

<http://poloclub.gatech.edu/cse6242>

CSE6242 / CX4242: **Data** & **Visual** Analytics

# Group project

# Heilmeier questions

# Using existing libraries/code

Duen Horng (Polo) Chau

Associate Professor, College of Computing

Associate Director, MS Analytics

Machine Learning Area Leader, College of Computing

Georgia Tech

Partly based on materials by

Professors Guy Lebanon, Jeffrey Heer, John Stasko, Christos Faloutsos

# Forming Teams

You're welcome to look for teammates **NOW**.  
**See deadline on course schedule.**

- **4-6 people** in each team
- Polo **does not** dictate who teams with whom (since each team decides on their topic)
- Try Piazza's **Search for Teammates** pinned post
- Polo suggests teaming with students in the **same section**, but OK to mix if really needed. See "Teaming" section on project description:

# Example Projects

- Two example projects next week
  - **Firebird:** Predicting Fire Risk in Atlanta
  - **PASSAGE:** Safe Path Recommendation
- Q&A after presentations

# Challenges in Larger Teams

- Can we form a 7-person team?
  - I highly recommend **not** to. Only with my permission.
  - **Larger groups are harder to run.**
  - **Higher coordination, motivation and intellectual costs**

<https://www.cmu.edu/teaching/design/teach/design/instructionalstrategies/groupprojects/challenges.html>

# Tips for Working Successfully in a Group

<https://www.cs.cmu.edu/~pausch/Randy/tipoForGroups.html>

1. Meet people properly
2. Find things you have in common
3. Make meeting conditions good
4. Let everyone talk
5. Check your egos at the door
6. Praise each other
7. Put it in writing
8. Be open and honest
9. Avoid conflict at all costs
10. Phrase alternatives as questions



# Requirements

## 3 core requirements

1. **Large real** dataset  
(see project description for rule-of-thumb meaning of “large”)
2. **Non-trivial** analysis/algorithms/computation
3. An **interactive user interface** that **interact** with the algorithms

## Grading & Schedule

- Proposal: writeup + in-class presentation
- Progress report (mostly as a “checkpoint”)
- Final report: writeup + poster presentation

# How to Come Up with Project Ideas?

## Polo's recommendations

- Work on something that you are **excited** about (e.g., NetProbe for eBay fraud detection)
- Is it **interesting**? e.g., computationally? visualization-wise?
- Is it **impactful**? (e.g., save lives? fight crime? shorten commute? save resources?)
- Work on something **interestingly challenging**, so you will **learn more**
- Browse the list of datasets on course homepage



**George Heilmeier**  
Former Director of DARPA



# Heilmeier Questions

Preflight checklist for successful projects

1. **What** are you trying to do?  
Articulate your objectives using absolutely no jargon.
2. **How** is it done today; what are the **limits of current practice**?
3. **What's new** in your approach; **why** it will be successful?
4. **Who** cares?
5. If you're successful, **what difference and impact** will it make?  
How do you measure them (e.g., via user studies, experiments, groundtruth data, etc.)?
6. What are the **risks and payoffs**?
7. **How much** will it cost?
8. **How long** will it take?
9. What are the midterm and final "exams" to **check for success**?

[http://en.wikipedia.org/wiki/George\\_H.\\_Heilmeier](http://en.wikipedia.org/wiki/George_H._Heilmeier)

<http://smlv.cc.gatech.edu/2010/10/17/heilmeiers-questions/>

# Using Existing Libraries/Code

## **GPL** (General Public License)

In purely private (or internal) use —with no sales and no distribution— the software code may be modified and parts reused without requiring the source code to be released. **For sales or distribution, the entire source code need to be made available to end users**, including any code changes and additions— in that case, **copyleft** is applied to ensure that end users retain the freedoms defined above.

# 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).

## 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, we have developed a mobile application, PASSAGE, that uses real-time 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 to view recommended paths.

