Detecting and Classifying Social Events Based on Formation Process of Human Crowds

Takuro Yonezawa, Jin Nakazawa
Graduate School of Media and Governance
Keio University
Fujisawa, Japan
Email: {takuro,jin}@ht.sfc.keio.ac.jp

Tomohiro Nagata
Research Laboratories
NTT DoCoMo, Inc.
Yokosuka, Japan
nagatat@nttdocomo.co.jp

Hideyuki Tokuda
Graduate School of Media and Governance
Keio University
Fujisawa, Japan
hxt@ht.sfc.keio.ac.jp

***Abstract*—This paper proposes a method to detect and distinguish the types of social events based on formation process of crowds. We often form crowds according to social events such as "car accidents, baseball games or music concerts". Since such kinds of crowds sometimes harm our social activities (e.g., bad traffic), detection or prediction of crowds and social events are important. This paper focuses on detecting and predicting crowds and social events using location information, such as GPS and WiFi. The aim of this paper is two folds; (1) discussing and classifying social events based on crowds' characterization and (2) providing design and implementation of discrimination algorithm to detect social events. Through simulator-based implementation and evaluation, we present the effectiveness of our approach and discuss the pros and cons of the system.**

***Keywords*-Social Computing, Social Context, Mobile Sensing**

I. INTRODUCTION

People in cities gather locally and instantly, forming a crowd, for such a special event as a music concert, car accident, or new year's countdown. Crowds in cities often effect our social activities negatively in terms of social efficiency. For example, a crowd of people waiting for a new year's countdown, that of curious bystanders around a car accident, or a mass demonstration would generate a traffic jam. Detection of such human crowds are promising ways to realize pervasive computing environment in large scale. A car navigation system would benefit from real-time crowd detection for automatic rerouting. A city planner would leverage crowd information of a longer period of time to re-design the social infrastructure in the city. Therefore, we need a sophisticated system that is aware of crowd characteristics, e.g., assembly and disassembly of people, number of people, and their source events.

The purpose of our research is to detect and classify *social events* which are one of the triggers to generate crowds. There are many types of social events in urban area such as sport events, festivals or concerts, what we call *planned social events*. Additionally, there are another types of social events such as curious onlookers of car accidents or street performances, what we call *incidental social events*. We present a social events analysis method based on formation process of crowds on using location information which are generated from GPS or WiFi localization technologies[1], [4]. In this paper, we firstly classify social events into four kinds according to two indices: time until crowds are formed, and travel distances of people who compose the crowds. Then, we propose social events classification methodology and implement it on a simulator which simply models people's movement to compose crowds in accordance with generation of various types of social events.

The rest of this paper is organized as follows. In the next section, we present motivation of the research with related work. Then, we discuss the characterization of social events and classified social events according to their properties in section 3. Based on this classification, social events distinguishing methodology is explained in section 4 with description of its implementation and experiment on a simulator. We also discuss limitation of our methodology in section 4. Finally, we conclude the paper, and state the future direction of this research.

II. CROWDS AND SOCIAL EVENTS

In this section, we firstly show our motivation to analyze relationship between crowds and social events. Then, we present related works for further understanding of the motivation and context of this research.

A. Motivation

Crowds, a large number of people gathered closely together, often bring on troubles in social activities. Especially, since the population densities of some cities in Japan are very high, crowds often emerge and negatively affect social infrastructures. For example, in fireworks festivals in Japan, cell-phone services are often interrupted due to a huge number of people flocking to tiny area of the festivals. In another example, during the period of FIFA World Cup™2010, surge of fans onto the street broke in on traffic system in *Shibuya* area. Therefore, detecting or predicting crowds is important for people/city/company to plan or change their activities.

In many cases, crowds are organized by people who has same purpose (e.g., watching baseball games or shopping at a newly opened store). We call these factors as *social events* in this paper. There are many types of social events, whose scales and/or periods differ from each other. We also consider that the types of social events decides crowds' property. For example, social events such as concerts or baseball games are strictly scheduled. In those events, crowds are perhaps gently organized before events are started, and rapidly disappear after the events are finished. Meanwhile, social events such as car accidents or fire accidents are suddenly occurring ones, hence, crowds are organized rapidly compared to those of scheduled events.

Conversely, we consider that the types of social events can be identified by analyzing formation process of crowds. [2] proved that participants in social events are strongly correlated with their home location. Additionally, [2] presented that the same type of events show similar spatial distribution of participants' origins. Our challenge, proving relationships between social events and formation process of crowds, was also strongly motivated by these facts.

