Towards Intelligent Vehicle Intrusion Detection Using the Neural Knowledge DNA

Fei Li [a], Haoxi Zhang [a], Juan Wang [a], Yong Liu [a], Lulu Gao[a], Xiang Xu[a] ,

Cesar Sanin [b], Edward Szczerbicki [c]

[a] School of Cybersecurity, Chengdu University of Information Technology, Chengdu, China; [b] School of Mechanical Engineering, The University of Newcastle, Newcastle, NSW, Australia; [c] Faculty of Management and Economics, Gdansk University of Technology, Gdansk, Poland

Abstract. In this paper, we propose a novel intrusion detection approach using past driving experience and the Neural Knowledge DNA for in-vehicle information system security. The Neural Knowledge DNA is a novel knowledge representation method designed to support discovering, storing, reusing, improving, and sharing knowledge among machines and computing systems. We examine our approach for classifying malicious vehicle control commands based on learning from past valid driving behaviour data on a simulator.

Address correspondence to Fei Li, School of Cybersecurity, Chengdu University of Information Technology, No. 24 Block 1, Xuefu Road, Chengdu, China, 610225. E-mail: lifei@cuit.edu.cn

Keywords: Vehicle Intrusion Detection, Neural Knowledge DNA, Set of Experience Knowledge Structure, Knowledge Representation, Deep Learning.

INTRODUCTION

With the introduction of Tesla cars, the automobile industry has set off an upsurge of embracing information technology (IT). More and more automobile companies start to apply IT to their products. Apart from automobile companies, big IT software/hardware vendors, such as Apple, Google, Microsoft, and Nvidia are joining this trend.

However, security risks come along with the increase of informatization of vehicles. For example, Koscher K. et al. (2010) successfully attacked the vehicle information system through a simple software code, and closed the car's braking system. Users are faced with the threat of network attacks during driving, which may lead to the loss of vehicle control and life-critical problems. Therefore, it is vitally important having the in-vehicle information system that is robust and secure becomes indispensable to modern automobiles.

VEHICLE INFORMATION SYSTEM SECURITY

The convergence of automobile and IT is a rapidly rising paradigm of modern vehicles, in which an electronic control unit (ECU) controls the vehicle electrical

systems. The controller area network (CAN), an in-vehicle network, is commonly used to construct an efficient system of ECUs. Fatally, security issues have not been tackled properly in CAN, although CAN control messages could be life-critical. With the appearance of the connected car environment, in-vehicle networks are now well connected to the outside world, which gives the hackers chances to perform cyber-attacks using CAN vulnerabilities. This security problem has drawn great attention of industry and academic communities.

Schweppe et al. (2011) presented an approach that can be applied to automotive bus systems based on the use of symmetric key material protected with inexpensive hardware. Lin and Sangiovannivincentelli (2012) designed a security mechanism to help prevent cyber-attacks (masquerade and replay) in vehicles with architecture based on Controller Area Network. A broadcast authentication protocol based on the key-chains and time synchronization was designed, refined and implemented by Groza and Murvay (2013) to fit the CAN bus system. Woo et al. (2015) proposed a more efficient and lightweight security protocol for CAN as a countermeasure designed in accordance with current CAN specifications. Rather than just relying on encryption and authentication, Muter et al. (2010) introduced a structured anomaly detection method that allows the recognition of attacks during the operation of the vehicle without causing false positives. Similarly, Kang MJ and Kang JW (2016) proposed an intrusion detection technique using a deep neural network (DNN). In their proposed technique, in-vehicle network packets exchanged between ECUs, are trained to extract

low-dimensional features and used for discriminating normal and hacking packets. Gu et al. (2017) also introduced a Sybil attack detection method based on k-Nearest Neighbours (kNN) classification algorithm. In this method, vehicles are classified based on the similarity in their driving patterns.

In this paper, we introduce a novel intrusion detection method based on the Neural Knowledge DNA (Zhang et al. 2017) and the past driving experience gathered on the vehicle for abnormal commands detection in vehicle electronic control system.

THE NEURAL KNOWLEDGE DNA

The Neural Knowledge DNA (NK-DNA) was proposed to store and represent knowledge captured in intelligent systems that uses artificial neural networks as the central power of its intelligence (Zhang et al. 2017). It utilises the ideas underlying the success of deep learning (LeCun et al. 2015) to the scope of knowledge representation.

