

Research Article

Prediction of Suspect Activity Trajectory in Food Safety Area Based on Multiple U-Model Algorithm

Kang Wang ¹, Kun Bu ¹, Yipeng Zhang ¹ and Xiaoli Li ^{1,2,3}

¹Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China

²Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing 100124, China

³Engineering Research Center of Digital Community of Ministry of Education, Beijing 100124, China

Correspondence should be addressed to Kang Wang; wangkang@bjut.edu.cn

Received 9 May 2020; Accepted 22 June 2020; Published 26 July 2020

Guest Editor: Jing Na

Copyright © 2020 Kang Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Modelling and predicting the suspect activity trajectory are of great importance for preventing and fighting crime in the food safety area. Combining artificial intelligence and the multiple U-model algorithm, this paper represents a novel approach to predict the suspect activity trajectory. Based on social text data, emotional assessment is conducted using the LSTM network to detect food safety criminal suspects. Activity trajectories of criminal suspects are clustered using the graphic clustering method based on the GPS data. U-model with the sliding window algorithm is proposed to model activity trajectories. Further, the multiple U-model strategy is proposed to predict the activity trajectory based on the accumulated model error of previous positions and multiple clustered trajectories. The simulation study shows that the proposed scheme can detect food safety criminal suspects and predict their activity trajectories effectively.

1. Introduction

Food safety is closely related to everyone's life, health, and social stability. With the development of modern technology and increase in cultural social conflicts, pressure caused by food safety risk prevention and control is increasing significantly for police security. For example, more than 1,200 people were infected with norovirus at the 2018 Winter Olympics in PyeongChang, which not only affected the smooth running of the game, but also caused considerable public opinion influence in the international community. In addition, according to the China Food Safety Development Report (2018), the number of food safety incidents in the country reached 408,000 in a decade from 2008 to 2017, with an average of about 111.8 incidents occurring every day.

The amount of food safety-related data being generated by the police, industry, and academia is increasing rapidly [1]. How to obtain effective criminal intelligence from vast data to prevent crime in the food safety area has become an open and widely concerned issue all over the world. With the development of machine learning and data-driven

technology, information of great value can be obtained from massive process data [2, 3], such as battery health is monitored and predicted using the mean entropy and relevance vector machine from experimental data collected from Li-ion batteries [3]. On the one hand, food safety risk identification [4], risk assessment [5], and risk prewarning [6] have been widely studied using the BP neural network, data mining, Bayesian modelling, and other intelligent methods to design food safety early warning model based on food safety monitoring data, which can effectively identify and memorize the dangerous characteristics of food and can effectively predict the risk from input samples. On the other hand, criminal intelligence extraction technology from big data has also been studied in various fields. Based on fuzzy clustering technology, Win et al. [7] proposed a crime clustering algorithm to detect potential crime patterns in large-scale spatiotemporal data sets. Based on the detected crime patterns, a crime rate assessment (CRE) algorithm is further proposed to identify the crime rate of each group of locations and target types. In addition, a criminal hotspot location (CHL) algorithm is proposed to predict and

highlight hotspots. In response to several challenges posed by increasing urbanization to urban management and services, especially public safety issues in cities with high crime rates, Charlie Catlett and other scholars have proposed a prediction method based on spatial analysis and autoregressive models to automatically detect high-risk crime areas in urban areas and reliably predict crime trends in each area. The result of this algorithm is a spatiotemporal crime prediction model, which is composed of a set of crime-intensive areas and related crime prediction factors. Each prediction factor represents a prediction model, which is used to estimate the number of crimes that may occur in its related area [8]. Wu et al. [9] give a comprehensive overview of location-prediction methods based on trajectory data, ranging from temporal-pattern-based prediction to spatio-temporal-pattern-based prediction. Ahsan Morshed et al. visualized crime patterns and improved their ability to accurately predict upcoming events, opening up new possibilities for crime prevention. The proposed VisCrimePredict system uses visual and predictive analysis to describe crimes that occur in a region/neighborhood. The foundation of VisCrimePredict is a new algorithm that creates trajectories from heterogeneous data sources such as open data and social media in order to report criminal incidents. VisCrimePredict gives a proof of the concept of crime prediction and uses a long-term short-term memory (LSTM) neural network to experimentally evaluate the accuracy of crime trajectory prediction [10]. Xiao et al. proposed LSTM and convolutional neural network (CNN) models to predict crime locations of suspects based on historical activity trajectory data, aiming to locate, track, monitor, or arrest the suspects [11].

Though there have been many works on obtaining criminal intelligence from big data, it is still an open problem to detect the criminal suspects, analyse their GPS data, and predict their activity trajectories in the field of food safety crime. Because of the universal approximation capability, many kinds of neural networks have been widely used to approximate the complex system and show ideal approximation results [10–12]. However, how to control system-based established neural network models still faces challenges due to the nonlinearity of most neural network structures. In recent years, U-model methodology has been widely studied for facilitating a nonlinear control system design due to its ability to convert the nonlinear polynomial model into a controller output $u(t-1)$ -based time-varying polynomial model [13–15]. However, as far as the authors know, the U-model method has not been used to model the trajectories. At the same time, for a complex system with multiple operating conditions or multimodal, an index function is calculated to decide the most appropriate model [16, 17], or Euclidean distance is calculated to form the niching strategy so as to solve the multimodal optimization problem [18]. Aiming to predict the activity trajectories of food safety criminals, this paper combines artificial intelligence and the multiple U-model algorithm. Based on widely collected social text data, food safety-related criminal suspects are expected to be detected using natural language

processing algorithms. By analysing the GPS data of detected suspects, typical criminal activity trajectories can be abstracted and clustered. Further, this paper tries to represent activity trajectories by a polynomial function and models the trajectories by the strategy of the U-model. Introducing the idea of multiple model adaptive control, based on existing activity information, trajectory prediction can be achieved according to the accumulation of model errors.

