The Unbiased & Interactive View of Tourism Data to Tourists for Making Informed Decisions

(Project Proposal – Team #31)

Xiaoqi Zheng
zhengxiaoqi@gatech.edu

Brian Ng
brian.ng@gatech.edu

Yiufung Leung
yleung8@gatech.edu

Si-Onn Kwok
skwok8@gatech.edu

Youjung Kim
ykim691@gatech.edu

Ankur Agarwal
ankur74@gatech.edu

ABSTRACT

Making sense of the vast amount of online digital data is always challenging, even more so when a tourist starts planning her/his next vacation. Travel web-sites and social media forums either have commercial interests or biased personal opinions presented as advice. We propose to provide prospective tourists with an unbiased view of tourism related data through an interactive web-based platform for making informed travel decisions.

1. INTRODUCTION

Tourism is one of the fastest growing industries impacting multiple socio-economic dimensions such as transportation, hospitality, local arts and craft and related e-commerce. Any individual can be classified as tourist when she/he voluntarily leaves normal surroundings of residence to visit another environment. The internet has had a major impact on tourism both for the service providers and the consumers. Tourists today have access to a vast amount of digital information distributed across the internet on commercial, public and government travel-websites, social media-websites and personal or public blogs/forums.

However, the information available on these digital platforms is usually biased based on the underlying intent. We propose to provide tourists access to the tourism data through an unbiased visual and interactive interface. The proposed solution will source data available through freely available public APIs and aggregate the information in five major categories viz, tourist attractions, overall budget (subdivided into travel, lodging and food), social media references, seasonal tourist density and government/publicly issued travel advisories.

2. RELATED WORK
2.1 Internet's Impact on Tourism

After proliferation of Internet usage, travel and tourism section has seen dramatic changes in its business model [1][2] and also in the form of user generated travel contents through blogs and review [5][6]. The consumer has access to travel related information, competitive price offers while sitting in the convenience of her/his home. The downside of this convenience is the biased view of the data presented by these commercial sites based on consumer behavior [3]. Our proposed solution intends to present data in an unbiased form that

can be directly consumed by a tourist through interactive interface.

2.2 Big Data Analytics

Tourists today do not want to be a passive consumer; they seek for the co-creation and personized services. Big data analytics allows to push classic electronic consumer relationship management into new levels [7] [8] [9].

Ping et al. [10] proposes tourism demand forecasting hybrid system combining FCM clustering method with logarithm least-squares support vector regression with genetic algorithms (LLS-SVRGA).

Ulrich et al. [11] presents a comparison of predictive accuracy between univariate and multivariate models. While another method [12] proposes an ANN-based forecasting method named C-C-ANN that considered the sequence patterns and near characteristics of a data series.

Woodside et al. [13] explores relationship between values and holiday preferences through questionnaires. The result shows correlations between individual values and types of destinations. The finding is useful to us in providing an evidence in the connection between user characteristic and holiday preference.

Joshi et al. [14] describes a travel system meeting tourism need in India. We will further expand the idea globally by looking at global data source such as Expedia.

Stefanidis et al. [15] introduce the paradigm shift in the Web from a document-centric infrastructure to an entity-centric organization of data. In a recommendation system, entity-centric data is required.

Lim, K.H., Chan, J., Leckie, C. et al. [16] present in their paper same problem set as our project but in a different approach. Their paper is focused on concepts of ideas and implementation of recommendation algorithm, while our project will mostly focus on data collection and visualization.

2.3 Social Media and Tourism

Matsuo et al. [17] proposes two ideas for tourism informatics, which inspires the constitute of our project backend system. The first one is to make use of geotagged user-generated content, which can replace tourism information manual maintenance by collecting location-based data from millions of Tweeter users. The second one is using data from Foursquare for tourist spot categorization. We will look at another social platform Reddit, extracting tourist information through its community, comment, and voting, to deduce the popularity and potential target audience.

Teare, Richard et al. [18] highlights the importance of the use of user-generated content from social media from the aspects of millennials growth, mobile's popularity, dynamic pricing and marketing shift. While it further roadmaps our development with the use of social media and any sites with UGC like Expedia, we will also look at statistic data from officially recognized international source to avoid commercial leaning. World Bank is the major source we will look at for deducing density of a country.