The NK-DNA is constructed in a similar fashion of how DNA formed (Sinden 1994): built up by four essential elements. As the DNA produces phenotypes, the Neural Knowledge DNA carries information and knowledge via its four essential elements, namely, States, Actions, Experiences, and Networks (see Figure 1).

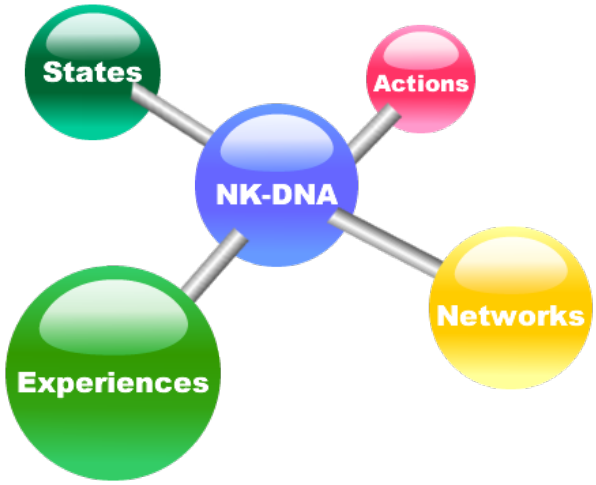

Fig. 1. Conceptual structure of the NK-DNA.

The NK-DNA's four-element combination is designed to carry detailed information on decisions. States are situations in which a decision or a motion can be made or performed. Actions are used to store the decisions or motions which can be selected in the given domain. Experiences are domain's historical operation segments with feedbacks from outcomes, and Networks store the description of neural networks for training and using knowledge about the network structure, weights, bias, and deep learning framework.

Generally, knowledge is acquired as models after training in deep learning systems. The model usually stores detailed information about weights and biases of the connections between neurons of the neural network, and its hierarchy. Once the neural network has been trained, the network will give results straightforward through the computation of its network layers after feeding it with inputs. Similarly, the NK-DNA stores knowledge using the same idea. Figure 2 shows the concept of knowledge captured and carried by the NK-DNA.

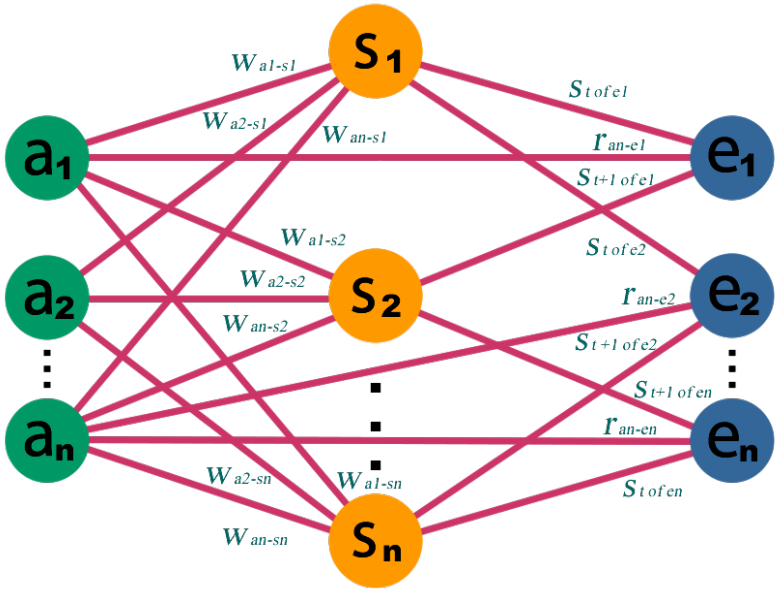

Fig. 2. Concept of the NK-DNA-carried knowledge.

In the NK-DNA, a neural network is used to carry the relation between actions and states: as we can see in the Figure 2, each state (represented as S_1, S_2 … S_n) can have connections with a set of actions (represented as a_1, a_2, … a_n). If an action is connected with a state, it means the connected action is an available action in that state; in other words, the agent can choose the action to perform if it is in that state. The trained neural network provides the knowledge of which action is the best choice to a specific state. The states here are the inputs, which can be the raw sensory data, or data describing the current situation of the agent.

