ISSN: 0974-5823 Vol. 7 No. 2 February, 2022

International Journal of Mechanical Engineering

A Review on Indoor Evacuation Model and Clustering Techniques in Developing Evacuation Assessment Algorithm

Amir Haikal Abdul Halim

Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Cawangan Melaka Kampus Jasin, Melaka, Malaysia

Khyrina Airin Fariza Abu Samah*

Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Cawangan Melaka Kampus Jasin, Melaka, Malaysia

Mohd Nor Hajar Hasrol Jono

Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Cawangan Melaka Kampus Jasin, Melaka, Malaysia

Lala Septem Riza

Department of Computer Science Education, Universitas Pendidikan Indonesia, Indonesia

Abstract

Evacuation models were developed to assist in any emergency evacuation. Despite many evacuation models available, it focuses more on evacuation behavior and time. The different features lead to users' hard decision to choose the best evacuation models, and validation is required. Thus, this review study proposed to answer the research question:1) how can the user determine the best indoor evacuation model be selected? 2)how can the indoor evacuation models have a standard assessment? The main scope is an indoor evacuation in a high-rise building, vertical evacuation, and fire disaster. The review starts with the existing evacuation models. Since we focus on the specific evacuation routes, we also review the existing clustering algorithms. As a result, we presented our review outcome using a knowledge chart to determine the research objective. The microscopic model is more advantageous for evacuation models than the macroscopic and mesoscopic models. The K-Mean partitional cluster is selected for the clustering algorithm because it is suitable for large datasets with relatively small computational requirements. Also, the partitioning technique's time complexity is almost linear. Thus, we set the research objective to formulate an indoor evacuation assessment algorithm for critical incidents through clustering the features of the developed indoor evacuation model using the K-Mean algorithm.

Keywords : Evacuation Model, Cluster Algorithm, Indoor Evacuation, Microscopic Model, Partitional Cluster

Introduction

Evacuation is defined as the retreat, dispersal, or withdrawal of individuals from locations of risk or threat, as well as their reception and treatment in secure environs, which were all organised, regulated, and monitored (Ventura, 2017). The large public buildings increased significantly with the economy's growth and the population's advancement. (Samah et al., 2015) identified that many people gathered in major buildings, such as libraries, shopping centres, stages, and metro stations. The complicated interior design and the large number of people involved make evacuation management more difficult and necessary for evacuation planning (Tan, 2014). Therefore, it is essential to have the right emergency evacuation in every high-rise building.

In the event of an emergency, adequate contingency plans will facilitate egress and save lives. Evacuation is not the same as the pre-designed evacuation plan, considering different aspects, including the enclosed layout, environmental change, and evacuees' features and likely encounters with evacuees' crowds. Evacuees with a misperception of the building environment may display significant rounding or even be trapped, resulting in an extremely longer evacuation time. Ventura (2017) claimed that people generally follow a route of self-estimated rapid escape based on a sense of the current situation.

Furthermore, intense hysteria and stamping can contribute to multiple individuals evacuating in emergencies. The layout of escapes from structures, human psychology and behaviour, and numerous social, behavioural patterns can heavily impact evacuation performance, leading to being trapped (Chen et al., 2020). Thus, it is a need for a high-rise building to have an evacuation model to allow evacuees to exit the building without any harm safely.

Luo et al. (2018) stated that numerous proposed intelligent indoor emergency evacuation preparedness techniques have as an "actively-drive" and "passively-drive". Also, many developed evacuation models focus on studying different evacuation behavior and evacuation time. The models differ in terms of the features and lead to users' hard decision to choose the best evacuation models.

However, the evacuation model's validation and comparison are missing before applying the suitable evacuation model. Thus, an evacuation assessment algorithm is needed before selecting the best evacuation model to be used.

This review poses the research questions:

1) How can the user determine the best indoor evacuation model be selected?

2) How can an indoor evacuation model have a standard assessment?

Both questions need to be answered before developing the evacuation assessment algorithm. The selection of both is based on the case study scope: indoor evacuation in a high-rise building, vertical evacuation and fire disaster. This paper's organization begins with a brief introduction to the problem in the first section. The second section explained the indoor evacuation and followed the evacuation model in the third section. The fourth section elaborates on the clustering algorithm, and the fifth section is the research knowledge chart. Finally, the last section concludes the study and briefly mentions future enhancement.

Indoor Evacuation

Larger and more sophisticated buildings have been built over urban areas in the last few decades. The complicated indoor layout and many people create more emergency management problems and boost higher evacuation preparation requirements (Tan, 2014). Evacuation is regarded as an emergent reaction to decrease the damage through transferring or moving people away or the act of leaving any danger zone as quickly as possible to a safer place. However, due to the enormous demand for floor space, high-rise buildings' complexity leads to uncertainty and urban disaster management complications. It causes the indoor evacuation wayfinding to the nearest exit during emergencies such as fire more challenging.

Ma and Guo (2012) stated three issues involved in the fire risks in high-rise buildings: the rapid-fire and spreading of the smoke, difficulties in firefighting and rescue, and difficulties in the safe evacuation of the occupants. There are two types of evacuation in indoor evacuation, which are horizontal evacuation and vertical evacuation. Horizontal evacuation is the evacuation of evacuees from a threatening environment to a safe space at a different building level. Vertical evacuation is the evacuation of evacuees from a threatening environment to a safe space within the same building (Ventura, 2017).

Many evacuation models have been developed to overcome the problem, especially virtual spatial. Each type differs in its approaches to the analysis and the different means of representing the occupants' space and behavior (Hoekstra & Montz, 2017). Nevertheless, only a few assessments on the literature's evacuation models have been done by focusing on such models. Ko et al. (2007) validate EvacuatioNZ; a coarse network evacuation model developed to simulate the occupant movement times and human behavior before and during the evacuation process. The effectiveness of different total evacuation strategies in a high-rise building has been investigated using egress modeling. Indoor evacuation relates to the evacuation model; thus, the next sub-topic will further explain the evacuation model.

