Prediction of Electric Load Neural Network Prediction Model for Big Data

Teng Tu[*,1,#], Yihao Ji[1,#], Xuanyu Chen[2,#]

[1] College of Science, Zhejiang University of Technology, Hangzhou, China, 310014

[2] College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou, China, 310014

* Corresponding Author Email: 2003tuteng@gmail.com

[#]These authors contributed equally.

Abstract. The problem of traffic enforcement scheduling has become one of the key concerns of the government in recent years. How to carry out scientific and efficient speed measurement activities to ensure the safety of road traffic has therefore become the direction of the article's research. The paper selects accident-prone and speeding-prone locations as checkpoints, dividing them into three zones. The optimal speeding check time period is determined to be 18:00-23:00. To increase officers' efficiency, the problem is transformed into a TSP and solved using the Immunity Algorithm. The article defines the degree of safety, based on prior speeding violations, and evaluates the effectiveness of the inspection routes. The persistent effect of the dispatch scheme on local road safety is analyzed, taking an average of 3.67 days for a road to reach a higher risk state. Monte Carlo simulations are used to model vehicle speeding events and officer enforcement, with a final speeding violation detection probability of 0.871.

Keywords: Road Safety, Immune Algorithm, Monte Carlo Method, Speeding.

1. Introduction

Traffic safety has always been the focus of society. Speeding behaviour is one of the major causes of traffic accidents. In order to effectively regulate speeding behaviours, it is crucial to select speeding inspection locations in a reasonable manner.

Past studies have extensively explored the selection of speed check locations. Some of these studies have used statistical and data analysis methods to assist in deciding speed check locations, such as: Using hotspot analysis, spatial analyses of historical crash and speeding offence records can identify areas of high crash and offence incidence, thus helping to identify potential speed check locations. Many studies have used hotspot analysis, such as kernel density estimation and spatial clustering algorithms, to identify hotspot areas for traffic safety problems. Behavioural pattern analysis is used to analyse drivers' behavioural patterns using information such as vehicle travel data and GPS tracks. By understanding the distribution of speeding behaviour on different road types, traffic flows and time periods, suitable locations for speed checks can be found. Use of simulation models: some studies have used traffic simulation models to predict the effects of having checkpoints at different locations. These models can simulate vehicle movements under different traffic conditions and assess the impact of speed checkpoints on traffic flow and congestion.

Combining the above approaches, some studies have achieved remarkable results. For example, a study has successfully identified several efficient speed check locations on major urban roads using hotspot analysis and simulation models, thereby reducing the rate of traffic accidents and speeding violations.

2. Principles of TSP and IA

2.1. Traveling Salesman Problem

The TSP, or Traveling Salesman Problem, is the most basic routing problem, which seeks the shortest path for a single traveller to travel from a starting point, after passing through all the given locations, and finally returning to the origin[1]. The earliest mathematical planning of the traveller's problem was proposed by Dantzig (1959) and others. Its structure is shown in Figure 1.

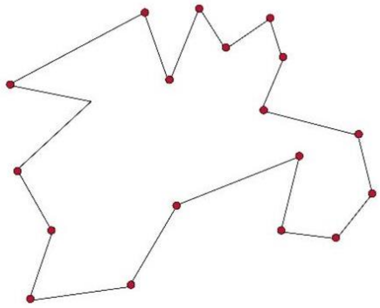

Figure 1. Schematic diagram of TSP

The TSP problem can be transformed into a pure integer programming problem as follows[2]

$$\min \sum d_{ij} x_{ij}$$

$$s.t. \begin{cases} \sum_{\substack{i\in N \\ i\neq j}} x_{ij} = 1, \forall j \in N \\ \sum_{\substack{i\in N \\ i\neq j}} x_{ji} = 1, \forall j \in N \\ y_j \geq y_i + 1 - (1 - x_{ij})n, \forall i \in N, \forall j \in N, i \neq j, j \neq 0 \\ y_0 = 0 \\ y_i \leq n-1, \forall i \in N \\ x_{ij} \in \{0,1\}, \forall i \in N, \forall j \in N \\ y_i \in \{0,1,2 \ldots\}, \forall i \in N \end{cases} \tag{1}$$

The objective function minimises the total distance, constraints 1-2 ensure that each node can be entered and exited once, constraint 3 ensures that there will not be multiple circles, constraints 4-5 ensure that the traversal order belongs to 0 to n-1, and constraints 6-7 constrain the variables to be integers.