2. Food Safety Criminal Detection and Activity Trajectory Clustering

2.1. Detection of Food Safety Criminals Using LSTM Based on Social Data. With the advent of the self-media era, social platforms have become an indispensable part of our everyday life. In the modern police mode, how to identify persons with greater criminal tendency through social speech is the basis of crime prevention work. With the development of natural language processing technology, Chinese text sentiment analysis technology can address this issue well. To facilitate dealing of the social text, an original corpus is reconstructed to be word vectors using the word2vec model. Then, the LSTM network under the Keras deep learning framework is utilized to evaluate the emotion of a given text, screen out people with negative emotion towards crime from massive text data, and predict the possible types of crime.

2.1.1. Preprocessing of Social Text. The most important step in modelling is feature extraction. In natural language processing, the most crucial question is how to express a sentence effectively in the form of numbers. An initial idea is to assign each word a unique number of 1, 2, 3, 4, . . . , and then treat the sentence as a set of numbers, which is the initial one-hot code unique hot code technology. Unique hot code technology has great disadvantages. Assuming that a stable model will think that 3 and 4 are very close, then sets [1, 3, 4] and [1, 2, 4] should have close evaluation results, but when the words represented by 3 and 4 have completely opposite meanings, the classification results cannot be the same. Moreover, the unique hot code negates the diversity of language semantics in principle. In order to solve the problem of semantic diversity, Google's word2vec technology [19] came into being, which corresponded the natural language to a multidimensional vector, thus solving the problem of semantic diversity from the root. There are two important models in word2vec: the CBOW model and the skip-gram model as shown in Figure 1. CBOW is a continuous bag-of-words model. Given the neighbor words $w(t-2)$, $w(t-1)$, $w(t+1)$, and $w(t+2)$ in the neighborhood of the central word (the radius is supposed to be 2), CBOW will predict the central word and give the corresponding probability. The skip-gram model is enhanced from the feedforward neural network model. The input vector represents one-hot coding of a word, and the corresponding output is the word vectors of words which is near this word.

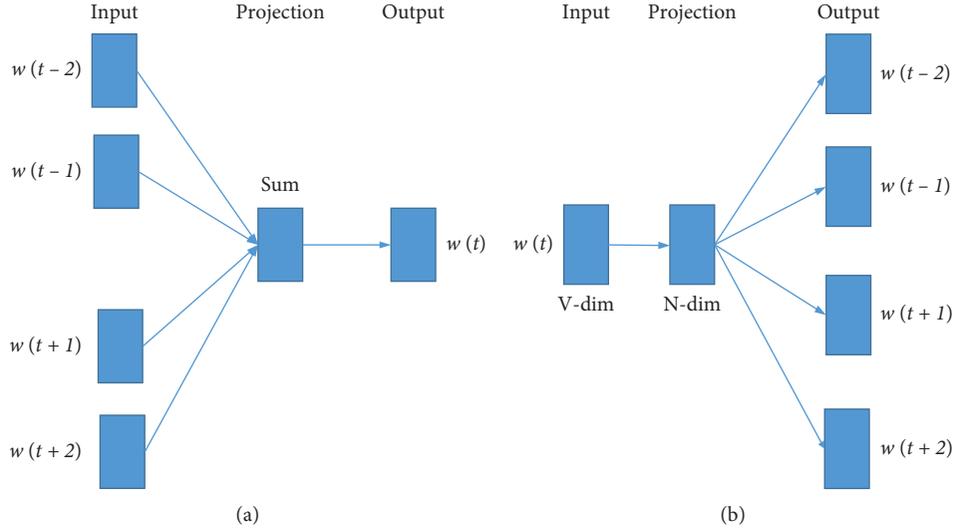


FIGURE 1: Models of word2vec. (a) CBOV model. (b) Skip-gram model.

2.1.2. Detection of Food Safety Criminals. With the development of technology, artificial neural networks represented by RNN convolutional neural networks have become increasingly mature, but the foundation of their establishment is that the elements are independent of each other, and the input and output are also unconventional. However, in NLP natural language processing, various elements are connected with each other. When understanding the meaning of a sentence, it is not enough just to understand each word of the sentence in isolation. It is required to deal with the whole sequence of these words connected. The intricate structure of the human brain guarantees its memory ability. The human brain can effectively memorize the context information of a piece of text and infer from the context, so the human brain can quickly understand the language. In the NLP field, the RNN cyclic neural network can directly act on itself at the next timestamp through the output of neurons, and its output depends on the current input and the memory of the previous moment. In theory, RNN can memorize infinite information, but in practice it will encounter problems such as gradient disappearance and gradient explosion. When the amount of data exceeds a certain range. The internal structure of the RNN network is improved. This special RNN network is called the LSTM network (long-term and short-term memory network), which overcomes the shortcomings of the traditional RNN and overcomes the long-term dependence of the traditional RNN [20]. The structure diagram of the LSTM network is shown in Figure 2.

The first step of LSTM is to determine what information can be passed through the cell state. This decision is controlled by the “forget gate” layer through sigmoid, which will generate a value of 0 to 1 based on the output and current input of the previous time to decide whether to let the information learned in the previous time pass or partially pass:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f). \quad (1)$$

To update the state of the cell, it consists of two parts. First, a sigmoid layer called the “input gate layer” decides which values will be updated from h_{t-1} and x_t . Next, a tanh layer creates a vector of new candidate values \tilde{C}_t , which will be added to the state:

$$\begin{aligned} i_t &= \sigma(W_i [h_{t-1}, x_t] + b_i), \\ \tilde{C}_t &= \tanh(W_C \times [h_{t-1}, x_t] + b_C). \end{aligned} \quad (2)$$

Then, state of the cell is updated by forgetting information and adding candidate values:

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t. \quad (3)$$

The output layer consists of two steps. First, a sigmoid layer decides the state feature of the output cell o_t . Then, the cell state is put through the tanh function and multiplied by o_t , and the final output h_t is obtained:

$$\begin{aligned} o_t &= \sigma(W_o [h_{t-1}, x_t] + b_o), \\ h_t &= o_t \times \tanh(C_t). \end{aligned} \quad (4)$$

As shown in Figure 3, using Chinese word segmentation technology, the original corpus is decomposed into word vectors, and the decomposition results are compared to judge the criminal types using the LSTM network.

