Modelling Crowds in Urban Spaces

Isaac Rudomín[1], Genoveva Vargas-Solar[2,3,4], Javier A. Espinosa-Oviedo[1,4],
Hugo Pérez[5,4], José-Luis Zechinelli-Martini[6]

[1] Barcelona Supercomputing Center (BSC),
Spain

[2] French Council of Scientific Research (CNRS),
France

[3] Laboratory of Informatics of Grenoble (LIG),
France

[4] French-Mexican Laboratory of Informatics and Automatic Control (LAFMIA),
France

[5] Universitat Politécnica de Catalunya (UPC),
Spain

[6] Fundación Universidad de las Américas, Puebla (UDLAP)
Mexico

genoveva.vargas@imag.fr, joseluis.zechinelli@udlap.mx
{isaac.rudomin, javier.espinosa, hugo.perez}@bsc.es

Abstract. Large scale crowd simulation and visualization combine computer graphics, artificial intelligence and high performance computing among other areas. Crowd sourced location data is used to compute spatio-temporal people and vehicle flows, while map and geometric data describe specific real places. With all this data, we can visualize both real trajectories and data driven on-line crowd simulation. We have some initial results using vehicle trajectory data.

Keywords**.** Crowd simulation, urban computing.

1 Introduction

In this era of big cities, one is confronted to emergency situations that require the intervention of qualified personnel to generate an orderly and safe urban experience (e.g., massive exodus, natural disasters, concerts, sports events, protests). In this context, data management and visualization techniques can support real-time observation for understanding the behavior of people and thus help in the development of security strategies (e.g., by supporting on-line and post-mortem analytics for guiding the process of recommendation and decision making in real-time). This paper presents one approach for simulating *crowd behavior* for supporting crowd behavior control in public spaces. Our objective is to visualize and predict the behavior of individuals and groups moving and evolving within real environments. For this purpose, we use geo located data produced by mobile devices and other sources of information (e.g., security cameras, DRONES) to predict individual and crowd behavior and detect abnormal situations in presence of specific events. We also address the challenge of combining all these individual's location with a 3D rendering of the urban environment. Our data processing and simulation approach are computationally expensive and time-critical, we rely thus on a hybrid Cloud-HPC architecture (CPUs + GPUs) to produce an efficient solution. Accordingly, the paper is organized as follows. Section 0 gives an overview of related work concerning individuals' location harvesting and crowd simulation techniques. Section 0 gives an overview of our approach consisting of two
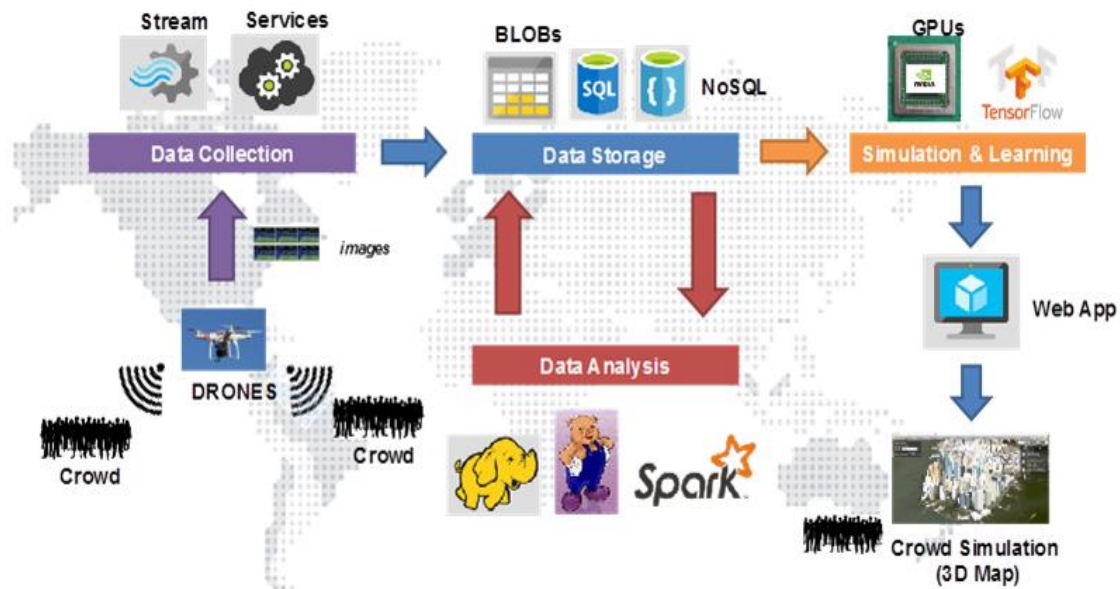
Fig. 1. General overview of the approach

