Movement pattern analysis using Voronoi Diagrams

David Amores
Infrastracture Engineering
The University of Melbourne
Parkville, VIC 3010, AU
damores@student.unimelb.edu.au

Maria Vasardani
Infrastracture Engineering
The University of Melbourne
Parkville, VIC 3010, AU
mvasardani@unimelb.edu.au

Abstract

Analysis of clusters and movement patterns during emergency scenarios has the potential to provide vital information to key decision stakeholders. This paper presents a method using Voronoi Diagrams for defining both clusters and movement patterns in terms of voronoi cell properties: size, elongation, orientation and neighbourhood. Initial experimentation and testing against baseline methods using an evacuation trajectory dataset show promising results.

1 Introduction

Movement pattern analysis during disasters aids in hotspot detection, revealing the effects of the crisis and the performance of evacuation measures. Identifiable patterns in people's behaviour display their cognitive biases and can affect emergency response efforts, hindering use of the (potentially) limited resources, and the efficient broadcast of emergency messages. Such biases — e.g., the *bandwagon effect* where people follow others in their surroundings without a clear reason — have been discussed in psychology research (Kinateder et al., 2014).

Previous approaches for cluster and movement pattern analysis failed to provide uniform definitions and methods, with the exception of attempts in Dodge et al. (2008). Moreover, analysis of specific sets of movements has been limited, especially due to the time complexity involved. The current research aims to (1) provide a consistent method for defining movement patterns, and (2) enable their fine-grained analysis. Our approach uses Voronoi Diagrams (VD) for defining and identifying dynamic clusters and movement patterns. Specifically, we identify two pattern types using properties of VD regions. The first pattern is clusters people form when exiting a building during an evacuation. Detection of such agglomerations and where they frequently happen is useful for better emergency staff positioning and for advising future building layout design. The second pattern is the transition of people from one cluster to another. This movement pattern could be used to recognise throughput in merging evacuation lanes. Initial results are promising and the suggested method displays consistently higher performance than baseline approaches. A number of potential method enhancements are also discussed.

The remainder of this paper first presents relevant work on movement patterns and Voronoi Diagrams in Section 2. Section 3 introduces a pattern detection methodology based on VD cell properties, along with the description of experiments for evaluation. In Section 4, the evaluation results are discussed. We provide some concluding remarks in Section 5.

2 Related work on movement patterns during disasters and Voronoi Diagrams

Previous research used data from tracking devices on people (e.g., in Zheng et al. (2008)), or georeferenced data from social media platforms such as Twitter, Gowalla or Brightkite (Cho et al., 2011) to study patterns during disasters. Qualitative data, such as interviews and reports, have allowed the creation of disaster simulations

ranging from cellular automata (Wei-Guo et al., 2006) to complex system modelling (Adam and Gaudou, 2016), and multi-agent based systems (Pan et al., 2006). Fine-grained analysis of movement patterns comprises an array of statistical methods and modeling. For instance, Lu et al. (2012) use data from mobile users during the 2010 Haiti Earthquake to build a model that predicts mobility patterns. Yabe et al. (2016) built a framework for hotspot detection using a Gaussian Kernel Density Estimation method, and then anomaly value detection by comparing to the usual density distribution.

A Voronoi Diagram is a data structure that partitions space around a set of points (generators) such that each generator is the nearest object to any point inside its corresponding partition (Gold et al., 1997). Although VD generalise to higher dimensions, for disaster scenarios we only consider them on a 2D space, the plane. The partitions created by the VD are called Voronoi cells. Given the high time complexity of building VDs and its variants (higher order, generalised), most uses of the VD are performed in a static environment and with a small number of generator points. Early work on dynamic VDs by Guibas et al. (1991) detects topological changes (called topological events) in the VD relations for low cost updating rather than rebuilding. Sud et al. (2008) make innovative uses of first and second order VDs to build a structure used for route planning in environments with a large number of moving agents. Additionally, a number of optimisation techniques are used when dealing with dynamic settings, such as using hierarchical (Choset and Burdick, 1996), discrete (Hoff III et al., 1999), or approximation VDs (Vleugels and Overmars, 1998).

