Empirical Study of Focus-Plus-Context and Aggregation Techniques for the Visualization of Streaming Data

Eric D. Ragan
University of Florida
eragan@ufl.edu

Andrew S. Stamps
Mississippi State University
andrewstamps@gmail.com

John R. Goodall
Oak Ridge National Laboratory
jgoodall@ornl.gov

ABSTRACT

Analysis of streaming data often involves both real-time monitoring of incoming data as well as contextual awareness of data history. A focus-plus-context approach can support both goals, with variable levels of visual aggregation making it possible to provide a high level of detail for incoming and recent data while providing contextual information about recent history. Visual aggregation reduces data resolution in order to show the context of data over large periods of time within a limited display space. With a controlled experiment, we evaluated the effectiveness of different types of aggregation for four types of stream-analysis tasks. Overall, the results show that a focus-plus-context design has little negative impact on the ability to successfully monitor and analyze streaming data, making it possible to show longer periods of time than other approaches. However, visual aggregation can be problematic for trend recognition tasks. This research demonstrates how the effectiveness of the visualization depends on the specifics of the analysis task.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in visualization**.

KEYWORDS

Visualization, human-computer interaction, streaming data, information visualization

ACM Reference Format:
Eric D. Ragan, Andrew S. Stamps, and John R. Goodall. 2020. Empirical Study of Focus-Plus-Context and Aggregation Techniques for the Visualization of Streaming Data. In *International Conference on Advanced Visual Interfaces (AVI '20), September 28-October 2, 2020, Salerno, Italy.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3399715.3399837

1 INTRODUCTION

Across a wide variety of domains, it is becoming possible—and necessary—to monitor and respond to events in real-time. Real-time fraud detection of financial accounts, intrusion detection in cyber security, monitoring of critical sensors within critical infrastructure, and monitoring social media for intelligence analysis are just a few examples of domains that require the timely analysis of complex data. Streaming solutions are becoming popular for processing high-volume, high-velocity data (e.g., [20, 26, 29]). Streaming analytics are attractive to reduce latency in identifying insights while reducing the volume of data that needs to be stored.

In some cases, reacting to events in streaming data can be a completely autonomous process, but often human intervention is desirable or required [11]. For example, automated response to streaming data can lead to undesirable effects; in cyber security, for example, automated response can lead to denial-of-service attacks that can effectively shut down a server. Some actions are simply too important to take without some kind of human supervision, such as decisions to shut down energy systems or to divert defense operations. Other times, uncertainty exists in the data or analysis that makes automated response infeasible, such as automatically denying a customer's credit card payment because of an anomalous, but legitimate, charge. In such cases, humans need tools to rapidly make sense of the events in the data streams.

Visualization has been shown to be effective for supporting analysis of time-based data (e.g., [10, 15]), but streaming data introduces unique challenges for visualization and interaction [16]. Effective real-time visualization and analysis of streaming data requires methods that allow analysts to examine dynamic data sets and monitor new data. Such methods should emphasize recent events while providing contextual awareness over longer time periods. One approach for supporting such contextual awareness is visual aggregation, which reduces data resolution in order to be able to show large periods of time within a limited display space. A focus-plus-context approach [4] varies levels of aggregation: more recent data can be shown in higher detail (i.e., for focus) while older data can be aggregated (i.e., for context). Aggregation allows visualizations to scale to cover large volumes of data over extended periods of time, but empirical evidence of how perceptual design options affect human interpretation of data is lacking.

This paper presents a controlled experiment testing how different visual aggregation designs can influence stream analysis for different visual analysis tasks. The study serves as a foundation for visualization research of the perception and understanding of streaming data, and it demonstrates how the effectiveness of focus-plus-context methods for working with streaming data can depend on specifics of the analysis task.

2 RELATED WORK

Many researchers have studied visualizations for temporal data (e.g., [1, 3, 14, 27]). As a type of temporal data, streaming data is dynamic and continually added over time. For some types of time-based data analysis, small-multiple visualizations that show temporal change through a series of "snapshots" are sometimes preferred over animated designs (e.g., [22]), but which approach is best may depend on whether the goal is to monitor for changes or to compare

AVI '20, September 28-October 2, 2020, Salerno, Italy

ACM ISBN 978-1-4503-7535-1/20/09…$15.00
https://doi.org/10.1145/3399715.3399837

specific temporal snapshots [24]. Some studies show evidence that animated visualizations are superior for detecting temporal changes and patterns [12, 24, 25]. In the case of streaming data, analysis often includes monitoring the continually updating the state of the data [11, 21]. Thus, for stream visualization, animated designs prioritize updating incoming data as quickly as possible when it arrives and helping to show changes as they happen.