contributions: modelling and locating individuals within crowds and simulating crowd behavior in natural and urban spaces. Section 0 gives the general lines of experimentation we conducted for following and predicting individual and crowd behavior in urban spaces. Finally, Section 0 concludes the paper and discusses future work.

2 Related Work

Studying the crowd (i.e., "mass or multitude" of people) is in the heart of research in different computer science disciplines. This paper combines results from visualization and database domains that have tackle the concept in different perspectives.

2.1 Data Visualization and Crowd Simulation

We can use simple visualization of trajectory data to aide decision taking. However, we believe that combining data analytics with the more complex crowd simulation and visualization techniques that have been used in the entertainment industry, where one must create populated environments that seem realistic, are an important way forward.

One of the challenges in crowd simulation is modelling realistic behavior of crowds within an environment (e.g., an army within the battle camp in a shooter game, fans moving to attend a concert or a football match). Rendering, visual variety, character animation, artificial intelligence and motion planning are common problems tackled in visualizing crowd simulations [1].

Modelling and visualizing crowd movement requires the virtual characters to be aware of their nearest neighbors for avoiding collision and communication. A naive neighbor search algorithm has a complexity of $O(n^2)$, where n is the number of simulated characters.

Crowd simulation is an important research direction in computer games, movies and virtual reality [2], urban planning, education and security, among others [3]. Large crowd simulation is in general coupled within the execution of another system, for instance a video game.

Thus, it is often restrained by the limited compute time available. Thus, the use of GPUs, has been proposed for achieving real-time crowd simulations. Realistic simulation of crowd movement can be achieved by observing and mimicking "real crowds".

2.2 Data Harvesting and Analytics for Monitoring Crowds

We are interested in techniques that use crowdsourcing (explicit and implicit) techniques for collecting data that contain information about the way people evolve in public and private places. These data collections can be used as input for learning crowd behavior and simulating it in a more accurate and realistic manner. The advance of location-acquisition technologies like GPS and Wi-Fi has enabled people to record their location history with a sequence of time-stamped locations, called trajectories. Some work has been carried out using cellular networks for user tracking, profiting from call delivery that uses transitions between wireless cells as input to a Markov model [4]. Wolf and others [5] used stopping time to mark the starting and ending points of trips. The comMotion system [6] used loss of GPS signals to detect buildings. When the GPS signal was lost and then later re–acquired within a certain radius, comMotion considered this to be indicative of a building. This approach avoided false detection of buildings when passing through urban canyons or suffering from hardware issues such as battery loss [7], introduces a social networking service, called GeoLife, which aims to understand trajectories, locations and users, and mine the correlation between users and locations in terms of user-generated GPS trajectories. GeoLife offers three key applications scenarios: (1) sharing life experiences based on GPS trajectories; (2) generic travel recommendations, e.g., the top interesting locations, travel sequences among locations and travel experts in a given region; and 3) personalized friend and location recommendation.

Existing work in robotics and autonomous vehicles applies automatic learning techniques for making them autonomous while they evolve in open spaces. They often use collected data for example, for training classifiers in order to reproduce the behavior of the crowd in synthetic environment.

One of the challenges in data analytics is to do prediction by deducing behavior models of the observed subject. A model is a collection of data on some particular aspect of a subject's behavior that, when associated with a limited set of contextual clues, yields predictions on what behaviour the subject will engage in next. Based on this notion, there is work similar to that described in [8], that use location as context to infer other data such as the presence of other people.

Predestination [9] is an approach that leverages an open-world modelling methodology that considers the likelihood of users visiting previously unobserved locations based on trends in the data and on the background properties of locations. Multiple components of the analysis are fused via Bayesian inference to produce a probabilistic map of destinations. The proposed algorithm was trained and tested using a database of GPS driving data gathered from 169 different subjects who drove 7,335 different trips.

The challenge is integrating computationally expensive data analytics with realistic character visualization within realistic environments giving the impression of reality to data analyzers who will be able to make critical decisions.

