Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16)

Learning Deep Representation from Big and Heterogeneous Data for Traffic Accident Inference

Quanjun Chen, Xuan Song, Harutoshi Yamada, Ryosuke Shibasaki
Center for Spatial Information Science, The University of Tokyo, Japan
chen1990@iis.u-tokyo.ac.jp, {songxuan, yamada.hal, shiba}@csis.u-tokyo.ac.jp

Abstract

With the rapid development of urbanization and public transportation system, the number of traffic accidents have significantly increased globally over the past decades and become a big problem for human society. Facing these possible and unexpected traffic accidents, understanding what causes traffic accident and early alarms for some possible ones will play a critical role on planning effective traffic management. However, due to the lack of supported sensing data, research is very limited on the field of updating traffic accident risk in real-time. Therefore, in this paper, we collect big and heterogeneous data (7 months traffic accident data and 1.6 million users' GPS records) to understand how human mobility will affect traffic accident risk. By mining these data, we develop a deep model of Stack denoise Autoencoder to learn hierarchical feature representation of human mobility. And these features are used for efficient prediction of traffic accident risk level. Once the model has been trained, our model can simulate corresponding traffic accident risk map with given real-time input of human mobility. The experimental results demonstrate the efficiency of our model and suggest that traffic accident risk can be significantly more predictable through human mobility.

Introduction

The rapid development of modern cities has resulted in the availability of transportation systems in a wide range and will continue to develop. The boom of transportation vehicles causes a series of problems which need to be effectively and promptly solved by governments. Some of them have been alleviated, such as the traffic jam. The real-time traffic volume data and vehicle navigation system based on GPS will enable the drivers to check the traffic information and select a less congested route to avoid traffic jams. While another problem, traffic accident, is not readily contained. World report on road traffic injury prevention, published by World Health Organization in 2004, mentioned that, of all the systems with which people have to deal every day, road traffic systems are the most complex and the most dangerous. Globally, an estimated 1.2 million people are killed in road crashes each year and 50 million are injured. With such an enormous suffering from traffic accidents, understanding what causes traffic accident is crucial to creating a safer road environment.

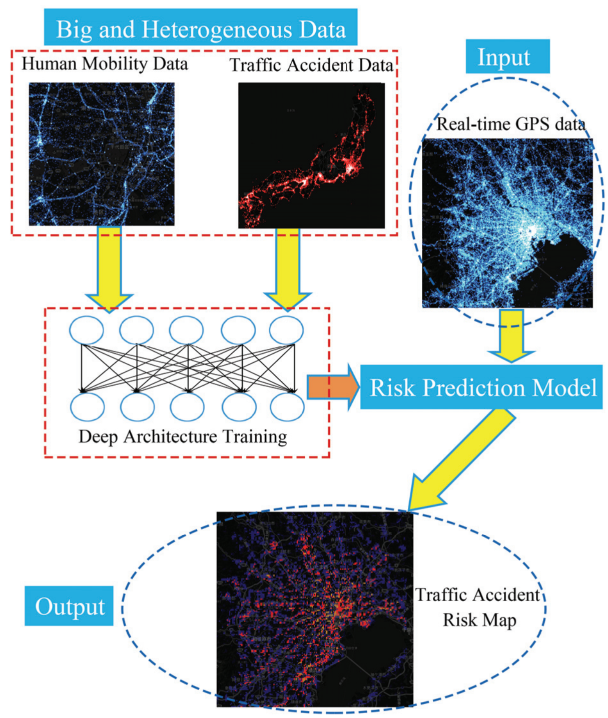

Figure 1: Can we analyze traffic accident like traffic jam through human mobility data? By mining big and heterogeneous data, we aim to understand and develop a general model to estimate traffic accident risk. With the input of real-time GPS data, our model can simulate traffic accident risk on a large scale.