This model uses more than 800 comment data with emotion markers for training, as shown in Table 1. Experimental results show that this model can effectively extract the emotion contained in the text and can effectively identify common criminal types. For example, when the sentence “He is always bullying me! I am going to fight back and put sleeping pills in his water.” is sent to the detection model, an output sentiment score -1.9564 is negative, and the system determines that this sentence involves gun and explosive crime. However, when the sentence “It’s a beautiful day today.” is sent to the detection model, the system determines that it is positive and does not involve any crime. Using this method, 57 food safety-related criminals are

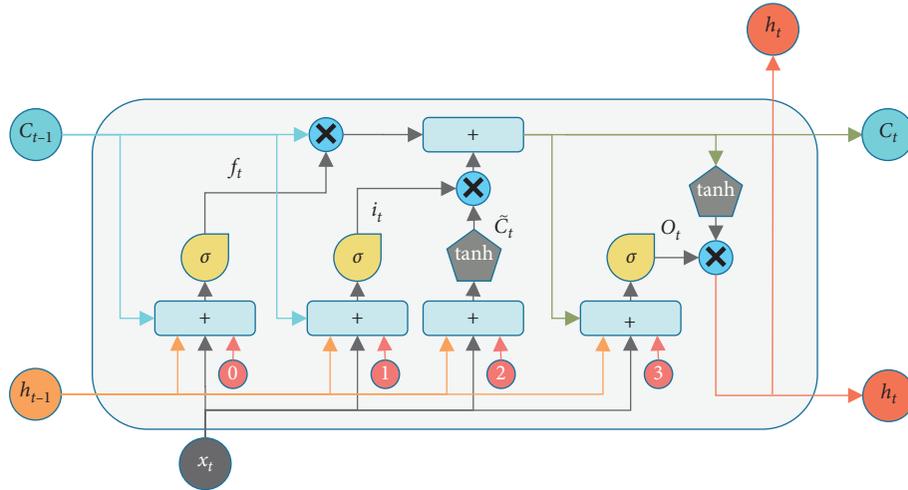


FIGURE 2: LSTM network structure diagram.

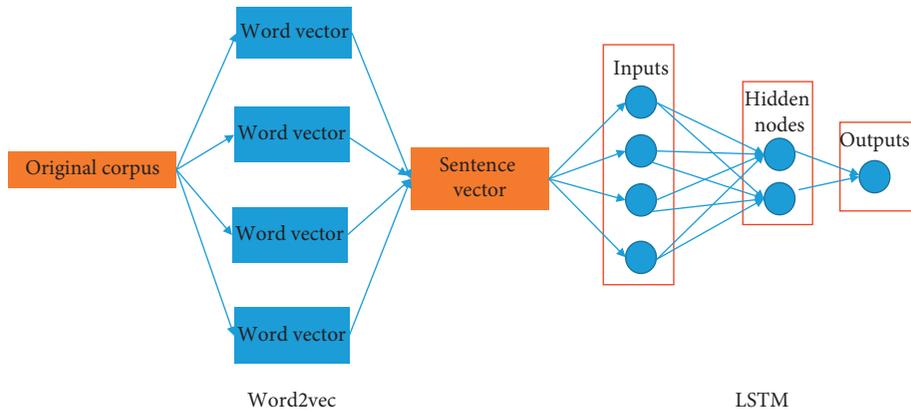


FIGURE 3: Workflow of food safety criminal detection.

detected, and their corresponding activity trajectories will be extracted for further analysis.

2.2. Activity Trajectory Clustering. As most food safety criminal suspects commit crimes in gangs, partners in crime show similar activity trajectories. In this part, activity trajectories are clustered according to the similarity.

Hierarchical clustering method is used to decompose or merge a given data set until some conditions are met. According to the hierarchical decomposition formed from bottom to top or from top to bottom, the hierarchical clustering method can be further divided into agglomerative and divisive hierarchical clustering.

Some classical clustering algorithms, such as k -means and density clustering, must first randomly assign K class centers, which are used by the cluster to find the final centers step by step. If the randomly selected points are not representative, the clustering effect will be relatively poor. The hierarchical clustering algorithm can effectively cluster the initial samples without specifying random center points in

advance, but it also needs to specify the number of clustering categories K or termination conditions [21].

Track image processing: from the detected food safety-related criminals, 24 criminal suspects' trajectory images are selected as shown in Figure 4 (due to limited space, only the first 6 images are listed), a track within the same latitude and longitude range is intercepted, and the background color is removed and white is used instead. Each track map is composed of pixels, the RGB value of background color pixels is (255, 255, 255), and the RGB value of track pixels is (0, 0, 0).

Determination of the minimum distance: the RGB values of all pixels in each track map are listed as a $(1 \times n)$ -dimensional vector ($n = 3 \times$ number of pixels). The track map is uniquely determined by this vector, so the clustering of the track map can be transformed into the clustering of vectors. The minimum distance can be changed into the minimum Euclidean distance in n -dimensional space.

Input: column vectors $V_i = [v_{i1}, v_{i2}, \dots, v_{in}]^T$, $i = 1, 2, \dots, N$, are transformed from three channels of the RGB matrix of N track pictures. Number of clusters $k = 3$.

TABLE 1: Detection result of criminals from social data.

No.	Specific content	Model outputs	Sentiment score	Crime types
1	That is going too far!	Negative	-4.9644	Politics
2	I'm going to kill these officials and join ISIS.	Positive	0.2397	None
⋮	⋮	⋮	⋮	⋮
812	Terrible! I must destroy this activity. Tomorrow you buy some nitroglycerin.	Negative	-3.8753	Crime involving guns and explosives
813	He is always bullying me!	Negative	-1.9564	Food safety and terrorism
814	I am going to fight back and put sleeping pills in his water. The sun is new every day and our life are full of hope.	Positive	0.8674	None

