale, and a traditional tree view at a small one.

Creating a native visualization involves a significant engineering effort. Although there is an obvious advantage in having direct access to data, integrating with an existing system adds a huge layer of complexity to a visualization project. Nonetheless, our experience suggests that creating a native visualization is worthwhile, and that users will welcome even an elaborate visualization if it appears in a familiar, convenient context. Our hope is that in other systems visualization can make a move from an "add-on" to an intrinsic part of the system.

7. ACKNOWLEDGMENTS

Many people contributed to the development of Ripples. Rick Borovoy led product management, along with Julie Farago. Jonathan Terleski provided critical design advice. Bill Strathearn provided technical assistance along with early inspiration. Marshall Gillson created the live version of Ripples. The visualization's smooth transitions owe much to Keith Ito. Alfred Spector's leadership was essential to the project. Many others on the Google+ engineering team helped bring the technology to launch.

8. REFERENCES

- [1] B. Bederson and J. Hollan, 1994. Pad++: a zooming graphical interface for exploring alternate interface physics. UIST 1994.
- [2] R. Boardman, 2000. "Bubble Trees: The Visualization of Hierarchical Information Structures". Proceedings of CHI 2000.
- [3] M. Bostock, V. Ogievetsky, J. Heer, 2011. D3: Data-Driven Documents. IEEE VisWeek 2011.
- [4] A. DiSilvestro, 2012. "google+ Ripples Explained," Search Engine Journal. <http://www.searchenginejournal.com/google-ripples-explained/48275/> Downloaded Nov. 11, 2012.
- [5] J. Donath, K. Karahalios, F. Viégas, 2002. Visualizing Conversation. Hawaii International Conference on Systems Science.
- [6] K. Etemad, S. Carpendale, 2009. Shamshtrees: Providing hierarchical context for nodes of interest. Proceedings of Bridges 2009, 293-300.
- [7] Facebook Friend Wheel. <http://thomas-fletcher.com/friendwheel/>
- [8] I. Herman, G. Melancon, M. S. Marshall 2000. "Graph visualization and navigation in information visualization: A survey." IEEE TVCG Jan-Mar. 2000.
- [9] Kleiberg, E., van de Wetering, H., J. van Wijk, J., 2001. "Botanical Visualization of Huge Hierarchies," IEEE InfoVis '01.
- [10] J. Leskovec, L. Backstrom, J. Kleinberg. Meme-tracking and the dynamics of the news cycle. Proc. 15th ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining, 2009.
- [11] J. Thorp and M. Hansen, 2011. Cascade. As described by Suzanne Labarre, "Infographic of the Day: 3D Model Unlocks Secrets of the Twitterverse", Fast Company Online. <http://www.fastcodesign.com/1663694/infographic-of-the-day-3d-model-unlocks-secrets-of-twitterverse-video>
- [12] A. Moore, 1991. "An introductory tutorial on kd-trees," tech. report Technical Report No. 209, Computer Laboratory, University of Cambridge.
- [13] T. Moore. Tim Moore Online, downloaded Feb. 26, 2012. <http://timmooreonline.blogspot.com/2012/01/case-study-google-vs-facebook.html>
- [14] P. Neumann, S. Carpendale, A. Agarawala. 2006 "Phyllotrees: Phyllotactic patterns for tree layout." EUROVIS-Eurographics/IEEE VGTC Symposium on Visualization, 59-66
- [15] Q. Nguyen and M. Huang, 2005. "EncCon: an approach to constructing interactive visualization of large hierarchical data," Information Visualization 2005 4, pp. 1-21.
- [16] SocialFlow, 2011. <http://socialflow.com>
- [17] B. Shneiderman, M. Wattenberg 2001. "Ordered Treemap Layouts" IEEE Infovis 2001.
- [18] A. Vande Moore (downloaded 2012) Information Aesthetics Blog. <http://infosthetics.com/>
- [19] F. Viégas, M. Wattenberg, J. Feinberg, 2009. "Participatory Visualization with Wordle". IEEE Infovis, 2009.
- [20] F. Viégas, M. Wattenberg, F. van Ham, J. Kriss, M. McKeon, 2007. "Many Eyes: A Site for Visualization at Internet Scale". IEEE Infovis, 2007.
- [21] F. Viégas, M. Wattenberg, M. McKeon, F. van Ham, J. Kriss, 2008. "Harry Potter and the Meat-Filled Freezer: A Case Study of Spontaneous Usage of Visualization Tools." Hawaii Intl Conf. Sys. Sci. (HICSS), 2008.
- [22] YouTube Insights. http://www.youtube.com/t/advertising_insight
- [23] <http://www.guardian.co.uk/media-network/media-network-blog/2012/feb/15/advantages-google-digital-marketing?newsfeed=true>
- [24] W. Wang, H. Wang, G. Dai, H. Wang, 2006. "Visualization of Large Hierarchical Data by Circle Packing," Proceeding of CHI 2006.
- [25] M. Wattenberg, 2005. "A Note on Space-Filling Visualizations and Space-Filling Curves". IEEE InfoVis 2005.
- [26] K. Wetzel. "pebbles." <http://lip.sourceforge.net/ctreemap.html>
- [27] R. Xiong and E. Brittain, 1999. LiveWeb: Visualizing live user activities on the Web SIGGRAPH.