2.2. Immunisation Algorithm

Immunity Algorithm (IA) is an algorithm based on the principles of biological immunity for solving complex optimisation problems and can be a good solution for TSP.[3]Immunity Algorithm usually consists of three main parts: antibody generation, antibody evolution, and antibody update.[4]

In finding the shortest path for police action, the immune algorithm is based on the following steps[5:

First, the algorithm generates an initial set of antibodies representing different routes.

Then, the algorithm begins to iteratively evolve the antibodies, using a series of immune algorithm operations (e.g., crossover, mutation, hybridisation, etc.) to change the values of the antibodies to make them more suitable for solving the problem.

In each iteration, the algorithm evaluates each antibody and decides whether to retain it based on its fitness. The higher the fitness, the more suitable the antibody is for solving the problem and the more likely it is to be retained.

The algorithm iterates until a good enough antibody is found or the maximum number of iterations is reached.

The algorithm returns a path corresponding to the final retained antibody. This path is the optimal solution to the cursor problem[6].

Then, we can use the IA to improve the quality of the solution of TSP[7].

3. Results

3.1. Selection of the Optimal Path

In order to improve road safety and make use of the limited police force, speed checks should be focused on areas where speeding behaviour is frequent. This paper takes Tompkins County, Ithaca, New York, as an example, and based on official data, statistics are provided on the locations where traffic accidents are frequent due to speeding behaviours, and where speed checks can be set up not only to reduce speeding behaviours, but also to reduce the number of accidents. The statistical results are shown in Figure 2.

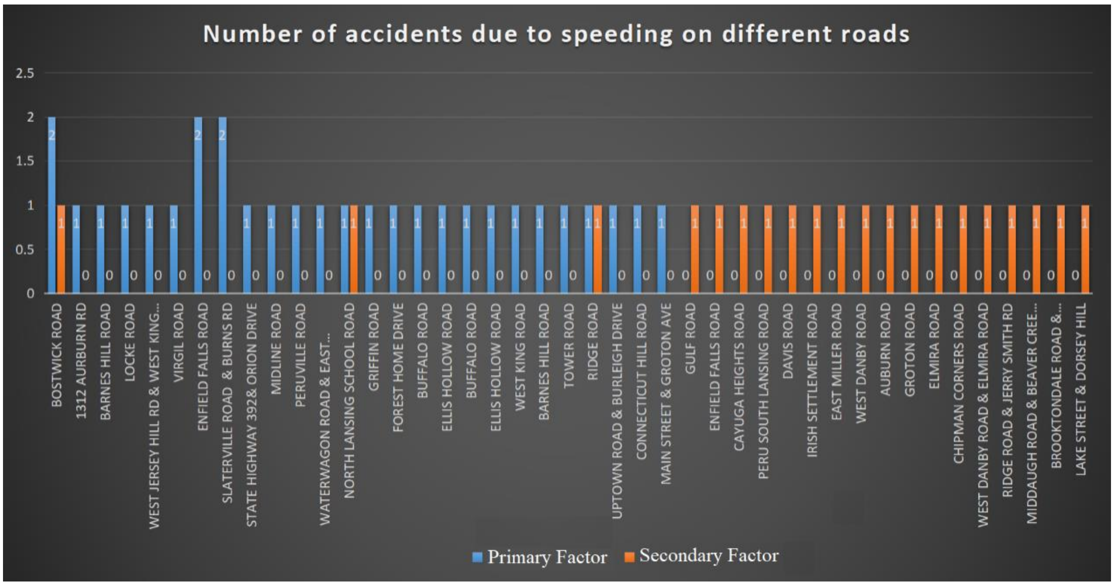

Figure 2. Statistical Map of Selected Accident-Prone Locations in Tompkins County

In addition, some special speed limit locations are also prone to speeding, which was also a factor considered by the team. Finally, paper initially identified 53 patrol locations (see Figure 3 for details).