Relevant to understanding animated designs, Cottam et al. [8] presented a taxonomy (using Bertin's spatial and retinal factors; see [5]) to describe different ways that animations and data presentations can be designed in dynamic visualizations. Specific to stream analysis, analysts and operators aim to respond to the most recent events quickly while preserving awareness of past events. Operators maintain this situation awareness [9] in order to put new events into context and to look for longer-term patterns.

To this end, many stream visualizations include methods for temporal aggregation and try to address the need to support both monitoring of incoming data as well as contextualizing a history of older data [21]. Using computing performance data as an example, Hao et al. [13] presented a design that balanced views from two periods of time. The design also provided data views with variable temporal resolutions to support temporal aggregation. Another tool, *StreamSqueeze*, presents events in a color-coded list with screen space prioritized by recency [18]. The list view shows more details for incoming events, and older events get reduced levels of detail.

Other examples are common for network monitoring and the cyber security domain (e.g., [6, 19]). *LiveRAC* [19] and CLIQUE [6] both use small-multiple views of line charts to show current and past network information, and both provide aggregate views by compressing chart sizes and using color coding to indicate density of combined items. *Traffic Circle* uses a circular interface that allows temporal aggregation of network flow data by expanding or zooming into periods of time [6]. Such examples demonstrate the common design approach of prioritizing detail in more recent streaming data while reducing granularity for older events through visual or semantic compression. While this basic design is employed in many applications (e.g., [6, 18, 19]), empirical knowledge of how such approaches impact fundamental stream analysis tasks is lacking, thus motivating our experiment.

3 EXPERIMENT

We conducted a controlled experiment to empirically study the effectiveness and limitations of focus-plus-context designs using visual and temporal aggregation for different stream-analysis tasks. While temporal aggregation and focus-plus-context designs offer the advantage of scalability with regard to the amount of time shown within limited screen space, it raises questions about how increased levels of aggregation affect performance on common analysis and interpretation tasks.

To facilitate an empirical study of a generalized version of the temporal aggregation approach used by many stream visualizations, the study tests the aggregation concept applied to a scatterplot representation. Scatterplots are commonly used and easily understood, and using one axis to represent time makes a straightforward design for temporal visualization. Visually encoding data values as positions provide high accuracy for graphical perception [7], and having an individual mark for each data point works well for displaying each new data point as discrete events arriving in a streaming scenario.

Because we wanted to explore how performance might be enhanced or constrained by display size in streaming tasks, we include two differently-sized scatterplots for baseline comparisons. We also tested a scatterplot with a power-scale on the horizontal axis. Finally, we included a hybrid heatmap/scatterplot design as an example of a focus-plus-context plot in alignment with existing stream visualizations (e.g., [6, 18, 19]).

3.1 Visualization Conditions

For the primary independent variable of our experiment, we compared four visualization designs (shown in Figure 1) for visual analysis of streaming data. The designs present different approaches to visual aggregation of scatterplot data. Each design in Figure 1 shows the same data at the same starting time. All designs were animated so that data points moved from the right to the left in accordance with the passage of time, and the horizontal axis updated accordingly. None of the visualization types in the experiment were interactive.

The *short linear-scale scatterplot* design (Figure 1A) was an animated version of a standard linear scatterplot. This representation provides a highly detailed view of each data point but has limited scalability due to the one-to-one mapping between time and location. With a linear-scale x-axis, the length of the visualization would have to grow linearly to match the length of the period of time. The size of the short linear-scale scatterplot was restricted to show 40 seconds of data at once, which was chosen as a baseline to restrict the width to match that of other conditions with an equal amount of screen space.

The *long linear-scale scatterplot* (Figure 1B) followed the same design as the short linear-scale scatterplot but was long enough to show 85 seconds of data. As the only visualization type that was wider than the three others, the long linear-scale scatterplot was included to demonstrate how much screen space would be needed to cover the same amount of time as the types with aggregation with a linear scale.

