**ABSTRACT**

**1 INTRODUCTION**

Time Sets consist of a timeline showing sequence of events displayed across a visualisation, while makings sense of sets relation among events in the timeline (Phong). Time set is very effective in identifying trends and sets relations within a large set of events.

The Study looked into extending Time Set to accommodate Visualisation of Trust and Uncertainty as parts of its variables for events displayed across the time. The above study is part of the UKDSC Autonomy and Big data sponsored by the ministry of defence through Defence Growth Partnership (DGP) Innovation Challenge with small and medium enterprises carried out by Middlesex University and Mass. The aim of the challenge is to build tools in the context of big data analytics that can be used to aid military operations through intelligence analytics and decision-making.

The study starts by looking into the related works carried out in the area of Uncertainty Visualisation and Variables that can be used to identify uncertainty in a visualisation, followed by trust models and perception in general principle. Extensive preliminary research, user observation and workshops where carried out to guide the design and implementation of effective Uncertainty and Trust visualisation variables while putting into consideration security and defence as the target users.

**2 RELATED WORKS**

Uncertainty and Trust has been the subject of less extensive research compared to data mining and extraction. While it is beyond the scope of this paper to provide a complete overview of research in Uncertainty and Trust, in the following section, we describe aspects of the research on Uncertainty and Trust propagation, we also review more recent literature on the role of Uncertainty and Trust and in Trust models and Finally, we briefly review finding from the research study carried out.

Some of the first attempts to quantify user’s perception of trust were carried out in a study by [Elliott and Yannopoulou -2007], although previous studies where geared towards identifying the sets of factors influencing trust, the author took the novel approach of quantifying the value of each trust factor in a given domain. The result of the study can enable web interface developers to focus on key elements that influence trust and increase the trustworthiness of a digital interface.

The Study carried out by [Representing Uncertainty in Graph Edges - An Evaluation of Paired Visual Variables.pdf] looks into visual variables used to represent uncertainty.

The paper reports on a study carried out about the perception of graph edge attributes when uncertainty associated with each edge and the main edge attribute are visualized simultaneously using two separate visual variables, Some of the results show that factors such as Grains, fuzziness and Transparency depict uncertainty effectively.

Visual Representation of uncertainty with focus on variables that can be used to depict uncertainty and trust in data is key to support the design study. An experiment carried out by [M. MacEachren] to determine the effectiveness of uncertainty variables such Grains, fuzziness and Transparency led to some generalized conclusion by the author, Fuzziness and Variable location work very well in uncertainty visualisation, Values and arrangement of variables are also an effective means of showing uncertainty and Transparency & variable sizes are theoretically valuable in variables for representing uncertainty.

Another study by [Uncertainty-Aware Multidimensional Ensemble Data Visualization and Exploration.pdf] looks into modelling and exploration in multidimensional data. The author presents an efficient visualization and exploration approach for modelling and characterizing the relationships and uncertainties in the context of a multidimensional ensemble dataset.

This author focuses more on simulation and analysis with some suggestion on using ensemble simulation to study uncertainty.

Theme Delta [Gad et al. - Unknown - ThemeDelta Dynamic Segmentations over Temporal Topic Models.pdf] is another timeline visualisation data similar to time sets but uses sinuous, variable-width lines to show this evolution on a timeline, utilizing colour for categories, and line width for keyword strength.

The study focuses on the visualization of data over time with focus on temporary data that has a life spam and can have reduced value over time. In relation to user observation study carried out, data life span affect the uncertainty level of that data over time.

**Methodology**

**3.1 Study Design**

**User Observation**

To determine users perception of Uncertainty and Trust in Data, We conducted an observational, exploratory study to observe the interaction between a subject and a controller (data analysis controller DAC) who assists the subject in an exploratory data analysis task: analyse data from VAST Challenge 2014 Mini Challenge 1 to provide suggestions as to how decision makers in the military sector handle uncertainty in data.

Subjects were exposed to a visual interface for visualization with the legacy version of the **TimeSets** for timeline data analysis, **Kebana** for quick search and **SenseMap** for hypothesis building. The subjects where also instructed to ask question directly to the DAE that they deemed relevant to the performance of their data analysis task.

The subjects were allowed to communicate and encouraged to think-aloud without any restrictions as they used the visualisations to analyse the challenge data to answer the challenge question.

Subject Users where given a complete description of the challenge data, problem description, access to computers with TimeSets, Kebana and SenseMap on double desktop display screen as seen in the figure below. The environment setup ensured DAC and the subjects had access to multiple visualisation and the necessary tools to address the questions set out in the VAST challenge 2014 Mini Challenge 1.

**Workshops**

The design workshops where aided with the use of Sprint Design to simply brainstorm design ideas in cooperation with agile environment.

**3.2 Participants**

3 pilot studies where conducted which included 1 user observation and 2 workshops to refine the refine the experimental design. The participating subjects in the User Observation are 2 Experienced data analyst with vast experiences in the military sector, both male aged between 40 and 55 which lasted for about 3 hours each. Participants had varied experience with the challenge and data analysis with one participant an ex-army analyst and ex-raf analyst respectively.