Evacuation Model

There are two main goals for the development of evacuation models. First, to determine whether the building safety efficiency (architectural layout and exit capability) is appropriate. Second, to ensure that evacuees may invest the least time in the safety zone in the event of a fire and other accidents, allowing an ideal escape path. Nonetheless, most current evacuation models either neglect a few human behavior characteristics in crowds or are computationally complicated, making it difficult to fully reflect evacuation outcomes that are more precise and practical (Xie & Li, 2014). Thus, it is imperative to implement a suitable evacuation model for the specific high-rise building. The evacuation model helps plan better because it imitates real evacuation situations. Henry (2015) claimed the model seeks to depict a scenario for the evacuation of natural hazards through principles extracted from transport science, risk analysis, sociology, and disaster management. An evacuation model was created to understand the effects of rapid flows and human experiences on the evacuation period (Gaire, 2017).

Moreover, an evacuation model concentrates on three main tasks: (1) move towards the nearest exit, (2) move into the area with small crowds, and (3) try to avoid evacuation obstacles. The three tasks can become complicated with the bigger size, capability, and architecture of a high-rise building floor plan (Xie & Weerasekara, 2016). Jiang et al. (2020) stated that the evacuation model is divided into microscopic, macroscopic, and mesoscopic models. Figure 1 shows the process involved in selecting the suitable indoor evacuation model. It starts with the evacuation preparedness, evacuation models, choosing suitable, and finally applying the chosen evacuation model. We found that one process is missing between the third and fourth elements: assessing a suitable evacuation model. The process is crucial since not all buildings can use the same evacuation model.

Kontou et al. (2018) said the microscopic models often define each individual's spatial and temporal actions. In this method, the individuals are considered interaction particles (Ming & Peng, 2017) and the individual's activity is often defined by their contact with other individuals (Ibrahim et al., 2017). This model by physical can be separated by discrete model and continuous model (Yu et al., 2014). Some examples of microscopic models include the social forces model, cellular automata models, agent-based models, NOMAD, and lattice-gas models.

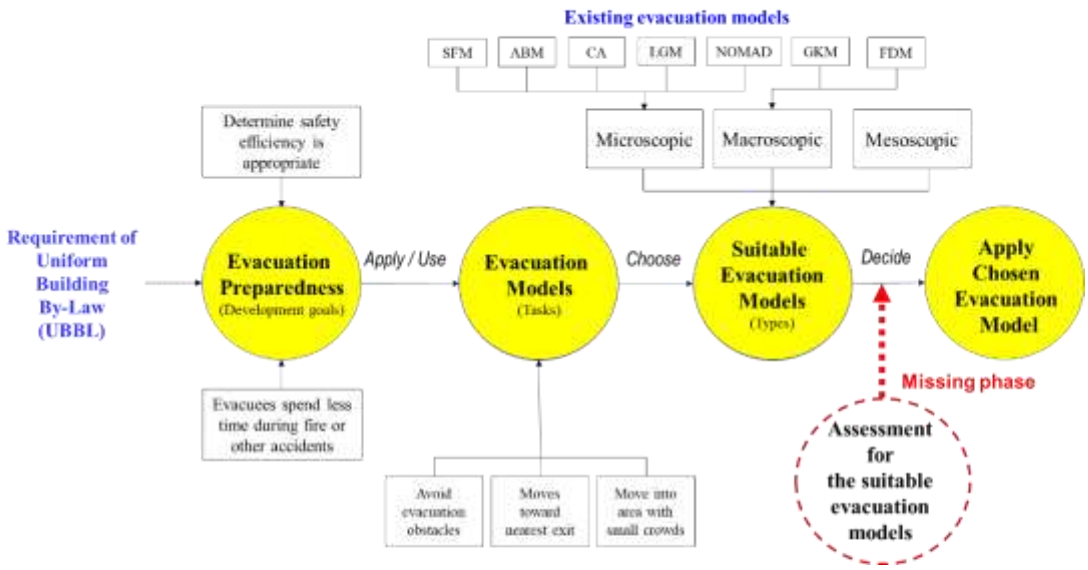

Figure 1 The process in choosing the evacuation model

Bakar et al. (2017) define the macroscopic model, also known as the continuum model, which combines variables and monitors parameters such as bottleneck densities. The model considers the evacuates to be an integer of the same properties, so an evacuation performance depends upon crowd flow, crowd density, and physical architectural factors (Wu & Huang, 2015). In other words, Liu et al. (2016) mentioned that it ignores the actions of individuals during evacuation and uses network stream models to address the evacuation issue. As evacuation research comes from traffic flow research, the perfect and mature fluid research method is naturally inherited; ready-made results from fluid mechanics research can quickly start.

However, the evacuation system swarms with several similar individuals with direct interactions. The simple overlay of the system's different properties is not a system's property as the system is nonlinear and limits the results from this approach (Ming & Peng, 2017). It also achieves excellent reliability and productivity to predict and rate overall while retaining the lack of detail and precise estimation of individual movement as a drawback (Kontou et al., 2018).

Lastly, the mesoscopic model uses probability distributions to represent the individual location and velocity of the individuals. This model usually focuses on the human activities elements in microscopic and macroscopic models in large crowd densities (Bakar et al., 2017), and the unit analysis is the individual. However, it does not consider the interaction between individuals (Ming & Peng, 2017).

1). Microscopic Model

Different types of existing approaches adapted to the evacuation models have been found in this subchapter. The existing evacuation models are discussed here and explained, starting with the microscopic model. The first to be explained is the Agent-Based Model (ABM). ABM is a microscopic simulation approach; in which a system is conceived as a collection of independent decision-making entities called agents (Poulos et al., 2018) based on an individual-based model. As elaborated by Hu et al. (2012), at each instant, each agent evaluates its immediate surroundings and makes decisions based on a set of rules and usually built in a "bottom-up" way. ABM is capable of coping with individual behavior to address crowd-control difficulties. Furthermore, it can model varied behavior and emulate individuals who may behave differently (Bakar et al., 2017).

The second microscopic evacuation model is the Social Force Model (SFM). The SFM is a microscopic model with constant time and space characterized by average speed, density, and position (Liu et al., 2018). It has frequently been used in engineering to represent the changing process of human behavior under the effect of a combined force of psychological and physiological elements since it accurately simulates people's situations. There will be a "faster is slower" effect and some conformity (Wang & Weng, 2014). It is a particle-based model in which humans are defined as particles impacted by social and physical influences from other barriers and people and inner drive to achieve the objective (Ratsamee et al., 2012).