The *power-scale scatterplot* (Figure 1C) was similar to the design of the linear scatterplots, but the scale for the horizontal axis followed a power scale instead of a linear scale. Using a power scale allows a larger period of time to be shown without requiring a linear increase in chart width. This type of representation could support a broader temporal view at the cost of reduced spatial resolution for older data points. This design supports of one the primary goals for visualizing streaming data: emphasizing the most recent values and changes. The size of the power-scale scatterplot was the same width as the short scatterplot, and it showed the same time period (85 seconds) as the long scatterplot.

The *focus-plus-context plot* (Figure 1D) uses a scatterplot of individual events for the most recent data, and older data are aggregated into a heatmap to increase scalability. For the experiment, the heatmap's bins used glyph size to encode the item count in each bin. The scale of the horizontal axis is non-linear in the heatmap, so events that are oldest are aggregated the most. From right to left, the each column of bins covers a period of 5, 10, 20, and 40

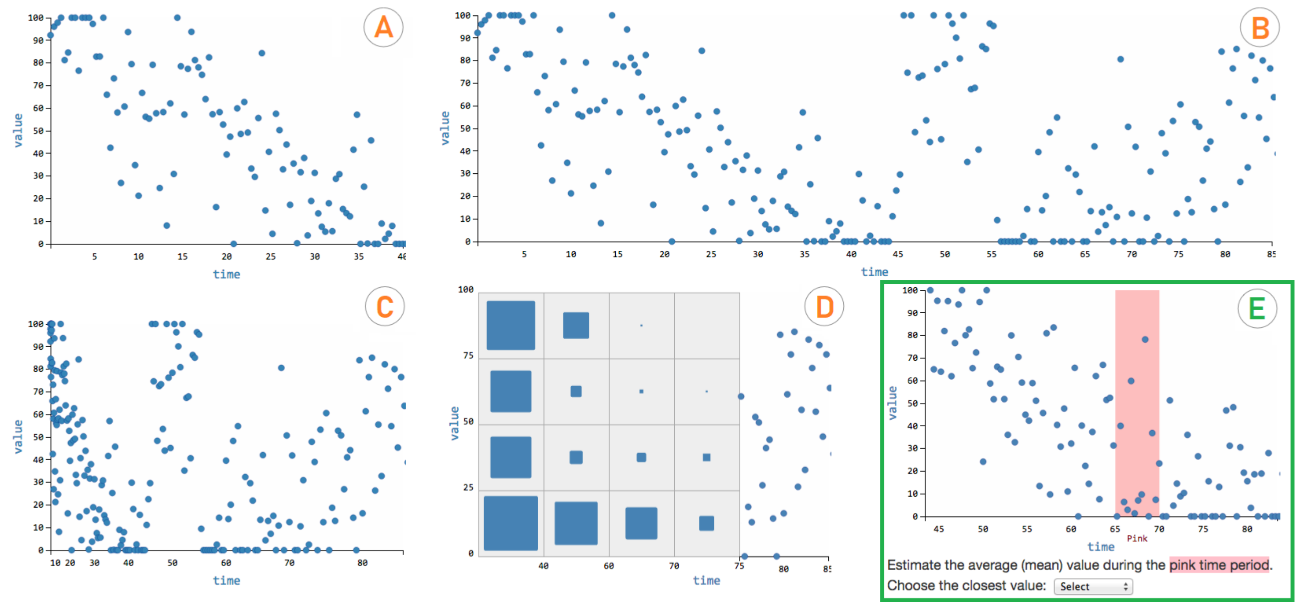

Figure 1: Subfigures A–D show the four visualization types tested: (A) short linear scatterplot, (B) long linear scale scatterplot, (C) power scale scatterplot, and (D) focus-plus-context plot. Image E shows an xample of an average estimation task with a small linear-scale scatterplot.

seconds, respectively. The practical benefit of this design is that aggregating older data into a heatmap allows the visualization to show longer periods of time. Because aggregation occurs in both the horizontal and vertical dimensions, spatial resolution of older data is lost; however, by adjusting the value ranges of the bins, it could be possible for the bins to provide a simple, easy-to-interpret overview of older events. For controlled comparison with the other visualization types in this study, the size of the focus-plus-context implementation was kept the same as the short linear-scale scatterplot and power-scale scatterplot. This visualization also showed 85 seconds—the same amount of time as the power-scale scatterplot and the long linear-scale scatterplot.

