

# Graphed-based K-Means and Shortest Distance Tree for the Construction of Elderly Safe Corridor Accident and Prevention Platform

Ying-Jhen Su<sup>a</sup>, Zhen-Kai Wu<sup>a</sup>, You-Jia Su<sup>b</sup>, Yu-Shan Lin<sup>b</sup>, Yao-Chung Chang<sup>a</sup>

<sup>a</sup> Department of Computer Science and Information Engineering, National Taitung University, Taitung, Taiwan

<sup>b</sup> Department of Information Science and Management System, National Taitung University, Taitung, Taiwan  
10811117@gm.nttu.edu.tw, 10811118@gm.nttu.edu.tw, 10912117@gm.nttu.edu.tw, ysl@nttu.edu.tw, ycc@nttu.edu.tw

**Abstract**—As medicine, science, and technology improve, the average age of the population on earth is also increasing. The World Health Organization (WHO) reported that the average age of the countries that had joined and surveyed was 68 years old, and the highest country had over 80 years. The care of the elderly population requires more attention and careful planning. There are a variety of social support systems available for the elderly today. Apart from encouraging them to participate in various social affairs and activities, it is also necessary to consider the safety and accident prevention needs of the elderly when they move. This is one of the critical factors for the elderly to obtain various opportunities and resources. Those areas or locations where older adults are densely concentrated require a more effective police presence to maintain the safety of moving routes. This research primarily deals with arranging police forces based on a simulation of elderly individuals' action trajectories. The level of care that the elderly require is weighted. Furthermore, the weighted K-means Clustering method is used to find the clusters of the elderly and make the cluster center close to the elderly who need high care to deploy the police force to provide higher priority services. Then the shortest paths between the elderly settlements (cluster center) and the activity venue are found based on the consideration that the movement patterns of the elderly are mainly walking, and they are most likely to arrive at the activity venue within the shortest distance. The activity venues of the shortest paths are taken as the root to construct the shortest path tree. During tree construction, the number of elderly in each settlement is summed in each intersection node as the number of elderly passing through. Then the breadth-first search method is applied to find the multi-path convergence points that the most elderly pass through as the reference for recommending police force arrangement. According to the proposed method, the police can plan safety corridors for the elderly so that they may receive the best safety care while away from home.

**Keywords**— elderly care, weighted K-means, movement patterns, shortest path tree, police force arrangement, safety corridors.

## I. INTRODUCTION

The aging society is defined as 7% of the population over 65, the advanced society as 14%, and the super-aged society as 20%. According to the National Development Commission of the Ministry of the Interior's Republic of China Population Estimation Timeline[1] and The Three-Stage Age Population

Trend of the Republic of China[2], Taiwan entered an aging society in 1993 and turned into an aging society in 1818. It is estimated that the country will join a super-aged society by 2025. The Ministry of the Interior announced in 1818 that Taiwan's elderly population had reached 14.05%, meeting the standard of the World Health Organization for a "senior society." In 2026, this percentage will get 20% of the total population, making the country officially a "super-aged society." In Figure 1, the aging trend of Taiwan's population is illustrated over time, and Figure 2 shows the changing trend in Taiwan's people and their aging. Taiwan's population structure has become progressively aged due to the rapid aging caused by aging. As Taiwan's elderly population overgrows, it should focus more on their needs and services.

