TimeSets: Temporal Sensemaking in Intelligence Analysis

Kai Xu[1], Saminu Salisu[2], Phong H. Nguyen[3], Rick Walker[4], BL William Wong[1],
Adrian Wagstaff[5], Graham Phillips[6], and Mike Biggs[6]

[1]Middlesex University, UK, [2]Vienna University of Technology, Austria, [3]Red Sift Ltd, UK,
[4]King, UK, [5]Hawkesford Consultancy Ltd, UK, [6]MASS Ltd, UK

ABSTRACT

TimeSets is a temporal data visualization technique designed to reveal insights into event sets, such as all the events linked to one person or organization. In this paper we describe two TimeSets-based visual analytics tools for intelligence analysis. In the first case, TimeSets is integrated with other visual analytics tools to support open-source intelligence analysis with Twitter data, particularly the challenge of finding the right questions to ask. The second case uses TimeSets in a participatory design process with analysts that aims to meet their requirements of uncertainty analysis involving fake news. Lessons learned are potentially beneficial to other application domains.

1 INTRODUCTION

Visualization techniques are effective at revealing temporal patterns [1], such as long-term trends and repeating occurrences, in collections of events. One common task in temporal analysis is to understand the development of narratives around an *entity*, such as person or organization, in the context of other entities they connect to. When visualized, such entities are often represented with color, shape, or glyph, which are not always easy to follow, especially when events involve multiple entities. There are a few set visualization techniques [2, 3, 4] that visually connect events of the same entity through lines or enclosing shapes. However, this is often achieved at the cost of visual clutter. A visualization component we have developed, TimeSets, was previously introduced to address this challenge by visually grouping events involving the same entity while minimizing the visual clutter (Figure 1). It has been shown to out-perform existing techniques mentioned earlier [5].

In this paper we describe the development of two visual analytics tools that use TimeSets. Both tools are designed to support temporal analysis in intelligence investigations, particularly Open-Source INTelligence analysis (OSINT). As far as the authors are aware, this is the first time that TimeSets has been used in software developed for such analysis. The first tool suite, SAVI (Social Analytics VIsualization), consists of three visual analysis tools to investigate Tweet-like data from multiple perspectives to support hypothesis generation. The second tool, TimeSets Uncertainty, extends TimeSets with uncertainty-related features that resulted from a participatory design process with intelligence analysts. While both tools are designed for intelligence analysis, we believe some of the design ideas and the design approach/process easily applies to other domains with similar analysis needs.

2 TIMESETS

Before going into the case studies, this section summarizes the design of TimeSets. More details can be found in our original paper [5].

Visual Representation of Events. An event is visualized as a line of text, starting with a visual glyph. The position of the glyph along the horizontal time axis marks the event time, either circles for time-point events (most cases) or horizontal bars for interval events (such as the "Libby's criminal trial" in the "Courts and Judges" set in the bottom right quadrant of Figure 1). To accommodate a large number of events, text can be trimmed (showing only the first few words) or multiple events can be merged into one *aggregated event* (example in section 4).

Visual Representation of Sets. A *set* contains all the events relating to a particular entity and has a unique color as shown in the legend (top-right corner of Figure 1). Sets are stacked vertically, and each set is further divided into a maximum of three layers: the top layer showing shared events with the upper set, the bottom layer showing shared events with the lower set, and the middle layer showing the remaining events in the set. In this paper, TimeSets uses color blending for the intersections of two sets, which is different from the color gradient method in the original TimeSets paper [5] and arguably better. Besides set background color, color-coded circles at the beginning of each event show all the entities involved, one for each entity. When two entities are not next to each other, an event is replicated in each set. When the mouse is over one such event, all its replications are highlighted for easy discovery.

Interactive Exploration. TimeSets supports standard panning and zooming. Zooming out compresses the time line and often results in more aggregated events men-