3 Using Big Data for Observing Crowds

Figure 1 illustrates the general overview of our approach that addresses three main problems: (i) data harvesting, (ii) crowd simulation and analytics and (iii) visualization. With respect to data harvesting, we can use different data sources like social networks, cellular communications and mobile devices data, and public cameras. Right now, we use existing temporal geo located observations concerning individuals' trajectories. In the future, we will combine these collections with drones for harvesting individuals and crowd observations. DRONES will fly specific zones and transmit observations that are correlated, processed, analyzed and visualized in 3D environments.

Building and storing representative data collections about individuals and crowd behaviour in urban spaces can be useful for performing offline analytics to discover patters, relations, and those members that might not belong to the group or that might have suspicious behaviour. *Data and analytics results must be stored in different supports according to the conditions in which they are shared and exploited*. We propose data

storage tools that use data sharing, geographical position and other context related strategies to distribute data across different stores, index with respect to different characteristics (time aspects), and efficiently deliver data to data analytics processes.

3.1 Identifying and Profiling the Crowd

We apply data analytics techniques (temporal and spatial reasoning) for computing trajectories and for identifying crowds. That is, people grouped in a proximity that adopt a specific "behaviour" referring to four well known naïve crowd patterns: (i) casual crowd which is loosely organized and emerges spontaneously, people forming it have very little interaction at first and usually are not familiar with each other; (ii) conventional crowd results from more deliberate planning with norms that are defined and acted upon according to the situation; (iii) expressive crowd forms around an event that has an emotional appeal; (iv) acting crowd members are actively and enthusiastically involved in doing something that is directly related to their goal.

The objective of our analytics study is to identify a triggering "symptom" that can evolve into the constitution of a crowd. For example, someone showing a banner in the middle of a plaza, in front of some monument or government office; people density increasing in some area. In our approach and since for the time being we do not use images recognition we address crowd creation identification by measuring people density in specific spatial regions during a time interval. This requires a continuous analysis of the evolution of the status of the areas of an urban space in order to measure the population density.

Density measuring is done using different data collections: (i) the continuously harvested observations of the geographical position of individuals (that accept sharing their position) along time; (ii) the images stemming from cameras observing specific "critical" urban areas, like terminals, airports, public places and government offices; (iii) data produced by social networks and applications like Twitter, Facebook, Waze and similar. The occupation density of specific urban regions is measured separately according to the political organization of the space (quarters, areas)

in every database. We use sliding windows in order to partition continuous data flows with respect to time intervals and we use the political division of the urban space for filtering, grouping the data and computing density per urban region. This is a straightforward yet somehow costly computation not because of the amount of data but because it should be continuously computed, and both data and density results are stored for performing other analytics processes.

Having different visions of the density of urban spaces given different data collections enables to perform other types of data analytics on the n-tuples region-density and to cluster regions both taking into consideration their density and their geographic position. Accordingly, we generate a "crowd heat thermometer" showing the crowd dynamic distribution view of the urban space that evolves along time.

Having an insight within the crowd. Not all crowds need to be managed within urban spaces, there are some that happen every day in public transport and others need particular attention and must be better profiled. The first challenge is to be able to discriminate. Using data collections harvested from social networks, some correlations are computed to identify connected people that were in the same urban place at some time interval, and that might have been part of a crowd event. Not all individuals sharing the same spatial region necessarily participate in a crowd. Thus, the first operation to solve is given a set of individuals located in the same geographic region at the same time interval, whether an individual located at the same space-time belongs to the crowd. A naïve way of evaluating this predicate belongs-to, is knowing whether there is a "social" connection between an individual and at least one of the crowd members. Using this analytics operation, it is possible to draw an urban occupation map and propose some connections among individuals occupying the same urban region at the same period of time. The objective of this classification is to identify possible outsiders, that is, individuals that do not really belong in the group. Those are the ones in which we are particularly interested because our guess is that within those outsiders we might find human traffic dealers, for example.

We adopt two complementary strategies for defining possible networks hidden in a crowd. First,

we retrieve the contacts graphs of individuals social networks accounts which have been identified in a crowd. We compute relations with other contacts using well known relations discovery techniques as the one proposed by [10]. Filtering strong related contacts, we verify whether they are themselves present in the crowd. The process is computationally expensive because we are looking for graphs intersections of possibly thousands of individuals. Yet, it is incremental in the sense that graphs are stored and they are enriched with new information coming from the observation of other crowds in other moments. These graphs versions are used for profiling and predicting the crowd behaviour.