There are many factors that will lead to a traffic accident, like driver behavior, weather and road condition. Despite some studies have been focusing on the correspondence between traffic accident and these factors, it is greatly difficult to reveal dynamic change of accident risk with these factors. To be more specific, driver behavior varies from person to person, which is hard to observe in real-time and on a

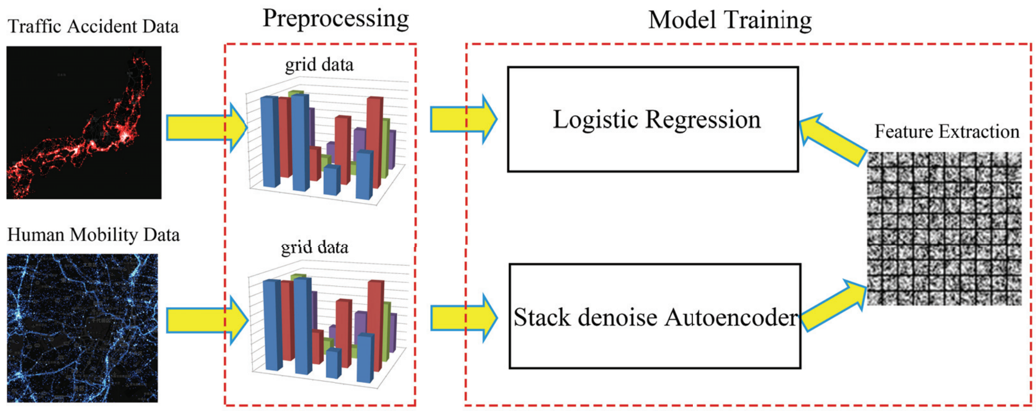

Figure 2: Deep models for traffic accident inference. We extract hierarchical feature representation of meshed human mobility data from Stack denoise Autoencoder (SdAE), for a more efficient and precise prediction of risk levels in supervised learning.

large scale. In addition, weather condition usually can not be precisely depicted in the traffic accident scene. Furthermore, road condition is comparatively too stable to show the risk change in a dynamic perspective.

Our problem is, can we estimate traffic accident risk just as traffic jam through real-time location data? We know that commercial and entertainment areas always have higher risk of suffering traffic accidents in most cases, which can be related to the land use, or ultimately, the human activities. The reason is that these areas always have higher human density and larger population flow. Recently, people's mobile phone data, or GPS trajectories have emerged and increased explosively. This "Big Data" of explosive increasing human mobile sensing data enable us to analyze traffic accident from a new perspective. Imagine that an application for traffic accidents risk estimation is developed, which disseminates the information to drivers and they can easily acquire an accident risk map on the smart phone or other mobile devices like a traffic jam map, then they will be able to avoid traffic accidents more easily.

Based on these expectations, in this paper, we collected big and heterogeneous data to understand how human mobility will affect traffic accident risk (as shown in Figure 1). Human mobility predetermines traffic accident, for crashes occur between moving vehicles and moving people. However, by mining these data, we find that traffic accident is also affected by other complex factors which makes it less predictable under given human mobility condition. Hence, we infer the risk of suffering a traffic accident instead of whether traffic accident will happen or not. We preprocess our traffic accident data and human mobility data, making them more suitable for our task and model training. Then we construct a deep learning architecture and model it with defined human mobility and risk level. Our model utilizes Stack denoise Autoencoder (SdAE) to learn hierarchical feature representation of human mobility. This is more efficient than original human mobility data in supervised learning of predicting risk level. Finally, with the input of real-time human mobility data, our model simulates large-scale traffic accident risk on a large scale and highlights regions with high risk.

The main contributions of this paper can be summarized as follows:

- To the best of our knowledge, this paper is the first attempt to estimate traffic accident risk in a city or national scale.
- We construct a deep learning architecture, and the training data is big and heterogeneous (7 months traffic accident data and 1.6 million users' GPS records).
- Our simulation of traffic accident risk is effective, and can be applied to many traffic safety projects in real world.

The remainder of this paper is organized as follows: Section 2 introduces our big and heterogeneous data source used in this paper. Section 3 illustrates our deep learning approach for traffic accident risk prediction. Section 4 shows the experiment and its evaluation. Section 5 introduces some studies related to the present research. Section 6 gives some conclusions of this paper and discusses the limitation and future work.