The third microscopic evacuation model is Cellular Automata (CA), a common example of a discrete mathematical model. It is a dynamic system that's distinct in terms of time, place, and status. The CA model consists of four parts: a cell, a cell space, a neighborhood, and an evolutionary rule. At each time step, the values of the variables in nearby cells determine the value of the variables in one cell. Moore's Neighbourhood consists of the eight cells encircling the central cell, which might be used. It considers individual traits, psychological responses to catastrophes, interpersonal interactions, and the impact of structure and environment on evacuation behaviors (Yang et al., 2014).

The Lattice Gas Model (LGM) is the fourth microscopic evacuation model. It does not directly establish differential equations for modelling the system's nonlinear physical events. Simulation and programming are clear, scalable, dynamic, and simple to simulate and is based on a discrete model. Random serial updating guarantees that everyone in the system has the same right to select and that movement conflicts are avoided (Yu et al., 2014). The drift random walk model is another name for LGM. It is based on the

CA framework and treats the individual as a particle. It is pointed out that everyone has a favorite direction at any given time, and the drift probability simulates the individual movement process (Li et al., 2019).

The fifth microscopic evacuation model is Normative Pedestrian Behavior Theory (NOMAD). Walking effort is defined using a more generic walking expenditure and activity usage concept. Pedestrians profit from events and are charged a price for walking. As a result, it is a normative concept that nomad pedestrians maximize the balance between the two (Campanella et al., 2014). The technique is based on a three-level pedestrian hypothesis. Such levels divide the main parts of pedestrian behavior into clearly defined tasks, reducing the model's complexity. The operational level outlines the walking activity or how walkers navigate to execute the plan. In contrast, the strategy level contains the activities to be done before the journey begins with tactical models, choices and decisions, and adjustments to the original plan throughout the trip (Campanella et al., 2009).

2). Macroscopic Model

The gas kinetic model (GKM) is the first macroscopic evacuation mode. GKM employ partial differential equations to describe how density and velocity vary over time using an analogy with gas dynamics (Pelechano & Malkawi, 2008). Henderson demonstrated how, using Maxwell-Boltzmann theory to describe crowd motion, some crowds (often less dense) might be described as homogenous crowd gas. Henderson used the following assumptions to apply the Maxwell-Boltzmann equation to the gaseous crowd phase (Isenhour, 2016):

1) the surface was continuous in both position and velocity at any time;
2) the crowd is homogeneous, so each particle has the same mass and probability of velocity;
3) particles could be composite (people walk in pairs and/or groups) or prime (single individuals);
4) regardless of composite or prime, the particles are statistically independent in position and velocity; and
5) the crowd is in equilibrium, so velocity is uncorrelated with the position.

The second macroscopic evacuation model is the fluid dynamic model (FDM). Yao et al. (2019) mentioned that partial differential equations for man-made systems are frequently employed to explain the dynamic features of crowds, such as continuous velocity and density. FDM recognized that crowds are more frequently logical than irrational and referred to pedestrian fluid mechanics as "thinking fluids," with crowd research focusing on the theoretical creation of scientific laws controlling collective pedestrian motion.

Table 1 compares two types of the existing evacuation model: microscopic and macroscopic, since the mesoscopic does not consider the interaction between individuals. The model selection is based on the case study of the specific research, such as individual, particle-based, discrete, walker, continuum-based and gas dynamic. Another factor, which is features, also contribute to the selection weightage. In evacuation assessments, microscopic and macroscopic models are commonly employed to depict pedestrian traffic (Shi et al., 2018). However, the Microscopic model has been much relied on within the research community in crowd simulation studies to better understand crowd behavior in emergencies (Sakour & Hu, 2017). Macroscopic models have high computational performance but cannot represent interaction and human heterogeneity. In contrast, the computing of microscopic models' output is relatively lower, yet the human movements and activities can be more reliable and generally defined; thus, Li et al. (2020) claimed that the microscopic models had been used widely in recent years. The following section reviews the clustering algorithms since we focus on the specific evacuation routes.

Table 1 Comparison of existing evacuation model

Evacuation Model	Model Type	Model Used	Individual/Groups	Features
ABM	Microscopic	Individual-based Model	Heterogenous	Based on agents' decision-making
SFM	Microscopic	Particle-based Model	Homogeneous (Zheng et al., 2009)	Defined by average speed, density, and location but in continuous time and space
CA	Microscopic	Discrete Model	Homogeneous and Heterogeneous (Zheng et al., 2009)	Composed of the cell, cell space, neighbourhood, and evolutionary rules
LGM	Microscopic	Discrete Model	Homogeneous (Zheng et al., 2009)	Individual movement process simulated by the drift probability
NM	Microscopic	Walker Model	Heterogeneous	Based on activity area, route choice (tactical level) and walking behavior (operational level)
FDM	Macroscopic	Continuum-based model	Homogeneous (Zheng et al., 2009)	Uses Navier–Stokes equations to model the motion in crowds
GKM	Macroscopic	Gas dynamics	Homogeneous	Uses the Maxwell-Boltzmann Equation to the gaseous crowd phase

Clustering Algorithm

The clustering algorithm plays an important role in finding a suitable evacuation model for a particular building. In analyzing data throughout various fields, such as modeling recognition, machine learnings, bioinformatics, and image processing, the problem of clustering is apparently, fundamental (Tzortzis & Likas, 2014). Most researchers focus on clustering the evacuees in specific evacuation routes (Fragkos et al., 2019) and dividing the problem into smaller subproblems in managing the crowds (Vogiatzis et al., 2013). It is identified that clustering attempts to group an unlabeled data set into related object clusters (Han et al., 2017). Each cluster consists of objects identical to other clusters within the cluster. Nevertheless, it is computationally unworkable to look at the best possible assignment of instances to clusters, and usually, an excellent local optimum of the clustering goal is required.

The clustering algorithm is divided into two types: partitional and hierarchical clustering algorithm. A partitional clustering algorithm separates the data set into various classes dependent on a criterion known as fitness. Nanda and Panda (2014) claimed that the partitioning task would be turned into an optimization problem with a proper fitness measure. In-depth, the partial clustering assigns k-clusters with iterative processes to collect data points, and n data would be listed in k-clusters in these processes. The j criterion function predefines the datum to k^{th}, which is set according to the calculation of maximization and minimization in k sets (Kutbay, 2018).