Profiling and predicting the crowd behaviour. Once the crowd has been formally identified and modelled in terms of its participants, it is possible to observe its behaviour as if it was an individual: (i) characterize its elements (find the leaders, the followers and eventually the outsiders); and (ii) predict the evolution of its behaviour for example, probability of conflict, space utilization and risk maps.

Once we have created a first connection graph describing how people possibly participating in the crowd are connected among each other, we apply methods to determine which is the degree of influence of each participant of the crowd. This is done by computing the influence of the elements of the crowd towards other elements according to their contacts network and that can eventually be participants in the crowd. The analysis ends up with groups of outsiders that have to be inspected to understand their presence in the crowd, and their possible role in the event. Again, since the inclusion of participants in the crowds evolves along time, the contacts' network represented by graphs and the influence of participants varies too. The computational cost is important considering semi-post-mortem computations. Of course, the ultimate objective is to be able to observe this evolution in real time, which introduces scalability problems that must be addressed with GPU architectures.

The final task is to predict the behaviour of the crowd. Therefore, we use the notions of space occupation and probability of conflict, by searching behaviour patterns in the evolution of the crowd status: it emerges with some individuals, increases

its size, in achieves the maximum of participants and then it fades. This "life cycle" happens within a space occupation process that can be controlled by its inherent behaviour but that can also be determined by external factors e.g., police, troublemakers. We define the status of the crowds with a set of attributes including, the approximate amount of participants, the main trajectories of the group, the spatio-temporal region occupied by the group, the possible outgoing directions in which participants can move within the urban space, triggering and termination event.

Instead of delivering textual or graphical results of these analytics operations, we aim to provide 2D and 3D visualizations that can reproduce the observations and simulate the behaviour of the crowd according to real data. The following section explains how.

3.2 Simulating and Visualizing Large Crowds in Real Time

There are several steps in the process of simulating and visualizing large and varied crowds in real time for consumer-level computers and graphic cards (GPUs). Animating varied crowds using a diversity of models and animations (assets) is complex and costly. One has to use models that are expensive if bought, take a long time to model, and consume too much memory and computing resources. We propose methods for simulating, generating, animating and rendering crowds of varied aspect and a diversity of behaviours.

In general, the principle consists in mapping human perception of the space stemming from cameras and expressed in geographical coordinates (latitude, longitude), for example, into pixels. For instance, as shown in Figure 2, "*Give me the GPS coordinates of the users evolving in Beijing ordered by time*". Once this query has been evaluated by the appropriate data processing infrastructure (in the work presented here Pig Latin [11] execution environment), results are transformed into the appropriate format. Textures and maps are retrieved in order to create the 3D space where individuals' movements will be visualized (simulated) according to the observed information. The visualization of a dynamic situation requires computing capacity in order to be sure that the rendering is realistic. Thus, for
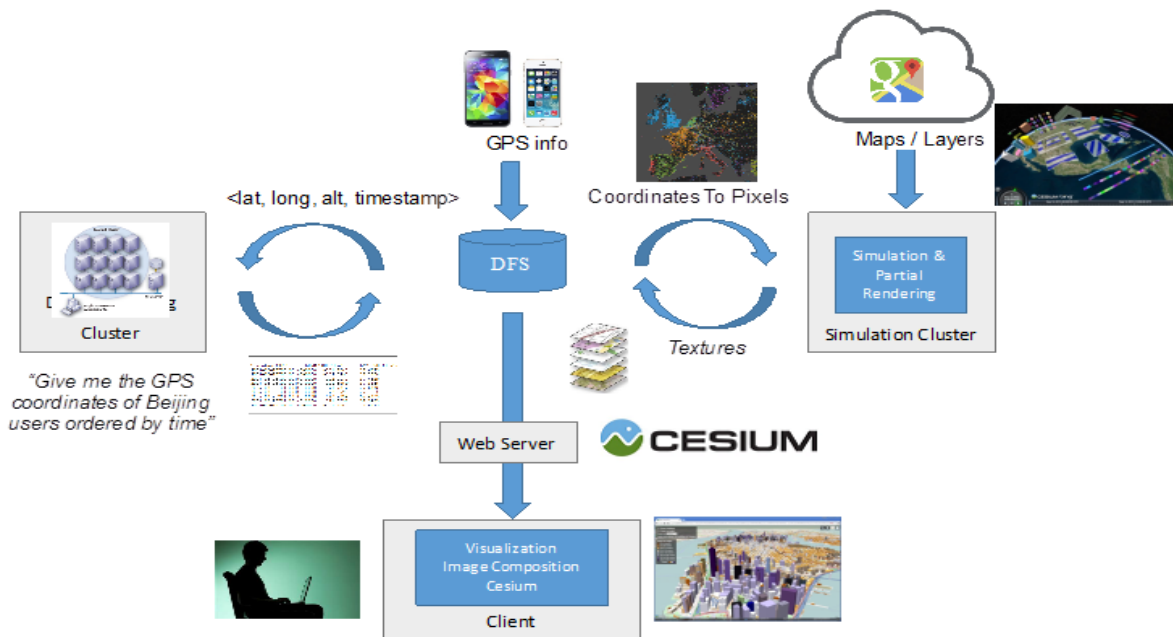
Fig. 2. Visualization process of individuals' movement within urban 3D spaces

