Narrative Visualization: Crime Against Women in India

Abhijeet Kumar, Saurabh Verma, Adil Shaik, and M. Sai Kavikondala

**Abstract**- We propose a crime data visualization tool for the purpose of spreading more awareness about the crime scenario in India. Although there are many crime visualization tools, the task of deriving insights is left to the user who may not get all the useful insights on their own. Our system is aimed towards solving this problem by following a narrative style to assist the user in getting useful insights about various types of crimes. To demonstrate the usefulness of this tool, we apply it to visualize crime data on kidnappings and compare the performance of test subjects using our actual narrative style tool and a stripped-down version containing only the visualizations.

**Index Terms**— Narrative visualization, annotations, crime data, design methods, scenes

Introduction

There are many crime data visualization tools for exploring crime. Most visualization tools let the user explore the data on our own and draw observations from it which takes a lot of time, effort and expertise. There are chances that the users might miss some of the key information of the crime dataset if left to explore on their own without any guidance. As a result, the general public may not be aware about the seriousness of many crimes. This is especially true of the crime scenario in India. In order to bring more awareness of crimes to the general public, we propose a system for crime data visualization based on a narrative style, which guides the user through the crime trends and also allows exploration of more facts about the crime in a self-discovery fashion through annotations and links to external sources.

# Related Work

Narrative visualizations are usually characterized by a series of techniques used to sequentially direct a users’ attention. These techniques hold the user’s attention towards the story through user-oriented transitions. Psychologists for years have studied the phenomena of depicting narrative stories through outliers among visual attributes like colour, orientation, size to orient and hold users attention [1,2,3]. Many different cultural factors like reading order from lift to right introduce several biases for people to look at a place and scan information. Several transition attributes and their classification with relation to *object continuity* and extra pictorial elements like annotations and callouts have been used to add richness to a narrative in McCloud and Taversky [4, 5]. For this study however we delved into the details of contemporary narrative visualizations with the primary source being online journalism with visualizations from the Washington Post, the Guardian and the Financial Times.We looked at the intricacies of Narrative Visualization from Segel & Heer [6] and applied the techniques learnt to Crime dataset for crimes against women in India from 2001-12.

In our research we did not find tools with sufficient visual narrative elements as required in a contemporary narrative visualization tool. Significant efforts have been made in [A7, A8], however their attempts lack the richness and comprehensive demands of a modern narrative visualization to adequately convey the intended storytelling. There are many visualization tools available for crime dataset [7,8] but they do not have any storytelling style associated with them. Our focus is on visualizing Indian crime data and in that context, visualizations lack at clear visual encodings and interaction mechanisms. The design requirements and design choices based on them are also not well justified in many of these tools. In the American context there are visualizations like Crime data Visualization in Chicago PD [9] that uses a narrative style and has visualization categorized in terms of scenes as in our design but each of these scenes use different datasets. Hence ours is a novel visualization that incorporates a rich dataset of reported crimes from 2002-12 and has different functionalities associated with each of the categories of the data. Other design tools

.

# Dataset and Data Analysis

We analyzed the data pertaining to crime against women in India and studied the design patterns of many modern narrative visualizations to arrive upon this design of the tool that we claim is a novel tool incorporating