The participants for the workshops consist of 1 ex-army chief, 3 researchers, 1 project manager and 1 developer with each participant showing moderate familiarity with uncertainty and information visualisation.

**3.3 Data Analysis**

3.1.1 Approach

Due to the limited number of participants, inferential statistical approach was adopted for this research. Inferential statistics analyses data that researches have limited access to, as a result they use procedures to infer the meaning on collected data from a given population, it is used when it is essential to analyse behavioural statistics of a given population

3.1.2 Model: Coding

Coding method involves the transcription of data collected from the user observation and interviews through audio / video recording to short written scripts that can be used to make analysis. Data collected using qualitative method such as the interviews with the Participants during the User Observations are analysed in coding by segmenting the response into meaningful variables and assigning those variables into categories know as a code.

The following codes where created to enable easy analysis of the data collected from the interview with the participants. The analysed result can be found in the next section of this report

1. Types or Variations of uncertainty  
2. Marking or Recording uncertainty  
3. Benefits of uncertainty to Reasoning and thinking  
4. Challenges in dealing with uncertainty

3.1.3 Tools: Thematic

Using thematic to monitor, examine and record theme patterns across the collected data from the qualitative research analysis. Thematic is best used to emphasise on the subjects use and awareness of the key elements that derives uncertainty and trust in data presented to the analyst. Contribution by the researchers in the data analysis stage using thematic method that requires looking out for patterns/themes in the data. Few factors has to be put into consideration: What is the size of the theme, what counts as a pattern/theme, Identifying the key themes, what are the key questions of interest in relation to uncertainty visualization.

To determine the answer to the above questions, the following themes and patterns where identified from the analysis of the result in interviews carried out with the participants during the user observation.

Source of Data is key to trust and certainty

2. Experienced MOD users are comfortable with dealing with un- certainty

3. New source of information increases confidence level overtime

4. Internal and known source of information more trusted than external source data

5. Aligning high confidence hypothesis on the top of low confidence hypothesis

6. Icon or component for marking confidence level in hypothesis

7. Familiarity impact the approach and use of tools

**3.2** INITIAL FINDINGS

3.2.1 Data source as a key factor to confidence

The observation study of the two participants enabled the researchers to identify data source as a key element that affect uncertainty and trust as agreed by both participants.

1. Both participants agree that the source of data is key to confidence in that data
2. Different data sources have different levels of trust

3.2.2 Updated source of information affects confidence level overtime

As the study progressed, continuous discovery of new and updated in- formation changed the participants confidence in that data. Both participants regardless of their approach to the solving the data challenge constantly referred back to time sets and Kebana for new information sources that will further support or refute the derived hypothesis.

3.2.3 Internal and known sources are more trusted than external sources

The response of increased confidence in a data from participant (James) when he discovered a source of information with the label (psycops) indicates that a form of label or logo showing who reviewed an information can significantly increase the level uncertainty and trust in that data. Some of the key points to further prove these points are below

1. Both participants agree news report form specific companies where more trusted than others just by identifying with the companies logo or brand

3.2.4 Uncertainty and Trust variables

Both participants at some point during the study wanted to identify the level of uncertainty by marking the confidence in hypothesis so as to remember, which further indicates the importance of showing confidence levels using variables such as transparency, hue, saturation, on hypothesis and data used to derive the hypothesis.

3.2.5 High Confidence before Low Confidence Arrangement

The user observation also highlighted the importance uncertainty arrangement by confidence level with high-level confidence information placed above on-top and low level information placed on the bottom. One of the participants consistently reflected this by placing all high- level hypothesis above less confident hypothesis in Sense Map.

**3.3 CONCLUSION**

The user observation has enabled the researchers to gain valuable in- sight into users perception of uncertainty different ways in which un- certainty is being handled during decision-making in relation to data analysis. The observation also identified and recommends some effective methods and techniques of communicating uncertainty in data during analysis by identifying key elements that constitute uncertainty and ways they can be communicated across the decision making circle in the military data intelligence and information consumption.

**ACKNOWLEDGMENTS**

The authors wish to thank A, B, C. This work was supported in part by a grant from XYZ.

Reference Papers

http://vis.ninja/papers/ShowMeData\_vis15.pdf

<http://vis.stanford.edu/files/2010-Narrative-InfoVis.pdf>

<http://lsr.nellco.org/cgi/viewcontent.cgi?article=1476&context=nyu_plltwp>

Many ways of representing uncertainty have been studied such as Encoding uncertainty in visual variables. MacEachren et al [8] studied how uncertainty could be represented in a number of different variables. Their results give evidence to support using fuzziness and location of elements on the screen rather than colour. Gschwandtner et al [9] carried out experiments to find out the most effective ways of encoding temporal uncertainty, comparing six types of visual representation. They concluded that error bars can cause confusion. Whilst the representation chosen should vary to meet the particular task they found that gradient plots are generally superior to other representations. Guo et al [10] studied the encoding of uncertainty in the representation of graph edges, which is particularly pertinent to SenseMap. They list a set of design guidelines which emphasise the robustness of fuzziness, grain and transparency when encoding uncertainty.