3.2 Analysis Tasks

Different types of visualization designs can be better suited for different types of analysis tasks. Therefore, we also studied different types of analysis tasks in our experiment. We rely on a subset of common analysis tasks and user objectives characterized by others for visual data analysis [2, 23, 28] and temporal data analysis [1, 17]. The study included four distinct types of tasks: *average estimation*, *comparison of averages*, *outlier comparison*, and *trend recognition*. In addition, two versions of each type of analysis task were created—*short time scope* and *long time scope* versions—with the difference being the length of time analyzed in the task.

The *average estimation* task asked to estimate the average value of data in a highlighted region of the visualization (see Figure ??). Participants could select the closest integer value from a dropdown list where values range from 0-100, the minimum and maximum values of the vertical axis. The highlighted period was 5 seconds in the *short time scope* version of this task and 15 seconds in the *long time scope* version.

In the *comparison of averages* task, the plot showed five highlighted periods of time. Participants had to pick which of the later four periods has an average value that is closest to that of the first highlighted period. In the *short time scope* version of the *comparison of averages* task, the highlighted periods were each five seconds long, and the periods were separated by five seconds. In the *long time scope* version, the highlighted periods were also five seconds long, but they were separated by 15 seconds.

In the *outlier comparison* task, the plot showed four highlighted periods. The participant had to choose the highlighted period with the greatest outlier from four options. The highlighted periods were each five seconds long for both the *short time scope* and *long time scope* versions of the *outlier comparison* task. The difference was that the highlighted periods were separated by five seconds for the *short time scope* and 15 seconds for the *long time scope* version.

The purpose of the *trend recognition* task was to interpret the general shape or trend of the data stream. This task presented four multiple choice icons that showed general shape and fluctuations of data values over the entire period of time for the trial. Participants were asked to select the icon with the shape that most closely matches the data. For this task, the *short time scope* version covered 50 seconds, while the *long time scope* version covered 100 seconds.

3.3 Experimental Design

The experiment followed a 4x4x2 mixed design. The independent variables were *visualization type*, *analysis task type*, and *time scope*. The four visualization types were varied between subjects, while the four analysis task types and the two time scopes were both controlled within subjects. The order in which the analysis task types were presented was determined by a Latin square design with four orderings possible. Trials used synthesized data sets comprised

	Average	Comparison	Outlier	Trend
$\chi^2(3)$	5.80	5.75	0.09	9.08
p	0.12	0.12	0.99	* 0.03

Table 1: Kruskal-Wallis omnibus tests for each task type on the effects of visualization type on correct responses.

of x and y pairs. New data was streamed at a constant rate of one new event every 200 milliseconds.

The dependent variable was task correctness, which as binary for the multiple-choice responses. Since the average estimation tasks involved a numerical answer, the trial was judged to be correct if the selected answer was within five units from the correct answer.

The experiment was conducted as an online study taking approximately 20 minutes. For each condition, participants were first shown instructions and a practice trial before the trials for the four main analysis tasks. Eighty-one participants completed the entire study procedure, and data from 76 participants (21 female) were included in the results analysis after quality and outlier filtering.

4 RESULTS

To provide an approximate analysis of the effects of visualization type on task performance, we summed the number of correct responses for each task type to derive a numerical measure of accuracy. We performed a Kruskal-Wallis rank sum test for each task type to task for effect of visualization type on correct responses. Test outputs are shown in Table 1. The only significant main effect was for the trend recognition task; the test yielded $\chi^2(3) = 9.08$ and p = 0.03. A post-hoc Dunn test with Holm-Sidák correction for the *trend recognition* task found significant pairwise differences between the *short linear-scale scatterplot* and the *focus-plus-context plot* (p = 0.02) and between the *power-scale scatterplot* and the *bin-focus-plus-context plot* (p = 0.04). In addition, the difference was near significant between the long *linear-scale scatterplot* and the *focus-plus-context plot* (p = 0.08). Overall, these results indicate that performance was significantly negatively influenced for the trend recognition task with the focus-plus-context plot. No other differences were detected for the other tasks.

To better understand effects due to differences among the task types, we also used Pearson's chi-square tests to detect differences in the frequency of correct responses across visualizations types for each trial (i.e., each combination of task type and time scope). The test for the *long time scope* version of the *trend recognition* task detected a significant effect with $\chi^2(3) = 9.93$ and p = 0.02. Inspection of frequencies by visualization types reveals that the frequency of correct responses was clearly lower for the *focus-plus-context* condition (0.53) than for the other visualizations for this task (ranging 0.79 to 0.89).

