



DEEP
LEARNING
INSTITUTE



DLI Accelerated Data Science Teaching Kit

Lecture 1.4 - Learning Goals and Expectations



The Accelerated Data Science Teaching Kit is licensed by NVIDIA, Georgia Institute of Technology, and Prairie View A&M University under the [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).

Learning Goals

- Learn **visual** and **computation** techniques and use them in **complementary** ways
- Gain a **breath** of knowledge
- Learn **practical** know-how by working on **real data & problems**

Topics Covered

(Analytics Building Blocks)

Collection

Cleaning

Integration

Analysis

Visualization

Presentation

Dissemination

Building Blocks. **NOT Rigid Steps.**

Collection

Cleaning

Integration

Analysis

Visualization

Presentation

Dissemination

Can skip some

Can go back (two-way street)

- Data types inform visualization design
- Data informs choice of algorithms
- Visualization informs data cleaning (dirty data)
- Visualization informs algorithm design (user finds that results don't make sense)

“Caveats” for Students

- Learning data science can involve **a lot of programming**
- You need to be prepared to **learn many things** in short amount of time, e.g., visualization, JavaScript, CSS, Scala
- **Very common in industry**

From Previous Offerings ...

The course materials of *CSE 6242 Data and Visual Analytics* at Georgia Tech are based on this teaching kit

- Class projects had turned into research publications, e.g., KDD, IUI
- Projects as portfolio pieces on CV
- Increased job and internship opportunities
- Receive numerous “thank you” notes from former students

Aurigo: An Interactive Tour Planner for Personalized Itineraries

Alexandre Yahia*, Antoine Chassang*, Louis Raynaud*, Hugo Duthil*, Duen Horng (Polo) Chau

Georgia Institute of Technology

{alexandre.yahia, antoine.chassang, l.raynaud, hduthil, polo}@gatech.edu

ABSTRACT

Planning personalized tour itineraries is a complex and challenging task for both humans and computers. Doing it manually is time-consuming; approaching it as an optimization problem is computationally NP hard. We present Aurigo, a tour planning system combining a recommendation algorithm with interactive visualization to create personalized itineraries. This hybrid approach enables Aurigo to take into account both quantitative and qualitative preferences of the user. We conducted a within-subject study with 10 participants, which demonstrated that Aurigo helped them find points of interest quickly. Most participants chose Aurigo over Google Maps as their preferred tools to create personalized itineraries. Aurigo may be integrated into review websites or social networks, to leverage their databases of reviews and ratings and provide better itinerary recommendations.

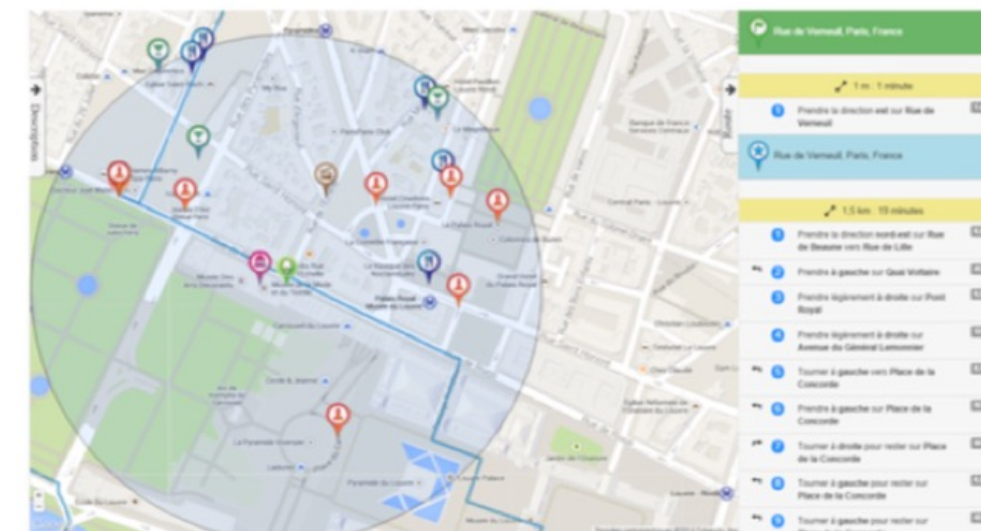
Author Keywords

User Interfaces; Visualization; Recommendation; Tour itinerary planning

ACM Classification Keywords

IUI'15 Full conference paper

(e.g. HCI): User interfaces



ISPARK: Interactive Visual Analytics for Fire Incidents and Station Placement

Subhajit Das, Andrea McCarter, Joe Minieri, Nandita Damaraju, Sriram Padmanabhan, Duen Horng (Polo) Chau
Georgia Tech
Atlanta, GA, USA
{das, andream, jminieri, nandita, sriramp, polo}@gatech.edu

ABSTRACT

In support of helping to reduce the response time of fire-fighters, and thus deaths, injuries, and property loss due to fires, we introduce ISPARK. The ISPARK system determines where fire stations should be located, analyzes the primary causes of fires, the existing infrastructure, and response times, by using visualizations which show the GIS mapping of fire stations on a dashboard. Incidents and response times are shown as additional layers, with clustering of fire incidents to determine predicted fire station locations, forecasting of fire incidents using regression, causal, infrastructure, and personnel analysis, creating an interactive, multi-faceted method for locating fire stations. A comparison of urban and rural fire incident response times is another dimension of this study. We demonstrate ISPARK's usage and benefits using a publicly available dataset describing 300,000 fire incidents in the states of Massachusetts and Maine. ISPARK is generalizable to other geographic areas