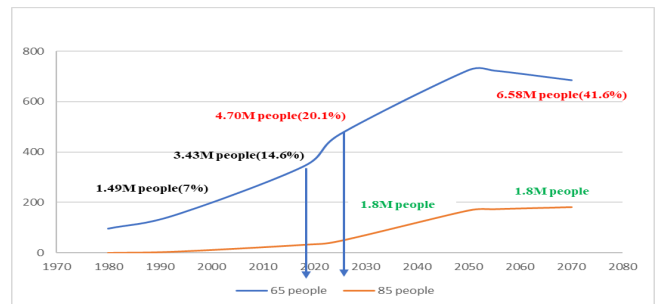


Figure 1. Time course of the aging of the country's population

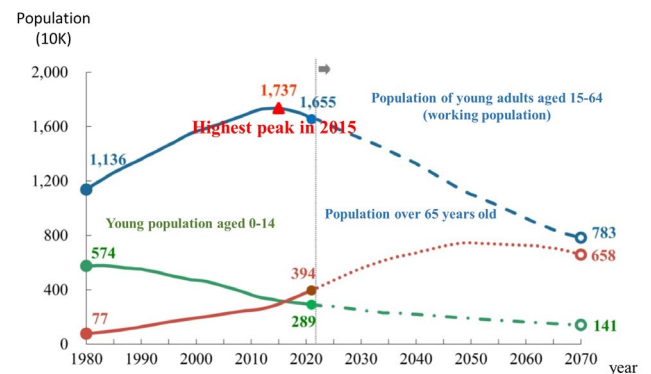


Figure 2. Trends in population age and population change in the country

It is crucial to consider the safety and accident prevention needs of the elderly during the relocation and encourage them to participate in various social activities and affairs. It will ensure that they do not need to worry about their safety while moving and will not suffer from social isolation. The restriction of the moving range of the elderly, which reduces their willingness to go outside, leads to social isolation. As social support systems for the elderly are constantly evolving, one of the essential factors in ensuring that the elderly have access to a wide range of opportunities and resources is the issue of safety. The elderly should be able to engage in ordinary life activities while maintaining their movement ability, autonomy, and independence. They could participate in daily activities, such as shopping, leisure, visiting relatives, seeking medical treatment, and communicating with neighbors, and don't need to concern about safety factors. Such events include gatherings for the elderly, community centers, medical institutions, community colleges, and other locations. As the elderly population grows, the conditions affecting their ability to participate in social affairs, mobility, and safety are of paramount importance.

As soon as the mobile network for the elderly has been constructed, they must fully consult the police force. The more densely populated an area or place is where the elderly are, the more appropriately they should deploy a police force to ensure their safety. Using an information system platform can facilitate a more effective decision-making process by collecting and analyzing data efficiently and effectively. A safety net platform is used to collect data about the elderly, and the data is analyzed and presented to determine the best police force arrangement. An appropriate amount of police force is assigned to patrol the elderly route, thus ensuring the security of the elderly and a streamlined police force configuration.

This study is motivated by the goal of thinking through the issue from the perspective of transportation infrastructure. It will be necessary to maintain and improve safety care and accident prevention for the elderly when they move in the future. As a result, establishing a safe mobile network enables the elderly to receive increased police protection when they move. To improve the quality of life of the elderly, it is imperative to enhance their willingness to move. In addition to improving physical and mental health, going outside can reduce medical and social costs for the elderly. Thus, creating a safe pedestrian environment for the elderly and enhancing their willingness to participate in society and care for their safety needs is a crucial and urgent issue.

Accordingly, the research topics for this study are as follows:

(1) Establish a safety corridor for the elderly and an accident prevention platform to collect information on the elderly.

(2) A police agency is simulated to collect and introduce movement data of elderly individuals, calculate movement paths, collect latitude and longitude data points for elderly activity venues, and calculate moving paths using the map application for research and analysis.

(3) Evaluate the movement trajectories of the elderly to a particular activity area, organize them according to their location, locate the elderly in a specific location, and calculate the nearest distance between the destination of the activity and the elderly living center. It combines these shortest paths into the shortest path tree. Calculating the number of elders that may converge at each node on the track will help determine the location of police duty. This is based on the number of elders at each node and the distance from the center where the elders reside.

(4) The moving route should be maintained with a more reasonable police force in areas where the elderly are frequently seen. The information platform will recommend the proper arrangement of police personnel so that the patrol police have a more appropriate and streamlined configuration.

(5) Provide police units with real-time information on the time, location, and intensity of elderly activities based on the collected statistics and analysis, which can be used as a reference for duty decision-making and increase decision-making accuracy.

## II. LITERATURE REVIEW AND DISCUSSION

It is not easy to accurately collect the location of human activities. A previous study observed the movement patterns of 160 volunteers aged 7 to 74. To determine the effect of human movement on physical activity and exposure to pathogens [3]. Using GPS data loggers and semi-structured interviews, the study collected location data. As mobile phones have become ubiquitous, [4] collect location information through their GPS module, and analyze patterns of mass gatherings and city travel based on where people stop moving. Deidentification is vital when location data is collected since personal information is often involved.