The method presented here is based on previous work on point pattern analysis, where VDs are used for pattern detection. Similarly to work described in Tuceryan and Jain (1990), Voronoi cell properties, such as area, elongation, eccentricity, or neighbourhood, form a feature vector, and then feature vectors of neighbouring cells are examined for possible patterns.

3 Method

Before describing how VDs are used to define and analyse movement patterns, a visual inspection of the evacuation data is performed. Observed patterns are defined in terms of Voronoi cell properties inspired by the work of Tuceryan and Jain (1990) on point pattern recognition. The results are compared to those of baseline methods.

3.1 Visualisation

The data used for exploratory analysis are taken from the VAST 2008 Challenge (Grinstein et al., 2008). This synthetic dataset contains trajectory information of 82 persons during 837 time steps of a building evacuation, as well as the topology of the building. An interactive visualisation allows viewing people's position at each time step. The layout of the building is drawn in the background. Peoples' locations are used as generator points for a dynamic VD displayed on top of these points, which changes at each time step. A snapshot is shown in Figure 1. For each step we collect the properties of cells of interest, as well as some general statistics. In particular, for person p and its respective voronoi cell V_p, we record the following properties:

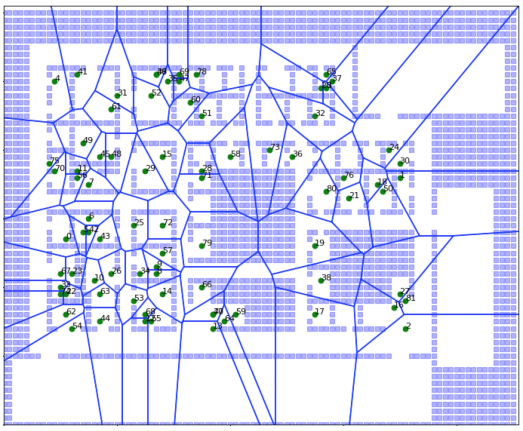


Figure 1: Voronoi Diagram drawn over people's location and building layout at a certain time step

- $A(V_p)$: The area of the voronoi cell.

- $E(V_p)$: Elongation of the voronoi cell. Calculated using $E(V_p) = 4\pi(\frac{A(V_p)}{P(V_p)})$ where $P(V_p)$ is the perimeter of the cell.
- $O(V_p)$: Orientation of the voronoi cell. A value from 0° to 180° obtained by calculating the orientation of the longest edge of the cell.
- $S(V_p)$: The number of sides (edges) of the cell.
- $D(V_p, V_q)$: The length of the Delaunay Triangulation edge from p to q, collected for every neighbour q of p.

The area $A(V_p)$, number of sides $S(V_p)$, and distance $D(V_p, V_q)$ to neighbours are standard geometrical measures. In contrast, the elongation $E(V_p)$ and orientation $O(V_p)$ of a polygon are not standard and the methods for computing them described above are empirically tested heuristics. Comparing or using other methods of computation are out of the scope of this paper, but are briefly discussed in future work (Section 5).

3.2 Voronoi Diagram cluster and movement patterns

During the visual analysis of the data, two types of patterns emerged in the VD structure: clusters and movement. The clusters are defined at corridors people use when exiting the building, especially when two such corridors meet. Accordingly, the movement pattern we focus on is the transition from one corridor cluster to the next. Both of these patterns are recurrently analysed in disaster evacuation related research (Pan et al., 2006).