B. State of the Art

Analyzing characterization of crowds has a great potential to provide many types of applications such as crowd management, public space design or emergency evacuation. For example, [8] proposes a "Social Coordination" system to provide limited service resources to mass users with considering their mass intentions. The traditional crowds or social events analysis approach is achieved by aggregating data in small scale area[3] or control point such as counting the number of tickets sold[9]. In [9], special events are classified in the viewpoint of the events' periods or size, and it analyzes the effects of crowds or events on social economics and transportation systems.

Recent progress of mobile devices which are integrated with various localization technologies, such as GPS, WiFi or Bluetooth, enables researchers to analyze practical human behavior in larger scale fields.To understand individual human mobility pattern, the trajectory of 100,000 anonymized mobile phones' position were tracked and analyzed for a six-month period[5]. Following [5], a mount of history of each trajectory showed that humans follow simple reproducible patterns, despite the diversity of their travel history.

Not only individual human behavior but also collective human behavior is also an interesting subject to be understood. Especially, collaboration research with mobile telecommunication companies to access to aggregate mobile phone data has been opened up recent methodology and experiences of analyzing urban dynamics[6], [7], [2]. In [6], social activity pattens were observed to visualize time modes of urban space use by monitoring the intensity of people chatting and population density based on cell-phone traffic data. As we mentioned in previous section, [2] also presents the relationship between social events and participants in terms of spatial viewpoint. The literature also proposes prediction methodology for the question of where people will come from for future events, and the prediction is partly succeeded.

Analyzing mass human behavior and bringing out relationships between the behavior and social events are one of key challenges to expand pervasive computing applications from small spaces such as indoors or buildings to larger spaces such as cities, countries or global fields. To encourage the research, modeling of crowds' behavior in various types of social events is an important issue.

III. Social Events Classification

Social events can be classified according to its contents, such as, sports, cinema, music or street performance. Obviously, by referring to place information (e.g., stadium, museum or park) or public announcement (e.g., public schedule on the premises' website), the types of social events could be identified easily. However, our question is "Is it possible to suppose the types of social events without knowing such contents information, but only by monitoring crowds' context ?". If it is possible, not only planned social events but also incidental social events, which have no preliminary information, could be identified. Additionally, the estimated types can complement to recognize correct types of events in cooperate with contents information.

To achieve the goal, we take a notice of an observation that social events can be classified according to formation process of crowds which are gathered following the occurrences of events. As we mentioned before, [2] presented relationship between humans' home locations and types of social events. Based on and extending the perspective, we introduce the relationship between temporal-spatial property of crowds and social events. We consider that audiences around street performance or customers to newly opened supermarket are rapidly gathered from neighbor location. On the contrary, fans of baseball games or audiences in concerts are supposed to gather gradually from farther locations compared with participants in the former case. Based on this observation, we introduce the following two indications to distinguish social events: *time before a crowd fully grows* and *travel distance of people who constitute the crowds*. Fig. 1 shows this classification, and its examples in accordance with above two indications. Social events of *Narrow-Rapid* type in the Fig. 1 means that participants from neighbor location organize crowds rapidly (several hours). On the contrary, social events of *Wide-Gradual* type means that crowds, which consists of participants from far locations, continues to be formed in longer time (e.g, more than several hours or almost a day).

In order to examine the adequacy of this observation, analysis of collective human mobility data is indispensable. As the first step to present this, the paper proposes a social

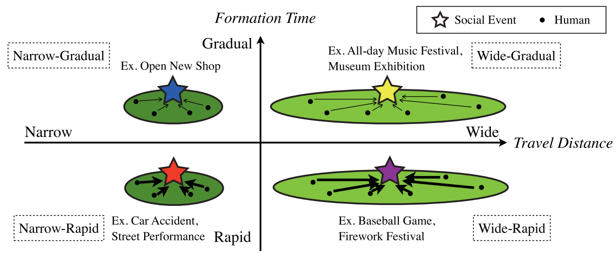

Figure 1. Assumption: social events are classified according to two indications, (1) time of crowds are being grown and (2) travel distance of participants who join the crowds

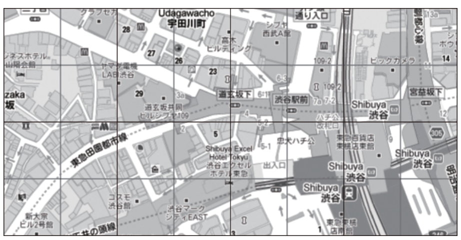

Figure 2. An example of a field and matrix sub-fields

events identification methodology by detecting above two indications.

