Oregon Traffic Accident Insights

Nic Desilets, Alex Schultz, and Bungo Takahashi

Fig. 1. Interstate 5 on the way to Portland, Oregon.

Abstract— This paper describes the problem that we addressing, why it is important, how visualizations can be effectively applied, potential users, and our general approach.

Index Terms—Traffic, accidents

1 Introduction

Traffic accidents happen to a certain extent every year, and the number of fatal accidents in Oregon has been increasing. There are detailed reports on traffic accidents online; however, those reports include only numbers and not easy to see what correlations exist among the data. Therefore, we would like to visualize and organize the data in order to make it easier to see correlations between variables. Hopefully this will lead to finding insights and solutions to help decrease the number of accidents.

It is important that the problem needs to be addressed with a visualization because its easy to get lost in black and white statistics. It is difficult for a person to identify trends just by looking at thousands and thousands of rows of raw data each with many columns. Transforming this raw data into easy to digest visualizations will allow people to more easily recognize and address any trends hidden inside all the numbers. Potential users could range from the average person interested in traffic statistics to ODOT crew members and law enforcement members interested in evaluating effectiveness of certain policies aimed at reducing traffic accidents.

Our approach is going to be to first process the data that we have gathered and put it into a SQL database so that we can more easily query the data for information that were interested in. From there, with the assistance of a user guide that is associated with our data set we will be able to identify trends within the data such as vehicular accidents per capita by county, amount of accidents resulting from drunk driving over time, and so on. Once we have identified the trends that we are interested in we will then develop a user interface containing visualizations of these trends so that people can interact with and understand them.

2 Visualization Tasks

Questions that we will address with the visualization are the following:

- How many accidents happened depending on vehicles (car, motorcycles etc.).
- How many people were injured or killed in the accidents.
- What are the demographics of those involved in accidents (in terms of age, sex, residence).

- *Nic Desilets is with Oregon State University. E-mail: desiletn@oregonstate.edu.*
- *Alex Schultz is with Oregon State University. E-mail: schulale@oregonstate.edu.*
- *Bungo Takahashi is with Oregon State University. E-mail: takahasb@oregonstate.edu.*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

- What caused the accidents (such as speed, road condition, light condition etc.).
- Amount of accidents caused by drunk driving over time (or other intoxicants).
- Amount of accidents per capita over time (e.g. are accidents increasing or decreasing?).

3 Data Source

Our primary data source is going to be the National Highway Traffic Safety Administration (NHTSA). They have an FTP server hosting very detailed accident data ranging from 1976-2015. For our purposes we will only be looking at the past five to six years, partially because of inconsistencies within the data by year. Data from 2010 and on can be easily compared against each other while data before 2010 is setup differently and will be more difficult to compare. Additionally, we may be able to leverage ODOT accident data to supplement the data coming from the NHTSA.

4 Data Organization

We will be organizing our data into three entities.

- Vehicle Crash: datetime, location, cause, fatalities, conditions.
- Person: personal info, intoxicants involved, role in vehicle crash
- Vehicle: make, model, year, status, issues, type

The relationships between these three entities are described below.

- Fatal/Injury vehicle crash to person: 1 mandatory to many optional
- Person to Vehicle: 1 mandatory to 1 mandatory
- Fatal/Injury vehicle crash to Vehicle: 1 mandatory to many optional

5 Implementation of ER Diagrams

Figure 2 shows the Entity-Relationship diagram detailing the relationships between our stored entities. We will be translating the ER diagrams into a schema for MySQL. Once the data is loaded into the MySQL database the application used to show the data will be able to query the necessary data.

6 Design of Visualization Interface

The primary component of the interface is going to be the map which shows accidents represented as a heat map. The data points used for the map will depend on what filters are set (e.g. number of all crashes, specific types of crashes, etc.). There will be a slider that can be used to adjust what year is shown for the data (e.g. all accidents, accidents in 2015, accidents in 2014, etc.). Additionally, a legend will complement the map figure to help describe the elements show on the map. For example, if a heat map is used, the legend could be used to map colors to accident values. Figure 3 details a mockup of our user interface.