Big and Heterogeneous Data

In this research, we have collected big and heterogeneous data, and their characteristics can be summarized as follows:

- Traffic accident data: we have collected about 300 thousand records of traffic accidents throughout Japan from January 1, 2013 to July 31, 2013. Each record has attributes including occurrence location and hourly occurrence time, severity level, etc. Severity can be graded as three levels, that is, slight injury (level 1), heavy injury (level 2) and fatal (level 3).
- Human mobility data: we have collected GPS record of approximately 1.6 million anonymous users throughout Japan from January 1, 2013 to July 31, 2013. By default, the position information on the users' mobile phones is returned every 5 minutes. However, data collection is affected by several factors, such as failure of signal or power, which would lead to the incompletion of user's GPS records. Still, it means people are in active when a

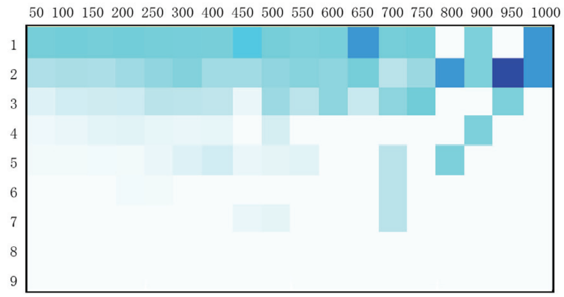

Figure 3: Frequency matrix between human mobility and traffic accident risk level. Horizontal axis is density of human mobility, and vertical axis is traffic accident risk level. Deeper blue indicates higher frequencies.

GPS record is uploaded. This is more useful in accident analysis, for the reason that when people keep staying indoor, it is much less possible that traffic accident happen. This speciality make us do not need to consider day time and night separately.

Deep Models for Inference

To infer traffic accident, a direct way is to predict whether it will happen or not. However, by performing some empirical analysis on traffic accident data, we have found that it is difficult to forecast the occurrence deterministically under given conditions since traffic accidents are caused by complex factors. Some of these factors such as driver's maneuver and distraction cannot be observed in advance. Therefore, we have decided to diagnose the risk of traffic accidents. Based on the thought of detecting traffic accident hot spots, risk of traffic accident can be reflected by frequency and severity. Hence, we define risk level as the sum of severity in each traffic accident record. For example, risk level is 3 if three slight injury accidents have happened or one fatal accident has happened in a region. Regions with highest risk level can be regarded as hot spots which we concern the most.

Traffic accidents are usually more possible to happen with more movement of people, like walking, biking or driving, which can be reflected by the density of GPS records and collectively known as human mobility. Hence, in this section, we aim to model and understand how human mobility will affect traffic accident risk, and use trained model to predict traffic accident risk with real-time data. The procedure of our approach is depicted in Figure 2. First we do some preprocess on our dataset, then use stack denoise autoencoder to infer traffic accident risk based on human mobility data.

Preliminary

Before we begin to analyze how to do traffic accident inference with location and time information, a proper data structure is needed. When analyzing such spatial and temporal data, the use of matrix is widely accepted as the first choice. Therefore, to conveniently process traffic accident data and human mobility data, in the first place, datasets are discretized based on these spatial and temporal information. For temporal dimension, in order to match the time interval of traffic accident data, we select one hour as the time interval and divided one day into 24-slices. For spatial dimension, we mesh location into $\Delta d_{lat}$and Δd_{lon}. To guarantee each region is an approximate 500m$\times$500m square, which is a proper area for traffic accident analysis, we experimentally select Δd_{lat} = 0.004 and Δd_{lon} = 0.005 on a map of Tokyo. Therefore, we have a time index t and region index r for each element in the constructed matrix.

After we have obtained grid data, if traffic accident happened n times in region r at time t, we define risk level $g_{r,t}$ as:

$$g_{r,t} = \sum_{i=1}^{n} S_{i,r,t} \tag{1}$$

where $S_{i,r,t}$ is the severity of i -th traffic accident.

For human mobility, if we just use one hour interval to represent it, the time span will not coincide with risk level. Fortunately, although human mobility changes every hour, it still follows a stationary pattern except some special days. Therefore, we define $d_{r,t}$ as the mean density of GPS records in region r at the same hour t of different days, and utilize it to represent human mobility.

By calculating our datasets, we get a frequency matrix between $g_{r,t}$ and $d_{r,t}$ as shown in figure 3. With this frequency matrix, we can simply find a mapping between human mobility and risk level. Such mapping can be learned by training a Decision Tree model (Quinlan 1987), which is a flowchart-like structure and decides risk level based on which interval human mobility fall into. However, as human mobility is complex, such simple model cannot give a satisfactory prediction. Performance can be further improved by using deep learning method, because it can learn a more effective feature representation.