## Author

Safe Path  
Pulse

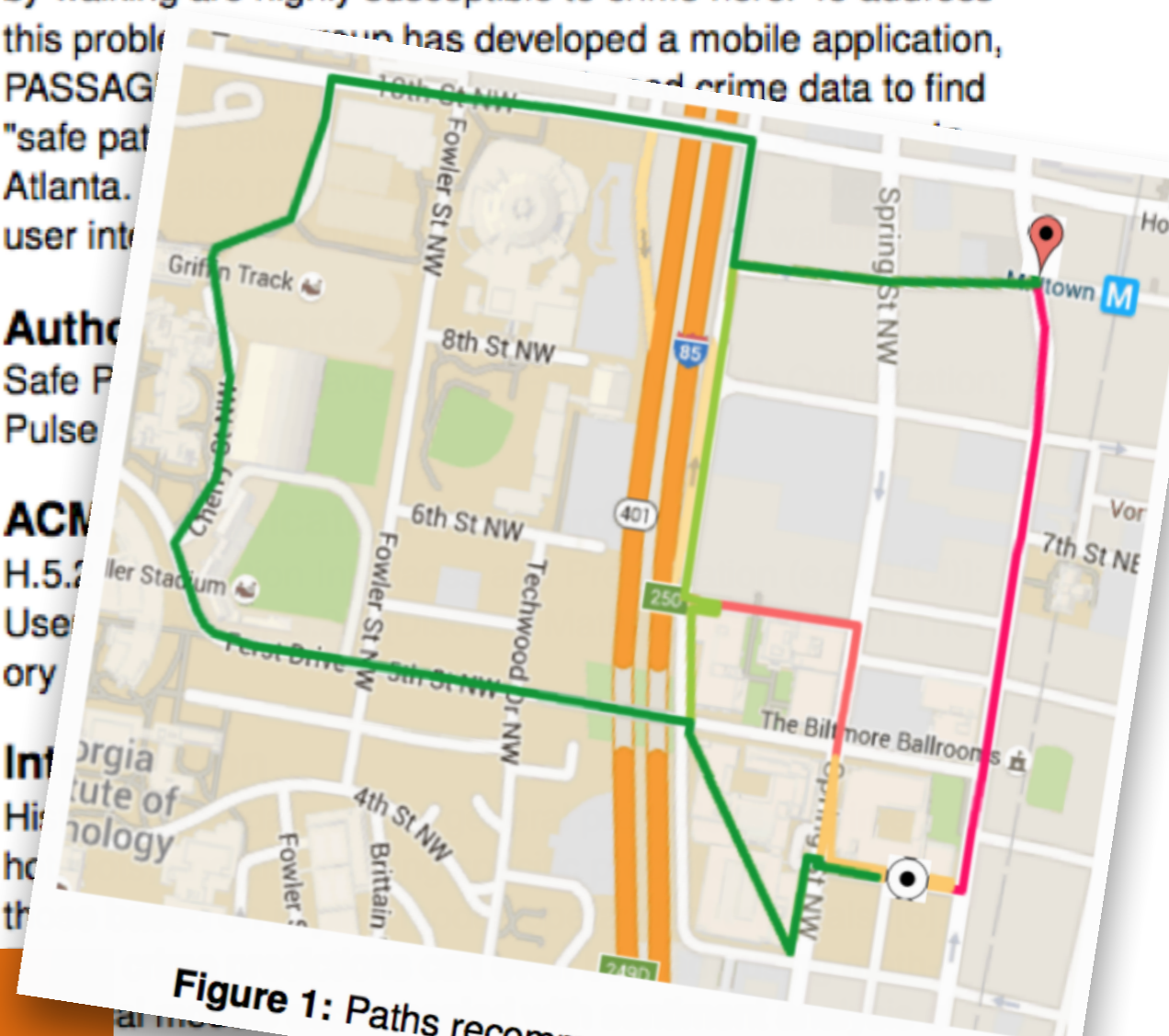
## ACM

H.5.2  
User-Centered  
ory

## Int

High  
h  
th

Figure 1: Paths recommended by PASSAGE



# Aurigo: An Interactive Tour Planner for Personalized Itineraries

Alexandre Yahi\*, Antoine Chassang\*, Louis Raynaud\*, Hugo Duthil\*, Duen Horng (Polo) Chau  
Georgia Institute of Technology  
{alexandre.yahi, 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.

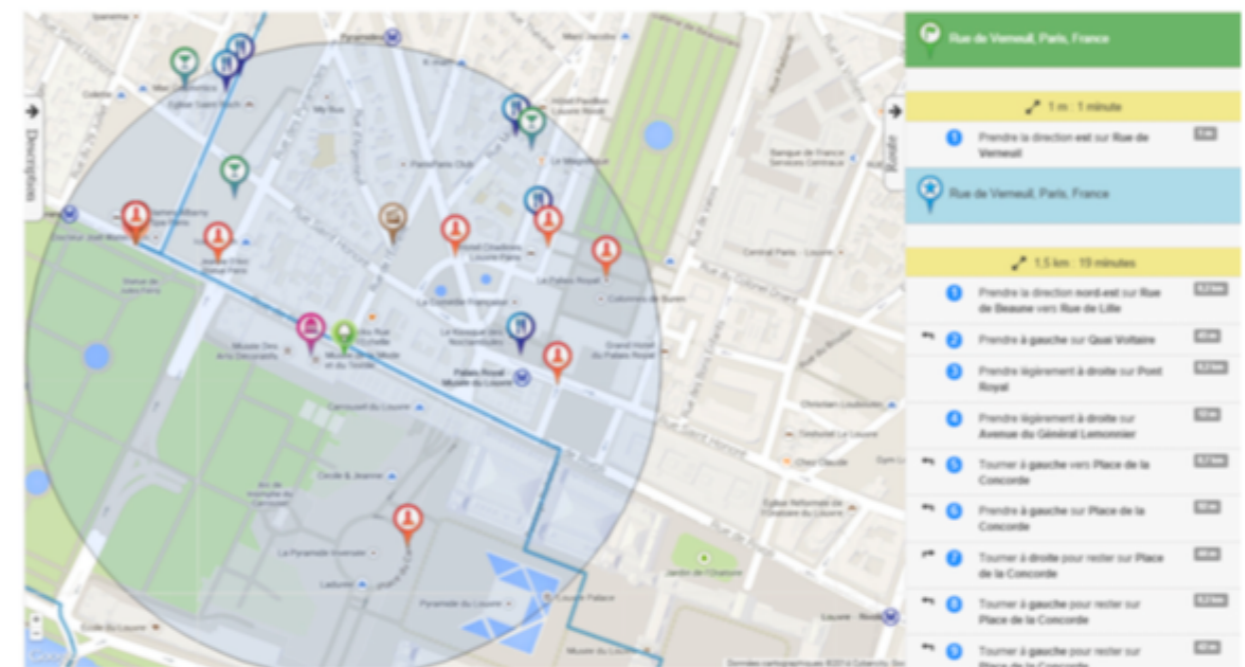
## Author Keywords

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



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