3. HEILMEIER QUESTIONS

3.1 What are you trying to do?

With the world getting smaller by the internet, travelling around the world is getting much easier. Although many trip planner applications are available freely, there is a lack of approaches in unbiased data visualization. Our objective is to collect information through various sources and provide interactive view of the data to tourists that can is customizable to their personal preferences.
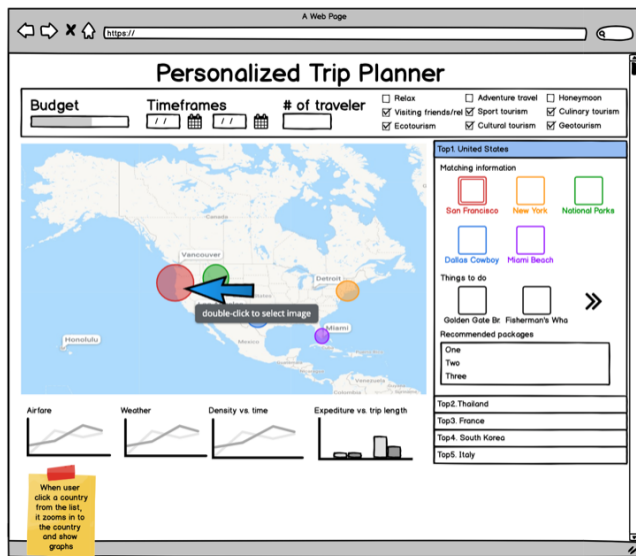
Figure 1. The proposed trip planner will present data from multiple sources in an interactive environment.

3.2 How is it done today?

Current tourism websites are information systems where users have to read and consolidate information from different sources in order to make an informed decision. These websites, although account for personal preference, introduce commercial bias when making recommendations. They also tend to introduce a small bias towards the preference of people who are more likely to be vocal on such platforms.

3.3 What's new in your approach?

Dynamic data interaction coupled with user input of their preferences such as budget, preferred activities, and pace of travel, the recommendation system we are trying to implement would present with personalized vacation options. Also, running sentiment analysis on social media platforms such as Reddit on potential destinations would help aggregate what users in general think about the destination instead of relying on ratings and reviews on travel websites.

3.4 Who cares?

Since the proposed solution is not limited to a geographical region and everyone is a tourist at some point, the proposed solution would appeal to an universal consumer base.

3.5 If you're successful, what difference and impact will it make, and how do you measure them?

Tourists can better plan their trips according to their unique demands and save time and money while getting an opportunity to find a travel destination closer to their personal preferences. The measure of success will be through review by volunteers and quantitative and qualitative feedbacks.

3.6 What are the risks and payoffs?

Complexity in data processing and data association sourced through multiple sources. Most of the commercial sites do not offer free APIs to educational/personal projects and may require data scrapping from these sites.

3.7 How much will it cost?

The project intends to utilize opensource and freely available data sources and APIs. We will also put the site in free hosting site for development so there's no additional investment in first phase.

3.8 How long will it take?

The project will follow the timelines assigned in the course and complete before the deadline set for final submission.

3.9 What are the midterm and final "exams" to check for success?

Our measure of success will focus on the voice from outside. For the midterm, we will conduct live review of draft version of our proposed solution and collect feedback via a short survey. The final version after incorporating some of the functionalities and suggestions received during

midterm review would be presented again to the same group of volunteers and also to a small percentage of first time users. The qualitative and quantitative feedback would constitute as measure of success.

4. Plan of Activities

Task	1/27	2/3	2/10	2/17	2/24	3/3	3/10	3/17	3/24	3/31	4/7	4/14	4/19
Project Brain Storming	▨	▨	▨										
Research		▨	▨	▨	▨	▨							
Proposal presentation					■								
Server setup				▨	▨								
User Requirement				▨	▨								
Design					▨	▨							
Progress Report								■					
Backend programming					▨	▨	▨	▨	▨	▨			
Frontend programming					▨	▨	▨	▨	▨	▨			
Integration										▨	▨		
Testing											▨	▨	▨
Final Presentation													■

Table 1. Proposed plan of activities for the project is shown above. The key milestones are marked in black colour.