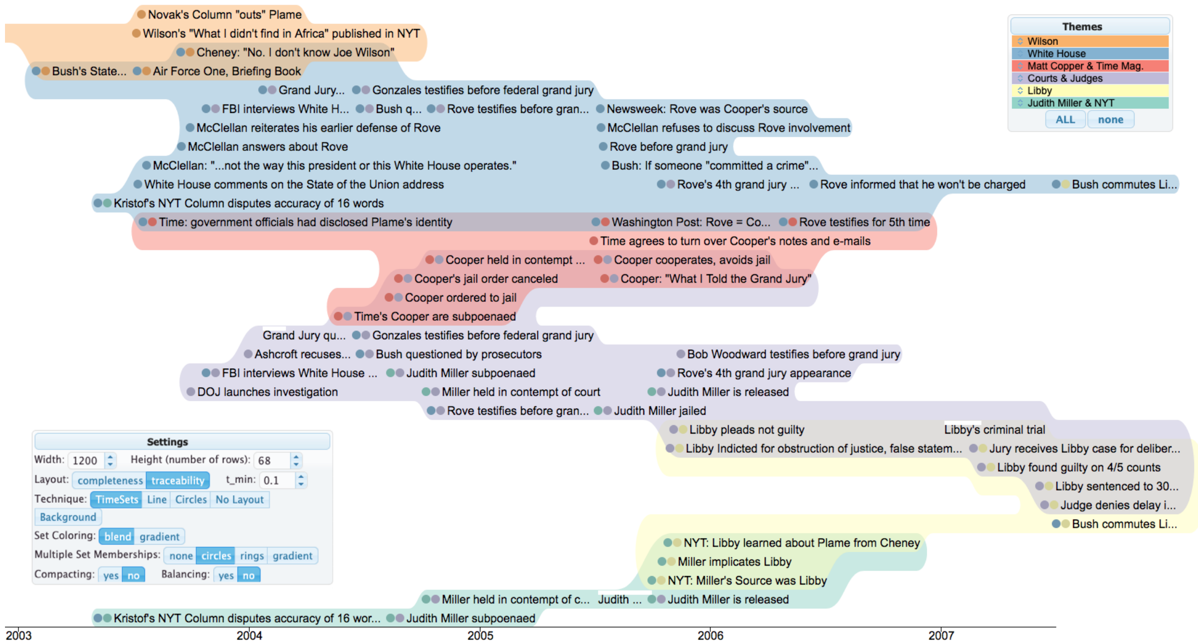


Figure 1: The CIA leak case event shown in TimeSets, with events belongs to the same entity grouped together.

tioned earlier. In addition, TimeSets provides interactive set filtering and reordering through the legend panel (top right corner of Figure 1): users can hide or show a set by clicking on its legend; dragging the set legend will change the vertical set order accordingly. The setting panel (bottom left corner of Figure 1) adjusts many other visual features of the TimeSets: such as canvas size, layout method, and event glyph. An interactive implementation of TimeSets is available online (`vis4sense.github.io/timesets/demo/`), showing the 'CIA leak case' dataset used in Figure 1.

3 Social Media Analysis

Social media analysis is an important part of intelligence analysis, and it is also known as OSINT or open-source intelligence analysis. Many visual analytics systems have been proposed to support the analysis of social media data [6], and our attempt focuses on the spatial-temporal aspect of understanding related messages over time, utilizing the unique capability that TimeSets brings.

3.1 Requirements and Design Goals

Our work started with an entry to the IEEE VAST Challenge 2014 (`www.vacommunity.org/VAST+Challenge+2014`) and won an award for "*Effective Support for Analytic Sensemaking*" for mini-challenge 3 [7]. Based on this work, we built a suite of visual analytics tools around TimeSets to support this type of social media analysis. We name it SAVI that stands for Social Analytics VIsualization. We will use the same challenge data and analysis questions to demonstrate SAVI, as it is a representative example of OSINT analysis.

The challenge is about piecing together information related to the disappearance of the executives of a fictitious energy company. The only data available is a collection of micro blogs (similar to Tweets) with the information about who (author), what (text content), and when (time stamp). A small percentage of messages have geographical locations, given as lat/long coordinates or street addresses. More details are available from the VAST challenge website.

The main challenges in such analysis include:

C1: Ill-defined analysis question. While the goal is clear, e.g., finding information related to the disappearance of company executives, how to achieve it through analytic means is not clear: where to start, what questions to ask, what analysis should be performed, in what order, etc.

C2: Deliberate hiding of information. Common in intelligence analysis, the adversary deliberately hides relevant information to avoid revealing themselves or being captured. This makes it not straightforward to discover salient information directly from data and requires more in-depth analysis.

C3: Large amount of data. The size of the data brings a few additional challenges: a) difficulty creating a meaningful overview that may help identify the starting point for analysis; b) the "needle in the hay stack" problem: only a very small percentage of the information is relevant, hiding among the large amount of 'noise'; c) such analysis often requires a team of

Figure 2: The messages (top-left part) are grouped according to their topics, indicated by background color. The histogram at the bottom and the entity list on the right can be used to filter the messages displayed by time interval and entity instance such as person and organization respectively.

analysts, which introduce new issues such as communication, sharing, and handover.

The design principle of the SAVI system is not to provide an automated answer, but complement analysts by addressing the challenges discussed earlier. This is achieved with the following design goals:

G1: Facilitate the discovery of unusual patterns. Such patterns can be the sudden changes in the attribute values or the unusual relationships between topics. Discovering unusual patterns will help analysts better understand the dataset (C3a) and identify the starting point and right analysis question (C1).

G2: Support multiple hypotheses. Given the open and exploratory nature of such analysis, users are expected to attempt many alternative hypotheses (C1 and C3b). It can be laborious to keep track of a large number of hypotheses together with related information, and the design of the SAVI aims to improve this.

G3: Support narrative building. With the relevant information deliberately hidden (C2), it will take significant effort to identify them and their connections. SAVI facilitates this process, making it easier to manage over a long period (intelligence analysis can last days, weeks, or even years).