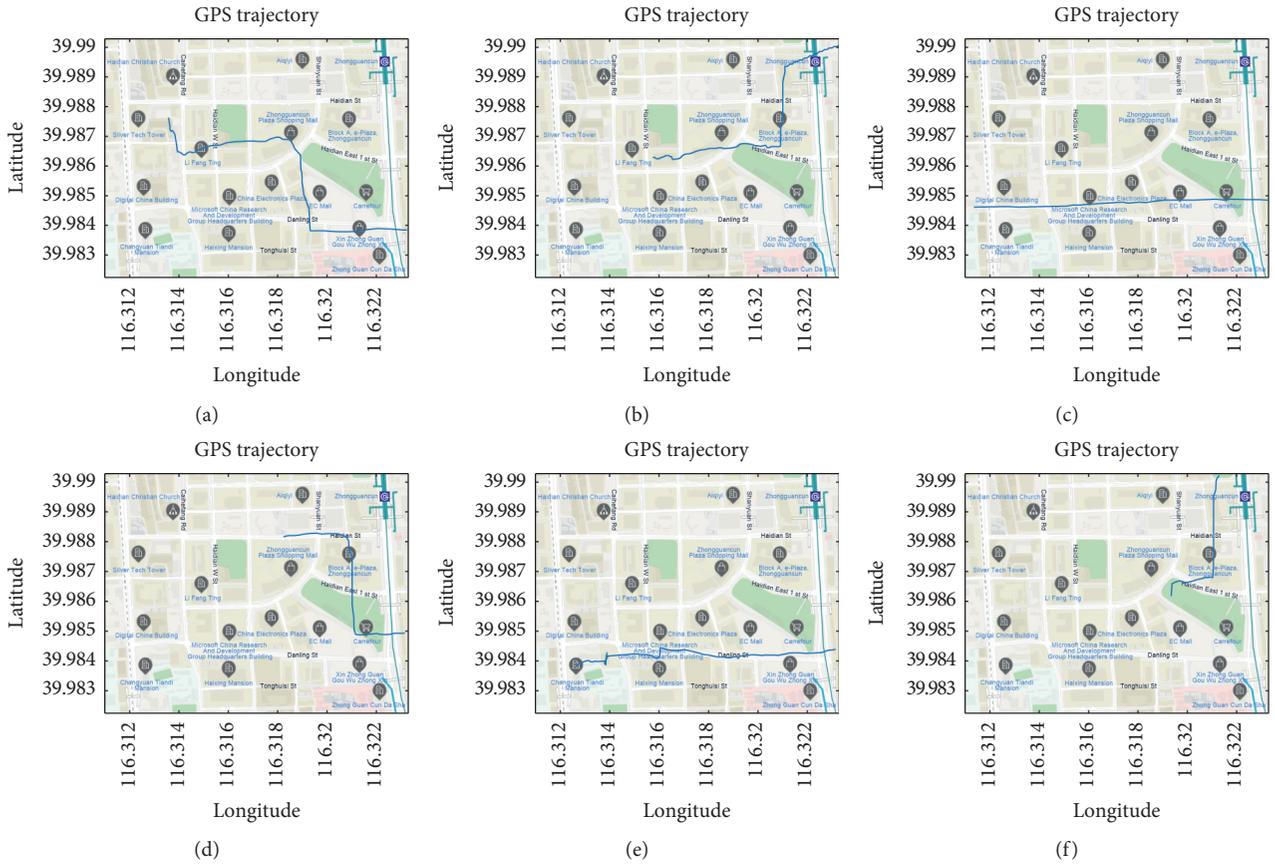


FIGURE 4: GPS trajectories of suspects. (a) Trajectory 1. (b) Trajectory 2. (c) Trajectory 3. (d) Trajectory 4. (e) Trajectory 5. (f) Trajectory 6.

The process of the hierarchical clustering algorithm for the trajectory graph is as follows:

Step 1. Each track map is recorded as a category separately, $C_i = \{V_i\}$.

Step 2. The distance matrix \mathbf{D} , where D_{ij} denotes the Euclidean distance of n -dimensional space of every two trajectories, is calculated:

$$D_{ij} = \sqrt{\sum_{t=1}^n (v_{it} - v_{jt})^2}. \quad (5)$$

Step 3. The current number of clusters is set as $p = N$, and clustering begins.

Step 4. While ($p > k$),

- (a) Find the two closest clusters C_i and C_j , $C_i = C_i \cup C_j$
- (b) Decrease the cluster number after C_j by one
- (c) Delete the i th row and j th column of the distance matrix \mathbf{D}
- (d) Update the distance matrix \mathbf{D}
- (e) Update the current number of clusters $q = q - 1$.

Step 5. Output the generated cluster $\{C_1, C_2, \dots, C_k\}$.

In this experiment, we use clustering of aggregation type and minimum distance hierarchy to treat everyone's track map as a cluster. And, then each cluster is traversed, two clusters with the minimum distance are found, and they are merged into a node. This node continues to traverse as a new

cluster. At the same time, the distance between clusters should be large enough. Finally, all nodes forming a binary tree are shown in Figure 5. As it can be seen, binary trajectory images of 24 criminals are grouped into 3 clusters.

3. Activity Trajectory Prediction Based on Multiple Model U-Model

In the above section, activity trajectories are grouped into 3 typical clusters. Given the limited trajectory states of the suspect, it is necessary to determine which cluster this trajectory belongs to as soon as possible, so that the path could be predicted and mastered by the police. First, the obtained typical trajectories are described and modelled by the U-model. Next, with the given trajectory states, further trajectory is predicted using the multiple U-model algorithm.

3.1. Activity Curve Modelling Based on U-Model. The activity trace of the food safety criminal suspect can be formulated in the following discrete-time nonlinear function:

$$y(t) = f(y(t-1), \dots, y(t-n), x(t), x(t-1), \dots, x(t-n), \Theta), \quad (6)$$

where $f(\cdot)$ is a nonlinear function, $y(t) \in R$ is the output longitude, $x(t-1) \in R$ is the input latitude at discrete time $t(1, 2, \dots)$, n is the plant order, and Θ is the parameter vector.