Another important feature of this approach is that the NK-DNA uses previous decisional experience as the main source to collect and expand intelligence for future decision making. Experience in the NK-DNA is stored as the Set of Experience Knowledge Structure (SOEKS) (Sanin & Szczerbicki 2004, 2006). Usually, the agent transitions from one state to another during its operation, and it makes decisions

(selects actions) in each state. It also receives feedbacks from its operation. All states, actions, feedbacks, and transitions form agent's experience.

THE NK-DNA BASED INTRUSION DETECTION

Overview of the NK-DNA based intrusion detection system

The model of vehicle states needs to consider a variety of parameters and situations. They would include for example information on whether the vehicle is driving abnormally, or whether the instruction the in-vehicle network receives is valid. It will also include characteristics of the driver and the external environment of the vehicle. For example, new drivers and experienced drivers are likely to behave differently, and driving behaviour in mountainous areas is also significantly different from that in cities. This means that if we want to classify an instruction on the vehicle CAN Bus as valid or not, we have to consider various conditions of driving, and link them together. For this reason, we designed three types of experience in the NK-DNA based intrusion detection method which are organized and stored as SOEKS. They are Driver SOEKS, Vehicle SOEKS, and Surroundings (see Figure 3).

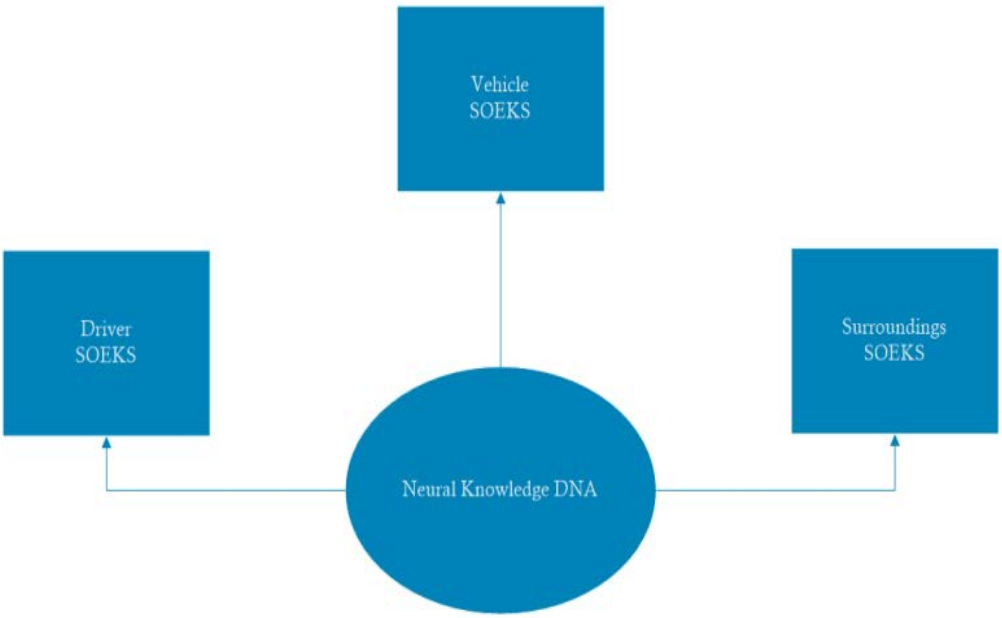

Fig. 3. Experience models in the NK-DNA method.

The Driver SOEKS stores knowledge of drivers, including the driver's age, gender, licence, etc. The Vehicle SOEKS keeps the data of the vehicle, such as make, model, and type. The Surrounding SOEKS carry the information about external environment, like camera images and radar cloud point data. All these SOEKS are sent to the NK-DNA for training.

Training process

Based on SOEKS knowledge representation, an intrusion detection model for automobile electronic control system can be determined through the NK-DNA training. In the NK-DNA, a neural network is used to carry the relation between actions and states. In our case, each state is defined by who is driving (Driver SOEKS), what kind is the vehicle and how is it moving (Vehicle SOEKS), and external environment

(Surrounding SOEKS). If an action is connected with a state, it means the connected action is an available and legitimate action in that state. Malicious actions (intrusions) will be assumed if there is no connection between them and states.

