The SIGSPATIAL Special

Newsletter of the Association for Computing Machinery
Special Interest Group on Spatial Information

Volume 10 Number 3 November 2018

The SIGSPATIAL Special

The SIGSPATIAL Special is the newsletter of the Association for Computing Machinery (ACM) Special Interest Group on Spatial Information (SIGSPATIAL).

ACM SIGSPATIAL addresses issues related to the acquisition, management, and processing of spatially-related information with a focus on algorithmic, geometric, and visual considerations. The scope includes, but is not limited to, geographic information systems.

For more details and membership information for ACM SIGSPATIAL as well as for accessing the newsletters please visit http://www.sigspatial.org.

The SIGSPATIAL Special serves the community by publishing short contributions such as SIGSPATIAL conferences' highlights, calls and announcements for conferences and journals that are of interest to the community, as well as short technical notes on current topics. The newsletter has three issues every year, i.e., March, July, and November. For more detailed information regarding the newsletter or suggestions please contact the editor via email at azufle@gmu.edu.

The SIGSPATIAL Special (ISSN 1946-7729) Volume 10, Number 3, November 2018.

Table of Contents

Message from the Editor

Andreas Züfle
Department of Geography and GeoInformation Science, George Mason University, USA
Email: azufle@gmu.edu

Dear SIGSPATIAL Community,

The newsletter serves the community by publishing short contributions such as SIGSPATIAL conferences' highlights, calls and announcements for conferences and journals that are of interest to the community, as well as short technical notes on current topics. In this November 2018 issue, the first section continues the tradition of highlighting the top papers selected for the 3rd ACM SIGSPATIAL Student Research Competition (SRC) held at the ACM SIGSPATIAL Conference 2018. Special thanks to the chairs of the student research competition 2018: Kyriakos Mouratidis and Fusheng Wang.

The second section consists of workshop reports of workshops that were co-located with SIGSPATIAL 2018 and held on November 6th 2018:

- 7th ACM SIGSPATIAL Workshop on Analytics for Big Geospatial Data (BigSpatial 2018)
- 4th ACM SIGSPATIAL Workshop on Emergency Management using GIS (EM-GIS 2018)
- 2nd ACM SIGSPATIAL Workshop on AI for Geographic Knowledge Discovery (GeoAI 2018)
- 2nd ACM SIGSPATIAL Workshop on Geospatial Humanities (GeoHumanities 2018)
- 12th ACM SIGSPATIAL Workshop on Geographic Information Retrieval (GIR 2018)
- 2nd ACM SIGSPATIAL Workshop on Analytics for Local Events and News (LENS 2018)
- 2nd ACM SIGSPATIAL Workshop on Recommendations for Location-based Services and Social Networks (LocalRec 2018)
- 2nd ACM SIGSPATIAL Workshop on Prediction of Human Mobility (PredictGIS 2018)
- 1st ACM SIGSPATIAL Workshop on Geospatial Simulation (GeoSim 2018)

I would like to sincerely thank all Student Research Competition authors and event organizers for their generous contributions of time and effort that made this issue possible. I hope that you will find the newsletters interesting and informative and that you will enjoy this issue.

You can download all Special issues from:

`http://www.sigspatial.org/sigspatial-special`

Yours sincerely,
Andreas Züfle
SIGSPATIAL Newsletter Editor

The SIGSPATIAL Special

Section 1: The 3rd ACM SIGSPATIAL Student Research Competition (SRC)

ACM SIGSPATIAL
http://www.sigspatial.org

Summit: A Scalable System for Massive Trajectory Data Management

Louai Alarabi
PhD Student, University of Minnesota, Minneapolis, MN, USA
Email: louai@cs.umn.edu, ACM Member Number: 5798098, SRC: Graduate, Advisor: Mohamed F. Mokbel

ABSTRACT

MapReduce frameworks, e.g., Hadoop, have been used extensively in different applications that include machine learning, and spatial processing. In meantime, huge volumes of spatio-temporal trajectory data are coming from different sources over sometime, raised the demand to exploit the efficiency of Hadoop, coupled with the flexibility of the MapReduce framework, in trajectory data processing. This work describes Summit; a full-fledged MapReduce framework with native support for trajectory data.

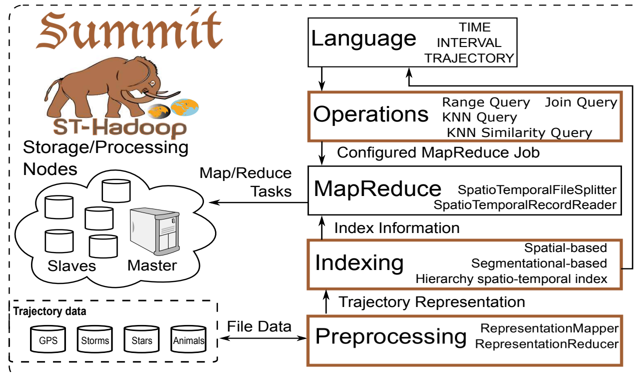


Figure 1: ST-Hadoop system architecture

1 INTRODUCTION

The importance of processing trajectory data is growing with the emerging and popularity of applications that produce them in large-scale [3]. For example, NASA generates over 4-TB of stars and asteroids moving in the space on a daily basis. Sloan Digital Sky project collects over 156 TB from millions of outer-space objects. MoveBank project archives more than 20 years of habitat trajectory data. New York City Taxi and Limousine Commission record over 1.1 Billion trajectories. National Hurricane Center stores comprehensive details of all storms' movement every year. Brain study in neuroscience model neuron fiber as a trajectory that creates petabytes of data. The explosive increase in data volumes raises the demand for managing and analyzing these mass archives of trajectories on big distributed platforms.

Scaling trajectory data received more attention, especially in utilizing big distributed platforms. The latest efforts for processing trajectory data are built on the top of big distributed frameworks, such as on-top of Hadoop [4], or *Heterogeneous* multiple platforms [2]. Using general purpose framework as-is will result in sub-performance for trajectory applications that require indexing, mainly because they store data as non-indexed heap files. For example, a most recent research study investigated the *k*NN join query on Hadoop employed five isolated map-reduce jobs to execute a single *k*NN join operation without indexing trajectory [4]. In our proposed Summit we achieve orders of magnitude better performance when spatio-temporally indexing trajectories.

Exploit the efficiency of Hadoop, coupled with the flexibility of the MapReduce framework for processing trajectory raised many challenges. Some of the most significant challenges of processing trajectory data is the inability of Hadoop to preserve the spatio-temporal locality, load balancing efficiency, and the capability to support various trajectory operations. We propose Summit system that overcome all these challenges by spatio-temporally loads and partitions trajectory data. We equipped Summit with fundamental operations such as Range, *k*NN, Similarity, and Join queries. Summit is an extension of Hadoop framework that built-in the spatio-temporal locality of trajectory in the base code of three layers inside ST-Hadoop [1], namely preprocessing, indexing, and operation layers. The key point behind the performance gain of Summit is the idea of indexing, where data are spatiotemporally loaded and divided as trajectory segments across computation nodes.

2 FRAMEWORK OVERVIEW

Figure 1 gives the high level architecture of our Summit; with a *built-in* support for trajectory data type and operations. Summit cluster contains one master node that breaks up a map-reduce job into smaller tasks, carried out by slave nodes. Summit adopts a layered design of ST-Hadoop briefly described below:

Language Layer: This layer provides a simple high-level SQL-like language that supports trajectory data types (i.e., TRAJECTORY, which described as a consecutive sequence of spatio-temporal points `ST-POINT`. Operations contains (e.g., OVERLAP and KNN).

Preprocessing Layer: This layer is responsible for reconstructing the raw representation of objects into trajectory segments, where each contains a continuous sequence of spatio-temporal points.

Indexing Layer: Summit employs a two levels index structure of *global* and *local* indexing. The *global* index partitions the data across the computation nodes, while the *local* index organizes the data inside each node. Space and time of trajectories are taken into consideration in each level.

MapReduce Layer: The primary task of this layer is to exploit the global and the local indexes, respectively, for data pruning. Summit utilized ST-Hadoop implementations to the MapReduce layer, and thus, it is not going to be discussed any further in this paper.

Operations Layer: This layer encapsulates the implementation of three common trajectory operations, namely, range query, *k*NN, and join queries. More operations can be added to this layer.

3 TRAJECTORY INDEXING

Summit organizes input files in the Hadoop Distributed File System (**HDFS**) in a way that preserves the spatio-temporal geometrical shape of trajectory. Hence, incoming trajectory operations can have minimal data access to retrieve the query answer, reduce the computation complexity, and allow applications to run more sophisticated operations on the entire trajectory or sub-trajectories; thus, more in-depth information gained. Summit indexes trajectory through the following four consecutive phases:

1. Preprocessing: Summit triggers this process to converts the trajectory into a segment representation. Each segment consists of a consecutive sequence of points in geographical space and time. This phase is necessary to assemble trajectories information in a singleton representation.

2. Sampling: The objective of sampling is to approximate the trajectory distribution and ensure the quality of partitioning. Due to the mass volume of data, Summit scans a representative sample that fit-in the main memory of the master node.

3. Bulkloading Partitioning: Summit manipulates the sample to constructs two-level indexing of temporal and spatial, respectively. Summit applies the temporal partitioning already equipped in ST-Hadoop to partitions the temporal dimension of trajectories into either equi-width or equi-depth [1]. As for the spatial level of indexing, Summit employs space or data partitioning algorithms for every temporal interval, namely *Spatial-based* or *Segmentational-based*. Figures 2 illustrates the logical design of both techniques. Rectangles represent the boundaries of the HDFS partitions. Dots and lines depict the trajectory information. Tables below list the contents of each partition. Described as follow:

• *Spatial-based*: This partitioning preserves the spatio-temporal locality closeness between sub-trajectories. The boundaries of the HDFS partition cut trajectory connectivity as shown in figure 2a. This organization of trajectory assists basic operations, such as range and join queries.

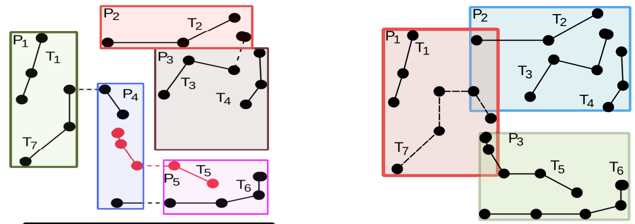
• *Segmentational-based*: This guarantee that the full information of nearby trajectories is organized in a single HDFS block, as shown in figure 2b. This technique preserves the spatio-temporal locality and shapes of trajectories. Such partitioning is more in favor for operations that not only need to process the spatio-temporal locality of trajectories but also their semantic or shapes over time, such as Similarity *k*NN query.

4. Physical Assigning: The objective of this phase is to scan through the whole data and assign each record according to the layout constructed from the previous phase.