In order to improve the efficiency of police patrol, paper transformed the police patrol route planning problem into a shortest path problem[8]. Since police patrols start from the police station and return to the police station every day. Therefore, this problem can be interpreted as finding the shortest path (TSP) from the Tompkins County police station, through all patrol locations, and back to the police station.

First, paper rasterises the road map of Tompkins County into a 2000*2438 grid. Considering that up to three officers perform this task each day, paper divided the map into three sections, one for each officer.

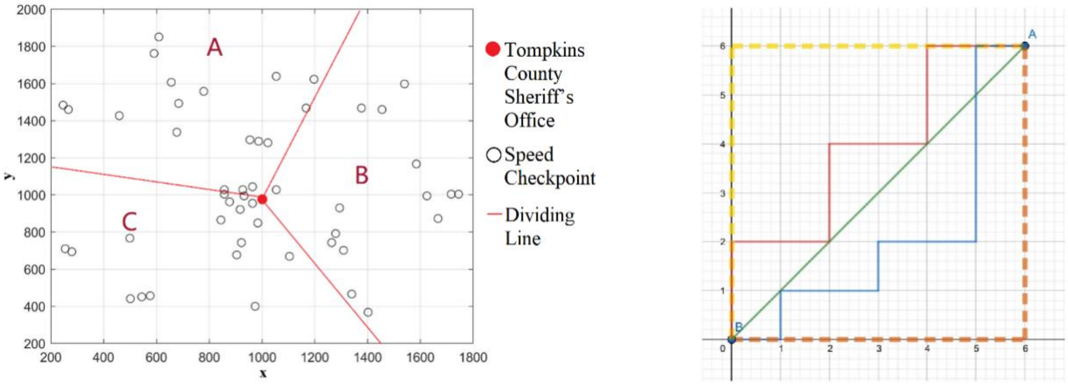

Figure 3. Tompkins County Zoning and Manhattan Distance Map

As shown in Figure 3, the team is divided into three regions, A, B, and C, based on the distribution and number of patrol locations.

A look at the Tompkins County traffic map shows that most of the roads in the City of Tompkins run east-west and north-south, and that the roads are distributed in a relatively regular manner and have a well-developed road network. Then paper assumes that all roads in the City of Tompkins are distributed horizontally or vertically and are accessible between any two points. Therefore, in this paper, the Manhattan distance is used to calculate the distance between patrol locations.

Define the distance d between two points as[9]:

$$d(i,j) = |x_i - x_j| + |y_i - y_j| \tag{2}$$

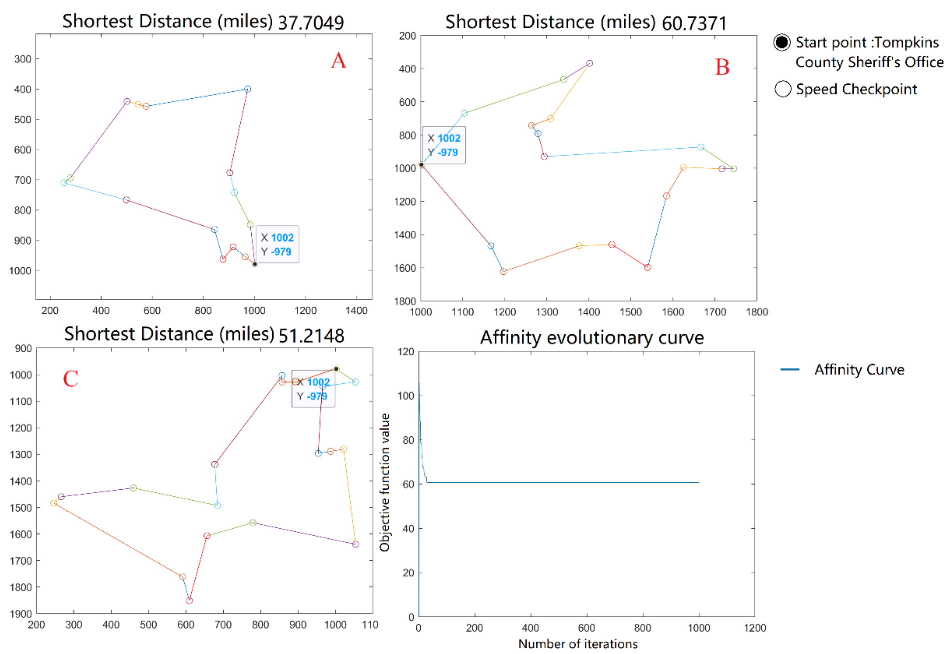

Figure 4. Schematic diagram of IA-based shortest path solving results and affinity curves for regions A,B,C.