It is necessary to analyze a large amount of location data to find where people gather frequently based on the collected location data. Location-based services should identify meaningful locations from large amounts of location information and predict where users will go. Using hidden Markov models, [5] identified the user's location using k-nearest neighbor analysis and decision trees. According to records of highway accidents within 24 hours, the K-means[6] grouping determines the place and time of accidents more likely to appear on the highway. As a simple and effective machine learning method, K-means can quickly group a group of comparable data and identify the cluster center without any prior knowledge of the data. However, when K-means is used to process large amounts of data, each iteration must be repeated using all the data, which can take considerable time. [7] discussed how a mini-batch K-means algorithm can improve clustering efficiency by using a fixed-size subsample at each iteration.

To ensure the safety and comfort of the elderly, police may be stationed at appropriate locations when events are held near the police station. As a result of limited police presence, police will be located either in the central cluster center or between several smaller cluster centers. For this study, it was assumed that older people would choose the most time- and

effort-saving method to travel along an established path. Using the shortest route on the map, the places where elders may pass are evaluated, and intersections where many elders may gather are recommended as police stations to serve them.

### III. RESEARCH METHODS AND PROCEDURES

The primary purpose of this research is to build a mobile safety net for the elderly to get maximum police care when they move. The research steps are

#### A. Building safety corridors and accident prevention platforms for the elderly

The platform's architecture is divided into four components (According to Figure 3): the network infrastructure, the website, the application side, and the library side.

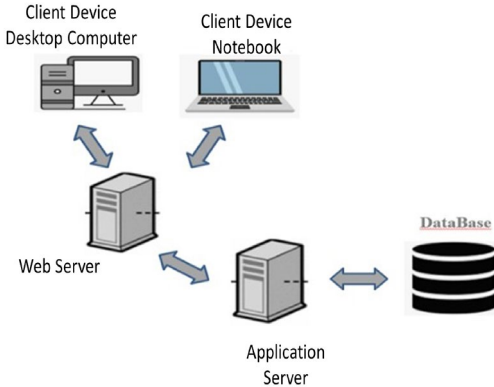


**Figure 3.** A safety corridor for the elderly and a platform for incident prevention

There are two sides to the platform, one for the manager and one for the police. On the manager side of the system, administrators manage the county police department, precincts, and police stations, while on the police side, officers perform their duties using the information system.

#### B. Data collection for elderly actions

In this study, longitude and latitude are used to mark haunting spots of the elderly on the leaflet map and store them in the database for further module development.

In the preliminary planning stage, information such as id, name, gender, village, street number, longitude, latitude, the status of the elderly, precinct, location, means of transportation, activity time, activity cycle, event location, longitude, latitude, and other information were collected.

To simulate the elderly data as close to the real world as possible, the elderly data is generated based on the number of men and women over the age of 60 in each township of Taitung County[9] in July 2022. We randomly selected the addresses and assigned them to the elderly from the address data of Taitung County available on the government's open data platform[10]. We generate random names as the primary data for the Taitung County police department to investigate the elderly population. Additionally, through the government address information query platform, the GPS location of the household address[11] is obtained to generate the distribution of the elderly population.

#### C. Determine the location or area where the elderly are haunted

Because of the limited number of locations the police can support, we will use the K-means cluster analysis method to centrally divide the elderly into k groups, where  $k \geq L_n$ . K-means is a popular data mining method to classify n samples into k groups. The examples at the site where the elders appeared are  $\{x_1, x_2, \dots, x_n\}$ . The criteria in cluster S will be classified into cluster S, defined as  $\{s_1, s_2, \dots, s_k\}$ , where the cluster's center significantly influences the samples in each cluster. The smallest sum of squares should be used to solve the following formula. (1):

$$\arg \min_S \sum_{i=1}^k \sum_{x \in S_i} \|x - \bar{x}_i\|^2 \quad (1)$$

This is the mean of all samples within  $s_i$ . Nevertheless, finding the optimal global solution with this method is challenging, which is why many iterative methods are used to find the optimal regional solution. According to the results of this study, if the longitude and latitude are used to group clusters and the average of longitude and latitude is used to calculate the cluster center, the cluster center is precisely the location of the center point of the cluster sample. Figure 4 illustrates the results of the K-means clustering method applied to the elderly residents of the Guanshan Police Station. In this paper, the blue flag represents the cluster center of the elderly residents, also called the cluster center, which means the average location of the group of samples.