4 CONTRIBUTION AND RESULT

The partitioning in Summit is the key feature of its superior performance over Hadoop. Preliminary experiments conducted on New York taxi dataset, to show that Summit achieves orders of magnitude higher job throughput. We present three case studies of trajectory operations that utilize Summit indexing, namely, range, *k*NN point-based, and *k*NN similarity-based queries.

• **Range query**: Given a spatio-temporal query predicates, the query retrieves all trajectories that belong to the query region in both space and time. For example, "*find taxi in downtown Seattle during time interval between January and March 2018.*



Partition	Trajectory data
P_1	T_1, T_7
P_2	T_2, T_3, T_4
P_3	T_3, T_4
P_4	T_7, T_5, T_6
P_5	T_5, T_6

(a) Spatial-based

Partition	Trajectory data
P_1	T_1, T_7
P_2	T_2, T_3, T_4
P_3	T_5, T_6

(b) Segmentational-based

Figure 2: Trajectory Partitioning

• ***k*NN point-based**: Given a query point and time interval, the goal is to find the k nearest trajectories to a point during some time interval. For example, "*Find the closest four animals to a Minnehaha waterfall between August and September*".

• ***k*NN Similarity-based**: Given a query trajectory, the objective is to find the *k*NN to the whole trajectory points for every time instance according to some aggregate or similarity function, such as MinMax. For example, "*Find the similar kNN Taxi to a given trajectory* Trj".

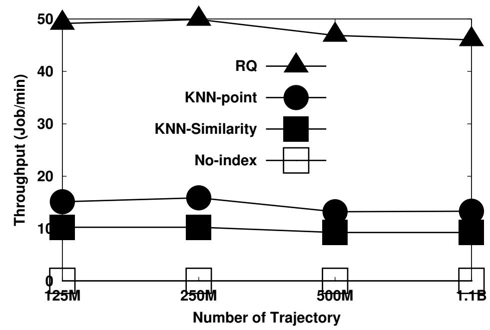


Figure 3: Preliminary Results

In figure 3, we increase the number of moving objects from 125 Million to 1.1 Billion, while measuring the job throughput per minute. We average the execution time of 20 randomly submitted queries for each operation. Summit index has more than two orders of magnitude higher throughput than on-top of Hadoop implementations, consistently, due to the early pruning employed by its index. We run the range and *k*NN point-based queries on the spatial-based trajectory partitioning. Meanwhile, we utilize the segmentational-based for *k*NN similarity query.

REFERENCES

[1] Louai Alarabi, Mohamed F. Mokbel, and Mashaal Musleh. 2018. ST-Hadoop: a MapReduce Framework for Spatio-temporal Data. *GeoInformatica* 22, 4 (2018), 785–813. https://doi.org/10.1007/s10707-018-0325-6

[2] Jie Bao, Ruiyuan Li, Xiuwen Yi, and Yu Zheng. 2016. Managing massive trajectories on the cloud. In *SIGSPATIAL*.

[3] Xin Ding, Lu Chen, Yunjun Gao, Christian S. Jensen, and Hujun Bao. 2018. UlTraMan: A Unified Platform for Big Trajectory Data Management and Analytics. *PVLDB* (2018).

[4] Yixiang Fang, Reynold Cheng, Wenbin Tang, Silviu Maniu, and Xuan S. Yang. 2016. Scalable Algorithms for Nearest-Neighbor Joins on Big Trajectory Data. *TKDE* (2016).

Automatic Intersection Extraction and Building Arrangement with StarCraft II Maps

Yuanbin Cheng
University of Southern California
Los Angeles, California
yuanbinc@usc.edu

Yao-Yi Chiang
University of Southern California
Los Angeles, California
yaoyic@usc.edu

ABSTRACT

In StarCraft, buildings arrangement near the intersections is one of most the critical strategic decisions in the early stage. The high time complexity of the buildings arrangement makes it difficult for AI bot to make the real-time decision. This paper presents an approach to analyze the intersection in StarCraft II maps. We propose a radar-like algorithm to automatically detect the intersection and use a designed heuristic search algorithm to arrange the building for building the wall. Our method can obtain the optimal solution while meeting the real-time requirement.

CCS CONCEPTS

• **Information systems** → *Spatial-temporal systems*;

KEYWORDS

Spatial Analysis, Heuristic Search, StarCraft

1 INTRODUCTION

StarCraft II is one of the most popular and successful real-time strategy (RTS) games in the world. From the artificial intelligence (AI) perspective, StarCraft II is the game with multiple agents (player can control hundreds of the units in game), imperfect information ("fog-of-war" cover the unvisited region of the map), vast and diverse action (Units can move to every point reachable in map, different unit has different abilities), which make StarCraft II a challenging playground for researchers to build the AI bot to automatically play the game [2].

Spatial analysis of the game map is important for the AI bot to make critical strategic decisions such as building planning and route planning. Building arrangement often plays a vital role in StarCraft II, especially for Terran and Protoss (two races in the game). It is a common strategy to arrange the buildings to block the narrow intersections for defending the opponent's army in the early stage of the game.

Blocking narrow intersections brings many challenges to the AI bot. First, it needs to determine the position of the critical intersection, which separates the different areas in the map. Second, AI bot need to make the arrangement decision for blocking the intersection in real-time. That means the bot should have the ability to determine the position of the critical intersection and arranging the buildings in short response time.

This paper presents an approach to determine the intersection in StarCraft II maps automatically in real-time. Our method initiates a radar-like algorithm to determine the intersection position. After finding the intersection, a building-arrange algorithm based on A* algorithm identifies the best locations of the buildings. The exposure is the boundary of the buildings that enemy units can directly attack, the shorter length of the exposure; the fewer enemy units can attack the block. The building-arrange algorithm can make sure the arrangement result has the shortest length of the exposure. Moreover, compared to the general search algorithm (BFS, DFS), the building-arrange algorithm can significantly reduce the searched space to shorten the response time.

2 APPROACH

During the game, some information can be derived from the StarCraft II in-game API,[1] including locations of the player and enemy's original base, resources locations on the map, and two grid maps. In Figure 1, the first grid map represents the buildable grids of the map. The second grid map in figure 2 indicates the grids that units can pass through it.

2.1 Intersection Detection

Our method first identifies two types of intersections that are the narrow passable place separate the different areas in the map.

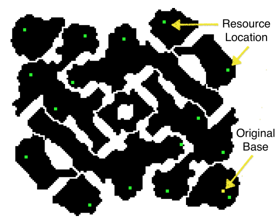


Figure 1: Buildable Grid Map

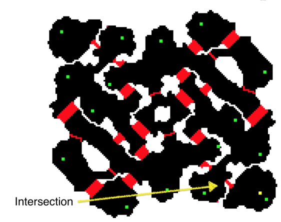


Figure 2: Passable Grid Map

2.1.1 Detecting the Intersections Between Different Heights. Many essential intersections appear between two planes with different heights. For detecting this type of the intersection, our method extracts the areas in the map that units can pass through while the buildings cannot be built, which is the common feature of this kind of intersection area, by comparing the difference between two grid maps. Figure 2 displays the result of this kind of intersection.

2.1.2 Detecting the Intersection of Narrow Terrain. Another kind of the critical intersections appears in the narrow terrain area connecting the different areas in the map. We propose a radar-like algorithm to detect this type of intersections. This algorithm iteratively searches for intersections by increasing the search radius from the start point, typically the locations of base and resource. This algorithm obtains the reachable boundary at each iteration step and recorded the minimize over the previous steps. After the total iteration, the algorithm finds the connected components of

[1] https://github.com/Blizzard/s2client-api

the minimize boundary as the candidate intersections. Figure 3 shows some internal states and Figure 4 present the result. In this situation, there is only one candidate intersection.

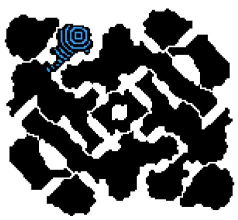


Figure 3: Radar-like Internal states

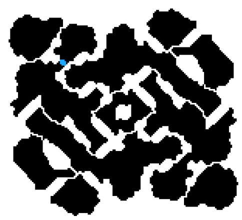


Figure 4: Radar-like Detection Result

2.1.3 Validity of Intersections. After getting the candidate intersections, we need to verify that the detected intersections can separate the map into multiple parts. Our method compares the number of the reachable grids from the start point before and after removing the intersections to determine whether the intersection is valid. If the number is significantly decreased, this intersection is valid.

2.2 Building Arrangement Search

After detecting of the critical intersections in the map, we utilize the building-arrange algorithm based on A* algorithm to find the optimal building arrangement for blocking the intersection.

2.2.1 Border of the intersection. In order to block the intersection, the building must be close to the border, the unpassable areas and unbuildable areas.

Figure 5: Border and Possible Building Location

Figure 6: Building Arrangement Result from A* Search

We located the border areas near the intersection by finding all unpassable and unbuildable grids within three grids from the intersection. The green and yellow areas in Figure 5 show the two border areas of the intersection.

2.2.2 Building-arrange Algorithm for Building Arrangement. In StarCraft II, Terran usually use Supply Depot (2×2 building) and Barracks (3×3 building) for blocking the intersection, and Protoss usually use Pylon (2×2 building) and Gateway (3×3 building) for blocking.

The blocking arrangement of the building should have two properties. First, at least one building is close to each of the border; otherwise, the unit can pass through the gap between the border and the building. Second, each building must be adjacent to at least one building; otherwise, the unit can pass through the gap between the buildings.

We designed the building-arrange algorithm based on A* search algorithm to search the building arrangement (as the node) from one side of the border until reaching to another side of the border (goal). Algorithm 1 shows the algorithm process. The *candidate*() obtain the candidate building locations near the current border (Gray color areas in Figure 5). The *minManh*() gain the minimize Manhattan distance between buildings in the current state and another part of the border (the heuristic function). Moreover, the *boundary*() acquires the length of the exposure boundary of the current state's buildings (the cost function).

Algorithm 1 Building-Arrange Algorithm

Input: start, goal(), candidates(), boundary(),
if goal(start) = true **then**
  return arrangment(start)
$open \leftarrow start, closed \leftarrow \emptyset$
while $open \neq \emptyset$ **do**
  sort(open)
  $n \leftarrow open.pop()$
  $candidates \leftarrow candidates(n)$
  for all $cand \in candidates$ **do**
    $cand.f \leftarrow cand.boundary + minManh(cand)$
    if $goal(cand) = true$ **then**
      return arrangment(cand)
    if $goal(cand) = false$ **then**
      $open \leftarrow cand$
  $closed \leftarrow n$
return $\emptyset$

The minimize Manhattan distance is the lower bound of the exposure boundary for any blocked buildings arrangements, which can make sure that the algorithm obtains the optimal solution [1]. The time complexity in the worst case is $O(b^d)$, b is the number of the candidates and d is the number of buildings in the arrangement result. Figure 6 shows the building arrangement result from the building-arrange search.