G4: Support collaborative sensemaking. This will allow the team of analysts to work together on the problem and easily share and communicate their progress (C3c). This often entails the possibility of collaborating remotely and asynchronously, which is often missing from the existing systems.

3.2 SAVI: Social Analytics VIsualization

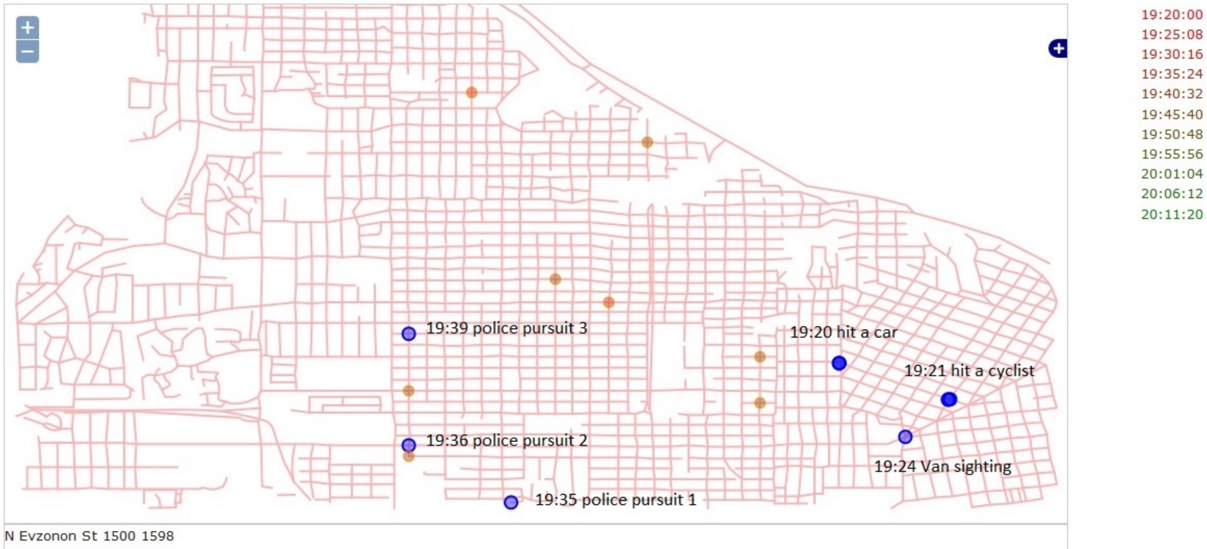
To achieve the design goals discussed earlier, we not only enhanced the TimeSets but also introduced new visual analytics tool. The result is the SAVI suite (Figure 2, Figure 3, and Figure 4).

3.2.1 Facilitate Discovery (G1)

To achieve G1, our main approach is to provide an overview of the data from as many perspectives as possible. For instance, we included a histogram (the bottom panel of Figure 2) to provide additional temporal information. We also added semantic analysis (the right panel in Figure 2) and a geo-spatial view (Figure 3).

All these views are designed to help analyst understand data and discover unusual patterns. For example, the histogram (bottom panel of Figure 2) shows the frequency of tweets and there appear to be two peaks, one around 6.45pm and the other at 7.45pm (as the investigation progresses these two turn out to indeed relevant events). This can be a clue that something important happened at these two time points. Analysts can select the period on the histogram, and all the tweets within the period will be shown in TimeSets (the histogram in Figure 2 shows the selection of a period from 6.40pm to 7.35pm).

The tweets within such a time period (usually hundreds or more) are often too many for TimeSets to visualize meaningfully (many tweets will be aggregated as a result). This is one of the motivations to introduce the *named entities analysis*, which automatically recognizes *entities*, such as “person” and “organization”, from the tweet message (the results are shown in the right pane of Figure 2). Under each type, such as “person”, entities are listed in descending order based on their frequency. The two numbers next to an entity are the number of occurrences within



Author	Message	Date	Location
ccdata	PURSUIT-POLICE UNIT #253 IN PURSUIT	2014-01-23T19:35:58.000Z	Egeou Ave / Nia Ave
ccdata	PURSUIT CONTINUES-POLICE UNITS #253 AND #357/SECOND UNIT JOINS	2014-01-23T19:36:00.000Z	Egeou Ave / Limnou St
ccdata	FELONY HIT & RUN-IN PROGRESS-WITH INJURY	2014-01-23T19:21:38.000Z	547 N. Schaber Ave

Figure 3: Geo-spatial view. Messages are shown as dots and colored according to time (legend in the top-right corner). The details of the selected messages are shown as a list at the bottom.

the current selection and the entire dataset respectively. For example, the first entity under ‘Facility’ is ‘Dancing Dolphin (14/38)’, which means it is the most mentioned facility and appeared 14 times in the current selection and 38 times in the entire dataset. Analysts may decide to further investigate ‘Dancing Dolphin’, which turns out to be another important piece of the puzzle.