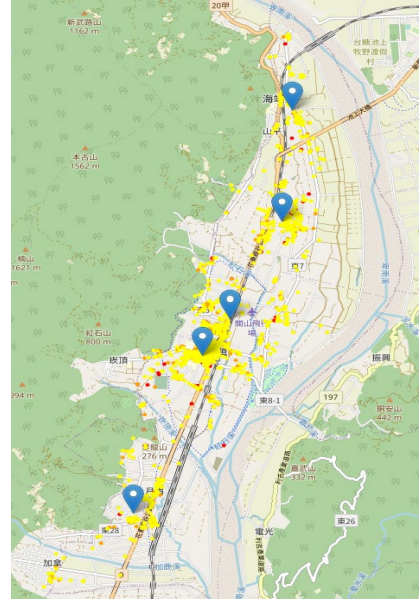


**Figure 4.** Based on the weighted K-means clustering method, elderly living centers have been identified within the vicinity of the Guanshan police station.

As many older adults are sick, unable to walk, or require assistance to move about, the police force should be arranged closer to these areas based on the required intensity and quantity of service. To make the police force more reasonable, this study classified the elderly population into three

categories based on their required assistance. According to the intensity of the necessary aid, they can be categorized as not requiring special assistance, requiring simple service, and requiring high support. It has been found that the proportion of elderly who do not require special assistance is higher than that of the elderly who need simple service, and the ratio of the elderly who require high aid is the smallest. In each state, elders are given varying weights according to their position. We provide the highest weight to elders who require special assistance, those who need simple assistance, and those who do not require special assistance.

Using the k-mean algorithm for clustering, we use pre-assigned weights, so formula (2) is as follows:

$$\arg \min_s \sum_{i=1}^k \sum_{x \in s_i} w_x \|x - \bar{x}_i\|^2 \quad (2)$$

In this case,  $w_x$  represents the weight of the position  $x$ . According to the above formula, the location is closer to the residence of the elderly in need of assistance. An example of a police station jurisdiction in Taitung County can be found in Figure 4. This study uses the weighted K-means clustering method. Yellow, orange, and red dots represent the residences of elderly residents who do not require special care or require high levels of care. The blue icon represents the weighted average of the elders' homes in the group.

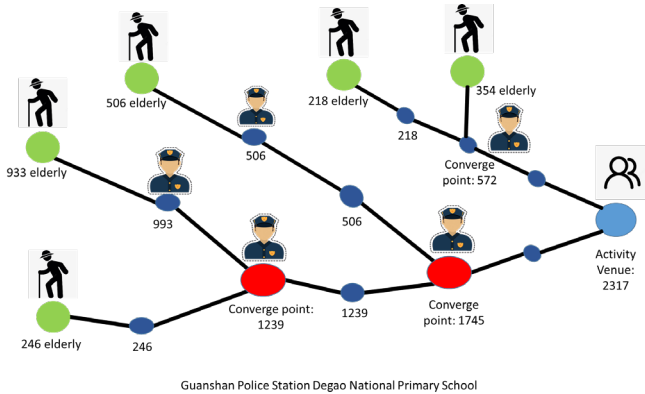


**Figure 5.** Schematic diagram of the shortest path from each settlement center to each activity venue.

Whether the elderly population center is close to the residence of those in need of high-level care is highly correlated with the selection of the number of clusters  $k$ . The deviation from the group's center is not evident when there are too many participants because everyone in the group needs high-level care. In this study,  $k$  is related to the jurisdiction's population, assuming an excellent proportional relationship exists between the police force and the people in need of care. There is a proportionality between the numbers. As a result, some areas have a wide range of jurisdictions, and the population is too small to be divided into one group, making it difficult to determine the possible convergence points during the movement of the elderly. Therefore, we set a minimum value greater than 1 for  $k$ .