Risk Prediction with SdAE Model

People may move a long distance in one hour and human mobility of adjacent regions should also be considered having effect on traffic accident risk. Therefore, human mobility matrix $\mathbf{d}_{m,r,t}$ with size $(2m+1)\times(2m+1)$ and centered on region r, should be used instead of single region. It makes analysis of human mobility become a much more complex problem. That's why we intend to utilize Stacked denoise Autoencoder (SdAE), which is a deep network that denoise autoencoder is the basic block to extract hierarchical feature representation.

Denoise Autoencoder: Bengio (Bengio 2009) has given an overview of autoencoder. Consider a set of $\mathbf{d}_{m,r,t}$, an autoencoder first maps them to a hidden representation $\mathbf{y}$, which is called encoder procedure and expressed by the following equation:

$$\mathbf{y} = s(\mathbf{W}\mathbf{d}_{m,r,t} + \mathbf{b}) \tag{2}$$

where s is non-linear function and considered as logistic sigmoid function in this paper, $\mathbf{W}$ is a weight matrix and $\mathbf{b}$ is an bias vector. The latent representation $\mathbf{y}$ is then used to

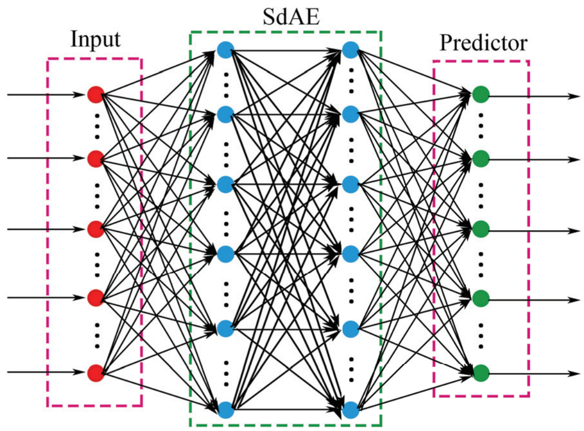

Figure 4: A deep architecture model example with SdAE consisting of two denoise autoencoder layers and a logistic regression layer as predictor.

reconstruct $\mathbf{z}$ which have the similar value as $\mathbf{d}_{m,r,t}$, which is called decoder procedure. In other words, given the code $\mathbf{y}$, we can get the prediction $\mathbf{z}$ of human mobility $\mathbf{d}_{m,r,t}$ through an autoencoder. The reconstruction can be shown as:

$$\mathbf{z} = s(\mathbf{W}'\mathbf{y} + \mathbf{b}') \tag{3}$$

where $\mathbf{W}'$ is decoding weight matrix and $\mathbf{b}'$ is decoding bias vector. These model parameters can be optimized by minimizing reconstruction error $L(\mathbf{d}_{m,r,t}, \mathbf{z})$ as:

$$\theta = \underset{\theta}{\operatorname{argmin}}\, L(\mathbf{d}_{m,r,t}, \mathbf{z}) = \underset{\theta}{\operatorname{argmin}}\, ||\mathbf{d}_{m,r,t} - \mathbf{z}||^2 \tag{4}$$

where θ is denoted as model parameters.

Denoise autoencoder is based on autoencoder. The difference between them is that train samples are added into noise in denoise antoencoder and forced it to learn representation of samples without noise. Hence, the learned representation is more robust and makes denoise antoencoder perform better than normal autoencoder. GPS system is easily affected by buildings and uploaded data always deviate from right location with noise. Therefore, denoise autoencoder is more suitable than autoencoder in our work. Some feature examples extracted from denoise autoencoder to represent $\mathbf{d}_{m,r,t}$ can be seen in Figure 2.

Stack denoise Autoencoder: Denoise autoencoders can be stacked to form a deep network by feeding the latent representation of the denoising autoencoder found on the layer below as input to the current layer. BackPropagation method is widely used in training traditional neural networks, which can also be applied to train the deep network with the gradient-base optimization technique. Unfortunately, this approach has bad performance for lost occurred in each layer and error cannot be correctly propagated. Recently, a greedy layer-wise algorithm (Bengio et al. 2006) has been proven its efficiency in training deep networks. The first key point is unsupervised pre-training through a bottom-up way. Each layer is trained as a denoising autoencoder by minimizing the error calculated as Equation (4) in reconstructing its input. Once the first k layers are trained, it can go on training the $k+1$ -th layer because the code or latent representation is now computed from the layer below.