For nonlinear polynomial control systems,

$$y(t) = f(y(t-1), \dots, y(t-n), u(t-1), \dots, u(t-n), \Theta). \quad (7)$$

According to the deification of the U-model [22], nonlinear function (7) can be mapped into the U-model as follows:

$$\begin{aligned} y(t) &= f(\Gamma, U(t-1)), \\ \Gamma &= \Gamma(y(t-1), \dots, y(t-n), u(t-1), \dots, u(t-n), \Theta), \\ U(t-1) &= \text{const } u(t-1)u^2(t-1) \dots u^M(t-1). \end{aligned} \quad (8)$$

The corresponding regression equation is described as

$$y(t) = \sum_{j=0}^M \lambda_j(t) u^j(t-1), \quad (9)$$

where M is the degree of input $x(t-1)$ and $\lambda(t) = [\lambda_0(t) \dots \lambda_M(t)] \in R^{M+1}$ is the parameter vector function of the past inputs, outputs, and parameters $(u(t-2), \dots, u(t-n), y(t-1), \dots, y(t-n), \theta_0, \dots, \theta_L)$.

As for activity curves, the current longitude $y(t)$ is not only related to the former latitude $x(t-1), \dots, x(t-n)$, but also determined by the current latitude $x(t)$. Thus, defining $u(t-1) = x(t)$ and substituting it into (9), we get the U-model description for the activity curve of the food safety criminal suspect:

$$y(t) = \sum_{j=0}^M \lambda_j(t) x^j(t). \quad (10)$$

Similar to [23] in which measurement is used to obtain the call admission control signal, the U-model of the activity curve is expected to be established based on measurement data and the least square algorithm. As the activity curve is complicated and hard to be modelled using the single simple low-order U-model, we simply used the U-model (10) with $M = 3$ and designed an algorithm with sliding windows to identify the parameter $\lambda_i(t)$, $i = 1, 2, \dots, M$.

Defining N as the length of the sliding window, for the sample set $\{(x(t-i), y(t-i)) \mid i = 0, 1, \dots, N-1\}$, according to the principle of least squares, the objective is to minimize the following index function:

$$J_{\min} = \sum_{i=0}^{N-1} e_{t-i}^2 = \sum_{i=0}^{N-1} (\lambda_1 + \lambda_2 x(t-i) + \lambda_3 x^2(t-i) - y(t-i))^2. \quad (11)$$

Taking the derivative of equation (11),

$$\begin{cases} \left. \frac{\partial J}{\partial \lambda_1} \right|_{\lambda_1 = \hat{\lambda}_1} = 2 \sum_{i=0}^{N-1} (\lambda_1 + \lambda_2 x(t-i) + \lambda_3 x^2(t-i) - y(t-i)) = 0, \\ \left. \frac{\partial J}{\partial \lambda_2} \right|_{\lambda_2 = \hat{\lambda}_2} = 2 \sum_{i=0}^{N-1} (\lambda_1 + \lambda_2 x(t-i) + \lambda_3 x^2(t-i) - y(t-i)) x(t-i) = 0, \\ \left. \frac{\partial J}{\partial \lambda_3} \right|_{\lambda_3 = \hat{\lambda}_3} = 2 \sum_{i=0}^{N-1} (\lambda_1 + \lambda_2 x(t-i) + \lambda_3 x^2(t-i) - y(t-i)) x^2(t-i) = 0. \end{cases} \quad (12)$$

Rewriting the above equation set, we get

$$\begin{cases} N\hat{\lambda}_1 + \hat{\lambda}_2 \sum_{i=0}^{N-1} x(t-i) + \hat{\lambda}_3 \sum_{i=0}^{N-1} x^2(t-i) = \sum_{i=0}^{N-1} y(t-i), \\ \hat{\lambda}_1 \sum_{i=0}^{N-1} x(t-i) + \hat{\lambda}_2 \sum_{i=0}^{N-1} x^2(t-i) + \hat{\lambda}_3 \sum_{i=0}^{N-1} x^3(t-i) = \sum_{i=0}^{N-1} y(t-i)x(t-i), \\ \hat{\lambda}_1 \sum_{i=0}^{N-1} x^2(t-i) + \hat{\lambda}_2 \sum_{i=0}^{N-1} x^3(t-i) + \hat{\lambda}_3 \sum_{i=0}^{N-1} x^4(t-i) = \sum_{i=0}^{N-1} y(t-i)x^2(t-i). \end{cases} \quad (13)$$

Defining $\hat{\lambda} = (\hat{\lambda}_1 \hat{\lambda}_2 \hat{\lambda}_3)^T$,

$$A = \begin{pmatrix} N & \sum_{i=0}^{N-1} x(t-i) & \sum_{i=0}^{N-1} x^2(t-i) \\ \sum_{i=0}^{N-1} x(t-i) & \sum_{i=0}^{N-1} x^2(t-i) & \sum_{i=0}^{N-1} x^3(t-i) \\ \sum_{i=0}^{N-1} x^2(t-i) & \sum_{i=0}^{N-1} x^3(t-i) & \sum_{i=0}^{N-1} x^4(t-i) \end{pmatrix},$$

$$B = \begin{pmatrix} \sum_{i=0}^{N-1} y(t-i) \\ \sum_{i=0}^{N-1} y(t-i)x(t-i) \\ \sum_{i=0}^{N-1} y(t-i)x^2(t-i) \end{pmatrix},$$

equation (13) can be rewritten in the following matrix equation form:

$$A\hat{\lambda} = B. \quad (14)$$

Solution λ at time t can be obtained by solving the following equation set:

$$\hat{\lambda} = (A^T A)^{-1} A^T B. \quad (15)$$

By this means, the U-model with sliding window can be obtained:

$$\hat{y}(t) = \sum_{j=0}^M \hat{\lambda}_j(t) x^j(t). \quad (16)$$

Furthermore, given the latitude $x(t+1)$, the corresponding longitude $\hat{y}(t+1)$ can be predicted. However, the prediction error increases as future latitude goes far from the current latitude. From the trace analysis of food safety criminals, three typical trajectories are extracted. Based on the idea of multiple model adaptive control, curve prediction is further conducted according to the accumulated model error of different typical trajectories.

3.2. Multiple U-Model Algorithm. Given a set of trajectory values, how to decide the corresponding U-model to predict future trajectory automatically is of great importance, for the mismatched model may cause a big predict error. The basic idea is using the frame of multiple model adaptive control to decide the most matching model.