Hierarchical clustering algorithm provides the output of a tree structure (dendrogram plot), representing the nestled group of data elements. It is not required to know the number of clusters beforehand (Nanda & Panda, 2014). Also, the hierarchical clustering algorithm is classified into agglomerative and divisive modes (Xu et al., 2020). In the agglomerative mode, each unit of data is initialized into its cluster by the agglomeration process. Xie (2013) claimed that the clusters are combined into bigger and bigger ones before each sub-cluster is fused with similarities mainly by standards. In divisive mode, it is often called the top-down technique, in which a broad data set is initiated. A variety of smaller subsets (called clusters) break up this data set before a threshold is achieved (Nisha & Kaur, 2016).

Thus, researchers in recent years have concluded that a partitional clustering algorithm is suitable to cluster large datasets due to relatively small computational requirements. The partitioning technique's time complexity is almost linear and is commonly used (Saini & Kaur, 2014). There are also several types of clustering algorithms found through research. These include density-based algorithms, graph-based algorithms, grid-based algorithms, and model-based algorithms (Lv et al., 2016).

Table 2 explain, in brief, the other groups of the clustering algorithm, including partitional and hierarchical. There are two types of partitional clustering algorithms: K-Medoid and K-Mean clustering algorithms. Guo and Xie (2018) mentioned that the K-Medoid clustering algorithm or K-Medoid is based on the measurement of medoids. It decreases the absolute distance between points and the preferred centroid rather than lowering the square distance. Thus, it is sturdier than the K-Mean clustering algorithm K-Mean for noise and outliers.

Table 2 The description of clustering algorithm

Clustering Algorithm	Description	Example
Partitional	• Generate one partition of data (Martins, 2015) • Some clusters are contained in the partitional clustering directly as a data partition without a hierarchical structure (Xie, 2013)	• K-Mean • K-Medoid (Xie, 2013)
Hierarchical	• Instead of only one partition, it shows a hierarchical arrangement of the data entry and aims to subsequently construct a hierarchy of clusters by two types of strategies: agglomerative and divisive • Creates a hierarchy of clusters based on the dissimilarity (distance), comparisons between rows, such as the Euclidean Distance, and the Pearson correlation coefficient and considers each row initially as an independent category and fuses the nearest pair of clusters at each point before all rows are combined into a single one (Zhang, 2011)	• MONA • DIANA • BIRTH • CURE (Xie, 2013)
Density-based	• A technique in which dense patterns are clustered together, mainly when the number of clusters is uncertain. (Gialampoukidis et al., 2019) • The process can be divided into two. First, estimating each sample's density should be established and used to classify the samples within dense areas. Second, add them to the same cluster for the corresponding dense areas (Li et al., 2020)	• DBSCAN • OPTICS (Li et al., 2020)
Graph-based	• It consists of a family of unattended grouping algorithms designed to organize the vertices and edges of a graph rather than objects in an interface (Hufnagl & Lohninger, 2020) • Two disadvantages: (1) the output of clusters is vulnerable to data graph building quality; (2) the configuration of clusters is not clearly expressed	• Spectral (Kim et al., 2020)

	in the cluster results, and a change to the processing stage is required to discover the clustering indicators (Nie et al., 2016)	
Grid-based	• Split the entire data space into many non-overlapping rectangular cells, and then cluster based on cells instead of data objects (Deng et al., 2018) • Represent Space Driven Algorithm and presents the advantage that the number of objects is fully independent and depends on the number of cells, allowing fast run time (Hireche et al., 2020)	• STING • CLIQUE (Hireche et al., 2020)
Model-based	• Able to explain the implementation of clustering mixture models or explain the use of a family of clustering mixing models (Ingrassia et al., 2014) • Most of the techniques are based on Normal multivariate models and their variants (Ingrassia et al., 2014)	• Gaussian (Ingrassia et al., 2014)

Each cluster is represented by one data point in the cluster in K-Medoid clustering, and these points are called medoids of the cluster. It has a time complexity of $O(k(n-2)^2)$ compared to K-Mean clustering algorithm with $O(n^{dk+1})$ (Soni & Patel, 2017). In more depth, k data objects are randomly chosen as medoids to represent k clusters. All remaining data objects are positioned in the nearest or most similar cluster to medoids. After all data items have been analyzed, new medoids can best represent the cluster, and the whole procedure repeats itself. Once more, all data items are connected to new medoids-based clusters. Medoids change their position step by step in each iteration. It lasts until there is no medoid activity. This way, a set of n data items is contained in the k clusters (Raj, 2017). The algorithm and explanation are as follows (Ushakov & Vasilyev, 2021): given an initial finite set of data objects $P = (p_1, \dots, p_m), p_i \in R^n, i = 1, \dots, m$, that is divided into k disjoint clusters. The K-Medoid problem is to select k medoids $C = (c_i, \dots, c_m)$ from within the set P so as to minimize the overall sum of distances $d(.,.)$ (or other dissimilarity measures) between data objects and their closest medoids as in Equation 1.

$$\min_{C \subset P} \left\{ \sum_{j=1}^{m} \min_{i=1,\dots,k} d(p_j, c_i), |C| = k \right\} \qquad (1)$$

Thus, the cluster members representing the clusters are the medoids (or specificities) selected from actual results.

Compared to K-Mean, K-Medoid has several advantages and disadvantages. García et al. (2013) claimed that K-Mean is robust against noise and outliers. Noise and outliers are defined as data that can disrupt the learning process as it causes data inaccuracies. The mean value is used in K-Mean as the cluster's center. Thus, it can suffer from noise and outliers.

Conversely, K-Medoid is less susceptible to noise and outliers because it utilizes the core value collection median. However, Lee et al. (2020) argued that it causes the mechanism of computation rendering in K-Medoid becomes more costly than K-Mean. Besides, for each run, K-Medoid does not have the same result. The cluster's result depends on the initial allocation, and it is hard to estimate the optimized number of clusters k. Thus, impossible for a user to calculate the value k without previous knowledge (Kaur et al., 2014). Lastly, Soni and Patel (2017) argue that K-Medoid is unsuitable for running large datasets than K-Mean. Table 3 shows the features comparison between K-Medoid vs K-Mean.