KDD'15 Workshop paper



Figure 1: Screenshot of ISPARK showing actual (pink) and predicted (green) fire station locations in Maine determined by our approach, using coordinates with actual driving distances from fire stations to actual fire incidents. Fire incidents are shown as small yellow dots. ISPARK reduces the average

PASSAGE: A Travel Safety Assistant With Safe Path Recommendations For Pedestrians

Matthew Garvey

College of Computing
Georgia Institute of Technology
Atlanta, GA 30332, USA
mgarvey6@gatech.edu

Meghna Natraj

College of Computing
Georgia Institute of Technology
Atlanta, GA 30332, USA
mnatraj@gatech.edu

Nilaksh Das

College of Computing
Georgia Institute of Technology
Atlanta, GA 30332, USA
nilakshdas@gatech.edu

Bhanu Verma

College of Computing
Georgia Institute of Technology
Atlanta, GA 30332, USA
bhanuverma@gatech.edu

Jiaxing Su

College of Engineering
Georgia Institute of Technology
Atlanta, GA 30332, USA
Jiaxingsu@gatech.edu

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

IUI'16 Poster paper

Abstract

Atlanta has consistently ranked as one of the most dangerous cities in America with over 2.5 million crime events recorded within the past six years. People who commute by walking are highly susceptible to crime here. To address this problem, our group has developed a mobile application, PASSAGE, which uses crime data to find "safe paths" for pedestrians in Atlanta. The application uses a user interface to allow users to input their starting and ending points and recommends a safe path based on the crime data.

Authors

Safe Path
Pulse

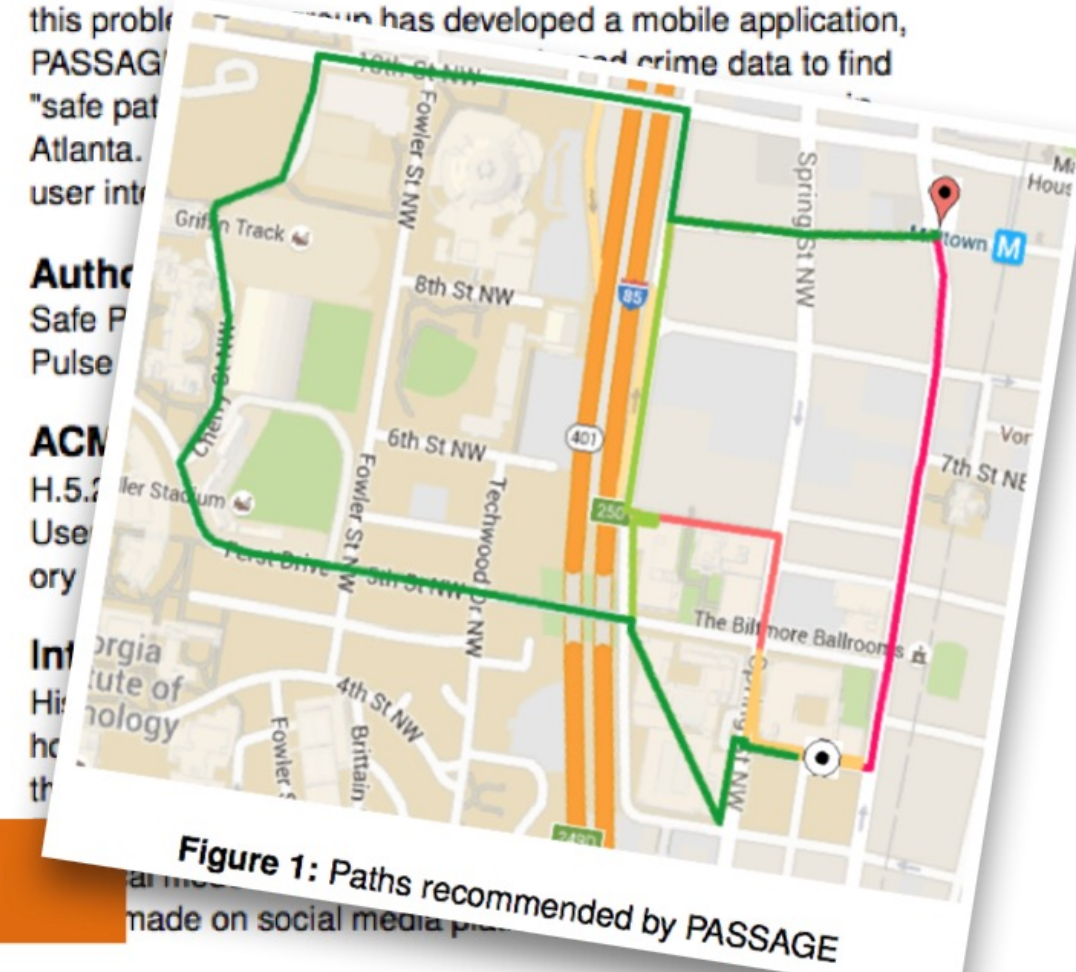
ACM

H.5.2
User
ory

Int

Georgia
Institute of
Technology

Figure 1: Paths recommended by PASSAGE



Student Reflections

*“I feel like the concepts from your class are like a **rite of passage for an aspiring data scientist**. Assignments lead to a feelings of accomplishment and truly progressing in my area of passion.”*

*“I really get more intuition about how to **deal with data with some powerful tools in HW3** [uses AWS]. That feeling is beyond description for me.”*

*“I would like to say thank you for your class! Thanks to the skills I got from the class and the project, **I got the offer.**”*



DEEP
LEARNING
INSTITUTE



DLI Accelerated Data Science Teaching Kit

Thank You