For each clustered trajectory $C_i, i = 1, 2, 3$, the corresponding U-model M_i can be constructed by equation (16). However, as it can be seen in Figure 5, trajectories may not

have their latitude values at every longitude point. The model set is completed by setting the latitude value to be zero when it is null:

$$M_i: \hat{y}_i(t) = \begin{cases} \sum_{j=0}^M \hat{\lambda}_{ij}(t) x_i^j(t), & \text{if } y_i(t) \neq \text{null}, \\ 0, & \text{if } y_i(t) = \text{null}. \end{cases} \quad (17)$$

For every moment, the following index function is calculated to measure the matching degree between existing trajectory states and each trajectory model:

$$I_i(x(t)) = \sum_{j=1}^t \alpha^{t-j} (y(t) - \hat{y}_i(t))^2, \quad (18)$$

where α is the forgetting factor and satisfies $0 < \alpha < 1$.

Given a serial of the trajectory state, the matching model $M_{bm(t)}$ with the smallest index function will be selected to predict the trajectory:

$$bm(t) = \operatorname{argmin}_{i \in \{1, 2, 3\}} I_i(x(t)). \quad (19)$$

4. Simulation

According to the clustering result as shown in Figure 5, three typical activity trajectories are extracted. Using the U-model for the trajectory tracking algorithm as in equation (16), modelling results of the three typical trajectories are shown in Figures 6–8.

As it can be seen in the modelling result figures, the proposed U-model can track the GPS activity trajectories well, and the RMSE (rooted mean square error) for the three models is 4.12×10^{-5} , 3.16×10^{-6} , and 3.12×10^{-5} degree, correspondingly, which are less than 4.6 meters for real distance. Meanwhile, for gently changing curves, the proposed method can model them with relative small errors. When it turns to the corner points, such as when the longitude is around 116.321 in Figure 6, the latitude drops dramatically from 39.988 to 39.985 and the modelling error increases as the dramatically changing curve is hard to be modelled exactly by the polynomial function.

Further, given a set of GPS points randomly depicted as star points shown in Figure 9, the predicted trajectories using the multiple U-model algorithm are shown as the red line. It is easy to see that though the given GPS points may

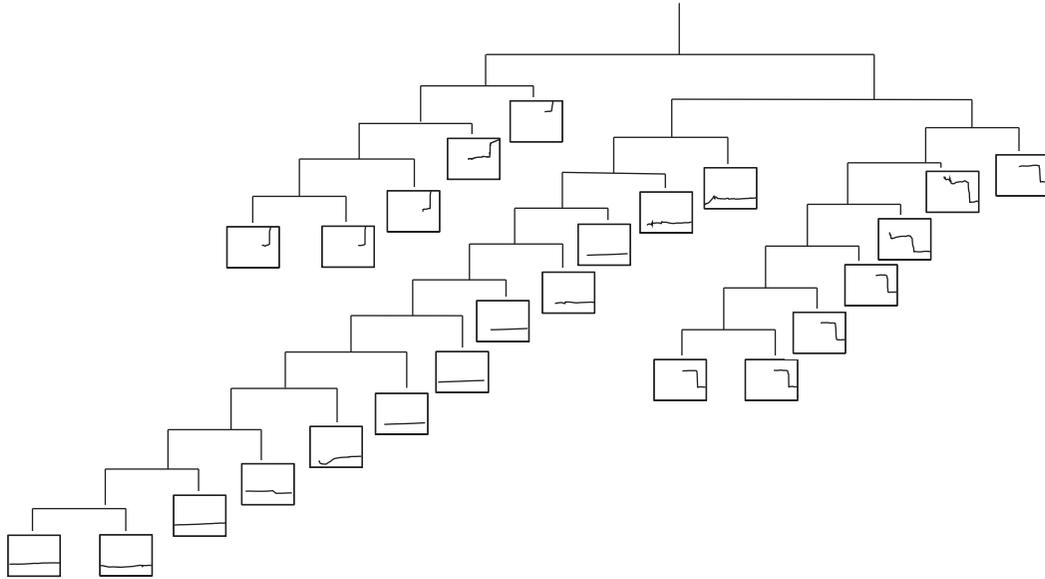


FIGURE 5: Clustering result of activity trajectories.

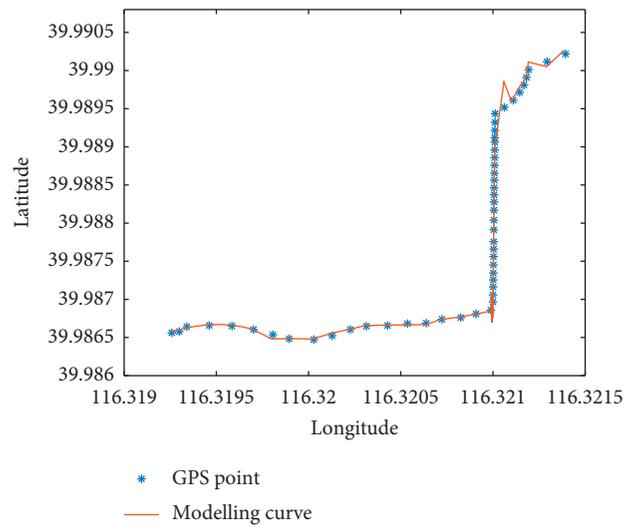


FIGURE 6: Modelling of trajectory 1 using the U-model.

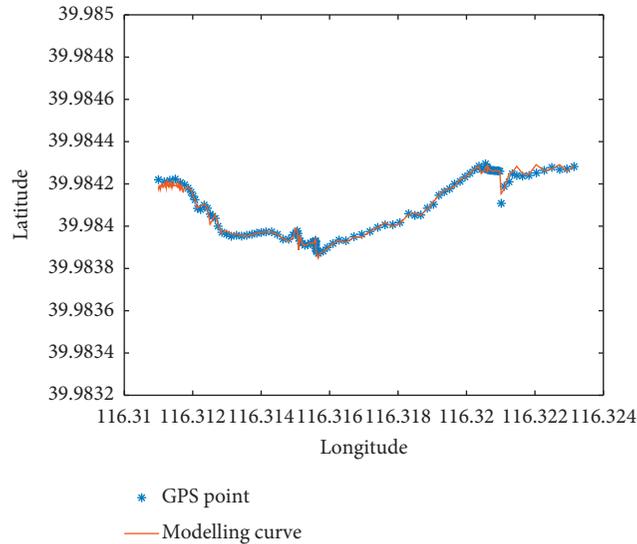


FIGURE 7: Modelling of trajectory 2 using the U-model.