Acknowledgments

None.

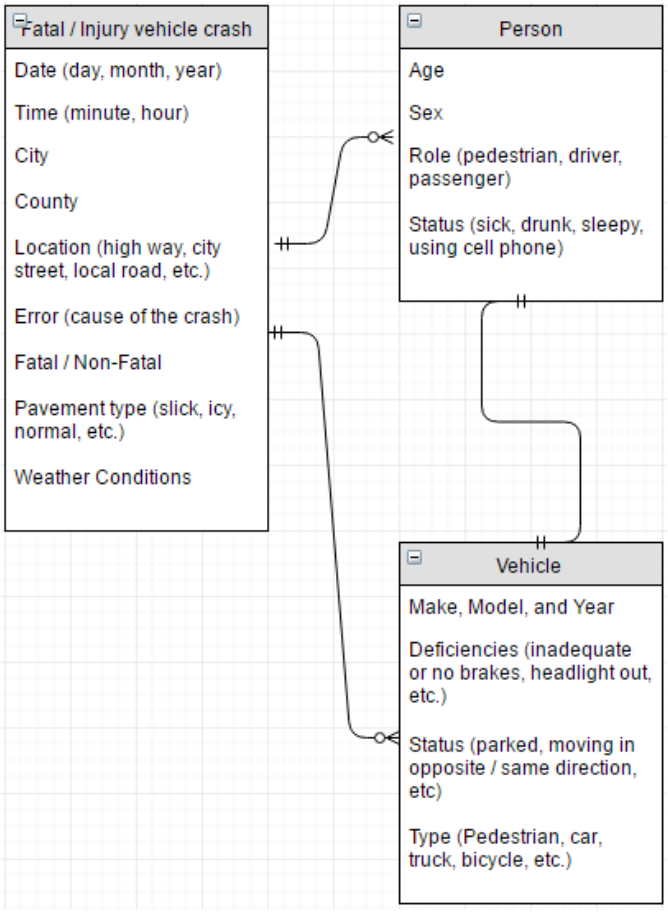

Fig. 2. Entity-Relation diagram showing the relationships between vehicle crashes, people, and vehicles.

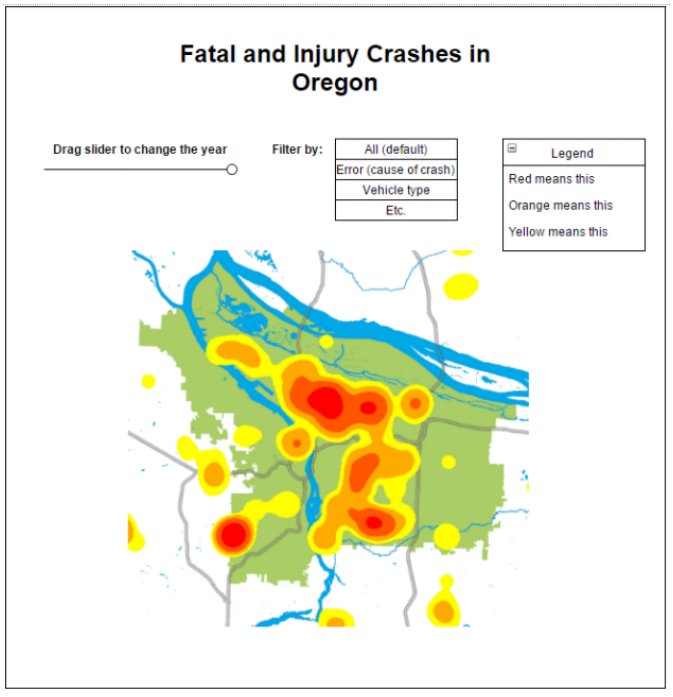

Fig. 3. Mockup of our proposed user interface for displaying data.