3 EXPERIMENTS

We conducted experiments on all maps from the StarCraft II Ladder 2017. For the original base location and the closest expand resource location, our approach could find the right intersection and plan the appropriate building locations for all maps. The number of nodes expanded in search for all map is less than one hundred, which meet the real-time requirement. For the start point in the middle of the map, due to the increase in connectivity paths and the variety of the terrain, our approach build lots of redundant buildings for blocking the location and need to expand hundreds of nodes to get the building arrangement.

REFERENCES

[1] P. E. Hart, N. J. Nilsson, and B. Raphael. 1968. A Formal Basis for the Heuristic Determination of Minimum Cost Paths. *IEEE Transactions on Systems Science and Cybernetics* 4, 2 (July 1968), 100–107.

[2] O. Vinyals, T. Ewalds, S. Bartunov, P. Georgiev, A. S. Vezhnevets, M. Yeo, A. Makhzani, H. Küttler, J. Agapiou, J. Schrittwieser, J. Quan, S. Gaffney, S. Petersen, K. Simonyan, T. Schaul, H. van Hasselt, D. Silver, T. P. Lillicrap, K. Calderone, P. Keet, A. Brunasso, D. Lawrence, A. Ekermo, J. Repp, and R. Tsing. 2017. StarCraft II: A New Challenge for Reinforcement Learning. *CoRR* abs/1708.04782 (2017).

SRC: An Intelligent and Interactive Route Planning Maker for Deploying New Transportation Services

Fandel Lin
Department of Electrical Engineering
National Cheng Kung University
Tainan, Taiwan (R.O.C)
e24046666@mail.ncku.edu.tw

Hsun-Ping Hsieh
Department of Electrical Engineering
National Cheng Kung University
Tainan, Taiwan (R.O.C)
hphsieh@mail.ncku.edu.tw

ABSTRACT

In this work, we propose a novel system, called Route Planning Maker (RPM) to help the government or transportation companies to design new route services in the city. The RPM system has a flexible user interface that allows users design the nearby areas of a new route and further deploying new stations. Moreover, based on user-designed arbitrary transportation routes and the expected locations of stations, the RPM system provides an intelligent function to infer passenger flows in certain time intervals so that the user can estimate the effectiveness of designed routes. To capture the spatial-temporal factors correlated with passenger flows, we propose to combine dynamic features such as human mobility, passenger volume of existing routes, and static features, including road network structure, point-of-interests (POI), station placement of existing routes and local population structure. Finally, to combine these features, we modified Deep Neural Network (DNN) for regression to derive the passenger flow for each given designated route. The experiments on the Tainan's bus-ticket data outperform baseline methods for 75%.

CCS CONCEPTS

• Information systems → Location based services;

KEYWORDS

Transportation Route Planning, Interactive Route Design, Passenger Flow Inference, Urban Planning

1 Introduction

The developments of traffic and human life are closely related. To reduce heavy traffic, the government usually deploys new services such as station and route construction. However, misestimating the demand for a certain route or a station could lead to serious impacts for both the government and people. Therefore, we aim to develop an efficient way to assess the benefit of a new route service in advance before deployment. We propose a novel system, Route Planning Maker (RPM), with two main functions. The first one is about the interactive route design and the second one is related to passenger volume inference. The system interface is shown in Figure 1.

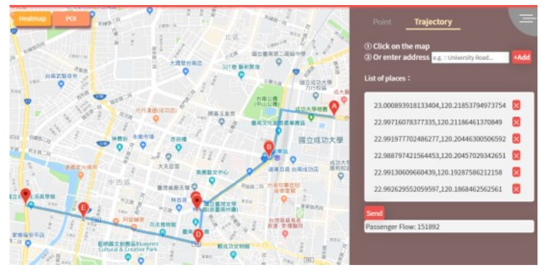


Figure 1: The proposed RPM system.

It is obvious our proposed RPM system could be exploited in many kinds of urban transportation, such as subway and taxi routes. The first function is that we allow users to design their arbitrary preferred routes segment-by-segment and the RPM system can visualize both new and existing routes in the map, which is quite useful for realizing the nearby area of a new route service. The second is to infer the passenger volume of a designated route. This can help to estimate the potential effectiveness. We propose several strategies to support the second function. To our best knowledge, no existing work has such contributions. First, we believe features observed in its route-affecting region (RAR), which indicates the influential range of a route, would influence the passenger flow (PF) of the route. Second, we exploit and modify the idea from DNN for regression to do PF inference. Third, to address the effects among new routes and existing ones, we divide each route into multiple segments and analyze the impacts by Dynamic Time Warping (DTW). The experiments show that our method outperforms state-of-art and baseline methods.

Related Works. Some works [2][6] focused on decreasing transportation time through route adjustment and shift. There are also some works [4][8] that optimized the route planning, in which distance, time, transference and passenger flow were considered. On the other hand, some work [1][3][5][7] studied the problem of predicting arrival time for public transportation based on regression methods. Moreover, some research [9][10] focused on the problem of predicting future passenger-flow. However, these works focus on existing routes, which are not our target problem.

2 Methodology

System Interface. The RPM system for users to schedule routes and stations is designed. The interface is shown in Figure 1. The system has three functions: (a) Map for users to schedule routes and stations. (b) Control panel for users to add or delete certain stations and adjust radius for RAR. (c) The PF of the route will be inferred to users.

Problem Definition. Given a set of trajectory for designated route with its stations labeled from users, our goal is to infer the passenger flow $PF(l_i t_j)$ for each segment l_i of the routes in certain time interval t_j. In other words, we devise the RPM system for users to schedule their own routes and stations. Then, the system derives the passenger flow of the user-designated trajectory and stations in a certain time interval.

2.1 Route-Affecting Region

The demand for public transportation is not only based on the origin and destination, but also the nearby geographical environment and urban functions of nearby area. Thus, we propose RAR for taking passenger-flow related features into account. A route can comprise multiple segments which contain successive points close to each other. Then we can draw a circle for each point, where we consider each point as the center of a circle, and then RAR formed by set of circles.

2.2 Feature Extraction Based on RAR

To correctly infer the PF of the trajectory, we consider six kinds of relevant urban features in RAR:

2.2.1 POI-related features. Various POIs (specific point location such as transportation spots or entertainment venue) and their density in RAR indicate the function of a route, which might have high correlation to PF of a route. We consider three kinds of POI features.

POI Density. The density of POI indicates the popularity of a certain activity type in *RAR*.

POI Entropy. The POI entropy in RAR shows the diversity of purpose for people to visit the nearby area of a route.

Competition of Similar POI. Similar transportation-related POIs might cause a competitive relationship between two routes.

2.2.2 Human mobility. Human mobility includes three ways, the transition density, incoming flow, and leaving flow in RAR.

2.2.3 Road network structure. We exploit road network structure since it might have correlation with real traffic condition. The degree and closeness centrality of road network in RAR are considered.

2.2.4 Competition and Transference with existing routes. Two routes might cause competitive relationship if their RAR is similar. However, intersected routes with considerable extended segments would encourage passengers to transfer between them. We analyze the impacts on passenger volume among the segments by DTW.

2.2.5 Population structure. People in RAR for different ages and genders have different intentions for taking public transportation.

2.2.6 Time Information and Granularity. Seasons and holidays can influence the passenger flow of the public transportation.

2.3 DNN Inference Framework

We adopt and modify DNN for regression to derive the PF for designated route. However, due to page limit, we skip the details here.

3 Experiments

We use the bus-ticket data from Tainan City Government for the evaluation, which contains 14,336,226 ticket records. The city bus system holds 104 routes and 6575 stations; meanwhile, each ticket record lists route and time information, starting and ending station. We have two kinds of experimental scenarios, the first one is querying only a route (trajectory-based), and the second one is a route with deployed locations (station-based). We develop four compared methods: (a) *Average value* and (b) *Median value* calculates the average and median value of the PF for all training routes respectively. (c) *XGBoost* and (d) *Support Vector Regression* considers only distances with existing routes. The performance result based on normalized RMSE with leave-one-out evaluation is shown in Figure 2.

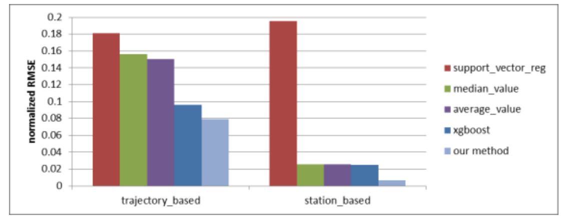


Figure 2: Normalized RMSE of PF results for different methods.

4 Conclusion

This work proposes an intelligent and interactive system called RPM to let users design novel routes and infer the passenger flows based on current bus routing and ticket data. The novelty of this work is two-fold. First, no existing work addressed the problem of arbitrary route design. Second, given heterogeneous features and faced with the competitive and transfer effects of existing routes, our proposed RAR and feature engineering methods are more effective for handling dynamic and static data. The experiments on Tainan City bus-ticket data outperform baseline methods for 75% in station-based scenario.

REFERENCES

[1] J. Biagioni, T. Gerlich, T. Merrifield and J. Eriksson (2011). EasyTracker: automatic transit tracking, mapping, and arrival time prediction using smartphones. Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems. Seattle, Washington, ACM: 68-81.

[2] H. Cancela, A. Mauttone and M. E. Urquhart (2015). "Mathematical programming formulations for transit network design." Transportation Research Part B: Methodological 77: 17-37.

[3] S. I.-J. Chien, Y. Ding and C. Wei (2002). "Dynamic Bus Arrival Time Prediction with Artificial Neural Networks." Journal of Transportation Engineering 128(5): 429-438.

[4] V. Guihaire and J.-K. Hao (2010). "Transit network timetabling and vehicle assignment for regulating authorities." Computers & Industrial Engineering 59(1): 16-23.

[5] Y. Lin, X. Yang, N. Zou and L. Jia (2013). "Real-Time Bus Arrival Time Prediction: Case Study for Jinan, China." Journal of Transportation Engineering 139(11): 1133-1140.

[6] M. Pternea, K. Kepaptsoglou and M. G. Karlaftis (2015). "Sustainable urban transit network design." Transportation Research Part A: Policy and Practice 77: 276-291.

[7] C. Shaowu, L. Baoyi and Z. Botao (2010). Bus arrival time prediction model based on APC data. 6th Advanced Forum on Transportation of China (AFTC 2010).

[8] Y. Yan, Q. Meng, S. Wang and X. Guo (2012). "Robust optimization model of schedule design for a fixed bus route." Transportation Research Part C: Emerging Technologies 25: 113-121. T.-H. Tsai, C.-K. Lee and C.-H. Wei (2009). "Neural network based temporal feature models for short-term railway passenger demand forecasting." Expert Systems with Applications 36(2, Part 2): 3728-3736.

[9] Y. Wei and M.-C. Chen (2012). "Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks." Transportation Research Part C: Emerging Technologies 21(1): 148-162.

[10] Z. Wenfeng, S. Zhongke and L. Zhiyong (2008). Prediction of urban passenger transport based-on wavelet SVM with quantum-inspired evolutionary algorithm. 2008 IEEE International Joint Conference on Neural Networks.