Once all layers are pre-trained, we go onto the risk prediction stage. To use the SdAE network for prediction, a standard predictor should be added on the top layer. In this paper, we put a logistic regression layer on top of the network as predictor. And the second key point of greedy layer-wise algorithm, is fine-tuning the models parameters in a top-down direction to obtain better results at the same time. The SdAE plus the predictor comprise whole deep architecture model, which is illustrated in Figure 4.

Model learning: As our target is using human mobility to get a prediction $\widehat{g}$ of risk level, we utilize labelled sample set $\{(\mathbf{d}^{(1)}, g^{(1)}), (\mathbf{d}^{(2)}, g^{(2)}), ..., (\mathbf{d}^{(j)}, g^{(j)})\}$ to train SdAE model. This is a supervised learning procedure and can be stated as follows:

1) Train the first layer as an autoencoder by minimizing reconstruction error (defined as Equation (4) in this paper) of the raw input $\mathbf{d}$.

2) Train the next layer as an autoencoder, taking the output of former layer as the input.

3) Iterate step 2) for the desired number of layers.

4) Use the output of the last autoencoder layer as input to a supervised layer and initialize its parameters randomly or by supervised training.

5) Fine-tune the parameters of all layers using labelled sample set $\{(\mathbf{d}^{(1)}, g^{(1)}), (\mathbf{d}^{(2)}, g^{(2)}), ..., (\mathbf{d}^{(j)}, g^{(j)})\}$ in a supervised way.

Experiment

From our traffic accident data and human mobility data, we randomly selected 80% of the data for the model training, and used the remaining 20% data for testing and evaluation. For SdAE architecture parameters, we chose the hidden layer size from 1 to 4, and the number of hidden units from {20, 40, 60, 80, 100}, finally we obtained the best architecture consisting of three denoise autoencoder layers, and the number of units in each layer is [40, 40, 40], respectively. In this section, we present experimental results and evaluation of model for traffic accident risk simulation.

Simulation Results

To evaluate the performance of our model, here we select human mobility data at different time, and figure 5 shows the visualization of human mobility data and corresponding results of our simulator.

In under figures of figure 5, inside the circles is the major commercial and business area of Tokyo. We can see regions which are highlighted as facing high traffic accident risk in our simulation are intensive in this circle. Outside the circle, high traffic accident risk regions distribute with regularities, which we can see outline along arterial roads. In particular, the roads connecting central Tokyo and the city of Yokohama have higher risks in the southern part. Because of the high land prices in Tokyo, many people who are working