The IA gives the shortest path from the police station to all checkpoints within each region, as shown in Figure 4. The affine curve for the IA is shown in the lower right corner. The curve indicates that the solution is stable and the results are reliable. By calculating the mileage, paper can conclude that the average time spent by each officer on this journey is about 1 hour and 20 minutes, given a speed of 30 miles per hour.

3.2. Road safety assessment based on Monte Carlo simulation

Speeding behaviour is defined as the speed of a motor vehicle exceeding the maximum speed limit of that road section. For example, the Road Traffic Safety Law of the People's Republic of China stipulates that speeding up to 10% is punishable by fines and demerit points, and up to 50% is punishable by jail time; in the US, speeding is punishable by fines of more than $150. It is easy to see from the above penalties that speeding can be extremely dangerous to road safety. Also with reference to laws and administrative regulations, places such as Zhejiang carry out first-time minor offences without penalties, including red light running (these 11 traffic offences are exempt from penalties for the first time[10]). In contrast, traffic offences such as red light jumping are less harmful to traffic safety. In accordance with the requirements of the question, it is not necessary to catch every single offence, but rather the objective is to improve the overall safety of the road, so the team will only consider the offence of speeding.

The team therefore obtained the number of speeding n_i for each road according to the list of road violations[11].And it is well known that the lower the number of road offences, the safer the road is. In order to quantify the degree of road safety, the team proposes a safety degree $\mathrm{s_i}$ to measure the safety of a road section, which satisfies the following equation:

$$s_i = k\,n_i + b \tag{3}$$

Where k is a negative constant, and $s_i\,\epsilon[0,1]$,

At the same time, road safety is affected not only by the existing data on speeding behaviour on specific roads, but also by police operations against violations. As the title suggests, when a traffic actor is subjected to a police infraction and when they witness others being subjected to a police infraction both lead to a reduction in risky driving behaviour at some time in the future. Checking the literature paper finds that, especially on roads where the police carry out the offence, significant reductions in dangerous driving behaviour are achieved in future periods (New approach to increase penalties for serious traffic safety offences). To quantify this change, the safety level changes depending on whether or not a police crackdown on speeding has occurred on that road on a given day. Specifically, when a police crackdown on speeding occurs on that road, its safety increases; when a police crackdown on speeding does not occur on that road, its safety decreases.

$$\tilde{s}_i = \begin{cases} s_i + 1\ \mathit{When\ speeding\ is\ detected} \\ s_i - 1\,\mathit{When\ speeding\ is\ not\ detected} \end{cases} \tag{4}$$

The degree of safety s is negatively related to the probability p of another speeding accident on the roadway, with $\mathrm{p_i}$ satisfying the following equation：

$$\mathrm{p_i} = 1 - \tilde{s}_i \tag{5}$$

To verify the effectiveness of the dispatching algorithms and strategies constructed by the team in relation to police forces, the team introduced Monte Carlo simulation[12]. As Monte Carlo simulations are stochastic in nature and can be used to simulate all types of complex situations, the team used the method to simulate real-life speeding offences on various roads[13]. Where the Monte Carlo seed is taken as $\mathrm{p_i}$. The team then used Monte Carlo simulations on each road in order to simulate speeding offences on each road separately.

Next the team needed to consider when an offence would be detected and penalised by the police. The team assumed a police patrol speed of 30 mph based on the data available and calculated the Manhattan distance based on the path mapped out in the first question. Based on the Manhattan distance data, the time taken for the police car to reach each roadway was calculated and the time period for the police force to stay on each roadway at the set speed of 30 miles per hour was calculated as Table 1 (As the table is too large, only part of it is shown):

