Chapter 30
An Interactive Tool for Visualizing and Analyzing Spatio-Temporal Data

Alican Turk
Sabanci University, Turkey

Selim Balcisoy
Sabanci University, Turkey

Burcin Bozkaya
Sabanci University, Turkey

ABSTRACT

Spatio-temporal data analysis has gained a degree of importance with the emergence of widespread location sensor technologies and applications generating data from activities such as location based GSM services and credit card transactions recorded by POS machines. By analyzing large-scale spatio-temporal data, companies can perform quality control for their services, and predict customer behavior. In this chapter, we introduce an interactive and configurable tool that effectively visualizes spatio-temporal data using two-dimensional cluster heat maps along with temporal histograms. We divide each dataset into geographical grids, cluster data within each grid, visualize clusters as heat maps, and finally calculate similarity scores between pairs of map images to help detect recurring patterns. We employ the tool in analyzing activity data of a GSM operator's friend-finder service and credit card transaction data of Akbank, a major bank in Turkey. We report examples of patterns observable using our tool, that are not otherwise observable.

INTRODUCTION

Visual exploratory data analysis, as first put forward by Tukey (1977), involved analyzing data sets to summarize their main characteristics via visual graphs. Data sets have become much larger ever since, and the means to visualize and analyze such data sets have also changed in order to adapt. According to a recent study by A.T. Kearney (2013), data generated all around the world is growing at an average

DOI: 10.4018/978-1-5225-0983-7.ch030

rate of 40% per year and is expected to reach nearly 45 zettabytes by 2020. Efficient visualization and thorough analysis of large-scale datasets containing spatio-temporal GPS (global positioning system) information has a great potential to draw accurate and credible conclusions for companies for developing present and future business strategies. Many companies today depend on large-scale data visualization and analysis to relate behavioral patterns of their customers with their products and services, and make business decisions based on the conclusions they draw.

Similar to the GPS technology, today's smartphones are also capable of collecting spatio-temporal data, which can be used for analytical purposes. An interesting example is a fall detection application that reports the location of an elderly person who has fallen and the severity of the fall in order to alert caregivers (Yavuz et al. 2010). Another example of data collection in smartphones involves monitoring user mobility during daily life in order to predict and provide contextual information about user mobility (Chon et al. 2011).

The interdisciplinary connection between the fields of visualization, cartography and geographic information systems (GIS) allows visualizing and exploring the geospatial data and the information derived from it. Analysis of geographic data in time and space is emerging as a very important subject with the increasing collection and use of location data, with many researchers looking for approaches to deal with the complexities of such datasets. Andrienko and Andrienko (2005) emphasize the need for sophisticated visualization techniques and analytical tools that facilitate spatio-temporal thinking. However, due to the complex nature of analyzing spatio-temporal data, current visualization and analysis tools are not fully efficient and they require improvements (Andrienko et al. 2010).

In this chapter, we introduce an interactive tool that allows analysts to effectively visualize spatio-temporal data in the form of two-dimensional heat maps and histograms. We also introduce a mathematical formula for measuring heat map similarities via image processing techniques, which help analysts manually detect patterns and anomalies in large-scale spatio-temporal datasets. We then compare our similarity scores with the existing method and other correlation-based formulas. The user-friendly and interactive interface of our tool encourages spatio-temporal thinking for users from all background levels. It allows users to clearly analyze large scale spatio-temporal data with its easy to understand visuals elements. The incorporated similarity formulas help users more quickly detect the existence of patterns or anomalies without having to repeatedly compare images pair by pair. Our tool adheres to the principle of maintaining a user-friendly interface that values clarity. Since not every user of the spatio-temporal data visualization tools can be a field expert, it is crucial for software tools like ours to contain comprehensible interfaces (Andrienko et al. 2010).

For the rest of our chapter, we present, In Section 2, a literature review of currently available techniques and tools for geo-visualization, spatio-temporal data analysis and image processing. In Section 3, we describe our mathematical model for similarity calculation and our heat map and histogram methodology. We also describe the structure of the proposed tool. In Section 4, we present the results we have obtained by applying our methodology on two datasets. Finally, in Section 5, we present our concluding remarks along with ideas for future research.

BACKGROUND

In this study, we primarily deal with visualization and exploratory analysis of spatio-temporal data. Our methodology derives from past work on image processing and especially the measurement of image