5 DISCUSSION AND CONCLUSION

Temporal aggregation is appealing due to its ability to provide contextual history over extended periods of time without linearly increasing the amount of screen space. When supporting streaming data, greater aggregation can provide analysts with a longer temporal window for more context to help interpret the most recent events. The focus-plus-context plot provided the highest level of visual aggregation of the visualization alternatives we tested. Results indicate that the only significant penalty of the increased aggregation in the focus-plus-context plot was for the trend recognition task with the long time scope.

The lack of detected differences for the other task types (i.e., average estimation, average comparison, and outlier comparison) suggests that the visualization types were similarly effective for determining approximate values and comparing values in different periods of time. The analysis does provide evidence that high levels of aggregation can make it difficult to recognize temporal patterns in data. The results are promising for the use of the focus-plus-context and power-scale designs due to their use of aggregation to show contextual information in a relatively small visual area. An intuitive interpretation of the response accuracy results is that no analysis penalties were detected for the power-scale design, and performance with the focus-plus-context design only suffered for the long time-scope version of the trend recognition task.

The focus-plus-context scatterplot can emphasize attention to recent events while also providing context to older events. In some domains, such as cyber security, this contextual awareness is crucial to interpreting the meaning of the visualization, and more context is typically better. The poor performance of participants in trend detection tasks over longer periods of time may be mitigated by complementary visualizations. The ability to scale the focus-plus-context plot to larger time spans than non-aggregated approaches and the good performance by participants in nearly all of the task types is promising. In future research, we will investigate alternative visual aggregation approaches that enable visual scalability and emphasize recent events while preserving context.

With data collection at an all time high, an increasing effort will be placed on both static and real-time data analysis. This trend is currently visible in research, government, and industry across a variety of domains. Though most data analysis tasks involve systems and algorithms to collect, aggregate, and filter data, all of this work is done to support human efforts and curiosity about data; thus, research like our study of stream visualization is important for providing guidelines of the best way to present data for human audiences and analysts.

ACKNOWLEDGMENTS

This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan http://energy.gov/downloads/doe-public-access-plan. This research is supported in part by the DARPA XAI program under Grant N66001-17-2-4031.

REFERENCES

[1] Wolfgang Aigner, Alessio Bertone, Silvia Miksch, Christian Tominski, and Heidrun Schumann. 2007. Towards a conceptual framework for visual analytics