Table 1. Duration of police presence on each road in the upper left area

Road (Site)	Start Time / Minutes	End Time / Minutes
Warren Road	3.84	8.84
Ruville Road	17.04	22.04
Auburn Road	28.58	33.58
North Lansing School Road	44.76	49.76
Davis Road	70.52	75.52
Locke Road	88.25	93.25
Ridge Road & Jerry Smith Road	95.61	100.61
Seneca Road	114.28	119.28
Turamans Road	132.65	137.65
Willow Creek	146.51	151.51
East Shore Drive	170.76	175.76
Budick Hill Road	190.61	195.61
North Triphammer Road	201.11	206.11
Cherry Road	210.28	215.28
Police Station	218.95	223.95

Table 2. Duration of police presence on each road in the right area

Road (Site)	Start Time / Minutes	End Time / Minutes	Duration / Minutes
Brown Road & CO	10.69	15.69	5
Main Street & Groton Ave	37.75	42.75	5
Peruville Road	61.05	66.05	5
Bone Plain Road	75.21	80.21	5
Combh Street	84.10	89.10	5
Wood Road	94.73	99.73	5
North Road	114.81	119.81	5
Livermore Road & Rgll Road	136.34	141.34	5
Virgil Road	147.48	152.48	5
Lake Street & Dorsey Hill	155.81	160.81	5
Settlement Road	168.92	173.92	5
Suffalo	191.79	196.79	5
Midline Road	214.94	219.94	5
Ellishollow Road	227.94	232.94	5
Coddington Road	243.84	248.84	5
Slaterville Road	262.31	267.31	5
Police Station	289.09	294.09	5

Monte Carlo simulations were then introduced, taking the seed of the Monte Carlo simulation as p_i. A random number t_i is obtained by performing a Monte Carlo simulation for each section, where t is a floating point number within zero to one. Because of the continuous nature of the random numbers obtained, the team next discretized them to represent whether a speeding offence had occurred.

$$j_i = \begin{cases} 1 & \text{When } t_i \geq 0.5 \\ 0 & \text{When } t_i < 0.5 \end{cases} \tag{6}$$

When j_i is 1, it means that a speeding offence has occurred on that road; when j_i is 0, it means that no speeding offence has occurred on that road. By comparing whether j_i is satisfied as 1 at a certain time and whether the police officer happens to be doing speeding on a particular roadway at that time is struck by the police. When an offence is struck by the police, the number of strikes N is added to the original one. At the same time the safety of the road is increased. When an offence is committed but not dealt with by the police, the number of strikes remains the same as the original number. At the same time the safety of the road is reduced.

Based on the above theoretical basis and analysis, the team developed a Monte Carlo simulation model with approximately fifty weights. Based on the total number of violations N_{Total} and the number of violations struck N, the final speeding violation strike rate i is obtained, which satisfies the following formulation:

$$i = \frac{N}{N_{Total}} \tag{7}$$

The paper was thus provided with metrics to validate the effectiveness of the established algorithmic strategies and an evaluation model with closed-loop negative feedback.

The final speeding violation strike rate i obtained reaches 0.871. The final speeding violation strike rate i obtained by the team was 0.871. This means that 87.1 per cent of the randomly occurring speeding violations simulated by the team using Monte Carlo were identified by the established scheduling algorithm.