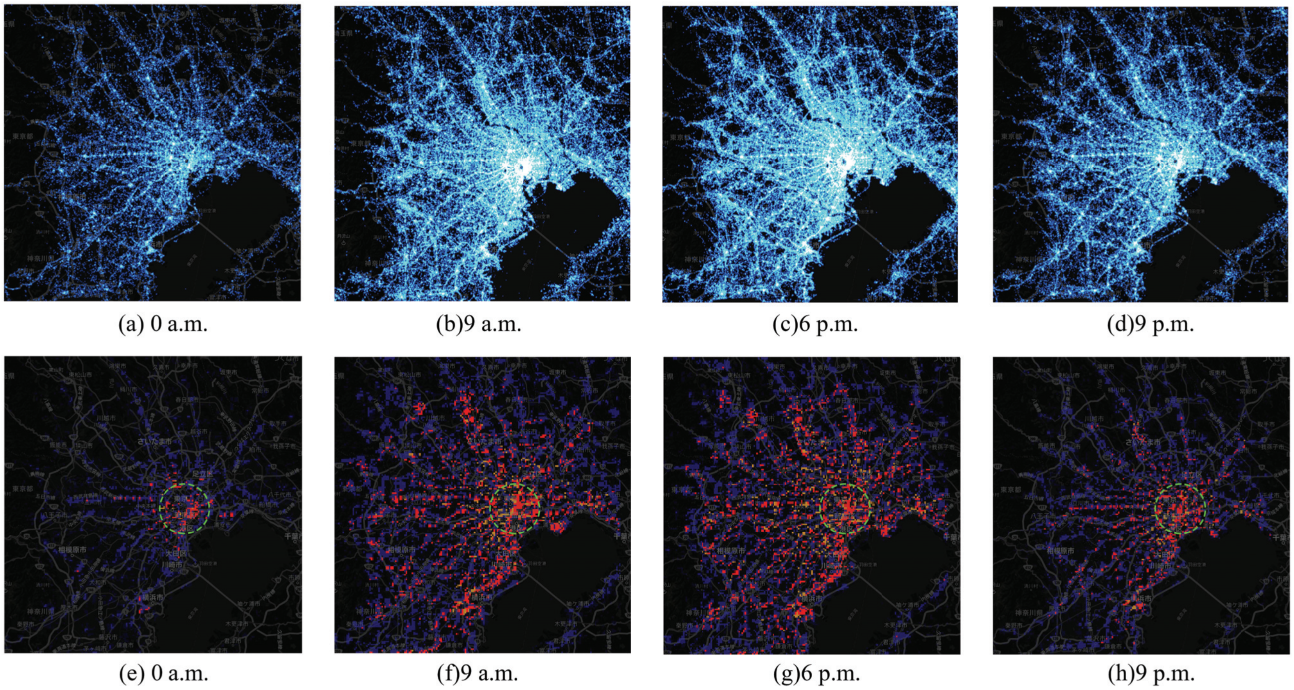

Figure 5: Visualization of simulation results. These figures show the example of input and our simulation results in selected time. Upper figure(a)∼(d) are the visualization of human mobility data, and under figure(e)∼(h) are the corresponding simulation results of traffic accident risk map. We can easily see the dynamic changes of traffic accident risk with human mobility data.

in the central business area of Tokyo are commuting from the city of Yokohama. From simulation results of 9 a.m. and 6 p.m., which are the rush hours and many people drive to work or drive home, we can notice that high traffic accident risk is more intensive in this connection region than other road network regions.

When time goes into the night, most people finish their public activities and stay at home. From the simulation results of 0 a.m. and 9 p.m., we can see that inside the circle, there are still many regions predicted as high traffic accident risk. The reason for this is that people are enjoying nightlife at pubs or nightclubs near the central district. Traffic accident risk dramatically decreases outside the circle, and we can see more obviously that these high traffic accident risk regions are along arterial roads. Interestingly, the accidents risk of the arterial roads connecting central Tokyo and Yokohama is low at 0 a.m because most people are at home at this time. Conclusively, our simulation results coincide with the characteristics of traffic accidents of Tokyo observed so far.

Performance Evaluation

Evaluation metrics: To evaluate the accuracy of the simulation results, we calculated prediction error with three different metrics, which are the mean absolute error (MAE), the mean relative error (MRE), and the root-mean-square error (RMSE). They are defined as:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|g_i - \widehat{g}_i| \tag{5}$$

$$\mathrm{MRE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|g_i - \widehat{g}_i|}{g_i} \tag{6}$$

$$\mathrm{RMSE} = \frac{1}{n}\left[\sum_{i=1}^{n}(g_i - \widehat{g}_i)^2\right]^{\frac{1}{2}} \tag{7}$$

Baseline models: Besides Decision Tree (DT) method mentioned in Section 3, we also considered other two methods which are widely used in classification for comparison. (1) Logistic Regression (LR): Logistic regression measures the relationship between input and target by estimating probabilities using a logistic function (Walker and Duncan 1967), while multinomial logistic regression deals with situations where the outcome can have three or more possible types. (2) Support Vector Machine (SVM): An SVM model treats samples as points in space, and maps samples into separate categories with a clear gap as wide as possible. And new samples are classified based on which side of the gap they fall on (Cortes and Vapnik 1995).

Performance evaluation: We have compared the performance of our model with the performance of the baselines, and Table 1 shows their MAE, MRE and RMSE values. This table indicates that our model performs better and the prediction error is smaller in comparison to these competing methods.