#### **D. Recommend the optimal arrangement of police personnel.**

To maintain the safety of the moving route, police forces should be oriented more toward areas where senior citizens often haunt and move. Analyzing the collected data recommends organizing police human resources and arranging appropriate police officers for patrols based on the collected information. A safer moving route for the elderly and a better configuration of the patrolling police force will ensure the safety of elderly citizens.

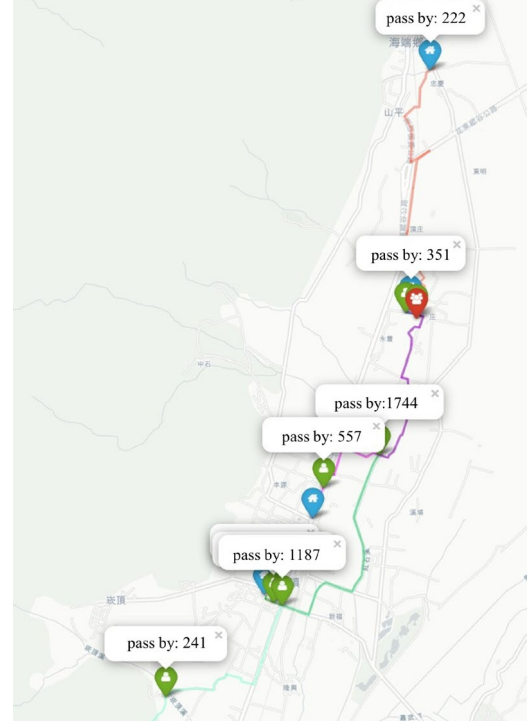


**Figure 6.** An analysis of the shortest paths between each settlement center and the activity site at Guanshan Police Station.

According to the leaflet map [12], the older adults and elderly gathering centers in the police station area, and the event locations in the area, the calculation is made at a specific speed, and the distance is used as an optimization index to determine the most probable convergence points when the elderly move. The shortest path is used as a reference, and the connection nodes are listed by their coordinate position on the map. Based on the walking speed in this study, the shortest path node sequence is calculated between the starting point (elderly population center) and the ending point (activity venue), with distance as the priority. Build the shortest path tree based on these shortest path nodes. Taking the map node closest to the endpoint as the root, reverse the sequence of the shortest paths from the settlement centers to the activity venue, and then build a tree structure for each shortest path one by one, as shown in Figure 5. When a non-overlapping path node is encountered, a bifurcated path is formed. The number of people passing through the path is determined by adding the population of the settlement centers of the overlapping nodes and according to whether the point in the path is a convergence point of different forks and how many elderly individuals are likely to pass through the convergence point, which can serve as a recommendation

point when recommending police dispatching positions. Based on the number of police officers they can deploy, the police station may choose the nodes with the highest number of elders that will pass through. According to Figure 6, the shortest path tree from each senior center of Guanshan Police Station to Degao National Primary School is shown on the map. A different color is assigned to each path. Red indicates the activity venue, and blue indicates the location of the senior center. It is recommended to dispatch police from the orange icon. The picture shows that the police dispatch points are all located at converging intersections where many people are present.

#### IV. RESULTS AND DISCUSSION OF THE RESEARCH

For this study, we focus on Taitung County, where the university is located. In Taiwan, Taitung County is a relatively remote county and city. Almost all residents are middle-aged or elderly because most young people work in urban areas.

We used public information from the government, including Taitung County's address data, GPS address data, and the number of men and women of all ages in Taitung's elderly population in July 2022. An elderly population was defined as one over the age of 60 in the study.

Based on each age and gender population, we generated the same number of simulated elderly personal data, including name, age, gender, address, and level of assistance required. According to the characteristics of Taiwanese households, an address is randomly assigned to 1 to 4 elderly.

We assigned 3 levels of assistance required for the elderly, namely, *I*1, no assistance requirement, *I*2, essential assistance required, and, *I*3, high assistance required. In each police station's jurisdiction, we designate the percentage of the elderly population to be assisted based on its elderly population. In this study, we set the number of people in *I*3 to be between 5% and 15%, while *I*2 is between 10% and 25%.