Centroid-Amenities: An Interactive Visual Analytical Tool for Exploring and Analyzing Amenities in Singapore

Jazreel Siew Xue Qian
School of Information Systems
Singapore Management University
Singapore Singapore
jazreelsiew.2015@sis.smu.edu.sg

Sean Koh Jia Ming
School of Information Systems
Singapore Management University
Singapore Singapore
sean.koh.2015@sis.smu.edu.sg

ABSTRACT

Planning for civic amenities in a fast-changing urban setting such as Singapore is never an easy task. And as urban planners look toward more data-driven approaches toward urban planning, so grows the demand for more flexible geospatial analytics tools to facilitate a more iterative and granular approach toward urban planning [1].

Such specific tools however, are not always readily available as plugins for traditional desktop GIS software, as numerous customizations must be made to model specific temporal planning scenarios for quick analysis, which could prove both costly and time-consuming. Hence, to address this need, open-source tools such as R Shiny could be used to rapidly prototype and test urban planning models, in an iterative fashion.

To demonstrate how this could be done, we developed a proof-of-concept that aims to provide urban planners with an open-source, interactive geospatial analytics tool to help optimize the placement of amenities and services through K-means clustering, making them as accessible as possible to the city residents they serve. The platform also allows planners to compare the accessibility of the existing amenities and services, against a suggested set of optimized amenity locations, using the Hansen Accessibility Score as a measure of accessibility. This allows planners a tangible grasp on how much of an impact and improvement a relocation of amenities could make for the residents served.

This paper details our research and development efforts to design and implement an open-source web-based geospatial tool for supporting the analysis of the accessibility of amenities and services.

SIGSPATIAL '18, November 6–9, 2018, Seattle, WA, USA

ACM ISBN 978-1-4503-5889-7/18/11.
https://doi.org/10.1145/3274895.3282798

CCS CONCEPTS

• Information Systems applications → Decision support systems → Data analytics

KEYWORDS

Amenities, Geospatial Analytics, K-Means Clustering, R Shiny, Urban Planning

1 APPROACH

Our sources of planning data include publicly available census data obtained though the Department of Statistics (Singapore), from a nationwide census taken in 2014[2]. Data on the locations of existing amenities were obtained from data.gov.sg[1].

1.1 Application Overview

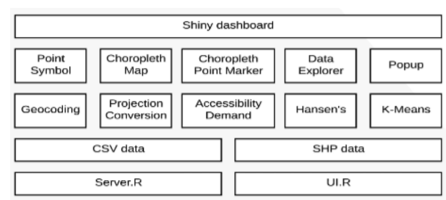


Figure 1: Application Architecture Diagram

The application[2] is developed using R with R Shiny platforms. Users will be using a Shiny dashboard which come with features available in the standard leaflet.js library functions. At the backend, CSV and SHP data are processed via geocoding, projection conversion, Hansen Accessibility Index and weighted K-Means clustering

1 Data.gov.sg, the government's one-stop portal to its publicly-available datasets.
2 The application code can be found in https://github.com/jazzywessy/Centroid-Amenities

1.1.1 Site Selection via K-Means Clustering. In the context of our amenities location optimization model, our 'optimal location' of any amenity would be the centroid location of each cluster of its beneficiaries, as it is the location of the least distance from their beneficiaries' residence.

K-Means clustering algorithm assigns each item to a cluster having the nearest centroid.

1.1.2 Hansen Accessibility Index. This index measures the accessibility of an area, by the sum of opportunities to reach the destination from other areas, weighted by a cost function. Where i is the index of Origin Zones, j is the index of Destination Zones, and $f\left(C_{ij}\right)$ is a function of the generalized cost of travel, where heavier weights are placed nearer places or less expansive places than further or more expansive places, and vice versa.

$$Accessibility_i = \sum_j Opportunities_j \left(C_{ij}\right)$$

For our application, the origin zones will be the public housing blocks and the destination zones will be the amenities. A higher Hansen Accessibility Index will indicate that the amenities within the planning zone are more accessible.

2 RESULTS

Planners first interact with our dashboard by selecting a planning subzone and by specifying an input number of amenities. The results will be displayed in an info panel and an interactive map layer for the ease of analysis.

To compare results of current accessibility index versus the suggested accessibility index, we chose two different subzones namely, Punggol and Pasir Ris, of similar population and area size with the same amenities category selection.

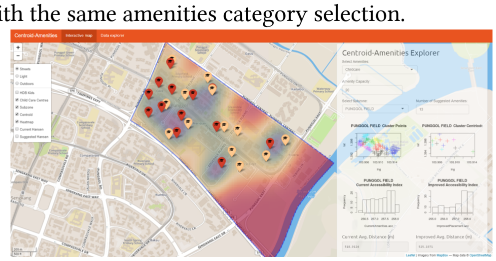


Figure 2: Clustering of Childcare facilities in Subzone of Punggol Field

The analysis is shown in multiple map layers. The kernel density map layer shows the density of the demand for childcare facilities in Punggol are much denser in at the left of the area. A high concentration of the target population will be displayed in blue whereas a low concentration of the target population will be displayed in red. Yellow pins represent the current locations of childcare facilities while red pins represent the suggested locations of amenities that are optimally placed based on the weighted k-means clustering. Based on the improved accessibility index, it shows a higher number of HDB blocks having better accessibility index with highest frequency of 40 and 259.0 index compared to the current accessibility index with highest frequency of 20 and 257.7 index. Visually from the map layers, it displays the suggested childcare facilities location are placed near the denser clusters to fill the demand of childcare facilities.

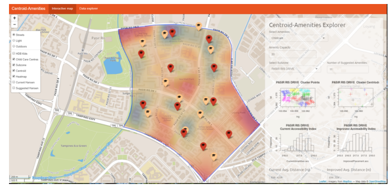


Figure 3: Clustering of Childcare facilities in Subzone of Pasir Ris Drive

Next, looking at the subzone of Pasir Ris Drive, the kernel density map layer shows the density of the demand for childcare facilities in Pasir Ris Drive are much well spread across its area. Even with such clustering pattern, the improved accessibility index, shows a higher number of HDB blocks having better accessibility index with more frequency of HDB Blocks have higher accessibility index of 217 compared to the current accessibility index with fewer frequency of HDB block having the same accessibility index of 217. par

In both cases, with different demand patterns, results show a better accessibility index in the suggested k-means optimized amenities locations. With the map layers of both current and suggested childcare facilities location displayed on the same map, it gives the urban planner a visual comparison of the difference in location. The same method could be applied to other amenities where the target population of beneficiaries can be clearly identified.

4 DISCUSSION AND FUTURE WORKS

We demonstrated how open-source platforms, can be used to rapidly prototype highly customized urban planning decision support tools, allowing for a more iterative approach towards spatial analysis. Some avenues for future works include using projected population trends and General Transit Feed Specification data to improve our k-means optimized amenities placement algorithm, and allowing users define their chosen area of analysis.

ACKNOWLEDGMENTS

The authors would like to thank Professor Kam Tin Seong from Singapore Management University, for his extensive guidance. The authors would also like to acknowledge Tay Wei Xuan for his assistance in developing our application prototype, and research help in this project.

REFERENCES

[1] Basu, M. (2016, June 2). Harnessing the Power of Digital Technologies. Retrieved August 17, 2018, from https://www.ura.gov.sg/Corporate/Resources/Ideas-and-Trends/Harnessing-the-Power-of-Digital-Technologies

[2] DOS. (2016, June). Singapore Residents by Subzone and Type of Dwelling, June 2016. Retrieved August 17, 2018, from https://data.gov.sg/dataset/singapore-residents-by-subzone-and-type-of-dwelling-june-2016

SRC: Transferring scale-independent features to support multi-scale object recognition with deep convolutional neural network

Xiran Zhou
School of Geographical Sciences & Urban Planning, Arizona State University
Tempe, AZ, U.S.
xrzhou@asu.edu

1 Introduction

Object detection is an essential remote sensing data application that attempts to detect and recognize semantically meaningful instances from complicated context and background [1]. Recently, a great number of works have reported that the state-of-the-art convolutional neural networks (CNNs), significantly facilitates automatically detecting a predefined class of object from high resolution aerial or satellite imageries [2]. However, different from the object in a photo, the objects in remote sensing imageries are always represented by largely varied scale or spatial resolution. Scale has become a critical attribute for recognizing an object from aerial or satellite imagery [3]. Up to now, the state-of-the-art DCNNs for object detection requires predefined scale, which produces a big challenge to automatically adjust the representative scale for an object on the ground [3-4].

Transfer learning is a branch of machine learning techniques that focuses on reusing the exiting data, knowledge or models to solve a new research problem [8]. Deep learning creates independent CNNs for different object detection tasks. Otherwise, transfer learning only builds one CNN, and attempts to transfer the features gained from this CNN to support to detect the small-scale airplane. It is agreed that fine-tuning the deep architecture of a CNN and collect the needed data are computationally-intensive and time-consuming. Transfer learning enables a machine to transfer the already gained knowledge to solve new tasks in an ad-hoc manner, and supports to conduct zero-shot and one-shot learning in an application with small-scale dataset.

A solution to effectively deal with cross-scale data features is important for successfully implementing CNNs to conduct efficient object detection from remote sensing imageries [1]. This paper focuses on transferring the pretrained knowledge derived from large scale object, to support to detect other objects represented within different scales. Based on the architecture of Faster R-CNN, this paper reports a new DCNNs for conducting scale-independent object detection from remote sensing imagery.

2 Methodology

The proposed CNN for object detection depends on the architecture of Faster R-CNN [6], which effectively supports to discover the representative region proposals through integrating the results from the Region Proposal Network (RPN) and those from the classification layer. This paper proposes a new algorithm called atrous region proposals to promote the efficiency of multi-scale region proposal searching, and exploits atrous convolution in the feature extraction layer.

2.1 Atrous convolution

Although pooling layer supports to reduce computational load and facilitates extracting rotational-invariance features, it may fail to support to automatically adjust the field-of-view of a filter, and always leads to the loss of spatial details being beneficial to the representation of an object.

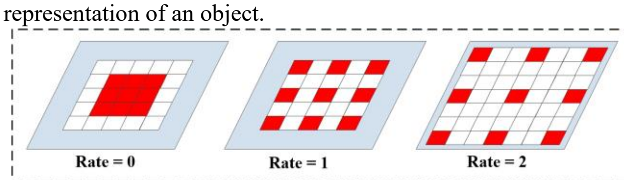


Figure 1: Illustration on atrous convolution layer.

Atrous convolution supports to filter the image through altering the size of the receptive field view in generating region proposal, to control the loss of spatial details [7]. Figure 1 illustrates the principle of atrous convolution. A pixel is filtered based on its neighboring pixels over different distances (rates). Then, the filtering results within a variety of distances are integrated. Due to page limit, more details about atrous convolution can be read in Reference [7].