IV. Methodology

In this section, we present social events identification methodology. Then, we describe prototype implementation with simulator which expresses human mobility patterns according to our assumption. We also present initial experiment to evaluate the methodology, and discuss limitations of it.

A. Social Events Property Recognition

To identify the types of social events, we define property of social events in following expression: *SocialEventProperty* = { *Location, Number OfTheParticipants, FormationTime, AverageTravelDistanceOfTheParticipants* } Each elements is mainly calculated by following two processes: (1) crowds detection and (2) property analysis. Let us introduce a supposition field, which is devided into matrix sub-fields (see Fig.2), for simple explanation. We describe each process below in order.

1) Crowds Detection: Crowds are detected by calculating the density of population stayed in each sub-field. Just acquiring density from one snapshot cannot judge high density means crowds are generated or not. To recognize crowds which is related to social events, movement information of participants in the crowds should be obtain. Thus, we analyze temporally transition of human location, and detect crowds based on the density of population staying at the same position.

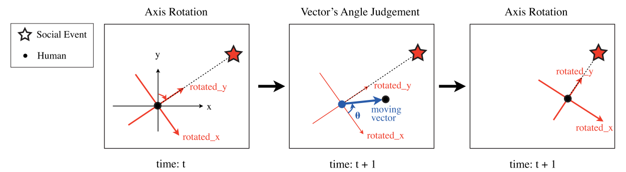

Figure 3. Calculating travel distance: if angle of each moving vector, θ, continues less than 180 degrees, a participant is judged to move towards a event place.

2) Property Analysis: Properties of social events, which contains formation time and travel distance of the crowds' participants, are derived based on detection of the crowds. Formation time is the time since the growth of crowds started and until it finishes. Notice that whole formation time of crowds can be divided into 3 categories: growing status, stable status and disappear status. These information is also useful for various applications, however, this paper simply focuses on using growing period to identify social events. This is because, we thought growing period can be adopted to various kinds of applications which need information about precognition of social events.

The methodology of detecting second indication, average travel distance of participants, needs discussion, because, it is difficult to distinguish the location where participants started to head the target place. The simplest way is to measure the distance between two points of each participa nt: from a location for certain period of time to a location where crowd is observed. However, in this way, we cannot deals the sudden situation (e.g., when a participant is suddenly attracted to the nearby event and change their behavior to get closer to the event). To deal such situations, we propose an origin recognition method by leveraging a spatial relationship between participants' moving vectors and the destination (i.e., location of the events).

Firstly, each participant's axis is rotated by following line connected between a participant's current location and the destination as right angle (see left figure in Fig.3). After the participant moves, the angle of moving vector on rotated axis (θ in center figure in Fig.3) is evaluated; if the angle is less than 180 degrees, it means that the participants is moving towards a event place. When a participant becomes a part of a crowd, this angle judgement is executed. The calculation is executed repetitively according to the participant's location history until the angle becomes greater than 180 degrees or the participant stays in a same location for a long time (we define these point as origin location). After finishing the analysis, travel distance of the participant is measured as distance between the origin location and the location of the event.

B. Initial Evaluation

Without subscribers permission, actual use of mobility data by a mobile phone company is prohibited by Japanese

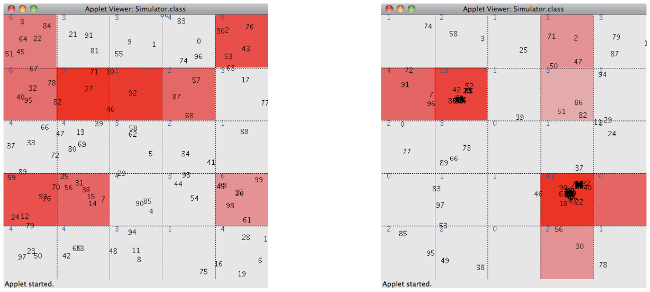

Figure 4. Simulator: Robots walks randomly (left), and compose crowds according with different formation processes (right).

law. Thus, as an initial evaluation, we checked validity of the proposed methodology on a simulator which simulates imaginary human behaviors by using walking robots(see Fig.4). In the simulator, robots normally moves 1 pixel as random walking in every per 100 milliseconds. When a point on the simulator is clicked (which means a social event is occurred), multiple robots creates crowd at the point by following the type of events which we assumed in section 3. Location histories of every robots are stored, and when a crowd is recognized, social event properties are calculated. In Fig.4, the number of blue colored number placed at upper left corner in each sub-field shows the density of population stayed in a certain time. Some sub-fields are filled with red: the brighter of the color shows the higher density of robots stayed.