Table 3 Features comparison between K-Medoid and K-Mean

Algorithm / **Features**	**K-Medoid**	**K-Mean**
Time-complexity	$O(k(n-2)^2)$	$O(n^{dk+1})$
Robustness to noise and outliers	Less	More
Computational cost	More	Less
Dataset size capability	Small	Large

In general, K-Mean randomly selects a k number of cluster centers from a data set (Rahman & Islam, 2014). The algorithm in clustering can be a feature, as an example in Figure 2. Training examples are shown as dots and cluster centroids as crosses. (a) original dataset. (b) random initial cluster centroids. (c-f) Illustration of running two iterations of K-Mean. Each training example is assigned to the closest cluster centroid in each iteration. It is shown by "painting" the training examples the same color as the cluster centroid to which it is assigned. Then, each cluster is moved from centroid to the mean of the points assigned to it.

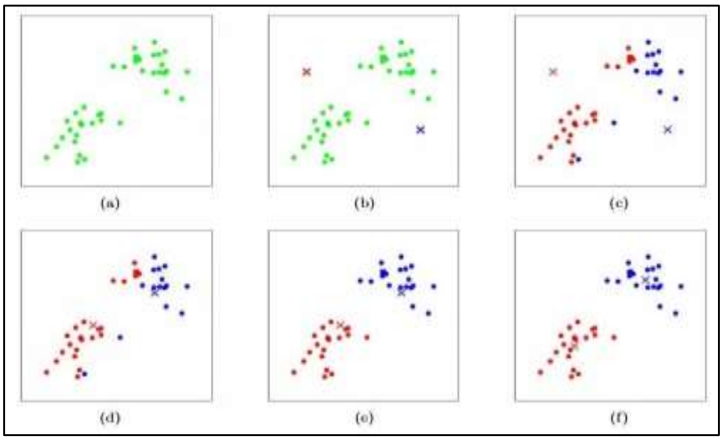

Figure 2 Clustering example of K-Mean

The algorithm usually ends when the centroids become stable or the points stop moving to other groups, but it depends on the type of grouped data to measure closeness and the objective function. Since K-Mean can have local optimal solution problems, a good initialization has proven to be an excellent way to overcome getting stuck in the wrong local optimal solutions (Nurhayati et al., 2019). The same goes for multiple algorithm executions with various random initialization (Scitovski & Sabo, 2014). Figure 3 shows the K-Mean pseudocode (Arora et al., 2016). The clustering aims to improve the objective feature (*f*) by measuring the range between entities and clusters (the most used measurement is the standard Euclidean Distance) as in Equation (2) (Serapião et al., 2016):

$$f = \sum_{i=1}^{K} \sum_{j=1}^{N} \left\| x_j - C_i \right\|^2 \qquad (2)$$
$$j \in G_i$$

where *K* is the number of clusters, *N* is the number of objects, x_j is the coordinate of object j, C_i is the coordinate of the cluster *i* and G_i is the group of objects that belong to cluster *i*. The algorithm shifts the cluster in space to reduce the square distances within the cluster. The positions of all objects belonging to each cluster are recalculated by averaging. Calculation of the center uses as in Equation (3):

$$C_i = \frac{1}{|G_i|} \sum_{j=1}^{N} X_j \qquad (3)$$
$$j \in G_i$$

where $|G_i|$ is the number of objects in the cluster *i*. The algorithm begins with a random set of the C_i cluster's initial *K* center points (*i* = 1, . . ., K), which are the present centroids.

Input: K_y: the number of clusters D_y: a data set containing n object
Output: A set of K_y clusters
Algorithm:
1. Input the data set and value of K_y.
2. If K_y = = 1 then Exit.
3. Else
4. Choose *k* objects from *D* randomly as the initial cluster centers.
5. For every data point in the cluster, *j* reissue and define every object into the cluster where the object matches, based on the object's mean value.
6. Update cluster means; after that, for each cluster, calculate the object's mean value.
7. Repeat from step 4 until no data point was assigned; otherwise, stop.

The satisfying criteria can be either number of iteration or the change of position of the centroid in consecutive iterations.

Figure 3 Pseudocode of K-Mean

Research Knowledge Chart

This study included both elements based on the review's outcome on the existing evacuation model and clustering algorithm for vertical indoor environments. We have mentioned the research questions posed: 1) how can the user determine the best indoor evacuation model be selected? 2) how can the indoor evacuation models have a standard assessment? Thus, we set the objective as to formulate the intelligent indoor evacuation assessment algorithm for critical incidents through clustering the features of the developed indoor evacuation model using the K-Means algorithm. Determining the appropriate selection is essential since the objective is to develop an indoor evacuation assessment algorithm for a critical incident. The focus for overall ideas is presented using a knowledge chart in Figure 4. The boxes highlighted in yellow indicate that the study's scope has gone through a systematic review and analysis to achieve the objective. We managed to see a clear picture of the research direction on specific methods and techniques. Therefore, we set our objective to formulate the intelligent indoor evacuation assessment algorithm for critical incidents by clustering the features of the developed indoor evacuation model using the K-Mean algorithm.