scalability we rely on a simulation cluster devoted to the execution of this costly task (see right side of Figure 2). Simulation and visualization is based on the work described in [12], modified to follow our data. The simulation algorithms have been modified such that when given a choice, simulated vehicles and pedestrians will use heat maps derived from the dataset in order to follow the most popular routes.

The details of these modifications are not within the scope of this paper and will be published later. Machine learning techniques for understanding data are to be used for visualizing the crowd and then for simulating its behaviour given some specific events. Visualization has also been modified to include vehicles as well as pedestrians, and to be rendered aligned with and combined with the scenery provided by Cesium.

The graphic vision does not only contain the view of the urban space it also draws individuals moving in it. The level of detail that a data consumer can see about people should depend on her access rights to personal information and on the privacy laws of the geographic space in which observed people is moving. For example, a consumer with few access rights would see people as points or avatars (like video game characters) moving in the street. Yet a consumer with enough access rights should see the actual person walking in a given street. Since there are no cameras everywhere, in our approach we would simulate characters in real time according to the physical characteristics of people that can be available (on social networks, files, seen in cameras, or explicitly provided by people or security institutions).

4 Initial Experimentation

We aim to exploit data collections about the way people move in public places for learning about transit and crowd behaviour and accurately and realistically simulating it. The graphic vision is to combine a 3D view of the urban space and draws individuals moving in it. A data analyst with enough access rights can see the actual person walking in a given street. Since there are no cameras everywhere, in a first experiment we infer characteristics to the characters from the movement, or if explicitly provided by people.

Fig. 3. Computing trajectories

We use the *GeoLife GPS trajectory dataset [7]* with data of 182 users, 17,621 trajectories of *ca.* 1.2 million Km. and 48,000+ hours. Thes*e* data is used to compute spatio-temporal people flows in real crowds to provide data driven on-line crowd simulation, enhanced with real places geometric

data running on GPU and HPC[1]. Since the data set for a given place and time is sparse, we will modify agent based microsimulation to complement the actual trajectories in the dataset by using all the trajectories in similar moments that are available in the dataset to derive the most probable trajectories for the simulated vehicles or pedestrians. Visualization of real and simulated

vehicles and/or pedestrians on the appropriate section of the planet uses Cesium, an open-source JavaScript library for 3D globes and maps (see www.cesiumjs.org), and custom software for playing multiple vehicles and animated pedestrians efficiently.

4.1 Computing Trajectories

The Geolife GPS trajectory dataset is used for computing the trajectories of the vehicle observations it contains. We used Big Data cleaning and processing tools particularly the language Pig Latin for computing such trajectories.

---

[1] See a video of our first experiment in
https://drive.google.com/open?id=0B16gYdkmxnihU2gzM0R0MC1WRnM

The process is done by defining four declarative expressions as shown in Figure 3.

– Q_1: Load the observations clean and prepare the data collection with respect to the initial meta-data of the observations. For computing trajectories three attributes were necessary, the initial and termination times and the transportation modes. The idea is that a trajectory is defined as a set of locations observed at a given moment identified by a time stamp. The PigLatin program presented in the figure implements this process.