The data for crime in India in general and for crimes against women in particular is published annually by the National Crime Records bureau (NCRB). The coverage of crime against women has been starting to get covered in great details in recent years. The reported data from each police station all over the country is then compiled and published. There are several issues with the data in terms of its completeness however it is adequate to study regional variations across states and linguistic belts. The dataset contains a mix of ordinal categorical and interval data. Nature of crimes is categorized into different crime types on the basis of Indian Penal Code (IPC) and State Local Laws(SLLs).Victims and perpetrators are divided into different categories on the basis of their demographics and the granular data is then accumulate into other data types for studying various characteristics of crimes. This data is mainly classified on the basis of overall reported cases and the overall crime rate. We use both the categories for the primary design classification in our tool. It has been broadly seen that there exist a correspondence between the total crime rate and overall crime against women excluding impact of low female population. Heteroscedasticity is significant here as the variation in crime against women rates increase when we move from states with low total crime rates to states with high total crime rates. Very clear spatial patterns emerge from the states and regions with high crime rates as different categories of crimes against women is prevalent in different regions of the country and regions with high crime rates most often adjoin regions of very high crime rates. This indicates that these are not merely figments of under reporting. A reader familiar with the subject matter may also deduce further significance in the spatial patterns of the very high crime rate regions. According to crime types almost all of Madhya Pradesh , Delhi, Chattisgarh, West Bengal fall in high crime rates against women for nearly all crime types . Other regions of adjoining states like Maharashtra, Assam, Rajasthan and Uttar Pradesh also figure in the prominent regions in high crime rate. Our visualization tool depicts all these in very intuitive manner. The reluctance of victims to report many of the cases has also been observed in the uniformly low proportion of the sexual harassment cases. This is primarily due to social stigma and complications associated with these cases. In some states like Assam as indicated by the data and the tool the higher proportion of rape cases among states with low overall crime rates is due to the low instances of reported cases and low female population as compared tot the major states.

# Visualization Design

Our visualization system is divided into scenes where each scene focuses on one important detail/observation that the author wants to bring to light about a particular crime. Our system can be plugged in with different types of crime data. For demonstration, we have taken crime data on kidnapping. The scenes represent the most prevalent and gruesome type of kidnappings. These scenes are backed up by annotations which help the user understand the scene and links them to articles and other resources for user to research and study about it more. For evaluation, two sets of participants are taken. One set of participants is given a version of the visualization tool without any narrative style and the other set of participants is given a version of the visualization tool with narrative style..

## Design Study Analysis

As detailed in Segel & Heer[6], We also follow the style of detailing the visual elements in representing the design space in a table that finally relates to specific design strategies observed across the dataset . We analyzed about 20 visualizations and selected about 16 of them for the final analysis based upon the case study method as detailed in [6]. We selected the visualizations that had clear sequence of narrative event. The examples of contemporary visualizations were mostly taken from the online journalism category. This organization mostly identifies three attributes -genre, visual narrative elements and narrative structure elements. The narrative structure is further divided into three subcategories which is ordering, Interactivity and messaging, the visual narrative category has been further divide into visual structuring, highlighting, and transition guidance. Applied to our visualization we observed that our tool has most of the elements in Narrative structure category and as detailed in Fig. 1

# Tool Design

This section details our narrative based design. The layout of our design consists of the following components

## Scenes

This is the key element of our design. Each scene focus on one of four prominent crimes - Domestic Violence, Rape, Kidnapping and Sexual harassment. The visualizations in the Overview Map, Detailed Graphs and Annotation box change based on the scene selected and they tell the user a story about the selected crime purpose.

## Overview Map

The map shows the total number of cases of the selected scene on a state level according to the year selected. It is a choropleth where the color intensity increases based on the number of cases in each state. Crime intensities are presented in two manners - crime rate (number of victims/ 100,000 women) and also the total reported number of cases. User can switch between these two crime intensity modes on the toggle of a radio button. When the user interacts with the map via a hover, a tool tip reveals the name of the state along with the crime intensity.

## Time slider

The time slider allows the user to adjust the overview map  to show the crimes for a particular year.

## Detailed Graphs

There are two detailed graphs - a grouped bar chart to show the number of cases reported, the number of people arrested and the

number of people convicted. There is another bar chart to show state wise comparison of crimes.

## Annotations Box

The annotations box is another key feature of our design. It contains text that provides a description of the scene. It introduces the  and talks about the key highlights for the selected scene. It is meant to serve as a guide to the user to discover the key highlights in the graphs. It is also meant to help users get more knowledge about the crime by providing links to external resources

## Highlighted States

Each scene consists of some highlighted states, which are states most prominent in that particular crime scene.

Our tool also includes an “Explore” tab which lacks narrative components.