Table 1: Performance Evaluation

Algorithm	MAE	MRE	RMSE
Our Model	0.96	0.39	1.00
DT	1.18	0.60	1.41
LR	1.21	0.40	1.41
SVM	1.40	0.43	1.73

Related Work

With the rise of urban computing in recent years, human mobility data have been widely used in various fields, such as human emergency mobility following disasters (Song et al. 2015; 2014), modeling population movements for very large populations (Song et al. 2010), and understanding basic life pattern of people flow (Fan, Song, and Shibasaki 2014). In addition, some researchers are focusing on social networks through human mobility (Eagle, Pentland, and Lazer 2009; Zhu et al. 2015) and recommend location-based services (Lian et al. 2014; Zhang et al. 2015). Some other interesting works, which are similar as understanding traffic accident risk from human mobility, are applied to traffic density prediction (Castro, Zhang, and Li 2012), diagnose noise in New York City (Zheng et al. 2014b), and infer gas consumption and pollution emission (Shang et al. 2014). Zheng (Zheng et al. 2014a) provided a comprehensive review on the concept, recent researches and challenges of urban computing.

Since the concept of deep learning has been proposed (Hinton, Osindero, and Teh 2006; Hinton and Salakhutdinov 2006), it has been widely used in image processing (Dean et al. 2012), acoustics processing (Hinton et al. 2012) and natural language processing (Collobert and Weston 2008). More recently, some researchers have tried to apply deep learning to intelligent transportation system, and proved its efficiency in traffic flow prediction (Huang et al. 2014; Lv et al. 2015). Another interesting application of deep learning is using deep hybrid model to do weather forecasting (Grover, Kapoor, and Horvitz 2015).

Recently, a number of researches on analyzing traffic accident have been proposed (Xie and Yan 2013; Anderson 2009; Bíl, Andrášik, and Janoška 2013), mainly focusing on hot spot detection of traffic accidents. However, they cannot meet requirements of knowing the real-time traffic accident risk in the neighboring roads in order to select a safer route. Research on dynamic prediction of traffic accident risk on a large scale is very limited due to the lack of support from human mobility data. Thus, in this research, we firstly propose a general model of traffic accident risk prediction that can be applied to simulate real-time risk with updates of human mobility data.

Conclusion

In this paper, big and heterogeneous data on traffic accidents and human mobility in Japan have been collected. By mining these big data, how human mobility affects traffic accident risk has been investigated. We have utilized a deep architecture to extract features from human mobility data, and trained a general prediction model for simulating traffic accident risk on large scale and in real-time, which can be applied to early warn people of possible traffic accident risk for the sake of a safer route. The experimental results demonstrate the efficiency of our simulation model.

However, our study has several limitations owing to the complexity of traffic accidents. Human mobility data are not enough to construct a satisfactory model for the prediction of risks. We will combine human mobility data with other data like land uses and POI (points of interest) data to improve the present model.

Acknowledgements

This work was partially supported by JST, Strategic International Collaborative Research Program (SICORP); Grant-in-Aid for Young Scientists (26730113) of Japan's Ministry of Education, Culture, Sports, Science, and Technology (MEXT); Microsoft Research collaborative research (CORE) program; and DIAS/GRENE project of MEXT. We specially thank ZENRIN DataCom CO., LTD for their support.

References

Anderson, T. K. 2009. Kernel density estimation and k-means clustering to profile road accident hotspots. *Accident Analysis & Prevention* 41(3):359–364.

Bengio, Y.; Lamblin, P.; Popovici, D.; and Larochelle, H. 2006. Greedy layer-wise training of deep networks. In *Advances in Neural Information Processing Systems*, 153–160.

Bengio, Y. 2009. Learning deep architectures for AI. *Foundations and Trends in Machine Learning* 2(1):1–127. Also published as a book. Now Publishers, 2009.

Bíl, M.; Andrášik, R.; and Janoška, Z. 2013. Identification of hazardous road locations of traffic accidents by means of kernel density estimation and cluster significance evaluation. *Accident Analysis & Prevention* 55:265–273.

Castro, P. S.; Zhang, D.; and Li, S. 2012. Urban traffic modelling and prediction using large scale taxi gps traces. In *Pervasive Computing*. Springer. 57–72.

Collobert, R., and Weston, J. 2008. A unified architecture for natural language processing: Deep neural networks with multitask learning. In *Proceedings of the 25th international conference on Machine learning*, 160–167. ACM.

Cortes, C., and Vapnik, V. 1995. Support-vector networks. *Machine learning* 20(3):273–297.