In our approach, experience is used to link states with actions. Experience, being some information gained from driving in the past, is the ideal source for learning and improving performance of agents. Every time a driver operates the vehicle, experience is collected and stored as $e_t = (s_t, a_t,)$: where s_t is the vehicle state at that time, while a_t is the action the driver performs at that time. After experience collection, we can train the intrusion detection model. Basically, each experience is labelled as invalid or valid according to whether the action is made for a malicious purpose or not. This labelled experience is treated as samples for performing supervised learning through a neural network. As a result, the trained neural network provides the knowledge of which action is invalid in a specific state.

INITIAL CASE STUDY

When hackers attack the vehicle, they might to accelerate, steer, or brake suddenly and abnormally. In this initial case study, we focus on acceleration operations. If a hacker wants the vehicle to accelerate, the malicious instructions will manipulate the engine's throttle value. However, the throttle value would never change suddenly if the driver does not press the gas pedal, and the normal throttle value is linearly related to revolutions per minute (RPM) and speed. The above helps to detect malicious

instructions of attacking the throttle. To test our idea, we collected the power system's CAN bus data of a 2010 Ford Escape (Figure 4). Around 400,000 sets of data were collected during a 60 minutes drive.

pam_prog.dat | abs_service_bleed.dat | autopark.dat | drive.dat

```
IDH: 00, IDL: 80, Len: 08, Data: FF FC 4E 20 01 C0 49 FF ,TS: 80,BAUD: 1
IDH: 02, IDL: 30, Len: 08, Data: DD 00 00 00 00 00 4F 10 ,TS: 97,BAUD: 1
IDH: 00, IDL: 41, Len: 08, Data: 00 61 00 00 00 00 00 00 ,TS: 105,BAUD: 1
IDH: 02, IDL: 15, Len: 08, Data: 27 10 27 10 27 10 27 10 ,TS: 109,BAUD: 1
IDH: 02, IDL: 16, Len: 08, Data: 00 00 00 00 AA 00 00 00 ,TS: 113,BAUD: 1
IDH: 04, IDL: 30, Len: 08, Data: F5 F8 00 02 00 10 00 40 ,TS: 117,BAUD: 1
IDH: 00, IDL: 81, Len: 08, Data: F9 5C 01 11 00 00 00 00 ,TS: 180,BAUD: 1
IDH: 00, IDL: 91, Len: 08, Data: 59 2F 7F 55 7F 09 A8 39 ,TS: 204,BAUD: 1
IDH: 00, IDL: 82, Len: 08, Data: 7E 08 40 00 00 00 00 00 ,TS: 208,BAUD: 1
IDH: 02, IDL: 00, Len: 08, Data: 27 1C 27 5C 27 5C 00 00 ,TS: 222,BAUD: 1
IDH: 02, IDL: 01, Len: 08, Data: 0B BC 00 00 27 10 00 00 ,TS: 226,BAUD: 1
IDH: 02, IDL: 30, Len: 08, Data: DD 00 00 00 00 00 4F 10 ,TS: 229,BAUD: 1
IDH: 00, IDL: 80, Len: 08, Data: FF FC 4E 20 01 C0 4A FF ,TS: 237,BAUD: 1
IDH: 00, IDL: 41, Len: 08, Data: 00 61 00 00 00 00 00 00 ,TS: 262,BAUD: 1
IDH: 00, IDL: 73, Len: 08, Data: 1C 25 17 65 17 4F 1C 37 ,TS: 266,BAUD: 1
IDH: 02, IDL: 11, Len: 08, Data: FF FE 00 64 00 4A 00 00 ,TS: 270,BAUD: 1
IDH: 02, IDL: 15, Len: 08, Data: 27 10 27 10 27 10 27 10 ,TS: 273,BAUD: 1
IDH: 04, IDL: 17, Len: 08, Data: 00 00 00 00 00 00 00 00 ,TS: 277,BAUD: 1
IDH: 02, IDL: 30, Len: 08, Data: DD 00 00 00 00 00 4F 10 ,TS: 347,BAUD: 1
IDH: 02, IDL: 50, Len: 08, Data: 80 46 38 7A 02 01 0C 28 ,TS: 386,BAUD: 1
```

Fig. 4. The 2010 Ford Escape power system CAN data.