The entity list can be used to filter the message in TimeSets: selecting ‘Dancing Dolphin’ will keep only tweets that contains the phrase. Multiple entities can be selected, with matching background color showing in the TimeSets. In Figure 2, three entities are selected (Aliba Police Department, Dancing Dolphin, and Aliba) and three entity sets are shown in TimeSets with matching background color. The color of the dot at the beginning of each message represents its sentiment: red for negative, green for positive, and white for neutral.

The map view (Figure 3) provides a geo-spatial perspective of the dataset. It is synchronized with the TimeSets view, i.e., only showing messages that are selected in the TimeSets (those with location information). Figure 3 shows an example about messages containing the phrase ‘black van’, which was first spotted at 19:20 when it hit a car (bottom right corner), then hit a cyclist, and eventually led to a police pursuit. It is reasonable to think these are the sightings of the same van, which appears suspicious.

3.2.2 Support Hypothesis, Narrative, and Collaboration

To achieve G2, G3, and G4, our approach again is not to automate the process, e.g. automatically generating a narrative, but instead provide the support needed during the investigation. We focused on the following features:

Bookmarking: allow analysts to save any insight as the investigation progresses, even when it is not immediately clear whether the insight is directly relevant. Such insights need to be recorded together with the (visual) analysis that led to it, so the analysts can later go back to verify or continue the investigation.

Connecting: support the discovery of connections among the insights, i.e., *connecting the dots*. This is important for hypothesis generation and narrative construction, as the newly discovered connection may lead to new hypotheses or contribute to the creation of a new narrative.

Collaborating: A team of analysts should be able to work together on the same investigation. They should be able to easily share not only the insights and narrative, but also the processes that lead to them.

The *sensemaking notebook* (Figure 4) is created to provide these features. Each analyst has their own list of insights and hypotheses, and can also browse those from other analysts. Figure 4 shows that ‘phong’ is the current user and the left pane lists all the insights he created. Only the text description of an insight is shown in the list, but the associated visualization (or other analysis information) will be shown once an insight is added to the main *hypothesis canvas*. Analysts can freely position the insights, for example by ordering them from left to right chronologically

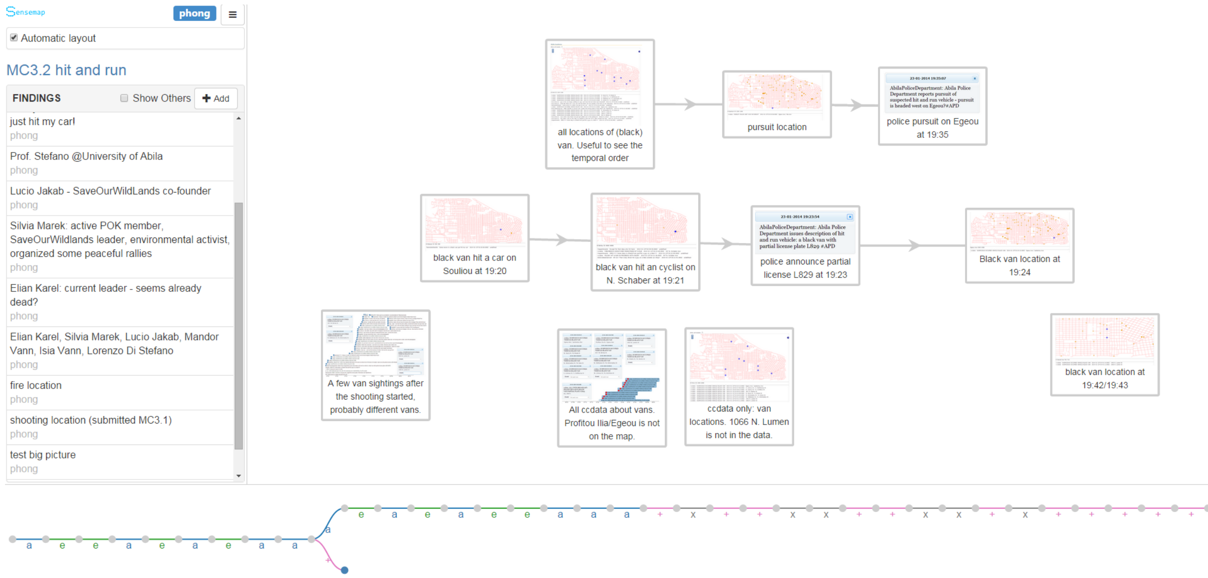


Figure 4: Sensemaking environment. On the left is a list of “findings” that users created. These can be used to construct narratives in the workspace on the right.

(essentially creating a narrative) or group all the insights related to the same person or organization together, which in our experience often lead to new hypotheses as the result of better understanding of a person or organization. The insights can be linked to form a narrative or an argumentation (i.e., connecting statement with evidence). As with insights, the hypothesis canvas can be accessed by all analysts.