3.2.1 Pattern 1 - Corridor Cluster pattern

While clusters formed by people queuing at an exit are visually easy to spot, to formally define them can be challenging. In particular, we want to define the two corridor clusters shown in Figure 2: the vertical (in brown) and horizontal (in purple). We refer to these as VC and HC, respectively. As observed, cells belonging to one of these clusters are elongated, have a certain orientation, and their generator points are close together. These Voronoi cell properties are used to formally describe the clusters.

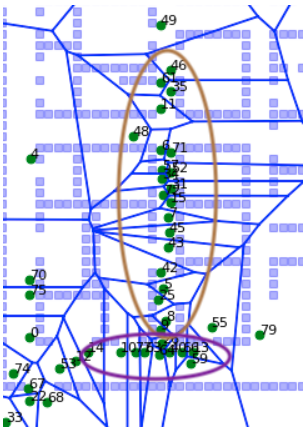


Figure 2: Visual identification of the VC and HC clusters

We define as **corridor cluster** the group of three or more neighbouring cells (cells that share an edge) that satifsy the following condition: Given any pair of neighbouring cells V_p and V_q, their $score(V_p, V_q)$ is less than a threshold η. That is, a generator p is part of the same cluster as q, iff $score(V_p, V_q) < \eta$. This score is the weighted average of the elongation $E(V_p)$, the normalised angle between orientations $O(V_p)$ and $O(V_q)$, and the normalised distance between the generator points $D(V_p, V_q)$:

$$score(V_p, V_q) = \frac{W_E \cdot E(V_p) + W_O \cdot \Delta O(V_p, V_q) + W_D \cdot d(V_p, V_q)}{W_E + W_O + W_D} \quad (1)$$

where $\Delta O(V_p, V_q) = \frac{|O(V_p) - O(V_q)|}{\vartheta}$ is the normalised angle difference, and $d(V_p, V_q) = \frac{D(V_p, V_q)}{\delta}$ is the normalised distance. As the maximum possible angle between two cells (in the current experiment) is 90°, it is used as the value ϑ for normalisation. Likewise, the maximum distance in the provided dataset between two points is

90, and is the value used as δ for normalisation. Having a single score that combines the three cluster defining properties allows for some cluster membership flexibility. For instance, a cell that is not so elongated, but has the same orientation and is very close to other members, can be in the same cluster. Likewise, an elongated cell with the same orientation as the cluster, even if a little distant (due to the underlying environment perhaps), can still belong to the cluster. An additional restriction is placed such that $W_E + W_O + W_D = 1$, so that the weights represent a trade-off between the properties.

The value of η roughly represents the property values of an 'ideal' cluster member cell. It is important, therefore, to define η appropriately. The ideal (trivial to identify both visually and computationally) corridor cluster would have members perfectly aligned one after the other, and as close as possible. In that case, cells are elongated, have the exact same orientation, and are minimally separated. Numerically, it would be:

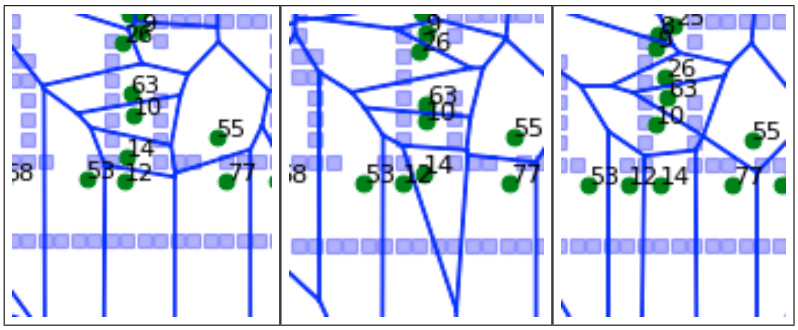
- $E(V_p) = 0.35$. Mildly elongated. This is the 80 percentile of elongation in the dataset used. It is equivalent to a rectangle of sides 7 and 1.
- $|O(V_p) - O(V_q)| = 0°$, same orientation
- $D(V_p, V_q) = 1$, minimum distance

Plugging these values to Equation (1) then results in $\eta = 0.12$, when all weights W_E, W_O and W_D are equal. As stated before, the unified score allows trade-offs to be made. For instance, if a cell were to have a slight rotation ($|O(V_p) - O(V_q)| = 5°$) and be further apart ($D(V_p, V_q) = 3$), but is more elongated ($E(V_p) = 0.26$), it would still be a member ($score(V_p, V_q) = 0.118 < 0.12$).