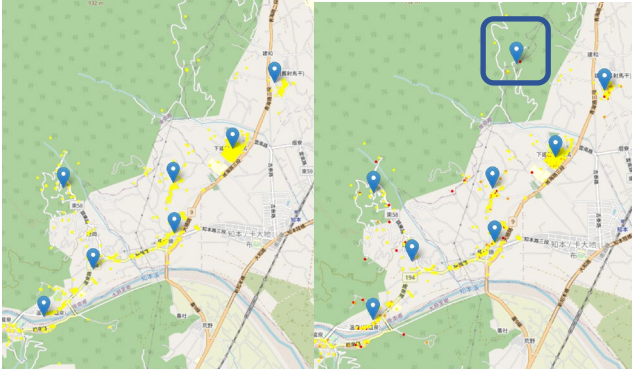


**Figure 7.** The location of the settlements in the hot spring police station jurisdiction (a) is based on the location of the population center. (b) is the location of the settlement center after setting the weight. The red dots are those who need high assistance.

By setting the weight of *I*1 to 1, *I*2 to 16, and *I*3 to 81, the path planning is biased towards the location of the elder *I*3. To prevent clusters from being too scattered, the number of clusters *k* in this study is set to

$$k = \max\left(\frac{P}{n}, 3\right) \quad (3)$$

where *P* is the population of the elderly in the police station jurisdiction, and *n* is the number of people cared for by each police officer.

Figure 7 shows the jurisdiction of the Taitung County Hot Spring Police Station. The flag in Fig. 7(a) shows the location of the settlement center when the weight is not used, and Fig. 7(b) shows the location of the settlement center after setting the weight. The settlement centers shown in the blue boxes are close to the addresses of the elderly needing high assistance, as shown in red points.

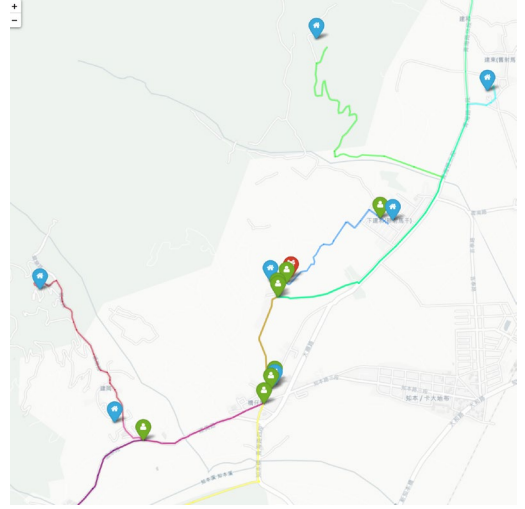


**Figure 8.** The settlement center (blue flag), the activity venue (red flag), and the suggested police deployment point (green flag) in the hot spring police station jurisdiction. Various colors are used on the map to indicate the shortest path from each settlement to the activity venue.

Each color route in Figure 8 is the shortest route from the center of the elderly to the activity venue. The red flag is the activity venue, and the blue flag is the settlement center. Green flags are suggested police deployment points. The different colored routes on the map represent the shortest path from a settlement center to the activity venue. If the number of people passing through the intersection is higher than the threshold, the intersection is suggested as a police deployment point. Finally, the recommended intersections are arranged according to the number of people passing through.

The number of the settlement clusters *k* should be carefully evaluated to ensure that the policies can be effectively located on the path that the elderly requiring high-level care will pass through on their way to the activity venue, and to create an optimum arrangement under a limited police force. The elderly settlement center should be as close as possible to the location of the elderly who need high-level care so that the path from the elderly population center to the activity venue would be closed to the elders who need help. The number of clusters that do not contain high care need elders should be restricted to avoid ineffective police dispatch.

In sparsely populated areas, elders who live in scattered places still need police forces to take care. Therefore, in this study, the minimum number of clusters in each police station's jurisdiction is set to 3.