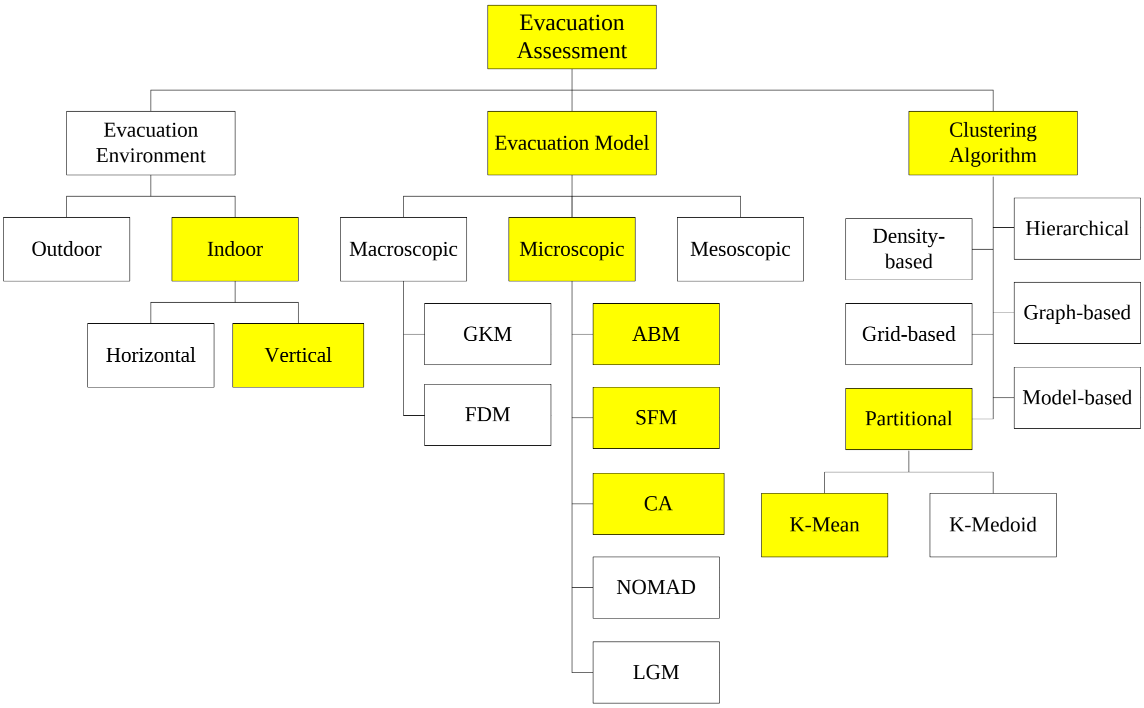

Figure 4 Knowledge chart in developing indoor evacuation assessment algorithm

Researchers have generally used the evacuation model for the three models: ABM, SFM, and CA under microscopic models (Fayez, 2017). In detail, these are more relied on because ABM can accommodate the high degree of system complexity and instantiate interaction between agents on one or another level. Jumadi et al. (2017) suggested it can be used to model social institutions, their behavior, social characteristics, and properties, which derive from their experiences. For SFM, the main contribution was to human behavior and associating motion with the laws of pedestrian behavior. This model has been generally adopted for pedestrian dynamics and primarily explains pedestrian activity and simulations (Zhou et al., 2019). Furthermore, CA has become an area of study at the national and international frontiers. It can involve human characteristics, the psychological reaction to disasters, relationships between people, and the effect on the evacuation behavior of the structure and climate (Yang et al., 2014).

The partitional clustering algorithm is the suitable clustering algorithm for this research due to being suitable for clustering large datasets. In comparing both clusters, the partitional clustering algorithm seems to benefit more. The reason is that hierarchical clustering algorithm lacks solidity, are computationally costly (complexity at least $O(n^2)$ with n number of objects) and susceptible to noise and outliers. Moreover, it will forever be permanent when a merging process is carried out, so corrections cannot be produced, and spherical clusters tend to form (Strazzeri, 2018).

The most often used partitioning-based clustering approach is the K-Means algorithm. It is a clustering algorithm that uses an unsupervised algorithm (Khalid et al., 2014). It is robust against noise and outliers, which the problem caused data inaccuracies. Thus, this research will be using the K-Mean clustering algorithm instead of K-Medoid. The purpose of the algorithm is to find the ideal set of points, which minimizes the average square distance from each information example to its nearest center, defined as square distortion and dependent on variance. Specific minimization targets include the number of distances, as with the Euclidean K-medians' question, and reduce the maximum distance from any point to its closest point in the Geometric K-center. Duwairi and Abu-Rahmeh (2015) revealed that another primary function is to define the property of clusters and maximize these functions, as in the case of clusters of papers, where the feature calculates the number of intra cosine similarities within clusters. K-Mean needs a certain requirement, which is:

1) determining a proximity measure
2) an assessment metric of the clusters' quality
3) the number of clusters (k)
4) the value of initial means a convergence condition

Conclusion

We presented the study's outcome in answering the two research questions set in developing the evacuation assessment algorithm. The literature review findings were conducted on evacuation models and clustering algorithms. In the evacuation model, microscopic and macroscopic models are commonly used to depict pedestrian traffic. However, the microscopic model has been much relied on within the research community in crowd simulation studies to understand crowd behavior in emergencies. We found that K-Mean has advantages, is suitable for our focus study with less computational cost and can cluster large datasets than K-Medoid. As a result, we developed a research knowledge chart to determine the objective based on the research questions. We would extend the simulation software for the selected evacuation ABM, SFM and CA in developing the evacuation assessment algorithm for future work. Also, we plan to identify the specific attributes and parameters involved.

Acknowledgement

The research is supported by the Ministry of Education Malaysia (MoE) and Universiti Teknologi MARA through the Fundamental Research Grant Scheme (FRGS-RACER) (600-IRMI/FRGS-RACER 5/3 (124/2019)).

References

- Arora, P., Deepali, & Varshney, S. (2016). Analysis of k-means and k-medoids algorithm for big data. *Physics Procedia*, *78*, 507–512.
- Bakar, N. A. A., Adam, K., Majid, M. A., & Allegra, M. (2017). A simulation model for crowd evacuation of fire emergency scenario. *2017 8th International Conference on Information Technology (ICIT)*, 361–368.
- Campanella, M., Hoogendoorn, S., & Daamen, W. (2009). Improving the nomad microscopic walker model. *12th IFAC Symposium on Transportation Systems*, *42*(15), 12–18.
- Campanella, M., Hoogendoorn, S., & Daamen, W. (2014). The nomad model: Theory, developments and applications. *Transportation Research Procedia*, *2*, 462–467.
- Chen, Y., Wang, C., Li, H., Yap, J. B. H., Tang, R., & Xu, B. (2020). Cellular automaton model for social forces interaction in building evacuation for sustainable society. *Sustainable Cities and Society*, *53*.
- Deng, C., Song, J., Sun, R., Cai, S., & Shi, Y. (2018). GRIDEN: An effective grid-based and density-based spatial clustering algorithm to support parallel computing. *Pattern Recognition Letters*, *109*, 81–88.
- Duwairi, R., & Abu-Rahmeh, M. (2015). A novel approach for initializing the spherical k-means clustering algorithm. *Simulation Modelling Practice and Theory*, *54*, 49–63.
- Fayez, L. (2017). *Modeling family behaviours in crowd simulation*. Qatar University.
- Fragkos, G., Apostolopoulos, P. A., & Tsiropoulou, E. E. (2019). ESCAPE: Evacuation strategy through clustering and autonomous operation in public safety systems. *Future Internet*, *11*(1), 1–17.
- Gaire, N. (2017). *A study on human evacuation behavior involving individuals with disabilities in a building* [Utah State University].
- García, L. P. F., De Carvalho, A. C. P. L. F., & Lorena, A. C. (2013). Noisy data set identification. *International Conference on Hybrid Artificial Intelligence Systems*, *8073*, 629–638.
- Gialampoukidis, I., Vrochidis, S., Kompatsiaris, I., & Antoniou, I. (2019). Probabilistic density-based estimation of the number of clusters using the DBSCAN-martingale process. *Pattern Recognition Letters*, *123*, 23–30.
- Guo, P., & Xie, J. (2018). A P system for k-medoids-based clustering. *International Journal of Advanced Computer Science and Applications(IJACSA)*, *9*(10), 41–48.
- Han, X. H., Quan, L., Xiong, X. Y., Almeter, M., Xiang, J., & Lan, Y. (2017). A novel data clustering algorithm based on modified gravitational search algorithm. *Engineering Applications of Artificial Intelligence*, *61*(September 2016), 1–7.
- Henry, K. D. (2015). *Development of a comprehensive network-based hazard evacuation model: A case study of Balboa Island, California*. University of Idaho.
- Hireche, C., Drias, H., & Moulai, H. (2020). Grid based clustering for satisfiability solving. *Applied Soft Computing*