At the bottom of the sensemaking notebook is the *provenance* of the reasoning process, i.e., the audit trail of the evolution of the hypothesis canvas. Each node is one state of the hypothesis canvas and the edge represents the “action” that changes the state, such as adding and re-positioning findings. Nodes and edges are added automatically when user works on the hypothesis canvas. Clicking on any node reverts the hypothesis canvas to that particular state. This allows branching from an earlier state and follow a different path of investigation, without losing any existing work. Again, the provenance is accessible by all analysts, so they can share not only the hypotheses and narratives but also their evolution process.

3.3 Implementation

SAVI is a web application with Node.js (`nodejs.org`) as the back end server and MongoDB (`mongodb.com`) as the data storage. Socket.io (`socket.io`) is used for the client-server communication, and the front end visualization is mostly done with D3.js (`d3js.org`). The client-server architecture allows multiple analysts to work simultaneously regardless of their location. The named entity analysis is achieved through web service provided by AlchemyAPI (now part of IBM).

4 Uncertainty Analysis

Uncertainty analysis becomes particularly relevant recently with the dramatic increase of misinformation, particularly in social media [8]. Existing work on visualizing the spread of the fake news mostly focuses on the topology of the social network involved [9], whereas our attempt targets the topics covered in the fake news and their connections. Together with the security experts from MASS Ltd (`mass.co.uk`), we developed our second tool, TimeSets Uncertainty, for analyzing uncertainty information in social media datasets.

4.1 Requirement Elicitation

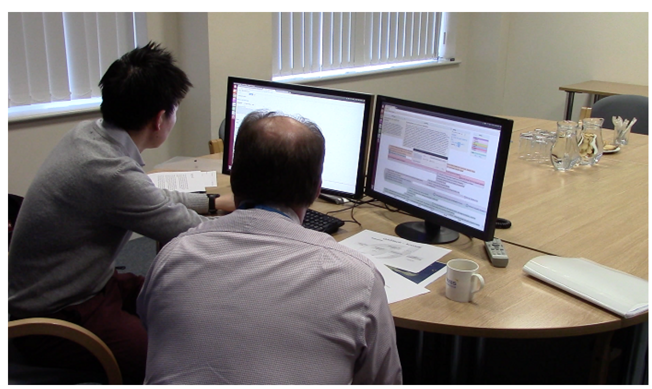


Figure 5: Requirement elicitation using Pair Analytics - The Subject Matter Expert (SME, right) works together with the Visual Analytics Expert (VAE, left) on the dataset from VAST Challenge 2014: the SME ‘drives’ the investigation and the VAE operates the visual analytics system.

Pair Analytics [10] was used to understand user requirements related to uncertainty analysis. It is carried out by a Subject Matter Expert (SME) and a Visual Analytics Expert (VAE). The VAE often plays the role of the "driver" (operating the software) and the SME plays the role of the "navigator" (directing the investigation). The interactions between the two parties help bring out insight about the analysis process that is more difficult to recover in a SME only observation.

The task used is the Mini Challenge 1 from the VAST Challenge 2014, which requires the investigation of reports with varying reliability. The two SMEs are ex-military analysts with extensive experience dealing with data quality issues. Each session lasted for about 2 hours, with a semi-structured interview at the end. Figure 5 shows the setup of the pair analytic session. Below are the results from the analysis of observation data:

R1: Visual representation of uncertainty is important. Both analysts commented that it is important to represent the uncertainty information visually and improvised when the visual analytics system does not support uncertainty natively. For example, one analyst ordered the reports vertically based on his confidence towards them, with the most reliable information on the top.

R2: Data source is a key uncertainty indicator. Both participants emphasized that the data source is one of the key pieces of information they use to judge the quality of the data. Information from well known sources, such as police announcement, was assigned with less uncertainty than information from sources such as social media. For example, one of the participants recognized that a report was produced by "psy ops" (psychological operations) and this knowledge significantly increased his confidence in that report.

R3: Source confidence change over time. The confidence level towards a new data source changes over time, as the analysts encountered increasing number of reports from the same source. Such confidence level increases if multiple reports were cross validated, and vice versa.

R4: Updating uncertainty level is needed. Both analysts mentioned that they want to manually change the uncertainty level as they see fit. This includes both assigning an uncertainty level to reports that do not have one and changing the uncertainty level of reports that have already been evaluated. The analysts arranged the reports vertically to show uncertainty changed the order a few times as the investigation progressed.

R5: Following existing standard. Given the diversity of the types of uncertainties and their measurements, both analysts thought it important that the uncertainty description follows the national standard, which is the $5 \times 5 \times 5$ model (`www.gov.uk/hmrc-internal-manuals/money-laundering-regulations-compliance/mlr3c14000`) for UK. The two uncertainty-related dimensions within this model are *source evaluation* and *intelligence evaluation*: both have five levels of uncertainty, ranging from '*A: Always reliable*' to '*E: Untested source*' and '*1: Known to be true without reservation*' to '*5: Suspected to be false or malicious*' respectively.