2.2 Atrous region proposals

This paper aims to extending the receptive field of RPN, which is a critical element used for bounding box searching with optimized dimension and shape. In the Faster R-CNN model, the dimension and the shape of region proposals are controlled by two parameters: archer size and ratio. However, the scales of the same class of objects might be significantly varied in remote sensing imageries [2]. This makes the range of archer size and ratio vary largely to characterize the objects represented by various scales, which posing challenging for object localization due to instability of representative region proposals search.

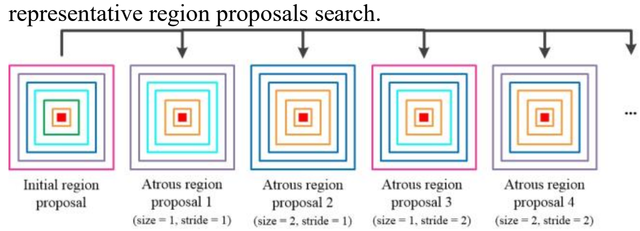


Figure 2: Illustration on atrous region proposals.

To address the challenge mentioned above, applying the idea of atrous convolution, this paper creates a new region proposal called atrous region proposal. Figure 2 shows the structure of atrous region proposals. Red box denotes a a pixel. Orange, green, cyan, blue, purple and pink line respectively denotes this pixel's neighboring pixels over multiple distances. Pixels in the red box and its surrounding lines constitutes an object. Based on the initial region proposal generated by Faster R-CNN, the proposed method creates a variety of atrous region proposals with the neighboring pixels over multiple distances. The atrous region proposals shown in Figure 2 selects the center pixel-oriented partial context to create a small-scale region proposal.

The dimension of atrous region proposal is controlled by size and stride. Size defines the number of neighboring pixels to be replaced. Stride controls the way how replacing the neighboring pixels. For example, when size equals 1, only the green line is replaced by the orange one in atrous region proposal 1. When size equals 2, the green and cyan line are replaced by the orange one in atrous region proposal 2. Additionally, stride is used to control how replacing around the neighboring pixels. When stride equals 1, after replacing the green line with the orange one, the cyan line is used to replace the blue one in atrous region proposal 1. In atrous region proposal 3, when stride equals to 2, after replacing the green line with the orange one, the proposed method jumps the cyan line and uses the blue line to replace the purple one.

3 Experiment

The dataset for evaluating the proposed DCNN for object detection is selected from a large-scale benchmark dataset called CSRS-SIAT [8]. SIAM-CSRS includes around 70 scenic and object categories of satellite imageries over scales, and each category contains 1000 images. To our knowledge, CSRS-SIAT is the only benchmark dataset that provides cross-scales imageries for every scene class. In this paper, we select dam, airplane, oil tank, football field, island, swimming pool, lake/pond and crater. 200 samples of every category were labeled, among them 120 samples have a similar scale, and other 80 samples have different scales. Figure 3 shows the selected samples of these eight categories. Red box denotes the labeled bounding box for training and test.

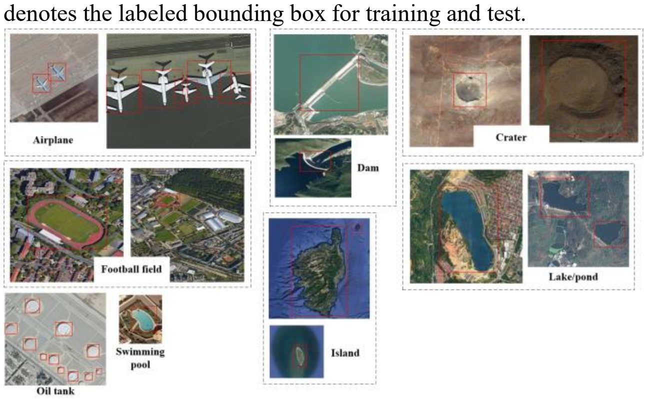


Figure 3: Illustration on experimental datasets.

The first experiment compares the result generated by similar scale object detection with the classic Faster RCNN and the proposed method. Another one compares the result of multi-scale object detection with the classic Faster RCNN and the proposed method. For each experiment, 80% samples were randomly selected for training, and the rest were used for evaluation.

Table 1: Comparison of object detection results.

Category	Similar scale		Cross-scale	
	Faster R-CNN [2]	Proposed method	Faster R-CNN [2]	Proposed method
Dam	90.00%	92.5%	81.5%	86.00%
Airplane	95.83%	97.5%	86.00%	90.00%
Oil tank	98.33%	98.33%	94.00%	95.00%
Football field	87.5%	90.83%	79.50%	82.00%
Island	100%	100%	93.00%	96.00%
Swimming pool	100%	100%	—	—
Lake/pond	90.83%	92.50%	86.00%	89.00%
Crater	91.67%	93.33%	81.00%	84.00%

Table 1 lists average precision (AP) results of object detection. AP is a commonly-used evaluation measure to assess the accuracy of CNN for object recognition [2]. The precision ranges from 87.5% to 100% due to the characteristics of object, the complexness of context, and the land cover where an object locates in. Faster R-CNN and the proposed method works well on similar scale object detection, and the proposed method slightly outperforms the classic Faster R-CNN. This is because atrous convolution performs better than pooling layer in features extraction [9]. Additionally, the results of similar-scale object detection prove that the atrous region proposals proposed remains the efficiency of object detection.

When the objects are represented by different scales, significant decreases of AP are observed in the results by using Faster R-CNN. This shows that it is challenging for Faster R-CNN to predefine appropriate archer size for precise object localization if the scale of objects varies significantly. Through creating extensive region proposals, the proposed method outperforms Faster R-CNN to a certain extent. The work above shows that the RPNs in Faster R-CNN can effectively support multi-scale object detection with modification considering spatial shape and scale.

REFERENCES

[1] G. Cheng and J. Han. (2016). A survey on object detection in optical remote sensing images. ISPRS Journal of Photogrammetry and Remote Sensing, 117, 11-28.

[2] H. Wu and Z.L. Li (2009). Scale issues in remote sensing: A review on analysis, processing and modeling. Sensors, 9(3), 1768-1793.

[3] J. Li, X. Liang, S. Shen, et al. (2018). Scale-aware fast R-CNN for pedestrian detection. IEEE Transactions on Multimedia, 20(4), 985-996.

[4] T. Y. Lin, P. Dollár, R. B. Girshick, et al. (2017). Feature Pyramid Networks for Object Detection. In IEEE Conference on Computer Vision and Pattern Recognition, 1(2), 4.

[5] S. J. Pan and Q. Yang (2010). A survey on transfer learning. IEEE Transactions on knowledge and data engineering, 22(10), 1345-1359.

[6] S. Ren, K. He, R. Girshick, et al. (2017). Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis & Machine Intelligence, (6), 1137-1149.

[7] L.C. Chen, G. Papandreou, I. Kokkinos, et al. (2018). Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(4), 834-848.

[8] Y. Shen, X Zhou, J. Liu, et al. (2018). CSRS-SIAT: A benchmark remote sensing dataset to semantic-enabled and cross-scales scene recognition. In IEEE International Geoscience and Remote Sensing Symposium, 2018.

The SIGSPATIAL Special

Section 2: Workshop Reports

The Seventh ACM SIGSPATIAL International Workshop on Analysis for Big Spatial Data Seattle, WA, USA - November 6, 2018

Varun Chandola
chandola@buffalo.edu
SUNY Buffalo, USA

Ranga Raju Vatsavai
rrvatsav@ncsu.edu
North Carolina State University, USA

(Workshop Co-chairs)

Big data is emerging as an important area of research for data researchers and scientists. This area has also seen significant interest from the industry and federal agencies alike, as evidenced by the recent White House initiative on "Big data research and development". Within the realm of big data, spatial and spatio-temporal data is one of fastest growing types of data. With advances in remote sensors, sensor networks, and the proliferation of location sensing devices in daily life activities and common business practices, the generation of disparate, dynamic, and geographically distributed spatiotemporal data has exploded in recent years. In addition, significant progress in ground, air- and space-borne sensor technologies has led to an unprecedented access to earth science data for scientists from different disciplines, interested in studying the complementary nature of different parameters. Today, analyzing this data poses a massive challenge to researchers.

The workshop series on Analytics for Big Geospatial Data (BIGSPATIAL), has become one of the key meeting points for researchers in the area of big geospatial data analytics, since 2012. Held every year, in conjunction with the annual ACM SigSpatial conference, this meeting has found strong support from researchers in government, academia, and industry.

Building on the success of the previous editions to bring together researchers from academia, government and industry, who have been working in the area of spatial analytics with an eye towards massive data sizes, the 7^{th} workshop on Analytics for Big Geospatial Data (BIGSPATIAL 2018) was held in conjunction with the 26^{th} ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM SIGSPATIAL 2018) on November 6^{th}, 2018. The main motivation for this workshop stems from the increasing need for a forum to exchange ideas and recent research results, and to facilitate collaboration and dialog between academia, government, and industrial stakeholders. We hope that this workshop provides a platform for researchers and practitioners engaged in addressing the big data aspect of spatial and spatio-temporal data analytics to present and discuss their ideas.

This year we received 21 technical submissions out of which 9 were selected for full presentations at the workshop. The technical program also consisted of a keynote talk by Professor Shashi Shekhar from University of Minnesota, who is widely regarded as one of the leading experts in the area of spatial computing. His talk titled, "Transforming Smart Cities and Communities with Spatial Computing" provided an excellent start to the workshop by laying out the challenges and opportunities for big data researchers in the area of smart cities and communities. The workshop was well-attended with 40 registered participants.

We would like to thank the authors of all submitted papers. Their innovation and creativity has resulted in a strong technical program. We are highly indebted to the program committee members, whose reviewing efforts ensured in selecting a competitive and strong technical program. We would like to express our sincere gratitude to Prof. Shekhar for his insightful keynote lecture. The BIGSPATIAL workshop series will continue to provide

a leading international forum for researchers, developers, and practitioners in the field of data analytics for big geospatial data to identify current and future areas of research.

EM-GIS2018 Workshop Report
The 4th ACM SIGSPATIAL International Workshop on Safety and Resilience

Hui Zhang[1], Yan Huang[2], Jean-Claude Thill[3], Danhuai Guo[4, 5], Yi Liu[1], Bin Chen[6], Wei Xu[7]

[1]Institute of Public Safety Research, Department of Engineering Physics, Tsinghua University

[2]University of North Texas

[3]Department of Geography & Earth Sciences University of North Carolina at Charlotte

[4]Computer Network Information Center, Chinese Academy of Sciences

[5]University of Chinese Academy of Sciences

[6]National University of Defence Technology

[7]School of Information and Smart City Research Center, Renmin University of China

Safety is crucial for people, and emergency management helps keep people safe. Emergency management involves four stages: Planning and Mitigation, Preparedness, Response and Recovery. Geospatial applications (including GIS) have been extensively used in each stage of emergency management. Decision-makers can utilize the geospatial information to develop planning and mitigation strategies. GIS models and simulation capabilities are used to exercise response and recovery plans during non-disaster times. They help the decision-makers understand near real-time possibilities during an event. Once disaster occurs, GIS will take effect in real time response and recovery activities

EM-GIS 2018(https://em-gis2018.github.io/CallForPaper/index.html) was held in conjunction with the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM SIGSPATIAL 2018) on November 6th, 2018 in Seattle, WA, USA. The purpose of the EM-GIS 2018 workshop is to provide a forum for researchers and practitioners to exchange ideas and progress in related areas. This workshop in the ACM SIGSPATIAL conference addresses the challenges of emergency management based on advanced GIS technologies. This workshop brought together researchers and practitioners in massive spatio-temporal data management, spatial database, spatial data analysis, spatial data visualization, data integration, model integration, cloud computing, parallel algorithms, internet of things, complex event detection, optimization theory, intelligent transportation systems and social networks to support better public policy through disaster detection, response and rescue.