**TABLE 1.** The relation between  $n$ , average distances from elders who need high care level to the settlement center (cluster center), the number of clusters containing high-care need elders, and the number of clusters does not have high-care need elders.

| $n$                                                    | 50  | 100   | 200   | 300   | 400   | 500   |
|--------------------------------------------------------|-----|-------|-------|-------|-------|-------|
| Average distance(meters)                               | 116 | 196.7 | 322.2 | 399.4 | 468.1 | 508.6 |
| No. of clusters containing high-care need elders       | 998 | 565   | 321   | 248   | 217   | 200   |
| No. of clusters does not contain high-care need elders | 138 | 27    | 12    | 11    | 11    | 11    |

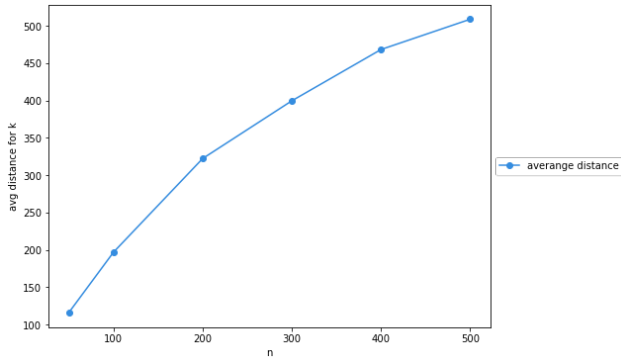


**Figure 9.** The relationship between  $n$  and average distances from elders who need high care level to the settlement center.

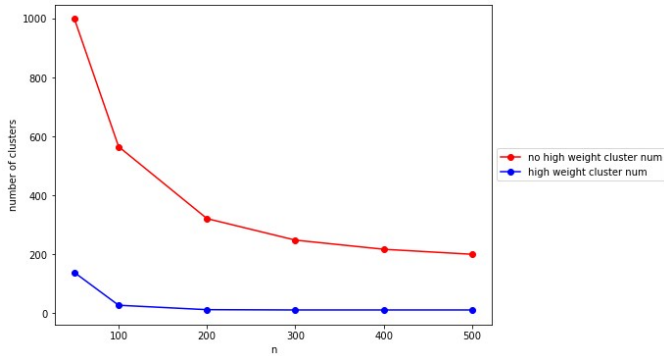


**Figure 10.** the relationship between  $n$  and the number of clusters containing and not containing high care need elders.

Table 1 shows that the average distances from elders who need high care level to the settlement center(cluster center) will be longer if  $n$  is greater. If the elders are far from the settlement centers, the policies will also be dispatched far from the elders. Based on the table, it is clear that if  $n$  is greater than 100, the average distances between elders and the settlement centers are longer.

From Figure 9 and Figure 10, when  $n$  is less than 100, the number of clusters is exponentially increased so that the time of path calculation will also increase exponentially.

## V. CONCLUSION

This study establishes a safety corridor planning and accident prevention platform for the elderly. With the help of the information platform, movement trajectory data of the elderly can be collected and analyzed. The data and data are analyzed to determine a model of the elderly's movement trajectory. As a reference. Research applications include the following features:

(1) Based on the address data of Taitung County on the government's open data platform, according to the number of older adults and the ratio of men to women at that time, randomly select the data as a simulation and obtain the GPS location of the household address through the government's address data query platform. On the map, indicate the distribution of the elderly population.

(2) A mobile model for locating the elderly by using a network map.

(3) Coupling the cluster center with the elderly who need high care using the weighted K-means clustering method.

(4) In addition, the planning is based on the movement patterns and timing of the elderly, with the shortest paths between the center and activity sites used as the calculation standard.

It is recommended that the shortest path tree be constructed, the places where the elderly may converge when moving by the breadth-first search method, the intersections the elderly pass, and recommendations and suggestions for police personnel placement be made.

The map displays the population center, the shortest path tree, and the recommended location for a police force arrangement.