R6: Uncertainty for report groups. Due to the large number of reports in the collection, TimeSets often needs to aggregate reports to make sure each shown report is legible. While analysts found such aggregation sensible, they would like to know the uncertainty level of an aggregated document group, just as it for a single report.

R7: Excluding highly uncertain reports. Another common desire from both analysts is to filter reports based on their uncertainty level. For example, an analyst may want to focus on reports with every low uncertainty level only when making a critical decision.

4.2 Design Workshop

After the requirement elicitation, a design workshop was held as part of our user participatory design effort. The workshop follows the process of *Design Sprint* [11], which enables quick iterations generating and improving design ideas that are closely linked to user requirements. The workshop participants consisted of a wide range of stake holders: two university researchers and three MASS staff members that include its research director, a project manager, and an ex-army analyst (potential user). One design sprint had three phases: *understand*, *diverge*, and *converge*:

Understand: clarify users' requirements and ensure all stake holders share a common understanding. In our case, the university researchers reported the results from the requirement elicitation, and MASS staff provided feedback from different perspective that related to their role. Then the whole team worked together to create a *scenario* that describes how the analysis is done currently.

Diverge: each participant is encouraged to work independently to envisage what the new system will look like and how users will interact with it. This is started by writing down a list of desirable features in the form of *situation*, *motivation* and *outcome*. Then, each participant sketched how these features would work,

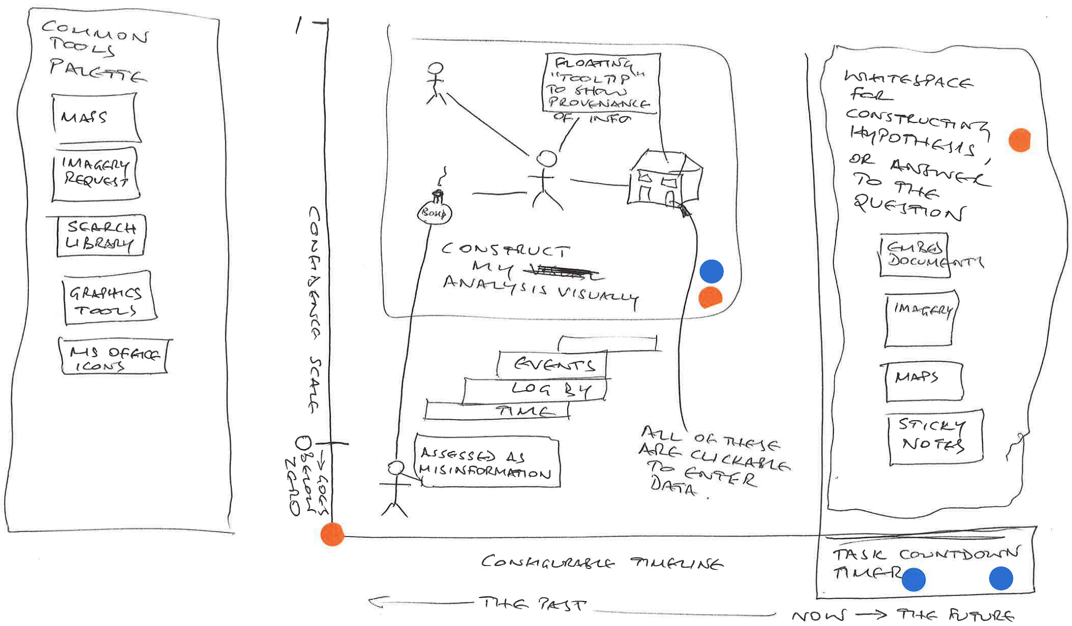


Figure 6: UI sketch created at the design sprint workshop. on the left is a list of tools available to the analyst; in the middle is the main analysis work space, showing the reports (lower half) and the connections among entities (upper half). On the right is the space for notes and hypothesis. The colored dots are the votes for favorite features.

with only the relevant UI components (i.e., not a complete UI wireframe). All these need to be based on the requirements and scenario agreed in the 'understand' phase.

Converge: bring the team together to produce the final design. It started with a critic and voting session of all the sketches created during the second phased. Then the team worked together to create a complete UI sketch that integrates the most liked ideas. Figure 6 shows the final UI sketch created at the end of the design sprint.

4.3 TimeSets for Uncertainty Analysis

Due to the project scope, the team decided to focus on the main work space described in Figure 6 and leave the two side panels to future development. Guided by the discussions and outcomes from the design workshop, TimeSets was extended to support uncertainty analysis with many new features (Figure 7),

To address R1, uncertainty is mapped to text transparency. This has been found to be the most effective among more than 10 alternatives in a previous study (described as 'fuzziness' [12]), and analysts found it intuitive during the design workshop. Low transparency indicates high confidence while high transparency represents low confidence (Figure 7A).