3.2.2 Pattern 2 - Transition Movement pattern

The second pattern defined and analysed is a movement behaviour during the merging of two corridor clusters, particularly when the clusters are orthogonal to each other. Analysing other configurations is out of the current scope, but is discussed briefly in future work. Whenever there is an intersection between two flows of people, there is a transitioning cell at the intersection of the two clusters. Using the same case shown in Figure 2, the transitioning cell goes through a specific set of states, outlined below:

1. State 1. The cell is elongated, it has a horizontal orientation, and has more than three sides
2. State 2. The cell becomes a triangle
3. State 3. The cell returns to being elongated with a vertical orientation, and having more than three sides



(a) Person 14 in VC (b) Person 14 transitioning (c) Person 14 in HC

Figure 3: Example of a person's cell (14) transitioning from one cluster to the other.

These states are shown in Figure 3. In general, we call a set of transitioning states a *movement pattern* and define it as a set of n sequential states $M = S_1(V_p), \ldots, S_n(V_p)$. Each state S_i is a set of conditions a Voronoi cell should meet to be in said state. We formally define the **transition movement** pattern as a movement pattern of $n = 3$, where each state has the following conditions (the threshold values for elongation (0.7) and orientation ($\pm 20°$) allow for flexibility in the membership of the vertical and horizontal corridor clusters):

- State 1 (S_1): (i) Is elongated: $E(V_p) < 0.7$, (ii) is horizontal: $-20° < O(V_p) < 20°$, and (iii) has more than three sides: $S(V_p) > 3$
- State 2 (S_2): Is a triangle: $S(V_p) = 3$
- State 3 (S_3): (i) Is elongated, $E(V_p) < 0.7$, (ii) is vertical, $70° < O(V_p) < 110°$, and (iii) has more than three sides, $S(V_p) > 3$

3.3 Experiment

In order to ascertain the veracity of the pattern detection methods described previously, two sets of experiments are performed using the data described in Section 3.1. For each set, manual pattern identification is used as ground truth. Then we apply the Voronoi cell properties method and measure its ability to identify the patterns and compare its performance to a baseline.

3.3.1 Pattern 1 - Corridor Cluster pattern testing

During the manual identification of the clusters, at time step i the members of each cluster C_i are recorded and named $M(C_i)$. The corridor cluster method (Section 3.2.1), retrieves a cluster CC_i with members $M(CC_i)$. These values are computed for clusters VC and HC, and are used to calculate the precision, recall and F_1 score as explained in Section 4.1. The same is done with a baseline for comparison. The baseline used is distance-based clustering, where all points within distance σ are clustered together.

This experiment was run with a number of different weight combinations using insights learned from the visual analysis. The most relevant combinations are summarised in Table 1. Experiments 1 and 2 are the baselines with two different σ values. Experiment 3 shows the method using equal weights, that is, using cell properties for the 'ideal' corridor cluster. As observed in the visualisation, the most important predictor of a cluster is still the distance between the generator points. Therefore, the distance weight W_D is assigned a higher value in both Experiments 4 and 5. Deciding for W_E and W_O proved more challenging as they seem equally important for the clustering. Orientation seems to prevail at times; therefore, two cases are tested: one with a higher weight for orientation (Experiment 4), and another with equal orientation and elongation weights (Experiment 5).