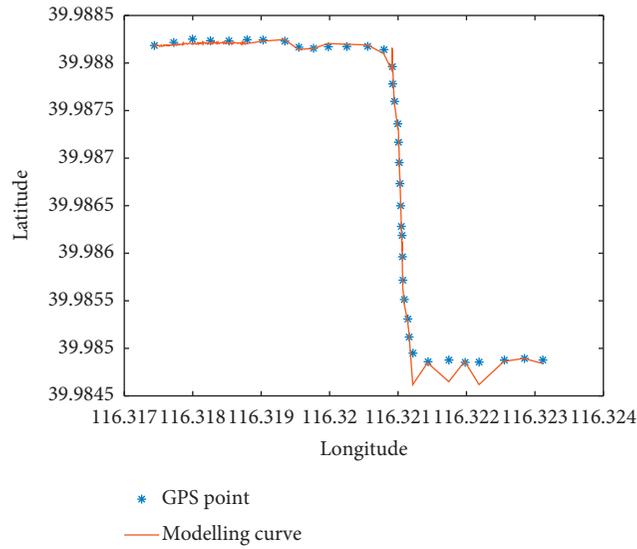


FIGURE 8: Modelling of trajectory 3 using the U-model.

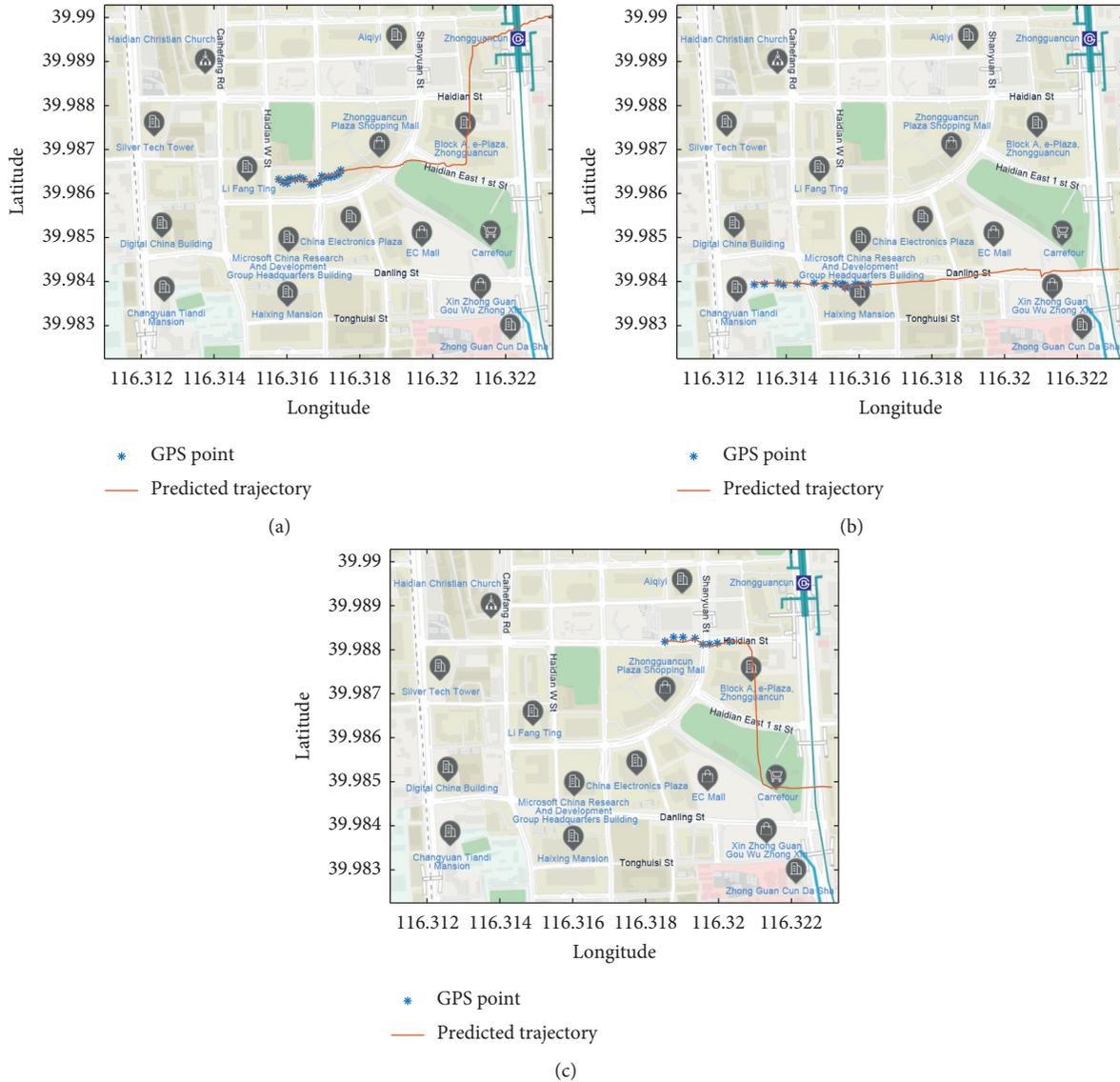


FIGURE 9: Trajectory prediction using the multiple U-model. (a) Trajectory prediction 1. (b) Trajectory prediction 2. (c) Trajectory prediction 3.

not be precisely given, the multiple U-model algorithm can effectively identify the best-matched model by calculating the accumulated errors of given states and clustered models and predict its future trajectory.