By training our neural networks of the NK-DNA, we can predict the speed from the RPM. Figure 5 shows the result of the prediction.

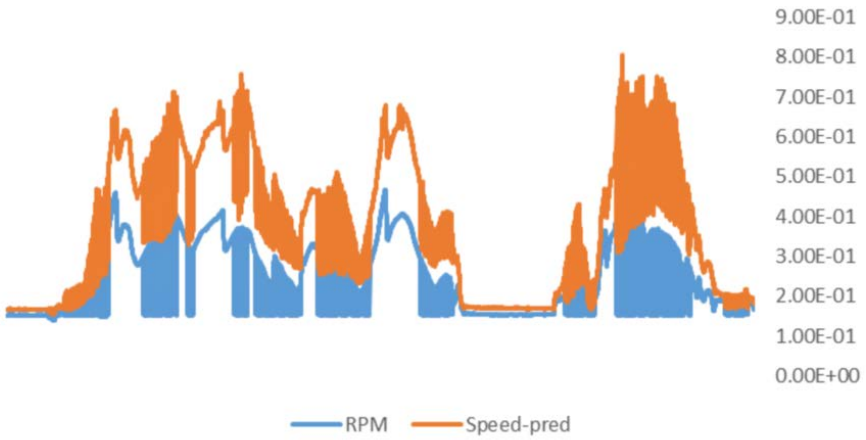

Fig. 5. Speed predicted by RPM.

When the observed speed values are not consistent with the predicted values, the exception is found, as shown in Figure 6.

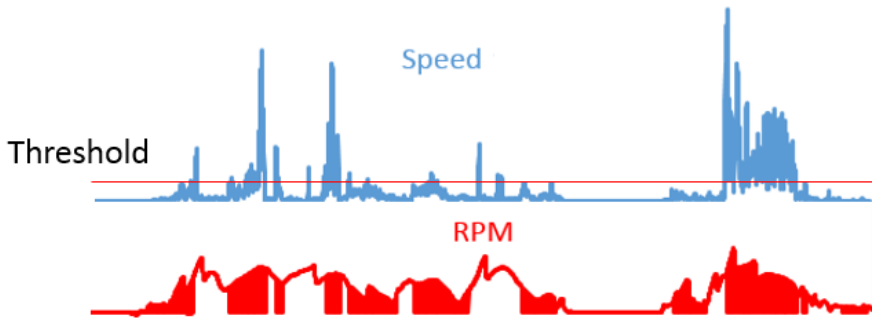

Fig. 6. Detecting abnormalities through correlation between RPM and Speed.

Finally, the neural network is able to classify the normal and the abnormal cases through values of throttle and gas pedal (Figure 7).

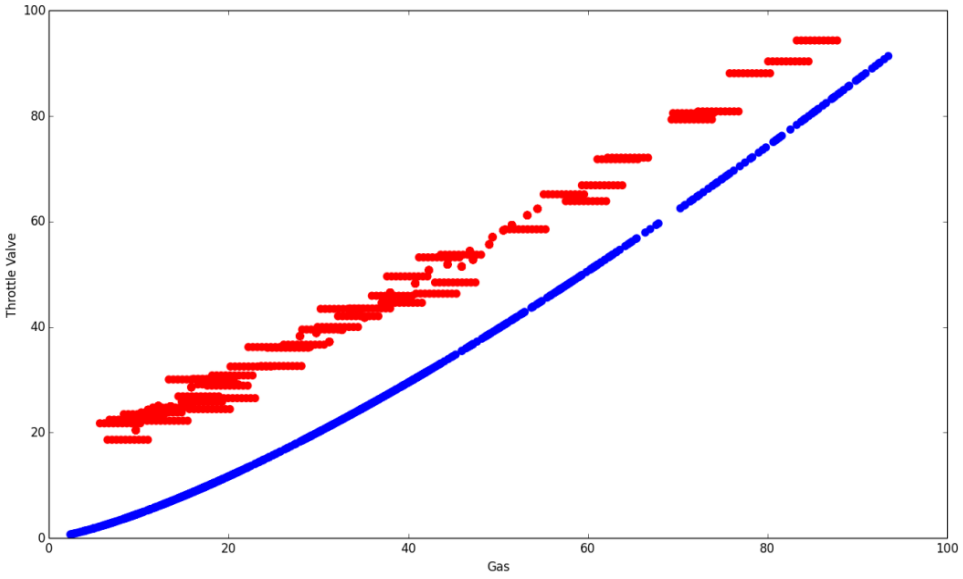

Fig. 7. Correlation between throttle and gas in normal (the regular lower line) and under attack condition (the irregular upper data)