– Q_2: Load the GPS logs filtering the latitude, longitude and time stamp according to a predefined schema defined for this purpose (see Figure 3).

– Q_3: Finally, the third query estimates the trajectories as sequences of locations where the sequence is determined by the time stamps. With these computed trajectories it is possible to perform other analytics operations, and reconstitute the movement of people in the corresponding urban space, in this case Beijing.

Even if they seem simple due to the expression power of Pig Latin, the execution of these queries can be computationally costly given the volume of data we processed. The Pig Latin execution environment was installed in a cluster of 8 machines and executed in parallel. We thus obtained results in reasonable execution time. The computed trajectories were stored in an integrated data collection on top of which we performed other operations as described in the following lines.

4.2 2D tracking of Individuals

Our objective is to track individuals and crowds in urban spaces. We compute heat maps in order to aggregate the trajectories of users during specific time intervals and we used a map service in order to visualize them. The result, as illustrated in Figure 4, shows those itineraries that are popular in red or thick lines. These heat maps concern the Beijing trajectories computed using the *Geolife GPS* trajectory dataset.

Heat maps enable tracking individuals within a specific urban space and see how they move during a specific period of time. The figure shows the simple code for computing heat maps that again serve as new data on which it is possible to perform more analytics. These analytics concern for example identifying the most visited regions, observe rush hours in certain regions, and eventually identify crowds. For the time being our experimentation is done in post-mortem.

4.3 3D Visualization of Individual Tracks

The visualization problem is state as follows. It is centered on the user that accesses a 3D virtual space and navigates in it, in our case, for following individuals or the crowd within the simulated urban space. Every interaction of the user with the 3D space triggers a process in which data are retrieved from a storage support (memory or disk) to feed the simulation module that will reproduce a realistic behaviour of the 3D space.

A realistic behavior in our case means that the individual or the crowd will move and behave smoothly as it is seen in "physical reality" and it will react to the environment if some event is produced (due to the interaction of the user). In the case of our experiment for individuals this implies maintaining the position of the individual on the road, respecting the traffic signals (e.g., semaphores, speed, direction), and the organization of the urban space (e.g. buildings, bridges, roads, parks and other open spaces).

We performed our tests using a workstation with an Intel Core i7-4820K CPU 4-cores at 3.70 GHz (hyper threading is disabled), running Linux operating system with 16 GB of RAM memory and 10MB of cache memory. It includes a GeForce Kepler GTX TITAN Black with 2880 cores and 6 GB of GDDR5 memory.

In the case of the crowd, it is similar, as the user interacts with the 3D environment the crowd must behave according to some of the patters we identified and defined in Section 0.

Of course, it would be very difficult to model every single possible behavior pattern both of individuals and the crowd, and this would lead non-realistic behaviors in the 3D virtual space. Thus, in the future we might need to use automatic learning techniques (deep learning) so that the system can "react" to events and simulate a synthetic behavior of the crowd. The computational cost is high so this
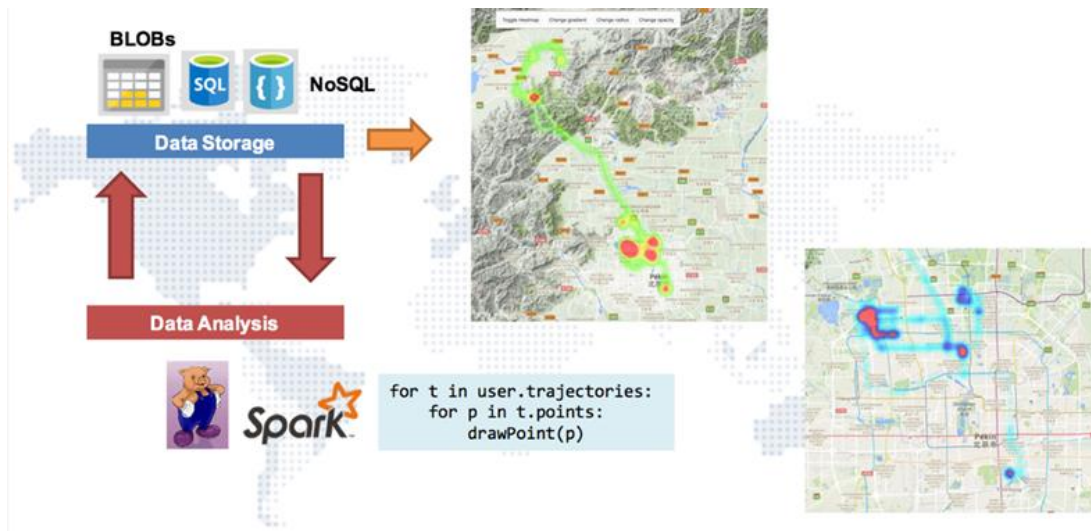
Fig. 4. Heat maps aggregating the trajectories of the same user during a time interval