To address R2, the source information is displayed in the bottom pane when a report is selected. For example, Figure 7B shows that the selected post ("Fox News sets another Trump town hall") is produced by "*POLITICO*", whose source uncertainty is "*mostly true*".

To address R3 and R4, a slider is provided so users can adjust the uncertainty level based on his or her judgment (Figure 7C).

To address R5, an *uncertainty (histogram) matrix* is added in the left pane (Figure 7D): the x-axis is the intelligence uncertainty and the y-axis is the source uncertainty. In this example, we use the post's uncertainly rating created by the dataset curator that has four levels (more details in subsection 4.5) as the intelligence uncertainty. The source uncertainty is calculated as the average rating for each publication source, i.e., the average rating of all articles published by a source. The uncertainty matrix also functions as a histogram, showing the number of posts for each rating combination (e.g., there are 219 posts with *B1* rating). The rating for individual post is shown in the bottom pane next to its details (Figure 7C).

To address R6, we set the uncertainty for an aggregated post as the average of the ratings of all the posts it contains. Similarly calculation is done for each entity, and shown in the legend (Figure 7E). The green, amber, and red triangle stands for low, medium, and high uncertainty level respectively, based on the domain experts' preference of a "traffic light" system.

To address R7, uncertainty filtering is added (Figure 7F). Analysts can selectively show (or hide) posts with low, medium or high uncertainty as discussed earlier.

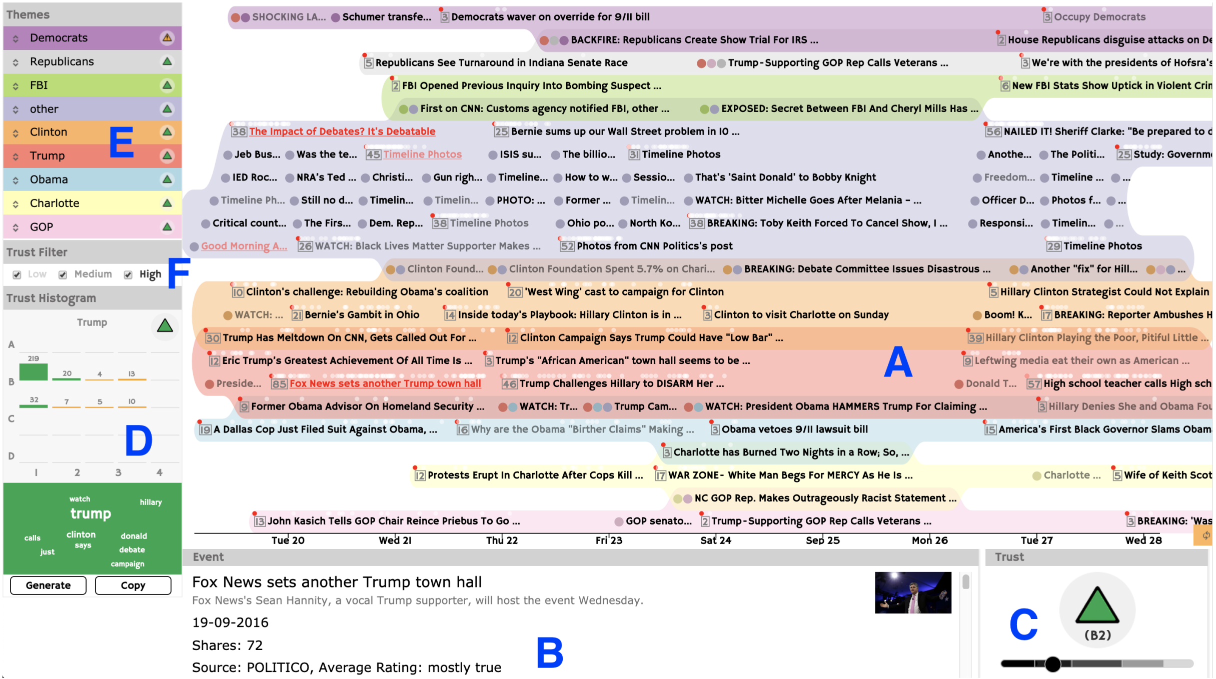


Figure 7: Extended TimeSets with new features to support uncertainty analysis.

4.4 Discussions

The initial feedback is largely positive from the two analysts who helped with requirement elicitation: the visual representation is intuitive and they envisage the improved version will provide much better support for uncertainty analyses. A formal user evaluation on the TimeSets Uncertainty is currently being planned.

Reflecting on the project, we think that Pair Analytics worked well and allowed effective elicitation of uncertainty-related requirements. The communications and interactions between the SME and VAE helped to bring these out. The Design Sprint workshop also worked well, but a longer duration (such as 2-3 days) may allow the exploration of additional design ideas and further development of the final UI sketch. Nevertheless, the participation of all stake holders and their contribution throughout the process certainly helped improve the final design.