Dean, J.; Corrado, G.; Monga, R.; Chen, K.; Devin, M.; Mao, M.; Senior, A.; Tucker, P.; Yang, K.; Le, Q. V.; et al.

2012. Large scale distributed deep networks. In *Advances in Neural Information Processing Systems*, 1223–1231.

Eagle, N.; Pentland, A. S.; and Lazer, D. 2009. Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences* 106(36):15274–15278.

Fan, Z.; Song, X.; and Shibasaki, R. 2014. Cityspectrum: a non-negative tensor factorization approach. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 213–223. ACM.

Grover, A.; Kapoor, A.; and Horvitz, E. 2015. A deep hybrid model for weather forecasting. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '15, 379–386. New York, NY, USA: ACM.

Hinton, G. E., and Salakhutdinov, R. R. 2006. Reducing the dimensionality of data with neural networks. *Science* 313(5786):504–507.

Hinton, G.; Deng, L.; Yu, D.; Dahl, G. E.; Mohamed, A.-r.; Jaitly, N.; Senior, A.; Vanhoucke, V.; Nguyen, P.; Sainath, T. N.; et al. 2012. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *Signal Processing Magazine, IEEE* 29(6):82–97.

Hinton, G. E.; Osindero, S.; and Teh, Y.-W. 2006. A fast learning algorithm for deep belief nets. *Neural computation* 18(7):1527–1554.

Huang, W.; Song, G.; Hong, H.; and Xie, K. 2014. Deep architecture for traffic flow prediction: Deep belief networks with multitask learning. *Intelligent Transportation Systems, IEEE Transactions on* 15(5):2191–2201.

Lian, D.; Zhao, C.; Xie, X.; Sun, G.; Chen, E.; and Rui, Y. 2014. Geomf: joint geographical modeling and matrix factorization for point-of-interest recommendation. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, 831–840. ACM.

Lv, Y.; Duan, Y.; Kang, W.; Li, Z.; and Wang, F.-Y. 2015. Traffic flow prediction with big data: a deep learning approach. *Intelligent Transportation Systems, IEEE Transactions on* 16(2):865–873.

Quinlan, J. R. 1987. Simplifying decision trees. *International journal of man-machine studies* 27(3):221–234.

Shang, J.; Zheng, Y.; Tong, W.; Chang, E.; and Yu, Y. 2014. Inferring gas consumption and pollution emission of vehicles throughout a city. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, 1027–1036. ACM.

Song, C.; Koren, T.; Wang, P.; and Barabási, A.-L. 2010. Modelling the scaling properties of human mobility. *Nature Physics* 6(10):818–823.

Song, X.; Zhang, Q.; Sekimoto, Y.; and Shibasaki, R. 2014. Intelligent system for urban emergency management during large-scale disaster. In *Twenty-Eighth AAAI Conference on Artificial Intelligence*.

Song, X.; Zhang, Q.; Sekimoto, Y.; Shibasaki, R.; Yuan, N. J.; and Xie, X. 2015. A simulator of human emergency mobility following disasters: Knowledge transfer from big disaster data. In *Twenty-Ninth AAAI Conference on Artificial Intelligence*.

Walker, S. H., and Duncan, D. B. 1967. Estimation of the probability of an event as a function of several independent variables. *Biometrika* 54(1-2):167–179.

Xie, Z., and Yan, J. 2013. Detecting traffic accident clusters with network kernel density estimation and local spatial statistics: an integrated approach. *Journal of Transport Geography* 31:64 – 71.

Zhang, F.; Zheng, K.; Yuan, N. J.; Xie, X.; Chen, E.; and Zhou, X. 2015. A novelty-seeking based dining recommender system. In *Proceedings of the 24th International Conference on World Wide Web*, 1362–1372. International World Wide Web Conferences Steering Committee.

Zheng, Y.; Capra, L.; Wolfson, O.; and Yang, H. 2014a. Urban computing: concepts, methodologies, and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)* 5(3):38.

Zheng, Y.; Liu, T.; Wang, Y.; Zhu, Y.; Liu, Y.; and Chang, E. 2014b. Diagnosing new york city's noises with ubiquitous data. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 715–725. ACM.

Zhu, W.-Y.; Peng, W.-C.; Chen, L.-J.; Zheng, K.; and Zhou, X. 2015. Modeling user mobility for location promotion in location-based social networks. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1573–1582. ACM.