Table 1: Parameters used for experiment runs

Exp.	σ	W_E	W_O	W_D
1	2	–	–	–
2	3	–	–	–
3	–	$1/3$	$1/3$	$1/3$
4	–	$1/6$	$2/6$	$3/6$
5	–	$1/8$	$1/8$	$6/8$

3.3.2 Pattern 2 - Transition Movement pattern testing

The number of points behaving in the way defined in Section 3.2.2 is first manually recorded for each time step, and used as ground truth. The set of manually collected points is named TP (transitioning points). Then, the described method to identify the movement pattern is applied , and a set of points AP (attempted points) is retrieved. These values are calculated for obtaining the precision, recall and F_1 scores, as described in Section 4.2. The same data is collected for the baselines, defined as follows: For every moving point, check if the point performed a vertical movement followed by a horizontal movement. For generator p at time step i, these movements are defined as:

$$vertical_{pmove} = y_{p_{i-1}} - y_{p_i} > \alpha \cdot (x_{p_{i-1}} - x_{p_i}) \tag{2}$$

$$horizontal_{pmove} = x_{p_{i-1}} - x_{p_i} > \alpha \cdot (y_{p_{i-1}} - y_{p_i}) \tag{3}$$

where α is a user-defined constant that represents how "straight" the movement is. The higher α is, the more "straight" the movement. Four experiments are realised. Experiments 1-3 are baselines and involve setting different values for α, namely 2, 4 and 10. Experiment 4 uses the suggested method and is contrasted to the other three in the results analysis.

4 Results analysis

Precision and recall values are frequently used for measuring the performance of pattern recognition methods and are, here, also compared to the values of the baseline methods. Additionally, we compute the F-measure, or F_1, as a way to represent the overall accuracy of precision and recall together.

4.1 Pattern 1 - Corridor Cluster pattern results

Precision, recall and F_1 values for the Voronoi cell properties method, as well as the baseline, are calculated using Equations 4, 5 and 6 and summarized in Table 2 for different weights and constant values. Experiments 1 and 2 are done using the baselines and, thus, a fixed distance metric σ. Experiments 3, 4 and 5 use the Voronoi cell properties method with different values for the weights W_E, W_O, and W_D.

$$precision = \frac{\sum_{i=1}^{n} |M(C_i) \cap M(CC_i)|}{\sum_{i=1}^{n} |M(CC_i)|} \tag{4}$$

$$recall = \frac{\sum_{i=1}^{n} |M(C_i) \cap M(CC_i)|}{\sum_{i=1}^{n} |M(C_i)|} \tag{5}$$

$$F_1 = 2 \cdot \frac{precision \cdot recall}{precision + recall} \tag{6}$$

The baseline methods have surprisingly acceptable precision and recall. The recall is high because, as they do not have additional conditions, the distance metric picks up everything nearby. Thus, members of both clusters are frequently considered as a single larger cluster, which is *precisely* what we sought to differentiate with the suggested method.

Experiment 3 ranks the highest in precision but has very low recall scores, much lower than the baselines. This confirms that the ideal definition of the corridor cluster is not achieved in practice and, therefore, variations in the weights are needed as proved by the next two experiments. Experiment 4 shows much better results for recall. Weights used in Experiment 5 prove to be superior to the baselines and the rest of the runs, with the highest F_1 scores for all VC and HC experiments. These two experiments demonstrate three points:

1. Distance remains the best predictor for cluster membership as shown by higher values of W_D
2. Combination of distance, cell elongation and cell orientation improves the performance, as shown by the consistent higher scores than the baselines.
3. Voronoi cell properties can in fact be used to effectively identify the corridor cluster pattern.

Table 2: Precision, recall and F_1 values for the five corridor clusters experiments.