4.5 Implementation

The dataset used in Figure 7 is a collection of Facebook posts curated by the BuzzFeedNews (www.buzzfeednews.com/article/craigsilverman/partisan-fb-pages-analysis). They analyzed over 2500 political posts on Facebook and rated the uncertainty of each as one of the following: 1) Mostly true; 2) Mixture of true and false; 3) Mostly false; and 4) No factual content (i.e., it is not about a fact). We applied topic modeling using spaCy (spacy.io), a natural language processing library, to identify popular entities, from which we manually selected the entities for TimeSets Uncertainty.

TimeSets Uncertainty extends the original TimeSets and share the most of technologies used. The latest version is available online (vis4sense.github.io/timesets/uncert) where user can try out the Facebook dataset discussed in the paper. A useful feature of TimeSets Uncertainty is that users can load their own data, which is achieved using Google Spreadsheet: user just needs to upload their data into a Google spreadsheet and append its sharing URL to the end of the TimeSets Uncertainty URL. The detailed instructions are available online (vis4sense.github.io/timesets#uncertainty-guide).

5 Conclusions

In this paper we introduced two visual analytics tools incorporating the TimeSets technique. Both tools are designed to support temporal analysis in the field of intelligence analysis, particularly Open-Source INTelligence analysis (OSINT). The resulting tool is well received by the domain experts and available online for users to experiment with their own data. While both tools are designed in the context of intelligence analysis, we believe the challenges they address, such as support hypothesis generation and uncertainty analysis, are applicable to other application domains with similar analysis needs.

References

[1] B. Bach, P. Dragicevic, D. Archambault, and S. Carpendale, “A Review of Temporal Data Visualizations Based on Space-Time Cube Operations - Semantic Scholar,” in *Eurovis STAR Report*, 2014, pp. 23–41.

[2] C. Collins, G. Penn, and S. Carpendale, "Bubble sets: revealing set relations with isocontours over existing visualizations," *IEEE Transactions on Visualization and Computer Graphics*, vol. 15, no. 6, pp. 1009–16, 2009.
[3] B. Alper, N. Henry Riche, G. Ramos, and M. Czerwinski, "Design study of LineSets, a novel set visualization technique," *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2259–67, Dec. 2011.
[4] W. Meulemans, N. Henry Riche, B. Speckmann, B. Alper, and T. Dwyer, "KelpFusion: A Hybrid Set Visualization Technique." *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 11, pp. 1846–58, Nov. 2013.
[5] P. H. Nguyen, K. Xu, R. Walker, and B. W. Wong, "Timesets: Timeline visualization with set relations," *Information Visualization*, vol. 15, no. 3, pp. 253–269, 2016.
[6] S. Chen, L. Lin, and X. Yuan, "Social media visual analytics," in *Computer Graphics Forum*, vol. 36, no. 3. Wiley Online Library, 2017, pp. 563–587.
[7] K. Xu, P. H. Nguyen, and B. Fields, "Visual analysis of streaming data with savi and sensemap." in *IEEE VAST*, 2014, pp. 389–390.
[8] S. Vosoughi, D. Roy, and S. Aral, "The spread of true and false news online," *Science*, vol. 359, no. 6380, pp. 1146–1151, 2018.
[9] M. Liu, S. Liu, X. Zhu, Q. Liao, F. Wei, and S. Pan, "An uncertainty-aware approach for exploratory microblog retrieval," *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 1, pp. 250–259, Jan 2016.
[10] R. Arias-Hernández, B. Fisher, and L. Kaastra, "Pair analytics," *Social Cognition and Interactive Expertise in Natural and Computational Environments*, 2011.
[11] R. Banfield, C. T. Lombardo, and T. Wax, *Design Sprint: A Practical Guidebook for Building Great Digital Products*. O'Reilly Media, September 2015.
[12] A. MacEachren, R. Roth, J. O'Brien, B. Li, D. Swingley, and M. Gahegan, "Visual Semiotics and Uncertainty Visualization: An Empirical Study," *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 12, pp. 2496–2505, Dec. 2012.

AUTHORS

Kai Xu is an Associate Professor in Data Analytics at Middlesex University in UK. Contact him at k.xu@mdx.ac.uk.

Saminu Salisu is a PhD student at Vienna University of Technology in Austria. Contact him at s.salisu@live.co.uk.

Phong H. Nguyen is a Data Scientist at Red Sift. Contact him at phong.nguyen@redsift.io.

Rick Walker is a Principal Data Scientist at King. Contact him at rick.walker@king.com.

B.L. William Wong is a Professor in Human-Computer Interaction at Middlesex University in UK. Contact him at w.wong@mdx.ac.uk.

Adrian Wagstaff is a Director at Hawkesford Consultancy Ltd in UK. Contact him at adrian@hawkesfordconsultancy.co.uk.

Graham Phillips is a Senior Consultant at MASS in UK. Contact him at gphillips@mass.co.uk.

Mike Biggs is a Program Manager at MASS in UK. Contact him at mbiggs@mass.co.uk.