5. Project Task Arrangement

Name	Time Allocation (hrs)	Tasks Assigned
Ankur Agarwal	120	Research, Architecture, Report
Brian Ng	120	Research, Application
Si-Onn Kwok	120	Research, Recommendation System
Xiaoqi Zheng	120	Research, Front End
Yiufung Leung	120	Research, Social Media Analysis

Youjung Kim	120	Research, Backend
Total	720	

Table 2. Proposed task allocation and approximate time estimates assigned to each member of the project team.

REFERENCES

[1] Wirtz, B. W., Schilke, O., & Ullrich, S. (2010). Strategic development of business models: Implications of the web 2.0 for creating value on the internet. *Long Range Planning*, *43*(2-3), 272-290

[2] Standing, Craig & Tang-taye, Jean-pierre & Boyer, Michel. (2014). The Impact of the Internet in Travel and Tourism: A Research Review 2001–2010. Journal of Travel & Tourism Marketing. 31. 82-113

[3] Scott A. Cohen, Girish Prayag & Miguel Moital (2014) Consumer behaviour in tourism: Concepts, influences and opportunities, Current Issues in Tourism, 17:10, 872-909

[4] Camilleri, M. A. (2018). The Tourism Industry: An Overview. In Travel Marketing, Tourism Economics and the Airline Product (Chapter 1, pp. 3-27)

[5] Huang, Ching-Yuan & Chou, chia-jung & Lin, Pei-Ching. (2010). Involvement theory in constructing bloggers' intention to purchase travel products. Tourism Management. 31. 513-526.

[6] Brey, E. T., So, S. I., Kim, D. Y., & Morrison, A. M. (2007). Web-based permission marketing: segmentation for the lodging industry. Tourism Management, 28(6), 1408–1416

[7] Leal, Fátima & Matos Dias, Joana & Malheiro, Benedita & Burguillo, Juan. (2016). Analysis and Visualisation of Crowd-sourced Tourism Data. 98-101.

[8] Mosch, M & Frech, I & Schill, Christian & Koch, Barbara & Stelzl, H & Almer, Alexander & Schnabel, Thomas. (2004). Tourism information based on visualisation of multimedia geodata – ReGeo

[9] Kádár, Bálint & Gede, Mátyás. (2013). Where Do Tourists Go? Visualizing and Analysing the Spatial Distribution of Geotagged Photography. Cartographica The International Journal for Geographic Information and Geovisualization. 48. 78-88.

[10] Pai, Ping-Feng & Hung, Kuo-Chen & Lin, Kuo-Ping. (2014). Tourism Demand Forecasting Using Novel Hybrid System. Expert Systems with Applications. 41. 3691–3702.

[11] Gunter, Ulrich & Onder, Irem. (2015). Forecasting international city tourism demand for Paris: Accuracy of uni- and multivariate models employing monthly data. Tourism Management. 46. 123-135.

[12] Jun, Wang & Luo, Yuyan & Lingyu, Tang & Peng, Ge. (2018). Modeling a combined forecast algorithm based on sequence patterns and near characteristics: An application for tourism demand forecasting. Chaos, Solitons & Fractals. 108. 136-147.

[13] Woodside, Arch G., and Kozak, Metin, eds. 2014. Tourists' Behaviors and Evaluations. Bingley: Emerald Publishing Limited. Accessed March 1, 2019. ProQuest Ebook Central.

[14] V. B. Joshi and R. H. Goudar, "Searching, categorizing and tour planning: A novel approach towards e-tourism," *2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, Bangalore, 2017, pp. 1002-1005

[15] Kostas Stefanidis, Vasilis Efthymiou, Melanie Herschel, and Vassilis Christophides. 2014. Entity resolution in the web of data. In Proceedings of the 23rd International Conference on World Wide Web (WWW '14 Companion). ACM, New York, NY, USA, 203-204

[16] Lim, Kwan Hui & Chan, Jeffrey & Leckie, Christopher & Karunasekera, Shanika. (2018). Personalized trip recommendation for tourists based on user interests, points of interest visit durations and visit recency. Knowledge and Information Systems. 54. 375-406

[17] Matsuo, Tokuro & Hashimoto, Kiyota & Iwamoto, Hidekazu. (2015). Tourism Informatics: Towards Novel Knowledge Based Approaches.

[18] Teare, Richard, Bowen, John, and Baloglu, Seyhmus. 2015. What is the current and future impact of social media on hospitality and tourism?. Bingley: Emerald Publishing Limited. Accessed March 2, 2019. ProQuest Ebook Central.