Exp.	Precision VC	Recall VC	F_1 VC	Precision HC	Recall HC	F_1 HC
1	0.85	0.76	0.80	0.72	0.52	0.60
2	0.60	**0.98**	0.74	0.74	0.18	0.23
3	**0.95**	0.14	0.24	**1**	0.15	0.25
4	**0.95**	0.62	0.75	0.99	0.67	0.80
5	0.91	0.92	**0.92**	0.87	**0.76**	**0.81**

4.2 Pattern 2 - Transition Movement pattern results

Precision, recall and F_1 scores are used to measure the performance of this method as well. They are calculated using Equations 7, 8 and 9. Table 3 shows the results from the transition movement pattern experiment. Experiments 1, 2, and 3 are the baselines with different values for α (Section 3.3.2). Experiment 4 implements the suggested movement pattern detection method, as described in Section 3.2.2.

$$precision = \frac{|TP \cap AP|}{|AP|} \tag{7}$$

$$recall = \frac{|TP \cap AP|}{|TP|} \tag{8}$$

$$F_1 = 2 \cdot \frac{precision \cdot recall}{precision + recall} \tag{9}$$

In the baselines, as the value for α increases, the number of retrieved patterns decreases (recall). Accordingly, Experiment 1 has a perfect recall but a very low precision. When increasing α, the precision improves, but the recall drastically drops. In contrast, the suggested method (Experiment 4) proves superior and acceptable overall, in both precision and recall, even if recall is lower than the highest observed. The method's better performance is also summarized with the highest F_1 score.

Table 3: Precision, recall and F_1 values for the four movement pattern experiments.

Exp.	α	Precision	Recall	F_1
1	2	0.13	**1**	0.23
2	4	0.26	0.88	0.4
3	10	0.3	0.38	0.33
4	–	**0.6**	0.75	**0.67**

5 Conclusions and Future Work

This research shows how elements of the VD structure can be used to determine and analyse two types of fine-grained spatio-temporal patterns: clusters and movements. The suggested methods allow for the flexible definition of these patterns by the combination of Voronoi cell properties — namely elongation, orientation, and distance to neighbours. These properties were selected as the most relevant for the testing dataset scenarios: elongation and orientation were paramount to the identification of corridor clusters, for example. The flexibility of the method, however, allows the use of different Voronoi cell properties that may be more appropriate for different datasets and environments. With consistently high scores in precision and improving recall scores, the suggested method proved better than baseline methods that are only considering distance factors for the formation of clusters.

The suggested method has advantages over previous approaches. Until now, there has not been a standard movement pattern definition and they are, thus, mostly analysed using general measures such as distance traveled (Lu et al., 2012) or cluster methods (Yabe et al., 2016). In contrast, the suggested method allows for very fine-grained definition of movement patterns by providing a sequence of states the Voronoi cells exhibit while 'moving' to form different patterns. The experiments show better overall scores than baseline methods. A shortcoming is the rather small number of instances of this dataset — people that move during an evacuation scenario and which are considered as the generator points for the dynamic VDs. This number is especially low for the transition movement pattern (8 instances). Even so, the promising experimental results suggest that with a larger evaluation dataset, the suggested method's advantages can be further verified.

This preliminary work shows some of the benefits of using the VD structure properties for the detection and analysis of various patterns, which can aid in emergency response and evacuation operations. The method's flexibility allows for the use of different or additional cell properties. For instance, eccentricity is another clustering indicator that has been used in digital image processing applications and could be considered in future research. In addition, patterns that use non-standard properties of a polygon such as elongation and orientation can benefit from different computing methods. For example, Tuceryan and Jain (1990) use mathematic moments in order to define elongation, orientation, area, and eccentricity, and such methods can be utilized in providing finer, more precise cell monitoring parameters for pattern detection.

In the analysis of movement patterns, the conditions of each cell transition state can also be more flexibly defined to allow for the detection of more and varied patterns. For example, state S_1 in the transition movement pattern could be something *close* to, rather than an exact triangle (i.e., $S(V_p) = 3$), such as having 4 sides with one of them being much smaller than the rest. This generalisation is especially important when aiming at identifying other corridor configurations. Finally, when dealing with large amounts of data, it is important to also consider time complexity for comparing with other methods. Building a VD has a time complexity of $O(n \cdot \log n)$ in the worst case scenario, and $O(n)$ in the average case. Therefore, any additional analysis would have to consider this as a base cost.