EM-GIS 2018 was a one-day workshop with 17 researchers and practitioners registered. Overall, 13 research papers were presented and discussed (20 minutes for each paper). The presentations were divided into three sessions:

(1) Emergency detection and prediction. In this section, authors presented their research in evaluating signage system (*A new method of evaluating signage system using mixed reality and eye tracking*), resource security management (*Design of risk monitoring and prediction system for resource security management*), computing the landslide degree of risk (*A tool to compute the landslide degree of risk using R-Studio and R-Shiny ∗*) and foodborn disease report and forecast system (A *VGI-based foodborn disease report and forecast system*).

(2) Transportation safety. In this section, the discussion focused on Traffic State Estimation (*Traffic state estimation with big data*), driving behaviors in disasters and risk perception (*How risk perception affect driving behaviors in disasters*), hotspots discovery (*Spatial data mining and O-D hotspots discovery in cities based on an O-D hotspots clustering model using vehicles' GPS data – a case study in the morning rush hours in Beijing, China*) and passenger flow (*Temporal-spatial Analysis & Visualization of Passenger Flow after Opening New Railway Lines in Shenzhen Metro*).

(3) Emergency evaluation. In this section, authors discussed topics on energy-optimal path planning (*Energy-Optimal Path Planning for Solar-Powered UAVs Monitoring Stationary Target*), atmospheric contaminants releases (*Playing a Chemical Cluster Environmental Protection Patrolling Game Addressing Atmospheric Contaminants Releases*), urban water distribution network pipes (*Risk evaluation of urban water distribution network pipes using neural network*) and geological disasters (*Spatial Distribution Analysis and Regional Vulnerability Assessment of Geological Disasters in China*).

The workshop had one best paper and one Best Student Paper. The best paper is *Traffic State Estimation with Big Data* authored by Han Xing, Ke Zhang, Zi Fan Yang, Jing Qian and Yi Liu. And the best student paper i*s A tool to compute the landslide degree of risk using R-Studio and R-Shiny* authored by Erica Goto, Keith Clarke and Edward Keller.

We would also like to thank the authors for publishing and presenting their papers in EM-GIS 2018, and the program committee members and external reviewers for their professional evaluation and help in the paper review process. We hope that the proceedings of EM-GIS 2018 will inspire new research ideas, and that you will enjoy reading them.

GeoAI 2018 Workshop Report
The 2nd ACM SIGSPATIAL International Workshop on GeoAI: AI for Geographic Knowledge Discovery
Seattle, WA, USA - November 6, 2018

Yingjie Hu[1], Song Gao[2], Shawn Newsam[3], Dalton Lunga[4]
[1]University at Buffalo, USA
[2]University of Wisconsin, Madison, USA
[3]University of California, Merced, USA
[4]Oak Ridge National Lab, USA

In today's era of big data, advanced algorithms, and immense computational power, artificial intelligence (AI) is bringing tremendous opportunities and challenges to geospatial research. Big data enable computers to observe and learn the world from many different perspectives, while high performance machines support the developing, training, and applying of AI models within reasonable amount of time. Recent years have witnessed significant advances in the integration of geography and AI in both academia and industry. There have already been many successful studies. Focusing on modeling the physical nature, research has shown that deep learning can improve the representation of clouds that are smaller than the grid resolutions of climate models. Examining the human society, AI and natural language processing methods, such as word embeddings, are helping quantify changes in stereotypes and attitudes toward women and ethnic minorities over 100 years in the United States. There are also many other applications that effectively integrate AI with problems in geospatial studies, such as vehicle trajectory prediction, indoor navigation, historical map digitizing, gazetteer conflation, geographic feature extraction, geo-ontologies, and place understanding. The 2nd International Workshop on AI for Geographic Knowledge Discovery (GeoAI 2018) builds on the success of the previous workshop in 2017. GeoAI is bringing together geoscientists, computer scientists, engineers, entrepreneurs, and decision makers from academia, industry, and government to discuss the latest trends, successes, challenges, and opportunities in the field of artificial intelligence for data mining and geographic knowledge discovery.

GeoAI 2018 received 19 paper submissions in total. After a rigorous peer-review process by the program committee, 10 papers were accepted by the workshop and selected for presentations. Dr. Rangan Sukumar, Senior Analytics Architect from Cray Inc., gave a keynote on "The AI Journey in Geospatial Discovery: Navigating Shapes, Sizes and Spaces of Data" (industry keynote), and Dr. Bruno Martins, Assistant Professor at the University of Lisbon, gave a keynote on "GeoAI Applications in the Spatial Humanities" (academic keynote). This year's workshop also featured one wrap-up discussion on "How can we make GeoAI better?" Many participants contributed constructive ideas, such as encouraging future submissions on dataset descriptions, privacy and data bias issues, transparency and reproducibility, and explicit spatial constraints on AI models. Participants also discussed the possibility of adding a panel discussion, and core skills for students who want to build expertise in GeoAI. There are 51 participants who officially registered to GeoAI'18. The workshop attracted in average 50 participants, with a maximum of about 70 participants in the room.

We sincerely thank our program committee members for their time and efforts in reviewing and evaluating the submitted papers. We hope that the proceedings of GeoAI'18 can stimulate new ideas and make a modest contribution to this fast growing field.

GeoHumanities 2018 Workshop Report

Bruno Martins[1], Patricia Murrieta-Flores[2]
[1]Instituto Superior Técnico - INESC-ID, University of Lisbon,Portugal
[2]Digital Humanities Hub, Lancaster University, United Kingdom
(Workshop Co-Chairs)

Abstract

This article reports on the 2nd ACM SIGSPATIAL Workshop on Geospatial Humanities, held in conjunction with the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. The article outlines the objectives of the workshop, and briefly describes the technical program.

1 Introduction to the ACM SIGSPATIAL Workshop on Geospatial Humanities

The 2nd ACM SIGSPATIAL Workshop on Geospatial Humanities (GeoHumanities'18) was held together with the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. It addressed the use of geographic information systems and other spatial technologies in humanities research.

Scholars in the humanities have long paid attention to spatial theory and cartographic outputs. Moreover, in recent years, new technologies and methods have lead to the emergence of a field that is now commonly known as the Spatial Humanities. Methods from the standard toolset of geographic information systems (e.g., computation of viewsheds and zones of influence, least-cost path analysis, mass-preserving areal weighting and dasymetric mapping, terrain classification according to land coverage or land use, different types of thematic cartography techniques, etc.) have been successfully employed to analyze the geographies of human cultures, both past and present, and to address research questions posed by humanities-based fields. However, many challenges persist in the application of more recent technical developments in the geographical information sciences, which have been showcased in venues such as the ACM SIGSPATIAL conference. The workshop is thus concerned with the use of geographic information systems and other spatial technologies in humanities research, placing a strong emphasis on new methodologies that leverage recent technical developments (e.g., the above-mentioned standard tools from geographic information systems, as well as more advanced methods such as text-based geographical analysis or spatial simulation, can all benefit from innovative approaches leveraging machine learning, parallel and/or distributed computation, semantic technologies, etc.).

2 The Workshop Program

The call for papers resulted in 5 submissions describing high quality research. A program committee of 32 members, which are listed on the workshop website[1], reviewed the submissions, and 4 papers were accepted for presentation. The workshop program featured two technical sessions, followed by a discussion at the end. Each of the sections featured one paper where the main contributions where related to technical advancements in the field, and one paper more concerned with innovative applications of established geo-spatial analysis methods.

After a brief opening address, the workshop started with the presentation of work by Vincent Mack and Tin Seong Kam, concerning the use of exploratory spatial data analysis methods (i.e., techniques for the identification of spatial outliers, and for the discovery of clustering patterns and hot-spots) for studying the distribution of

[1]`http://bgmartins.github.io/sigspatial-geohumanities/program-committee.html`

violence against civilians in Africa. The authors used Local Indicators of Spatial Association (LISA), together with indicators of spatial auto-correlation, to analyze data from 1997 to 2017 collected in the context of the Armed Conflict Location and Event Data Project (ACLED[2]). Tin Seong Kam presented tools, methods, and insights into temporal changes for violence hot-spots, arguing that exploratory spatial data analysis methods can offer an important supplement to more traditional analyses of political violence.

The second presentation covered research by Esko Ikkala et al. concerned with the development of the NameSampo linked data infrastructure, which aims to support the analysis of one of the main information sources for toponomastic research in Finland, i.e. the corpora of 2.7 million place names in the Names Archive database[3] of the Institute for the Languages of Finland. Esko Ikkala gave an overview of the NameSampo project and associated web-based application[4], created in collaboration between toponomastic researchers and computer scientists for supporting interactive search, analysis, and visualization of digitized toponomastic data. The presentation covered the overall system architecture and gave first results related to the vocabularies/ontologies used for data modeling, data transformation into linked open data, and interfaces for exploratory analysis.

After the coffee break, the second session started with the presentation of a paper by Barz et al., in which the authors described a novel method for quality assessment of crowd-sourced data. The authors leverage the intuition that average inter-user agreement can be indicative of quality, and they proposed using an adaptation of the PageRank algorithm to compute reputation scores over a graph in which edges encode the similarity between annotations provided by crowd workers. Using two real crowd-sourcing activities developed within the NYC Space/Time Directory[5] project of the New York Public Library, Thomas C. van Dijk showed that the obtained reputation scores are plausible, and that they can provide insights into user behavior.

Afterwards, Sohrab Rahimi described work developed in collaboration with Mallika Bose, towards the validation of Bourdieu's theory of distinction. Bourdieu argues that social classes are distinguished by their taste, as opposed to mere economic status. To assess this claim, the authors proposed to analyze two datasets characterizing taste and social class: restaurant reviews from Yelp, and social class information collected from the American community survey in 2016. Regression analysis was used to quantify the relationship between taste indices (i.e., keywords from the Yelp reviews) and indicators of social class. Results showed that target variables like education and race are highly related to food consumption patterns.

A discussion period followed the last presentation of the second session, covering common aspects between the different contributions at the workshop. In total, the workshop had 8 registered participants and, on average, 10 attendees were also present at each of the sessions. We believe GeoHumanities'18 was a successful event that, although small, allowed the participants to explore the contributions that modern GIS and geo-spatial analysis technologies can enable within and beyond the digital humanities.