## REFERENCES

- [1] National Development Commission of the Ministry of the Interior's Republic of China Population Estimation Timeline, [https://www.ndc.gov.tw/Content\\_List.aspx?n=D527207EEEF59B9B](https://www.ndc.gov.tw/Content_List.aspx?n=D527207EEEF59B9B), 2022.
- [2] National Development Commission of the Ministry of Internal Affairs of the Republic of China, "The three-stage age population change trend of the Republic of China," [https://www.ndc.gov.tw/Content\\_List.aspx?n=695E69E28C6AC7F3](https://www.ndc.gov.tw/Content_List.aspx?n=695E69E28C6AC7F3), 2022.
- [3] Paz-Soldan, V. A., Jr, R. C. R., Morrison, A. C., Stoddard, S. T., Kitron, U., Scott, T. W., Elder, J. P., Halsey E. S., Kochel, T. J., Astete, H., and Vazquez-Prokopec, G. M., "Strengths and Weaknesses of Global Positioning System (GPS) Data-Loggers and Semi-structured Interviews for Capturing Fine-scale Human Mobility: Findings from Iquitos, Peru," *PLoS Neglected Tropical Diseases (PLOS NEGLECT TROP D)*, vol. 8, no. 6, e2888, 2014. DOI: <https://doi.org/10.1371/journal.pntd.0002888>
- [4] Gong, L., Sato, H., Yamamoto, T., Miwa, T., and Morikawa, T., "Identification of activity stop locations in GPS trajectories by density-based clustering method combined with support vector machines," *Journal of Modern Transportation*, vol. 23, pp. 202-213, 2015.
- [5] Sung-Bae Cho, "Exploiting machine learning techniques for location recognition and prediction with smartphone logs," *Neurocomputing*, vol. 176, pp. 98-106, 2016. ISSN 0925-2312, <https://doi.org/10.1016/j.neucom.2015.02.079>.
- [6] Aljofey, A.M. and Alwagih, K., "Analysis of Accident Times for Highway Locations Using K-means Clustering and Decision Rules Extracted from Decision Trees," *International Journal of Computer Applications Technology and Research*, vol. 7, no. 1, pp. 1-11, 2018.

- [7] Sculley, D., "Web-scale K-means clustering," In Proceedings of the 19th international conference on World wide web, pp. 1177–1178, ACM, 2010.
- [8] Denzin, N. K., The research act: A theoretical introduction to the sociological method (2nd ed.). New York: McGraw-Hill, 1978.
- [9] Government Information Open Platform, "11107 Number of households in the village, single age population," Number of households in the village, single age population (Add area code) , 2022.
- [10] Website of the Department of Household Affairs, Ministry of Internal Affairs of the Republic of China, "Query by township, urban area, and compilation category," 2022.  
<https://www.ris.gov.tw/app/portal/3053>.  
<https://data.gov.tw/dataset/77132>.
- [11] Internal Affairs Geographic Information Map Cloud Integration Service Platform, "Batch house address comparison service," National house address location service, 2022.  
[https://www.tgos.tw/tgos/Web/Address/TGOS\\_Address.aspx](https://www.tgos.tw/tgos/Web/Address/TGOS_Address.aspx)
- [12] Agafonkin, Volodymyr, "Leaflet: an open-source JavaScript library for mobile-friendly interactive maps," leaflet official website, Oct. 13, 2022.  
<https://leafletjs.com/>



**Ying-Jhen Su** is a student in the Department of Computer Science and Information Engineering, National Taitung University, Taitung, Taiwan. His main research interests include Information and Network Security, AIoT, Deep Learning, Embedded System, and Algorithm.



**JheN-Kai Wu** is a student in the Department of Computer Science and Information Engineering, National Taitung University, Taitung, Taiwan. His main research interests include IoT, Network Application, and Embedded System.



**You-Jia Su** is a student in the Department of Information Science and Management Systems, National Taitung University, Taitung, Taiwan. His main research interests include Information & Network Security, AIoT & Machine Learning.



**Yao-Chung Chang** (Member, IEEE) is a professor in the Department of Computer Science and Information Engineering, National Taitung University, Taitung, Taiwan. He received a Ph.D. degree in Computer Science and Information Engineering in Computer Science and Information Engineering from National Dong Hwa University (NDHU), Hualien, Taiwan. Also, he received the Outstanding Teaching Award from Nation Taitung University. His main research interests include Information & Network Security, Cloud & Mobile Computing, AIoT & Machine Learning.



**Yu-Shan Lin** is a professor in the Department of Information Science and Management Systems, National Taitung University, Taitung, Taiwan. She received her Ph.D. degree from National Sun Yat-sen University, Kaohsiung, Taiwan. Her research interesting areas include e-Learning, Information Technology Education, and Internet Marketing. She has published about 20 journal papers, and some are highly-cited papers. Moreover, she received the Best Paper Award from IEEE ICASI 2018, IC3 2018, 2019 and TANET 2020, 2021.