Table I shows analyzed values of social event properties on the simulator of 500*500 pixels field when following 4 Patterns of robots' behaviors are executed:

P1: 20 of 100 robots near the clicking point rapidly gather.

P2: 20 of 100 robots near the clicking point gradually gather.

P3: 20 of 100 robots which are randomly chosen rapidly gather.

P4: 20 of 100 robots which are randomly chosen gradually gather.

"Rapidly" in Pattern 1 & 3 means that all selected robots move towards clicking point immediately. Meanwhile, "gradually" in Pattern 2 & 4 means that third part of robots start to move immediately when an events is occurred, next third part of robots start to move 1.5 seconds later, and last third part of robots start to move 3.0 seconds later. In this experiment, we confirmed that different 4 Patterns can be recognized from average travel distance and formation time, which are calculated by our proposed methodology: average travel distance of Pattern 3 & 4 is longer than it of Pattern 1 & 2, and formation time of Pattern 2 is longer than it of Pattern 1 (also time of Pattern 4 is longer than it of Pattern 3). Though we must take notice that the experiment is executed under particular assumption of human behavior,

Table I
CALCULATED SOCIAL EVENT PROPERTIES BY FOLLOWING 4 PATTERNS.

	Pattern 1	Pattern 2	Pattern 3	Pattern 4
Average Travel Distance (pixel)	94.52	78.26	201.73	223.99
Formation Time (sec)	1.89	6.53	8.55	12.06

the result of the experiment partially showed availability of the proposed methodology.

C. Discussion

Our methodology has two limitations. The density of staying population is expressed in static matrix sub-fields. More flexible crowd detection such as real-time clustering (e.g.,applying k-nearest neighbor algorithm) is necessary.

Secondly, the methodology of detecting travel distance needs to deal more complex human behaviors. While implemented simulator simulates simple behaviors of human, there are more diverse behaviors in real-world. For example, in some cases human needs to take a long way round to the destination due to road situations. Thus, the methodology needs to be improved in considering with more realistic environment.

V. CONCLUSION

Based on an observation of the relationship between crowds and social events, we classified social events in terms of crowds' two indices; time before a crowd fully grows and average travel distance of participants to the crowd. Additionally, we presented social events distinguish methodology and confirmed it's availability by using a simulator which simply simulates imaginary human mobility patters. Further studies considering larger real datasets of human behavior and concrete fields should be performed to evaluate and improve our observation and method.

ACKNOWLEDGMENT

REFERENCES

[1] Airsage: Live, real-time traffic information, real-time travel times, location, movement data. http://www.airsage.com/.

[2] F. Calabrese, F. C. Pereira, G. Lorenzo, L. Liu, and C. Ratti. The Geography of Taste: Analyzing Cell-Phone Mobility and Social Events. In P. Floréen, A. Krüger, and M. Spasojevic, editors, *Proceedings of the 8th International Conference on Pervasive Computing*, pages 22–37, Berlin, Heidelberg, May 2010. Springer Berlin Heidelberg.

[3] A. C. Davies, J. H. Yin, and S. A. Velastin. Crowd monitoring using image processing. *Electronics and Communication Engineering Journal*, 7:37–47, 1995.

[4] N. Eagle and A. (Sandy) Pentland. Reality mining: sensing complex social systems. *Personal and Ubiquitous Computing*, 10:255–268, 2006. 10.1007/s00779-005-0046-3.

[5] M. C. Gonzalez, C. A. Hidalgo, and A.-L. Barabasi. Understanding individual human mobility patterns. *Nature*, 453(7196):779–782, June 2008.

[6] R. C. T. E. Pulselli R M, Romano P. Computing urban mobile landscapes through monitoring population density based on cell-phone chatting. *International Journal of Design & Nature and Ecodynamics*, 3(2):121–134., 2008.

[7] J. Reades, F. Calabrese, A. Sevtsuk, and C. Ratti. Cellular census: Explorations in urban data collection. *IEEE Pervasive Computing*, 6:30–38, 2007.

[8] A. Sashima, N. Izumi, and K. Kurumatani. Consorts: A multiagent architecture for service coordination in ubiquitous computing. In K. Kurumatani, S.-H. Chen, and A. Ohuchi, editors, *MAMUS*, volume 3012 of *Lecture Notes in Computer Science*, pages 190–216. Springer, 2003.

[9] C. R. Skolnik J and W. M. Planned special events - economic role and congestion effects.