Fig. 5. Visualizing individuals in 3D urban spaces

complex data processing and classification tasks are to be executed using GPUs.

4 Conclusions and Future Work

Crowd sourced location data is used to compute spatio-temporal people flows in real crowds. We combine both to provide data driven on-line crowd simulation, enhanced with real places geometric data running on GPU and HPC.

This paper presented the general approach for simulating crowd behaviour and thereby supporting individuals' and crowd behaviour in public spaces. The main contribution is combining location based data collections previously harvested together with online geo-tagged data for

visualizing crowds at different levels of precision and detail according to access control and privacy constraints. Our data processing and simulation process are computationally expensive and critical; thus, we rely on hybrid cloud-HPC infrastructures for producing an efficient solution.

Acknowledgements

References

1. **Thalmann, D., Grillon, H., Maim, J., & Yersin, B. (2009).** Challenges in Crowd Simulation. *Conference on CyberWorlds (CW'09).* IEEE Xplore Digital Library.

2. **Guodong, R., Yang, L., Wenping, W., Xiaotian Y., Gu, X. D., & Xiaohu, G. (2011).** GPU-Assisted Computation of Centroidal Voronoi Tessellation. *IEEE Transactions on Visualization and Computer Graphics,* Vol. 17, No. 3.

3. **De Gyves, O., Toledo, L., & Rudomín, I. (2013).** Comportamientos en simulación de multitudes: revisión del estado del arte. *Research in Computer Science, Special Issue: Avances en Inteligencia Artificial,* Vol. 62.

4. **Bhattacharya, A. & Das, S. K. (2002).** LeZi-Update: An Information-Theoretic Framework for Personal Mobility Tracking in PCS Networks. *Wireless Networks,* Vol. 8, No. 2/3.

5. **Wolf, J., Guensler, R., & Bachman, W. (2001).** Elimination of the Travel Diary: Experiment to Derive Trip Purpose from Global Positioning System Travel Data. *Transportation Research Board, 80th annual meeting.*

6. **Marmasse, N. & Schmandt, C. (2000).** Location-Aware Information Delivery withComMotion, *Symposium of Handheld and Ubiquitous Computing.*

7. **Zheng, Y., Xie, X., & Ma, W. Y. (2010).** GeoLife: A Collaborative Social Networking Service among User, location and trajectory. *IEEE Data Engineering Bulletin*, Vol. 33, No. 2.

8. **Ashbrook, D. & Starner, T. (2003).** Using GPS to learn significant locations and predict movement across multiple users, *Personal and Ubiquitous Computing*, Vol. 7, No. 5.

9. **Krumm, J. & Horvitz, E. (2006).** Predestination: Inferring Destinations from Partial Trajectories. *8th Int. Conference on Ubiquitous Computing (UBICOMP'06).*

10. **Garcia-Gasulla, D., Alvarez-Napagao, S., Tejeda-Gómez, Oliva-Felipe, A. L., Gómez-Sebastian, I., Bejar, J., & Vázquez-Salceda, J. (2014).** Social network data analysis for event detection. *21st European Conference on Artificial Intelligence (ECAI'14).*

11. **Olston, C., Reed, B., Srivastava, U., Kumar, R., & Tomkins, A. (2008).** Pig latin: a not-so-foreign language for data processing. *Proc. of the 2008 ACM SIGMOD Int. Conference on Management of data (SIGMOD'08).*

12. **Hernandez, B., Perez, H., Rudomin, I., Ruiz, S., DeGyves, O., & Toledo, L. (2014).** Simulating and Visualizing Real-Time Crowds on GPU Clusters, *Computación y Sistemas,* Vol. 18, No. 4, pp. 651–664.