CONCLUSIONS AND FUTURE WORK

In this paper, we briefly introduced a novel intrusion detection approach using past driving experience and the NK- DNA for in-vehicle information system security. Using embedded in the NK_DNA deep learning and SOEKS technologies, our approach can detect malicious throttle control instructions employing experiential knowledge.

Future research and system refinement steps include:

- refining and coding of the proposed intrusion detection framework,

- expansion of the intrusion detection strategy,

- further elaboration of the intrusion detection algorithm.

ACKNOWLEDGEMENT

The authors would like to thank the editors and anonymous reviewers for their valuable comments and suggestions on this paper. This work was supported by the Scientific Research Foundation of Sichuan Province as part of the Project 2016GZ0343.

REFERENCES

Groza, B., & Murvay, S. 2013. Efficient protocols for secure broadcast in controller area networks. IEEE Transactions on Industrial Informatics, 9(4), 2034-2042.

Gu, P., Khatoun, R., Begriche, Y., & Serhrouchni, A. 2017. K-Nearest Neighbours classification based Sybil attack detection in Vehicular networks. International Conference on Mobile & Secure Services (pp.1-6). IEEE.

Haoxi Zhang, Cesar Sanin, Edward Szczerbicki, and Ming Zhu, Towards Neural Knowledge DNA. 2017. Journal of Intelligent & Fuzzy Systems, vol. 32, no. 2, pp. 1575-1584

Kang, M. J., & Kang, J. W. 2016. A Novel Intrusion Detection Method Using Deep Neural Network for In-Vehicle Network Security. Vehicular Technology Conference (pp.1-5). IEEE.

Koscher K., Czeskis A, Roesner F, et al. 2010. Experimental security analysis of a modern automobile. IEEE Symposium on Security and Privacy. Piscataway, USA: IEEE Computer Society, 2010:447-462.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. 2015. "Deep learning." Nature 521.7553: 436-444.

Lin, C. W., & Sangiovannivincentelli, A. 2012. Cyber-Security for the Controller Area Network (CAN) Communication Protocol. International Conference on Cyber Security (Vol.390, pp.1-7). IEEE Computer Society.

Müter, M., Groll, A., & Freiling, F. C. 2010. A structured approach to anomaly detection for in-vehicle networks. Sixth International Conference on Information Assurance and Security (Vol.29, pp.92-98). IEEE.

Sanin, C., and E Szczerbicki, 2004, Knowledge Supply Chain System: a conceptual model, in *Knowledge Management: Selected Issues*, A Szuwarzynski (Ed), Gdansk University Press, Gdansk pp. 79-97, 2004

Sanin C., Szczerbicki E., (2006) Using Set of Experience in the Process of Transforming Information into knowledge, International Journal of Enterprise Information Systems, 2(2), pp. 40-55.

Sanin, C., C. Toro, Z. Haoxi, E. Sanchez, E. Szczerbicki, E. Carrasco, W. Peng and L. Mancilla-Amaya. 2012. Decisional DNA: A Multi-Technology Shareable Knowledge Structure for Decisional Experience. Neurocomputing 88(0): 42–53.

Schweppe, H., Roudier, Y., Weyl, B., & Apvrille, L. 2011. Car2X Communication: Securing the Last Meter - A Cost-Effective Approach for Ensuring Trust in Car2X Applications Using In-Vehicle Symmetric Cryptography. Vehicular Technology Conference (VTC Fall), 2011: 1-5. IEEE.

Woo, S., Jo, H. J., & Dong, H. L. 2015. A practical wireless attack on the connected car and security protocol for in-vehicle can. IEEE Transactions on Intelligent Transportation Systems, 16(2), 993-1006.