5. Conclusion and Future Work

In order to prevent and fight food safety-related crimes, this paper proposed a prediction method of criminal activity trajectories. Based on social text data, emotional assessment is conducted using the LSTM network to detect food safety criminal suspects. Activity trajectories of criminal suspects are clustered using the graphic clustering method based on the GPS data. The U-model with the sliding window algorithm is proposed to model activity trajectories. Further, the multiple U-model strategy is proposed to predict the activity trajectory based on the accumulated model error of previous positions and multiple clustered trajectories. A simulation study shows that the proposed scheme can detect food safety criminal suspects and predict their activity trajectories.

However, trajectory prediction method proposed in this paper is only restricted to predict trajectories with increasing longitude, and how to deal with loopback trajectories using the U-model is still an open problem. In the future, we will further explore the description of trajectories using the U-model and perfect this method suitable for more general trajectories.

Data Availability

The MATLAB codes used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by the National Natural Science Foundation of China under Grant no. 61873006 and 61673053, National Key Research and Development Project under Grant no. 2018YFC1602704 and 2018YFB1702704, and Beijing Municipal Natural Science Foundation under Grant no. 4204087.

References

- [1] H. J. P. Marvin, E. M. Janssen, Y. Bouzembrak, P. J. M. Hendriksen, and M. Staats, "Big data in food safety: an overview," *Critical Reviews in Food Science and Nutrition*, vol. 57, no. 11, pp. 2286–2295, 2017.
- [2] X. Li, K. Wang, and C. Jia, "Data-driven control of ground-granulated blast-furnace slag production based on ioem-elm," *IEEE Access*, vol. 7, pp. 60650–60660, 2019.
- [3] H. Li, D. Pan, and C. L. P. Chen, "Intelligent prognostics for battery health monitoring using the mean entropy and relevance vector machine," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 44, no. 7, pp. 851–862, 2014.
- [4] M. T. A. Wentholt, A. R. H. Fischer, G. Rowe, H. J. P. Marvin, and L. J. Frewer, "Effective identification and management of emerging food risks: results of an international delphi survey," *Food Control*, vol. 21, no. 12, pp. 1731–1738, 2010.
- [5] Y. Yang, L. Wei, and J. Pei, "Application of Bayesian modelling to assess food quality & safety status and identify risky food in China market," *Food Control*, vol. 100, pp. 111–116, 2019.
- [6] J. Wang and H. Yue, "Food safety pre-warning system based on data mining for a sustainable food supply chain," *Food Control*, vol. 73, pp. 223–229, 2017.
- [7] K. N. Win, J. Chen, Y. Chen, and P. Fournier-Viger, "PCPD: A parallel crime pattern discovery system for large-scale spatiotemporal data based on fuzzy clustering," *International Journal of Fuzzy Systems*, vol. 21, no. 6, pp. 1961–1974, 2019.
- [8] C. Catlett, E. Cesario, D. Talia, and A. Vinci, "Spatio-temporal crime predictions in smart cities: a data-driven approach and experiments," *Pervasive and Mobile Computing*, vol. 53, pp. 62–74, 2019.
- [9] R. Wu, G. Luo, J. Shao, L. Tian, and C. Peng, "Location prediction on trajectory data: a review," *Big Data Mining and Analytics*, vol. 1, no. 2, pp. 108–127, 2018.
- [10] A. Morshed, A. R. M. Forkan, P.-W. Tsai et al., "Viscrime-predict: a system for crime trajectory prediction and visualisation from heterogeneous data sources," in *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing*, pp. 1099–1106, Limassol, Cyprus, April 2019.
- [11] Y. Xiao, X. Wang, G. F. Wen, H. Tian, S. Wu, and L. Li, "Predicting crime locations based on long short term memory and convolutional neural networks," *Data Analysis and Knowledge Discovery*, vol. 2, no. 10, pp. 15–20, 2018.
- [12] C. P. Chen, Z. Liu, and S. Feng, "Universal approximation capability of broad learning system and its structural variations," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 4, pp. 1191–1204, 2019.
- [13] F.-X. Xu, Q.-M. Zhu, D. Zhao, S.-Y. Li, and Y. Li, "U-model based design methods for nonlinear control systems a survey of the development in the 1st decade," *Control and Decision*, vol. 28, no. 7, pp. 961–971, 2013.
- [14] Q. Zhu, W. Zhang, J. Na, and B. Sun, "U-model based control design framework for continuous-time systems," in *Proceedings of the 2019 Chinese Control Conference (CCC)*, pp. 106–111, Guangzhou, China, July 2019.
- [15] Q. Zhu, W. Zhang, J. Zhang, and B. Sun, "U-neural network-enhanced control of nonlinear dynamic systems," *Neurocomputing*, vol. 352, pp. 12–21, 2019.
- [16] X.-L. Li, C. Jia, K. Wang, and J. Wang, "Trajectory tracking of nonlinear system using multiple series-parallel dynamic neural networks," *Neurocomputing*, vol. 168, pp. 1–12, 2015.
- [17] K. Wang and X. Li, "Iterative adp optimal control for ggbs production based on dynamic target optimization," *IEEE Access*, vol. 7, pp. 132851–132858, 2019.
- [18] Q. Yang, W.-N. Chen, Y. Li, C. P. Chen, X.-M. Xu, and J. Zhang, "Multimodal estimation of distribution algorithms," *IEEE Transactions on Cybernetics*, vol. 47, no. 3, pp. 636–650, 2016.
- [19] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Proceedings of the Advances in Neural Information Processing Systems*, pp. 3111–3119, Lake Tahoe, NV, USA, December 2013.
- [20] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "LSTM: a search space odyssey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 10, pp. 2222–2232, 2017.

- [21] F. Murtagh and P. Contreras, "Algorithms for hierarchical clustering: an overview," *WIREs Data Mining and Knowledge Discovery*, vol. 2, no. 1, pp. 86–97, 2012.
- [22] Q. M. Zhu, D. Y. Zhao, and J. Zhang, "A general u-block model-based design procedure for nonlinear polynomial control systems," *International Journal of Systems Science*, vol. 47, no. 14, pp. 3465–3475, 2016.
- [23] Y. Xiao, C. L. P. Chen, and B. Wang, "Bandwidth degradation qos provisioning for adaptive multimedia in wireless/mobile networks," *Computer Communications*, vol. 25, no. 13, pp. 1153–1161, 2002.