3 Acknowledgments and Final Remarks

The organizers would like to thank the authors for submitting and presenting their contributions, and also the program committee members for their commitment to the paper review process. We hope that the proceedings of GeoHumanities'18 will inspire new research ideas, and that you will enjoy reading them. A special issue in Springer GeoJournal, featuring extended versions of the best workshop papers, is currently under preparation.

The authors would also like to acknowledge the support provided by Fundação para a Ciência e a Tecnologia (FCT), through the project grants with references UID/CEC/50021/2013 and PTDC/CCI-CIF/32607/2017, and also by the Trans-Atlantic Platform for the Social Sciences and Humanities, specifically through the 'Digging into Early Colonial Mexico: A large-scale computational analysis of 16th century historical sources' project with reference HJ-253525-ES/R003890/1.

[2] http://www.acleddata.com

[3] http://www.kotus.fi/en/corpora_and_other_material/names_archive

[4] http://seco.cs.aalto.fi/projects/nimisampo

[5] http://spacetime.nypl.org

GIR'18 Workshop Report 12th ACM SIGSPATIAL Workshop on Geographic Information Retrieval Seattle, USA 6^{th} November 2018

Ross S. Purves[1]
[1] University of Zurich, Switzerland
ross.purves@geo.uzh.ch

Christopher B. Jones[2]
[2]Cardiff University, United Kingdom
c.b.jones@cs.cf.ac.uk

(Workshop Co-chairs)

Geographic Information Retrieval (GIR) addresses challenges of gaining access to documents and their content that relate to geographical locations. The field can be regarded as a fusion of Information Retrieval (IR), which is concerned with information in unstructured documents and associated methods of natural language processing, with Geographical Information Systems, which are oriented to structured data, and to a range of spatial analytical and data access methods that are relevant to GIR. This workshop is the twelfth of a series of workshops, the first of which was held in 2004. The topics covered by the workshops have have included methods for recognising and disambiguating references to place names in text (referred to as geo-parsing); determining the geographic scope of documents; developing gazetteers and ontologies to maintain knowledge of toponyms and geographic concepts; spatio-textual indexing methods that combine inverted file methods with those of spatial database indexing; managing vagueness and uncertainty in geographic terminology; extracting geo-spatial facts and events from documents; and evaluating the performance of geo-information retrieval systems.

The 12th GIR workshop was held on 6th November 2018 at the ACM SIGSPATIAL conference in Seattle, USA. Previous workshops have been held either in combination with the SIGIR and CIKM conferences or as stand-alone events in cooperation with ACM SIGSPATIAL. The stand-alone events have all been located in Europe, in particular in Zurich, Paris and last year in Heidelberg.

At GIR'18 there were 8 presentations, of which 2 were full papers and 6 were short papers. We received 12 submissions. The number of people who registered for the workshop was 14, while the typical attendance was slightly less at around 10 to 12.

The workshop was organized around four sessions, the third of which included a discussion session.

The papers cover a wide range of themes. These included the well-established challenge of geo-parsing of which there were two papers. One of these (Hu) presented a website for evaluating geo-parsers, with support for different corpora and various methods of geo-parsing. Akdemir et al applied deep learning with word embeddings (which characterise words by a representation of their contexts) to English language news documents published in India, which are distinctive in their use of English. Two papers were concerned with corpora and geospatial resources, with Yin et al looking at issues of the geographical variation in characteristics of social media and automatically categorising volunteered resources (with a view to automated quality control), while Chesnokova and Purves described methods for building a corpus that represented personal perceptions of landscape. Two papers in the programme were focused on question answering (QA) systems with Punjani et al using DBpedia, the semantic web (linked data) representation of Wikipedia, in combination with the Geonames gazetteer. The other paper by Mai et al was concerned with the creation of a test dataset for QA based on Point of Interest (POI) data from Yelp, that was evaluated using a machine learning method that again demonstrated an advantage in using text embedding methods (in their case for sentences). The programme also included a position paper by Adams that encouraged us to push the boundaries of GIR to take more account of processes, relations and inferred thematic concepts.

2nd ACM SIGSPATIAL Workshop on Analytics for Local Events and News (LENS 2018)
Seattle, Washington, USA - November 6, 2018

Amr Magdy
University of California, Riverside
amr@cs.ucr.edu

Xun Zhou
University of Iowa
xun-zhou@uiowa.edu

Liang Zhao
George Mason University
lzhao9@gmu.edu

Yan Huang
University of North Texas
yan.huang@unt.edu

The 2nd ACM SIGSPATIAL Workshop on Analytics for Local Events and News (LENS 2018) was held in conjunction with the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM SIGSPATIAL 2018). The workshop is intended to bring together experts in the workshop scope to exchange ideas on opportunities, challenges and cutting-edge techniques for local events and news analytics. The workshop has attracted 8 submissions, including 7 regular full-length research papers and one short paper, accepted 5 submissions for publications; with 62.5% acceptance rate. The papers are reviewed by 13 program committee members, where each paper is assigned to three reviewers. Total 7 attendants registered for the workshop. The actual number of attendants along the day fluctuated from 8 to 12, with an average of 9 attendants maintained almost all the time. The workshop has achieved its goal with bringing up together experts and conducted a set of productive discussions and talks.

The workshop has featured an excellent keynote speech by Prof. Feng Chen from the Computer Science Department at the University at Albany - State University of New York, where he directs the Event and Pattern Detection Laboratory. His talk discussed his work on "*A Unified Optimization Framework for Event and Pattern Detection in Attributed Networks*" where the main goal is providing an optimization framework for applications that exploits different types of networks such as transportation networks, social networks, diseases, and so on. The talks focuses on challenges that are imposed by big data that is often created by aggregating multiple data sources and modeled as large-scale networks. Many applications of big data analytics are concerned of discovering complex patterns (subnetworks) that are interesting or unexpected, such as the detection and forecasting of societal events (disasters, civil unrest), anomalous patterns (disease outbreaks, cyber-attacks), discriminative subnetworks (cancer diagnosis), knowledge patterns (new knowledge building), and storylines (intelligence analysis), among others. The talk presents a unified graph-structured optimization framework for solving a broad class of such problems that runs in nearly-linear time and at the same time provides rigorous guarantees on quality. This framework models the problems as non-convex optimization problems subject to combinatorial constraints, in which the objective function is defined based on attribute data and the constraints are defined based on network topology (e.g., connected or dense subnetworks). The key idea is to iteratively search for a close-to-optimal solution by solving easier sub-problems in each iteration: (1) identification of the subnetwork(s) that maximizes the objective function in a sub-space determined by the gradient of the current solution and the topological constraints; and (2) approximate projection of the identified subnetwork(s) onto the feasible space that satisfies the topological constraints. The talk has demonstrated the effectiveness and efficiency of the proposed approach using several real-world datasets. The talk has triggered a very interactive discussion with several workshop attendants.

LocalRec 2018 workshop report The Second ACM SIGSPATIAL Workshop on Recommendations for Location-based Services and Social Networks*
Seattle, Washington, USA - November 6, 2018

Panagiotis Bouros[1], Matthias Renz[2], Dimitris Sacharidis[3]
[1]Institute of Computer Science, Johannes Gutenberg University Mainz, Germany
[2]Department of Computer Science, Christian-Albrechts-Universität zu Kiel, Germany
[3]Institute of Information Systems Engineering, Technische Universität Wien, Austria
bouros@uni-mainz.de, mr@informatik.uni-kiel.de, dimitris@ec.tuwien.ac.at
(workshop co-chairs)

Driven by technological advances in hardware (positioning systems, environmental sensors), software (standards, tools, network services), and aided by various open movements (open, linked, government data) and the ever-growing mentality of sharing for the greater good (crowdsourcing, crowdfunding, collaborative and volunteered geographic information), the amount of available geo-referenced data has seen dramatic explosion over the past few years. Human activities generate data and traces that are now often transparently annotated with location and contextual information. At the same time, it has become easier than ever to collect and combine rich and diverse information about locations. Exploiting this torrent of geo-referenced data provides a tremendous potential to materially improve existing and offer novel types of recommendation services, with clear benefits in many domains, including social networks, marketing, and tourism.

Fully exploiting this potential requires addressing many core challenges and combining ideas and techniques from various research communities, such as recommender systems, data management, geographic information systems, social network analytics, text mining. The goal of the LocalRec 2018 workshop was to bring together researchers and practitioners from these communities providing at the same time a unique forum for discussing in depth and collecting feedback about challenges, opportunities, novel techniques and applications. The general theme in on making recommendations in which location plays a key role, either as part of the recommended object, or as part of the recommendation process.

LocalRec 2018 was held as a half-day workshop. The program committee received and evaluated 6 submissions (4 full papers and 2 short papers/demos), out of which 3 full papers and 3 short/demos were selected for publication and presented in the workshop. Among the main conference attendees, 10 registered specifically for our workshop, 19 people attended the workshop at peak time, while 15 was the average number of attendees; see Figure 1. The event was organized around two sessions. In the first session, Jose Macedo presented their study for trip and sightseeing tours planning [4]. The authors propose TrajectMe, an algorithm which extends the memetic-based state-of-the-art with hotel selection in several cities, taking advantage of the tourists' trajectories extracted from location-based services such as Foursquare and Flickr. In the same context, Madhuri Debnath presented their work on preference-aware travel recommendations with temporal influence [2]. The key idea of

*`http://www.ec.tuwien.ac.at/localrec2018/`

Figure 1: Group photo of LocalRec 2018 attendees.

this study is to first find interesting locations by considering user categorical preferences, temporal activities and popularity of location, and then, generate travel routes that include such locations while also specifying the visit time. Last, Vikram Patil presented their survey in the plain-text and encrypted domains for secured trajectory comparison [5]. This work also discussed potential methods for encrypted domain computing, which can be applied in the domain of trajectory similarity.

The second session opened with Keerti Banweer and their study on geotagging messages using techniques from recommender systems, more specifically, collaborative filtering [1]. Their work proposes a multi-stage iterative model based on the popular matrix factorization technique, which exploits the relationship of messages, location, and keywords to recommend locations for non-geotagged messages. Next, Suprio Ray presented their work on temporally relevant top-k spatial keyword search [6]. The authors focus on the parallel processing of the queries; for this purpose, they propose a novel parallel index, called Pastri. The index is built inside a system which offers persistent document storing and multi-threading functionality to exploit parallelism at various levels. Finally, Yuhao Kang presented their novel method for image positioning that combines spatial analysis and computer vision techniques [3]. The discussion revolved around their prototype that is based on large-scale Flickr photos and demonstrated a case-study for images taken of the Eiffel Tower in Paris.

In conclusion, we would like to thank the authors for submitting, publishing and presenting their papers in LocalRec 2018, and the program committee for their professional evaluation and help in the paper review process. We hope that the proceedings of the workshop will inspire new research ideas and that you will enjoy reading them.