Acknowledgements

The authors wish to thank Dr Egemen Tanin for his valuable comments during initial discussions on this work.

References

Adam, C. and B. Gaudou (2016). Modelling human behaviours in disasters from interviews: application to melbourne bushfires. In *Social Simulation Conference (SSC), Rome, Italy*.

Cho, E., S. A. Myers, and J. Leskovec (2011). Friendship and mobility: user movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 1082–1090. ACM.

Choset, H. M. and J. Burdick (1996). *Sensor based motion planning: The hierarchical generalized Voronoi graph*. Ph. D. thesis, California Institute of Technology Pasadena (California).

Dodge, S., R. Weibel, and A.-K. Lautenschütz (2008). Towards a taxonomy of movement patterns. *Information visualization 7*(3-4), 240–252.

Gold, C. M., P. R. Remmele, and T. Roos (1997). Voronoi methods in gis. In *Algorithmic Foundations of Geographic Information Systems*, pp. 21–35. Springer.

Grinstein, G., C. Plaisant, S. Laskowski, T. O'connell, J. Scholtz, and M. Whiting (2008). Vast 2008 challenge: Introducing mini-challenges. In *Visual Analytics Science and Technology, 2008. VAST'08. IEEE Symposium on*, pp. 195–196. IEEE.

Guibas, L. J., J. S. Mitchell, and T. Roos (1991). Voronoi diagrams of moving points in the plane. In *International Workshop on Graph-Theoretic Concepts in Computer Science*, pp. 113–125. Springer.

Hoff III, K. E., J. Keyser, M. Lin, D. Manocha, and T. Culver (1999). Fast computation of generalized voronoi diagrams using graphics hardware. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pp. 277–286. ACM Press/Addison-Wesley Publishing Co.

Kinateder, M. T., E. D. Kuligowski, P. A. Reneke, and R. D. Peacock (2014). *A review of risk perception in building fire evacuation*. National Institute of Standards and Technology.

Lu, X., L. Bengtsson, and P. Holme (2012). Predictability of population displacement after the 2010 haiti earthquake. *Proceedings of the National Academy of Sciences 109*(29), 11576–11581.

Pan, X., C. S. Han, K. Dauber, and K. H. Law (2006). Human and social behavior in computational modeling and analysis of egress. *Automation in construction 15*(4), 448–461.

Sud, A., E. Andersen, S. Curtis, M. C. Lin, and D. Manocha (2008). Real-time path planning in dynamic virtual environments using multiagent navigation graphs. *IEEE transactions on visualization and computer graphics 14*(3), 526–538.

Tuceryan, M. and A. K. Jain (1990). Texture segmentation using voronoi polygons. *IEEE transactions on pattern analysis and machine intelligence 12*(2), 211–216.

Vleugels, J. and M. Overmars (1998). Approximating voronoi diagrams of convex sites in any dimension. *International Journal of Computational Geometry & Applications 8*(02), 201–221.

Wei-Guo, S., Y. Yan-Fei, W. Bing-Hong, and F. Wei-Cheng (2006). Evacuation behaviors at exit in ca model with force essentials: A comparison with social force model. *Physica A: Statistical Mechanics and its Applications 371*(2), 658–666.

Yabe, T., K. Tsubouchi, A. Sudo, and Y. Sekimoto (2016). A framework for evacuation hotspot detection after large scale disasters using location data from smartphones: case study of kumamoto earthquake. In *Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pp. 44. ACM.

Zheng, Y., L. Liu, L. Wang, and X. Xie (2008). Learning transportation mode from raw gps data for geographic applications on the web. In *Proceedings of the 17th international conference on World Wide Web*, pp. 247–256. ACM.