Journal, *88*.

- Hoekstra, S., & Montz, B. (2017). Decisions under duress: factors influencing emergency management decision making during superstorm sandy. *Natural Hazards*, *88*(1), 453–471.
- Hu, Y., Liu, X., Wang, F., & Cheng, C. (2012). An overview of agent-based evacuation models for building fires. *2012 9th IEEE International Conference on Networking, Sensing and Control*, 382–386.
- Hufnagl, B., & Lohninger, H. (2020). A graph-based clustering method with special focus on hyperspectral imaging. *Analytica Chimica Acta*, *1097*, 37–48.
- Ibrahim, A. M., Venkat, I., & De Wilde, P. (2017). Uncertainty in spatial evacuation model. *Physica A: Statistical Mechanics and Its Applications*, *479*, 294–306.
- Ingrassia, S., Minotti, S. C., & Punzo, A. (2014). Model-based clustering via linear cluster-weighted models. *Computational Statistics and Data Analysis*, *71*(4), 159–182.
- Isenhour, M. L. (2016). *Simulating occupant response to emergency situations* [George Mason University].
- Jiang, Y., Chen, B., Li, X., & Ding, Z. (2020). Dynamic navigation field in the social force model for pedestrian evacuation. *Applied Mathematical Modelling*, *80*, 815–826.
- Jumadi, Carver, S., & Quincey, D. (2017). A conceptual design of spatio-emporal agent-based model for volcanic evacuation. *Systems*, *5*(4), 53.
- Kaur, N. K., Kaur, U., & Singh, D. (2014). K-Medoid clustering algorithm- A review. *International Journal of Computer Application and Technology*, *1*(1), 42–45.
- Khalid, M., Pal, N., & Arora, K. (2014). Clustering of image data using k-means and fuzzy k-means. *International Journal of Advanced Computer Science and Applications*, *5*(7), 160–163.
- Kim, Y., Do, H., & Kim, S. B. (2020). Outer-Points shaver: Robust graph-based clustering via node cutting. *Pattern Recognition*, *97*.
- Ko, S., Spearpoint, M., & Teo, A. (2007). Trial evacuation of an industrial premises and evacuation model comparison. *Fire Safety Journal*, *42*(2), 91–105.
- Kontou, P., Georgoulas, I. G., Trunfio, G. A., & Sirakoulis, G. C. (2018). Cellular automata modelling of the movement of people with disabilities during building evacuation. *2018 26th Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP)*, 550–557.
- Kutbay, U. (2018). Partitional clustering. *IntechOpen*, 20–33.
- Lee, M., Lee, S., Park, J., & Seo, S. (2020). Clustering and characterization of the lactation curves of dairy cows using K-medoids clustering algorithm. *Animals*, *10*(8), 1–14.
- Li, H., Liu, X., Li, T., & Gan, R. (2020). A novel density-based clustering algorithm using nearest neighbor graph. *Pattern Recognition*, *102*.
- Li, Y., Chen, M., Dou, Z., Zheng, X., Cheng, Y., & Mebarki, A. (2019). A review of cellular automata models for crowd evacuation. *Physica A: Statistical Mechanics and Its Applications*, *526*, 120752.
- Li, Y., Chen, M., Zheng, X., Dou, Z., & Cheng, Y. (2020). Relationship between behavior aggressiveness and pedestrian dynamics using behavior-based cellular automata model. *Applied Mathematics and Computation*, *371*.
- Liu, B., Liu, H., Zhang, H., & Qin, X. (2018). A social force evacuation model driven by video data. *Simulation Modelling Practice and Theory*, *84*, 190–203.
- Liu, C., Mao, Z. L., & Fu, Z. M. (2016). Emergency evacuation model and algorithm in the building with several exits. *Procedia Engineering*, *135*, 12–18.
- Luo, J., Qiu, J., & Han, J. (2018). Research status and prospect of fire rescue equipment in high-rise building. *International Conference on Sensing, Diagnostics, Prognostics, and Control, SDPC 2018*, 642–647.
- Lv, Y., Ma, T., Tang, M., Cao, J., Tian, Y., Al-Dhelaan, A., & Al-Rodhaan, M. (2016). An efficient and scalable density-based clustering algorithm for datasets with complex structures. *Neurocomputing*, *171*, 9–22.
- Ma, Q., & Guo, W. (2012). Discussion on the fire safety design of a high-rise building. *Procedia Engineering*, *45*(3), 685–689.
- Martins, P. (2015). *Contribution to the knowledge of hierarchical clustering algorithms and consensus clustering*. University of Porto.
- Ming, R., & Peng, X. (2017). Study on the social force model of personnel evacuation in large stadiums. *2017 International Conference on Service Systems and Service Management*, 1–5.
- Nanda, S. J., & Panda, G. (2014). A survey on nature inspired metaheuristic algorithms for partitional clustering. *Swarm*

and Evolutionary Computation, *16*, 1–18.