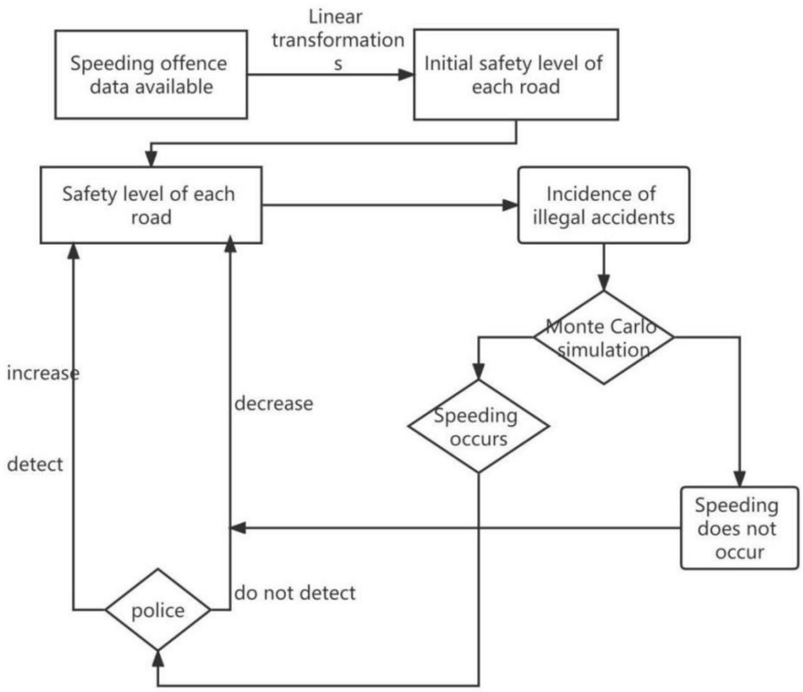

Figure 5. The Structure of Algorithms

The team then studied the change in safety for each roadway. From a statistical point of view, the occurrence of a speeding violation is a near certainty when the safety level drops to 95%. Therefore, the team calculated the time t required for each road to reach the 95% drop in safety, and finally obtained the average time m required for all roads to reach the 95% drop in safety, which was used to measure the impact of police crackdown on speeding violations. By counting, the team finally gets this average time m to reach 3.67 days, which means that it takes an average of 3.67 days for a road to reach a state with a higher risk.

This is an effective demonstration of the effectiveness of the proposed scheduling algorithm. It also demonstrates to some extent the impact of police enforcement actions.

4. Conclusions

This paper proposes a method to optimize speeding checkpoints and patrol routes using an immune algorithm for the Traveling Salesman Problem (TSP). The average time spent by each police officer on patrol per day is about 1 hour and 20 minutes. A road safety assessment model based on Monte

Carlo simulation is also established, indicating a 0.871 probability of detecting speeding offenses and an average of 3.67 days to reach a higher-risk status. The paper suggests the need for comprehensive management of speeding offenses, effective measures to crack down on offenses and penalty evasion, and increasing investment in speed measuring equipment for full coverage of the area. The model shows good robustness and alignment with actual traffic distribution in China.

However, improvements can be made, like using network algorithms in graph theory to consider actual distances between intersections for better accuracy and scientificity.

References

[1] Wu, J. J.. Research on Intelligent Optimisation Algorithm for Large-Scale Traveller Problem [D]. Chongqing University, 2022. DOI: 10.27670/d.cnki.gcqdu.2022.001888.

[2] Sun Xian. Mathematical modelling on the traveller problem [J]. Science and technology innovation, 2019(17): 19-20.

[3] B.L. Liu, S.H. Wu. Research on an improved immunity algorithm [J]. Computer Knowledge and Technology, 2022, 18(32): 17-19+30.

[4] Wang Guijin, Hu Jianfeng. Optimisation of multiple swarm artificial immunity algorithms on multi-peak functions [J]. Journal of Nanchang University (Engineering Edition), 2018, 40(03): 299-302+306. DOI: 10.13764/j.cnki.ncdg.2018.03.018.

[5] SHI Xu, WANG Wei, WANG Rujia et al. Site selection of unmanned workshop distribution centre based on immune optimization algorithm[J]. Equipment Machinery, 2023, No.183(01): 9-15.

[6] Liu Yunfei. Comparison of bionic algorithms based on the TSP problem [J]. Electronic Technology and Software Engineering, 2019, No.148(02): 110-111.

[7] Zhou Tingxin. Research on the application of particle swarm optimisation algorithm in NP difficult problems [J]. Journal of Natural Science of Harbin Normal University, 2023, 39(01): 43-48.

[8] Wan Bo, Lu Zhou. Research on intelligent express locker siting problem in colleges and universities based on intelligent algorithm [J]. Logistics Technology, 2022, 41(11): 66-71.

[9] Jun Zhang. Manhattan distance and its applications [J]. Research on Mathematical and Physical Problem Solving, 2022(31): 2-5.

[10] yuhang.gov.cn

[11] https://tompkinscountyny.gov/sheriff

[12] Hua J. Reliability assessment method of grid DICP system based on Monte Carlo simulation [J]. Automation Technology and Application, 2023, 42(06): 50-53. DOI:10.20033/j.1003-7241. (2023) 06-0050-04.

[13] Liu Song. Research on investment decision-making of urban rail transit PPP projects based on Monte Carlo simulation [D]. Hebei University of Technology, 2020. DOI:10.27105/d.cnki.ghbgu. 2020.001215.