# User Study

To evaluate the usefulness of our scene - based narrative tool, we took a comparative approach. We split our study into two timed tasks. The first task is of 15 minutes and the second task is of 5 mins. Participants were also split into two groups to perform the tasks. The first group got narrative style scenes as the first task and the Explore tab as the second task. The second group got the Explore tab as the first task and the narrative scenes as the second task

## Participants

We selected a total of 12 participants from different age groups, gender, demographic and occupation. The participants were aged 18-25 years, 25-40 years, 40+ years. Half of the participants were female. Our participants occupations ranged from school teacher, data scientist, software developer, marketing executive and retired government employee. They belonged to different states ranging from Uttar Pradesh, Delhi, Andhra Pradesh, Rajasthan and Bihar.

## The Experiment

Before starting the experiment, the participants were given a briefing about our motivation and then they were asked some questions related to their general awareness about the crime scenario of women, such as the most prevalent crime in their state and the whole country, the reported number of cases per year in their state as well as the whole country. Then they were asked to do the first task for 15 minutes followed by the second task for 5 minutes. The first task and the second task differed for the participant based on their group. After completion of both tasks, they were questioned for insights gained in each task, and asked to compare which aspect was better ( narrative scenes/ Explore tab) and the positives and downsides of each.

## Results

This section presents the results of a few selected participants

### Participant A - Data Scientist

This participant was a male data scientist in his early twenties. His first task was the narrative scene. He spent a majority of the time on the choropleth, and interacted a lot with the time slider to check out the crime intensity. He spent less time on the charts and the annotations. He derived insights on the most crime prone state and was surprised to learn that Madhya Pradesh ranked high in Sexual Harassment. He was also surprised to see Kerala, a state having the literacy rate in the country, featuring a sizeable number of reported cases. He was satisfied with the results that Uttar Pradesh ranked high in Kidnapping as it validates his previous knowledge. When asked to

compare both tasks, he said that the narrative scenes are a good fit for a detailed study in a long duration whereas the Explore tab was good for quick insights in a shorter duration.

### School Teacher

This participant was a female teacher in her late forties. Her first task was the Explore tab. She was interested to know about the crimes in Bihar. She spent time going through the line chart to see the trends of each state through the years. When she switched to the scenes task she was surprised to see that Bihar, which is prejudiced to be a highly crime prone state, faring better as compared to the rest of the states in all crime scenarios. For her, the scenes helped a lot in coming to this conclusion.

# Conclusion and Future Work

The users found our tool useful for getting information about specific scenes easily. On comparison of both tasks, they concluded that the narrative scenes are a good fit for a detailed study in a long duration whereas the Explore tab was good for quick insights in a shorter duration.

Acknowledgments

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

References

1. M. R. Garcia and P. Stark. Eyes on the News. The Poynter Institute, 1991.

2. L.Itti and C. Koch. Computational modeling of visual attention. Nature. Reviews Neuroscience, 2(3):194–203, 2001.

3. . A. M. Treisman and G. Gelade. A feature-integration theory of attention. Cognitive Psychology, 12(1):97–136, 1980

4. S. McCloud. Understanding Comics. Kitchen Sink Press, 1993.

5. B. Tversky, J. Heiser, S. Lozano, R. MacKenzie, and J. Morrison.Enriching animations. In R. Lowe and W. Schnotz, editors, Learning with animation. Cambridge University Press, 2007.

6. Segel, E., & Heer, J. (2010). Narrative visualization: Telling stories with data. IEEE transactions on visualization and computer graphics, 16(6), 1139-1148.

7. Ruchita Rathi, Janine Heiser, Ankita Bhosle http://courses.ischool.berkeley.edu/i247/s14/reports/CrimesAgainstWomen\_Rathi\_Heiser\_Bhosle.pdf

8. http://www.analyticsandvisualization.com/p/his-is-culmination-of-discussions-with.html

9.https://public.tableau.com/views/VisualizingChicagoPDcrimedatawithStoryBoard/Story1?:embed=y&:showVizHome=no&:display\_count=y&:display\_static\_image=y&:bootstrapWhenNotified=true

A screenshot of a computer

Description automatically generated

Figure 1 (Design Study Analysis)- Table of visual design elements in the selected contemporary visualizations vis-a-vis our visualization tool