of time and time-oriented data. In *Simulation Conference, 2007 Winter*. IEEE, 721–729.
[2] Robert Amar, James Eagan, and John Stasko. 2005. Low-level components of analytic activity in information visualization. In *IEEE Symposium on Information Visualization*. IEEE, 111–117.
[3] Natalia Andrienko and Gennady Andrienko. 2006. *Exploratory analysis of spatial and temporal data: a systematic approach*. Springer Science & Business Media.
[4] Patrick Baudisch, Nathaniel Good, Victoria Bellotti, and Pamela Schraedley. 2002. Keeping things in context: a comparative evaluation of focus plus context screens, overviews, and zooming. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 259–266.
[5] JB Bertin. 1983. Semiology of Graphics Diagrams Networks Maps. Berg, WJ. *Trans., Madison, Wis.: U. of Wisconsin Press* (1983), 269–283.
[6] Daniel M Best, Shawn Bohn, Douglas Love, Adam Wynne, and William A Pike. 2010. Real-time visualization of network behaviors for situational awareness. In *Proceedings of the seventh international symposium on visualization for cyber security*. ACM, 79–90.
[7] William S Cleveland and Robert McGill. 1984. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American statistical association* 79, 387 (1984), 531–554.
[8] Joseph A Cottam, Andrew Lumsdaine, and Chris Weaver. 2012. Watch this: A taxonomy for dynamic data visualization. In *IEEE Conference on Visual Analytics Science and Technology (VAST)*. IEEE, 193–202.
[9] Mica R Endsley. 1995. Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 37, 1 (1995), 32–64.
[10] Fabian Fischer, Florian Mansmann, and Daniel A Keim. 2012. Real-time visual analytics for event data streams. In *Proceedings of the 27th Annual ACM Symposium on Applied Computing*. ACM, 801–806.
[11] John R Goodall, Eric D Ragan, Chad A Steed, Joel W Reed, G David Richardson, Kelly MT Huffer, Robert A Bridges, and Jason A Laska. 2018. Situ: Identifying and explaining suspicious behavior in networks. *IEEE transactions on visualization and computer graphics* 25, 1 (2018), 204–214.
[12] Amy L Griffin, Alan M MacEachren, Frank Hardisty, Erik Steiner, and Bonan Li. 2006. A comparison of animated maps with static small-multiple maps for visually identifying space-time clusters. *Annals of the Association of American Geographers* 96, 4 (2006), 740–753.
[13] Ming Hao, Daniel A Keim, Umeshwar Dayal, Daniela Oelke, and Chantal Tremblay. 2008. Density displays for data stream monitoring. In *Computer Graphics Forum*, Vol. 27. Wiley Online Library, 895–902.
[14] Susan Havre, Elizabeth Hetzler, Paul Whitney, and Lucy Nowell. 2002. Themeriver: Visualizing thematic changes in large document collections. *IEEE Transactions on Visualization and Computer Graphics* 8, 1 (2002), 9–20.
[15] Samuel Huron, Romain Vuillemot, and Jean-Daniel Fekete. 2013. Visual sedimentation. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013), 2446–2455.
[16] D.A. Keim, J. Kohlhammer, G. Ellis, and F. Mansmann. 2010. *Mastering the Information Age - Solving Problems with Visual Analytics*. Eurographics Association.
[17] Jessica Lin, Eamonn Keogh, Stefano Lonardi, Jeffrey P Lankford, and Donna M Nystrom. 2004. Visually mining and monitoring massive time series. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 460–469.
[18] Florian Mansmann, Milos Krstajic, Fabian Fischer, and Enrico Bertini. 2012. StreamSqueeze: a dynamic stream visualization for monitoring of event data. In *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 829404–829404.
[19] Peter McLachlan, Tamara Munzner, Eleftherios Koutsofios, and Stephen North. 2008. LiveRAC: interactive visual exploration of system management time-series data. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 1483–1492.
[20] Leonardo Neumeyer, Bruce Robbins, Anish Nair, and Anand Kesari. 2010. S4: Distributed stream computing platform. In *2010 IEEE International Conference on Data Mining Workshops (ICDMW)*. IEEE, 170–177.
[21] Eric D Ragan, Andrew Pachuilo, John Goodall, and Felipe Bacim. 2020. Preserving Contextual Awareness during Selection of Moving Targets in Animated Stream Visualizations. In *Proceedings of the 2020 International Conference on Advanced Visual Interfaces (AVI)*. 1–9.
[22] George Robertson, Roland Fernandez, Danyel Fisher, Bongshin Lee, and John Stasko. 2008. Effectiveness of animation in trend visualization. *IEEE Transactions on Visualization and Computer Graphics* 14, 6 (2008), 1325–1332.
[23] Robert E Roth. 2013. An empirically-derived taxonomy of interaction primitives for interactive cartography and geovisualization. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013), 2356–2365.
[24] Terry Slocum, Robert Sluter, Fritz Kessler, and Stephen Yoder. 2004. A qualitative evaluation of MapTime, a program for exploring spatiotemporal point data. *Cartographica: The International Journal for Geographic Information and Geovisualization* 39, 3 (2004), 43–68.
[25] Tatiana Tekušová and Jörn Kohlhammer. 2007. Applying animation to the visual analysis of financial time-dependent data. In *Information Visualization, 2007. IV'07. 11th International Conference*. IEEE, 101–108.
[26] Ankit Toshniwal, Siddarth Taneja, Amit Shukla, Karthik Ramasamy, Jignesh M Patel, Sanjeev Kulkarni, Jason Jackson, Krishna Gade, Maosong Fu, Jake Donham, et al. 2014. Storm@ twitter. In *Proceedings of the 2014 ACM SIGMOD international conference on Management of data*. ACM, 147–156.
[27] Jarke J Van Wijk and Edward R Van Selow. 1999. Cluster and calendar based visualization of time series data. In *IEEE Symposium on Information Visualization, 1999*. IEEE, 4–9.
[28] Stephen Wehrend and Clayton Lewis. 1990. A problem-oriented classification of visualization techniques. In *Proceedings of the 1st Conference on Visualization'90*. IEEE Computer Society Press, 139–143.
[29] Matei Zaharia, Tathagata Das, Haoyuan Li, Timothy Hunter, Scott Shenker, and Ion Stoica. 2013. Discretized streams: Fault-tolerant streaming computation at scale. In *Proceedings of the Twenty-Fourth ACM Symposium on Operating Systems Principles*. ACM, 423–438.