References

[1] K. Banweer, A. Graham, J. Ripberger, N. Cesare, E. Nsoesie, and C. Grant. Multi-stage collaborative filtering for tweet geolocation. In *Proceedings of the 2nd ACM SIGSPATIAL Workshop on Recommendations for Location-based Services and Social Networks*, LocalRec'18, pages 4:1–4:4. ACM, 2018.

[2] M. Debnath, P. K. Tripathi, A. K. Biswas, and R. Elmasri. Preference aware travel route recommendation with temporal influence. In *Proceedings of the 2nd ACM SIGSPATIAL Workshop on Recommendations for Location-based Services and Social Networks*, LocalRec'18, pages 2:1–2:9. ACM, 2018.

[3] Y. Kang, S. Gao, and Y. Liang. Utilizing reverse viewshed analysis in image geo-localization. In *Proceedings of the 2nd ACM SIGSPATIAL Workshop on Recommendations for Location-based Services and Social Networks*, LocalRec'18, pages 6:1–6:5. ACM, 2018.

[4] E. Oliveira, I. R. Brilhante, and J. A. F. de Macedo. Trajectme: Planning sightseeing tours with hotel selection from trajectory data. In *Proceedings of the 2nd ACM SIGSPATIAL Workshop on Recommendations for Location-based Services and Social Networks*, LocalRec'18, pages 1:1–1:10. ACM, 2018.

[5] A. C. Pesara, V. Patil, and P. K. Atrey. Secure computing of gps trajectory similarity: A review. In *Proceedings of the 2nd ACM SIGSPATIAL Workshop on Recommendations for Location-based Services and Social Networks*, LocalRec'18, pages 3:1–3:7. ACM, 2018.

[6] S. Ray and B. G. Nickerson. Improving parallel performance of temporally relevant top-k spatial keyword search. In *Proceedings of the 2nd ACM SIGSPATIAL Workshop on Recommendations for Location-based Services and Social Networks*, LocalRec'18, pages 5:1–5:4. ACM, 2018.

PredictGIS 2018 Workshop Report
Held in conjunction with ACM SIGSPATIAL 2018

Akihto Sudo[1], Yoshihide Sekimoto[2], Lau Honng Chuin[3],
Xuan Song[4], Takahiro Yabe[5]

[1]Department of Computer Science, Shizuoka University, Japan
[2]Institute of Industrial Sciences, University of Tokyo, Japan
[3]School of Information Systems, Singapore Management University, Singapore
[4]Center for Spatial Information Science, University of Tokyo, Japan
[5]Lyles School of Civil Engineering, Purdue University, USA

The prediction of human and vehicle mobility in a city is becoming an attracting field. This topic attracts researchers from broad fields including behavioral sciences, where understanding the complexity of the human mobility behavior is one of the hot topic, and also to the industrial partners, who apply such results to many beneficial applications. Recent progress to sensing human mobility via smartphones is boosting this trend. However, due to the complexity and context-dependence of human behavior and the incompleteness and noise of geospatial data collecting from various sensors, the prediction of human and vehicle mobility is still far from solved. This workshop aimed at collecting contributions on the cutting-edge studies in human mobility description, modeling, intelligent computational method which can advance the human and vehicle prediction research. Potential topics included, but were not limited to 1) The next location prediction of individual mobility, 2) The crowd or population mobility prediction, 3) Dynamics of pedestrians, 4) commute flow and migration flow, 5) Traffic congestion, road usage forecast and optimal vehicle routing, 6) Social event forecast using geospatial data, 7) Novel agent mobility simulators, and 8) Case studies of mobility estimation in academia as well as in industrial field.

The second PredictGIS workshop was held on 6th November 2018 at the ACM SIGSPATIAL conference in Seattle, USA. At PredictGIS 2018, there were 7 presentations, of which 5 were full papers, 1 was a review paper, and 1 was a keynote presentation. The average number of attendees were around 10 people, with a maximum of around 15 people. There were 12 people who registered for the workshop. Overall, the workshop attracted papers with various topics, methods, and datasets. The variety of papers increased the number of topics covered in the workshop, and triggered an intense discussion between attendees on the current trends, issues, and also future research opportunities related to the prediction of human mobility.

Our keynote presentation was delivered by Dr. Naoya Fujiwara*, an Associate Professor from Tohoku University, Japan with the title, "Towards prediction of complex geospatial phenomena". Professor Fujiwara, an expert on complex networks and non-linear dynamics, covered various interesting cutting edge topics that utilized mathematical models and large scale data, including prediction of evacuation under severe disasters using mobile ad hoc wireless networks, spread of infectious diseases in cities, and urban level clustering of human mobility patterns. The theme of Professor Fujiwara's topic was on the fusion of geospatial big data (e.g. mobile phone data, social media data) with mathematical models for better understanding and prediction of complex phenomena. His keynote talk provided a novel direction of future research on human mobility

*https://sites.google.com/site/nfnetz/home?authuser=0 prediction, and inspired many of the attendants in this session.

In our first presentation, Dr. Christian Schreckenberger from University of Mannheim gave a talk on the review of next place prediction models. He was able to organize the large collection of works very well, focusing on the 4 points: Which features are used? (2) Which input data is required? (3) Which technique is used? (4) How is the prediction evaluated? This talk provided a good overview of the works in the field, and worked very well an an introduction to the workshop. The second talk was Sungha Ju's paper on understanding student characteristics from smart card data.

After the coffee break, Abdeltawab Hendawi gave a talk on a novel system named SimilarMove to predict the future paths of moving objects on road networks without relying on their past trajectories. Then, Douglas Teixeira gave a talk on the predictability of a user's next check-in using data from different social networks, and presented the interesting conclusion that the use of data from different social networks does not necessarily increase the predictability of a person next check-in, and that user behavioral characteristics play an important role on the predictability of the next check-in. Yuqin Jiang gave a talk on the analysis of Twitter geo-tag data to analyze the inter city connectivity. Applying this study to more detailed data such as mobile phone data could increase the impact of the work. Finally, Sung Bum Yun gave a talk on the implementation of floating population analysis which contains more information than traditional census population such as hourly based population and weekly based population. They presented a case study on Songdo area in South Korea.

As a whole, this workshop had very fruitful discussions along with very interesting and cutting-edge talks from the presenters. We would like to thank the presenters and attendees of the workshop for making it a huge success, and also the organizing members of ACM SIGSPATIAL 2018 for giving us an opportunity to hold this workshop.

GeoSim 2018 Workshop Report
The 1st ACM SIGSPATIAL International Workshop on Geospatial Simulation

Joon-Seok Kim[1], Hamdi Kavak[1], Andreas Züfle[1], Andrew Crooks[1,2], Umar Manzoor[3], Paul M. Torrens[4]
[1]Department of Geography and Geoinformation Science, George Mason University, USA
[2]Department of Computational and Data Sciences, George Mason University, USA
[3]Department of Computer Science, Tulane University, USA
[4]Department of Computer Science and Engineering, New York University, USA
{jkim258,hkavak,azufle,acrooks2}@gmu.edu, umanzoor@tulane.edu, torrens@nyu.edu

Geospatial simulation is a powerful tool for many disciplines including urban analytics, social science and meteorology to understand, explain, and predict overly complex natural and human-made systems. For instance, hurricane simulation has been used to predict hurricane paths and to make a decisions such as evacuation plans. Traffic simulation have enabled us to forecast traffic congestion in cities around the world.

With the advances in computing and software technology, simulation is becoming a commonly accessible solution. The spatial information community has a huge potential to contribute to the geospatial simulation and *vice versa* as it provides many research avenues including creating methods for the ingestion of big spatial data, studying domain-specific problems, and using simulation-generated spatial data. In particular, synthetic data generated by plausible simulation provides many advantages over publicly available data sets that are sparse and noisy.

The 1st ACM SIGSPATIAL International Workshop on Geospatial Simulation was held in conjunction with 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems in Seattle, Washington, USA. It is meaningful that GeoSim 2018 takes the initiative to bring together researchers from the social sciences and engineering disciplines to disseminate their cutting-edge research in geospatial simulation and to provide a forum for cross-pollination between these communities. To maximize synergy between the SIGSPATIAL communities and simulation communities, GeoSim organizers invited experts in both the communities to the program committee for GeoSim 2018.

To draw a broad audience, the workshop focused on all aspects of simulation. Specifically as a general paradigm to model and predict spatial systems and generate spatial data. The workshop had 11 registered participants which was in the middle range of all workshops at ACM SIGSPATIAL 2018. In addition to these participants, at times the workshop had upwards of 30 participants (ranging from academia (e.g. LMU Munich, Colorado State University, Eindhoven University of Technology) to industry (e.g. Here, Apple, ESRI, Google).

This half-day workshop comprised two concise but dense sessions. The workshop featured a panel discussion regarding grand challenges in geospatial simulation and three invited talks. Dr. Joon-Seok Kim kicked off the first session with welcome and opening remarks, having brief time for participants to introduce each others. In the first session, Dr. Andrew Crooks gave a talk entitled "*Geosimulation: A Gallery of Applications*" which set the scene for the workshop. Dr. Sarah Wise from University College London presented her recent research entitled "*Using an Agent-based Model to Explore Alternative Modes of Last-Mile Parcel Delivery in Urban Contexts*" [2] which proposed a practical application in delivery industry. Taylor Anderson from Simon Fraser University presented the paper entitled "*Geographic Network Automata for Representing Complex Evolving Spatial Systems*" [1] which demonstrated how geosimulation models can be applied to real world networks.

Figure 1: Group Photo of GeoSim 2018 Attendees

In the second session, Dr. Andreas Züfle gave an invited talk entitled "*Location-Based Social Simulation: Using Agent-Based Simulation to Simulate Location and Friendship*" which outlined how the SIGSPATIAL community could leverage geosimulation models for their research. Dr. Umar Manzoor presented his work on scaling agent-based models using distributed computing under the title "*Distributed Scalable Geospatial Simulation: Challenges and Future Directions*". All entrants including organizers actively participated in panel discussion entitled "*Challenges and opportunities for Geosimulation from Model to Data and Data to Model*." During which participants and the panel covered a range of diverse topics from best practices for building models, validating such model to approaches to data collection and storage.

Based on the submitted papers received, the best paper award was given to Taylor Anderson and Dr. Suzana Dragicevic from Simon Fraser University with their paper entitled: "*Geographic Network Automata for Representing Complex Evolving Spatial Systems*."

References

[1] T. Anderson and S. Dragicevic. Geographic network automata for representing complex evolving spatial systems. In *Proceedings of the 1st ACM SIGSPATIAL International Workshop on GeoSpatial Simulation*, pages 5–8. ACM, 2018.

[2] S. Wise, K. Cheliotis, O. Bates, A. Friday, J. Allen, F. McLeod, and T. Cherrett. Using an agent-based model to explore alternative modes of last-mile parcel delivery in urban contexts. In *Proceedings of the 1st ACM SIGSPATIAL International Workshop on GeoSpatial Simulation*, pages 1–4. ACM, 2018.

The SIGSPATIAL Special

ACM SIGSPATIAL
http://www.sigspatial.org