- Nie, F., Wang, X., Jordan, M. I., & Huang, H. (2016). The constrained laplacian rank algorithm for graph-based clustering. *30th AAAI Conference on Artificial Intelligence, AAAI 2016*, *1*, 1969–1976.
- Nisha, & Kaur, P. J. (2016). Cluster quality based performance evaluation of hierarchical clustering method. *2015 1st International Conference on Next Generation Computing Technologies (NGCT)*, 649–653.
- Nurhayati, Sinatrya, N. S., Wardhani, L. K., & Busman. (2019). Analysis of k-means and k-medoids's performance using big data technology. *2018 6th International Conference on Cyber and IT Service Management, CITSM 2018*, 1–5.
- Pelechano, N., & Malkawi, A. (2008). Evacuation simulation models: Challenges in modeling high rise building evacuation with cellular automata approaches. *Automation in Construction*, *17*(4), 377–385.
- Poulos, A., Tocornal, F., de la Llera, J. C., & Mitrani-Reiser, J. (2018). Validation of an agent-based building evacuation model with a school drill. *Transportation Research Part C: Emerging Technologies*, *97*, 82–95.
- Rahman, M. A., & Islam, M. Z. (2014). A hybrid clustering technique combining a novel genetic algorithm with K-Means. *Knowledge-Based Systems*, *71*, 345–365.
- Raj, C. K. (2017). Comparison of k-means, k-medoids, DBSCAN algorithms using DNA microarray dataset. *International Journal of Computational and Applied Mathematics*, *12*(1), 344–355.
- Ratsamee, P., Mae, Y., Ohara, K., Takubo, T., & Arai, T. (2012). Modified social force model with face pose for human collision avoidance. *2012 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 215–216.
- Saini, G., & Kaur, H. (2014). A novel approach towards k-mean clustering algorithm with PSO. *International Journal of Computer Science and Information Technologies (IJCSIT)*, *5*(4), 5978–5986.
- Sakour, I., & Hu, H. (2017). Robot-assisted crowd evacuation under emergency situations: A survey. *Robotics*, *6*(2).
- Samah, K. A. F. A., Hussin, B., & Basari, A. S. H. (2015). Modification of dijkstra's algorithm for safest and shortest path during emergency evacuation. *Applied Mathematical Sciences*, *9*(29–32), 1531–1541.
- Scitovski, R., & Sabo, K. (2014). Analysis of the k-means algorithm in the case of data points occurring on the border of two or more clusters. *Knowledge-Based Systems*, *57*, 1–7.
- Serapião, A. B. S., Corrêa, G. S., Gonçalves, F. B., & Carvalho, V. O. (2016). Combining k-means and k-harmonic with fish school search algorithm for data clustering task on graphics processing units. *Applied Soft Computing Journal*, *41*, 290–304.
- Shi, M., Lee, E. W. M., & Ma, Y. (2018). A novel grid-based mesoscopic model for evacuation dynamics. *Physica A: Statistical Mechanics and Its Applications*, *497*, 198–210.
- Soni, K. G., & Patel, A. (2017). Comparative analysis of k-means and k-medoids algorithm on iris data. *International Journal of Computer Applications*, *5*, 899–906.
- Strazzeri, F. (2018). *A morse-theoretical clustering algorithm for annotated networks and spectral bounds for fuzzy clustering*. University of Southampton.
- Tan, L. (2014). *Wayfinding modelling using cognizing agent for evacuation simulation of multi-level buildings* [Chinese University of Hong Kong].
- Tzortzis, G., & Likas, A. (2014). The minmax k-means clustering algorithm. *Pattern Recognition*, *47*(7), 2505–2516.
- Ushakov, A. V., & Vasilyev, I. (2021). Near optimal large scale k-medoids clustering. *Information Sciences*, *545*, 344–362.
- Ventura, G. M. (2017). *Patient evacuation resource classification system (PERCS) for residential healthcare facilities: Patient classification system translatable to healthcare evacuation protocols, system modeling, and transportation resources* [The George Washington University].
- Vogiatzis, C., Walteros, J. L., & Pardalos, P. M. (2013). Evacuation through clustering techniques. *Models, Algorithms, and Technologies for Network Analysis*, 185–198.
- Wang, C. Y., & Weng, W. G. (2014). Study on evacuation characteristics in an ultra high-rise building with social force model. *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 566–571.
- Wu, G. Y., & Huang, H. C. (2015). Modeling the emergency evacuation of the high rise building based on the control volume model. *Safety Science*, *73*, 62–72.
- Xie, H., & Weerasekara, N. N. (2016). Modeling and simulation of the evacuation plan for Hancock Stadium. *Construction Research Congress 2016: Old and New Construction Technologies Converge in Historic San Juan*, *8*, 2453–2462.
- Xie, Q. Y. (2013). *K-centers dynamic clustering algorithms and applications* [University of Cincinnati].
- Xie, R., & Li, L. (2014). Simulation of optimized evacuation processes in complex buildings using cellular automata model.

Journal of Software, *9*(6), 1428–1434.

- Xu, Q., Zhang, Q., Liu, J., & Luo, B. (2020). Efficient synthetical clustering validity indexes for hierarchical clustering. *Expert Systems With Applications*, *150*.
- Yang, Y., Deng, J., Xie, C. C., & Jiang, Y. T. (2014). Design and implementation of fire safety evacuation simulation software based on cellular automata model. *Procedia Engineering*, *71*, 364–371.
- Yao, Z., Zhang, G., Lu, D., & Liu, H. (2019). Data-driven crowd evacuation: A reinforcement learning method. *Neurocomputing*, *366*, 314–327.
- Yu, X., Chang, R., & Zhang, C. (2014). Evacuation of pedestrians using lattice gas model and floor field model. *2014 International Conference on Audio, Language and Image Processing*, 812–816.
- Zhang, Z. (2011). *Clustering: Algorithm, optimization and inference*. University of California Riverside.
- Zheng, X., Zhong, T., & Liu, M. (2009). Modeling crowd evacuation of a building based on seven methodological approaches. *Building and Environment*, *44*(3), 437–445.
- Zhou, J., Guo, Y., Dong, S., Zhang, M., & Mao, T. (2019). Simulation of pedestrian evacuation route choice using social force model in large-scale public space: Comparison of five evacuation strategies. *PLoS ONE*, *14*(9), 1–25.