Technische Universität München
Lehrstuhl für Datenverarbeitung

Research Proposal

Deep Reinforcement Learning based Resource Management for Cellular Vehicle-to-Vehicle Communication(C-V2V)

Alperen Gündogan

September 2, 2018

1 Motivation

The biggest purpose of the vehicular communication system is to avoid the accidents in traffic and to provide a more safe environment for cars. 1.2 million deaths around the world are caused by road accidents according to the report of World Health Organization(WHO)[1]. Also, the expenditure of car collisions in the United States is $300 billion per year as stated in the American Automobile Association(AAA) study[2]. Those accidents can be significantly avoided with the communication between cars. Cars can deliver the current status of the road(i.e. accidents, icy road, detected objects, etc.) with Cooperative Awareness Message(CAM) to the other nearby cars. For example; all the processed data which are obtained by the sensors of an autonomous car, are shared with surrounded cars. Then, those informed cars can benefit from the knowledge of the other cars to prepare themselves for undesirable road conditions and to have a more reliable control mechanism for the objects around the car. Also, autonomous driving cars can operate only in line-of-sight condition with a limited range of 100-200m whereas vehicular communication can disseminate information over a larger range and promotes advanced driver-assistance system(ADAS) which now enables cars to see from "someone else's eyes". There are various standards that are currently exploiting for vehicular communication but standardized and the most famous one is WAVE which is known as IEEE 802.11p, enables vehicle-to-vehicle(V2V) and vehicle-to-infrastructure(V2I) communication by using the band of 5.9 GHz[3]. However, deployment of 802.11p into road infrastructure requires a huge amount of investment. Therefore, 5G Vehicle-to-Everything(V2X) communication would be a suitable candidate to reduce the cost of operations since current telecommunication infrastructures can be directly used.

The main idea is allocating resources for V2V communication between two cars by cellular network scheduler. Then, those cars can directly communicate with each other by using this assigned side-link e.g. without the cellular network operating as a relay. C-V2V provides lower latency, higher reliability than 802.11p and also can operate in higher loads[4]. The ultimate goal of this proposal is to design a deep reinforcement learning(DRL) algorithm for resource management of V2V communication assuming that vehicles are in the cellular cell and scheduled by E-UTRAN Node B(eNB). This research proposal is inspired from the several lectures of Topics in Multimedia Signal Processing course; "Reinforcement Learning and Artificial Intelligence" by Martin Knopp, "A Gauss-Newton algorithm for Deep Reinforcement Learning" by Martin Gottwald and "Autonomous Driving" by ITK Engineering.

2 State of the Art

LTE Direct Device-to-Device(D2D) which was advertised with release 12, was exploited first for public safety use cases but it did not provide ultra-low latency which is essential for V2V communication. Then, release 14 introduced C-V2X(Cellular-V2X) communication which was built upon LTE Direct D2D with improvements to fulfill low latency requirement of V2V communication. However, current standards of 3GPP release 14 do not specify the resource management algorithms for the case where vehicles are scheduled by eNB which is known as mode 3 scheduling. Operators can determine their own algorithm which should be under one of these two categories[5]; dynamic scheduling which may not be a good choice for V2V scenario since a vehicle requests resources from eNB for each transmission which results in huge signalling overhead in the network and increases latency. Another category is termed as semi-persistent scheduling(SPS) which reduces the signalling overhead for real-time applications[6]. Basically, eNB assigns the resources to a user for a specific time interval e.g. 20ms. However, the problem with resource management of V2V communication is channel condition and network topology constantly change due to the high mobility of cars. Furthermore, sensors such as radar, lidar, and cameras of a car generate a huge amount of data which require sufficient bandwidth and collision free communication to properly disseminate those data. Therefore, traditional resource management algorithms are having difficulties to cope with such a rich context and dynamic channels while supporting stringent QoS requirements of V2V links[7]. Furthermore, 99.999% transmission reliability and maximum 5 ms end-to-end delay should be guaranteed according to the EU project METIS[8]. Fortunately, deep reinforcement learning can adapt dynamic environments where there is extremely large state-action space.

Reinforcement learning techniques have been exploiting recently in wireless communication to fulfill the demands of mobile applications and services. The article[9] showed that state-of-art DRL approaches can be employable for resource management problems in large-scale systems. Gursoy *et al.* proposed DRL based framework for power-efficient resource allocation in cloud RANs [10]. State-Action-Reward-State-Action(SARSA) is exploited by Ferreira *et al.* to address resource management in cognitive space communication[11]. In [12], dynamic scheduling for data traffic management is proposed which provides surpassing delay and packet drop rate than traditional methods by exploiting DRL principles. Only one example exists where DRL is applied for V2V communication, resource allocation strategy in [13] has been published recently which deploy a decentralized approach to satisfy V2V latency constraints. However, resource management for 5G-V2X is a very novel topic in which no explicit algorithm is dedicated for this challenging task yet. Hence, further investigations are needed to straighten out the complexity of the system and deep reinforcement learning would be a perfect option for this.

3 Research Objectives and Approach

V2X control function is introduced in the V2X architecture with 3GPP release 14 which determines the radio sources of vehicles[14]. After the initial allocation of radio resources of vehicles which are assigned by eNB V2X control function, vehicle sends its current state and the reward obtained from that state periodically to the eNB for scheduling resources of next transmission. Reward value is determined by the reward function which based on the latency among V2V links.

3.1 Approach

V2X control function observes the state $s_t\,(s_t \epsilon S)$ where S is the all possible states and $t\,(t = 1, 2, 3...)$ is each time step. Then, action a_t is chosen from the set of available actions $A_{(s_t)} \subseteq A$ upon the state s_t. A reward $r_{t+1} \epsilon R$ associated with (s_t, a_t, s_{t+1}) is received from vehicle after expiration of the resource reservation period of SPS. Since the environment of vehicles constantly and rapidly change, it is not practical to store all (s_t, a_t) pairs with related rewards. Therefore, deep neural network can be effectively exploited to approximate the expected sum reward $Q(s_t, a_t)$, given (s_t, a_t) as an input.

One can exploits Veins[15] for vehicular networks and SimuLTE[16] for LTE network system-level simulation which are both in the frame of OMNet++ to simulate and test the designed algorithm. VeinsLTE[17] integrates those two frameworks however it is difficult to update when new features are added in either of those programs. Therefore, the article[18] could be a good guide by clarifying how to make SimuLTE and Veins interoperable.

3.2 Objectives

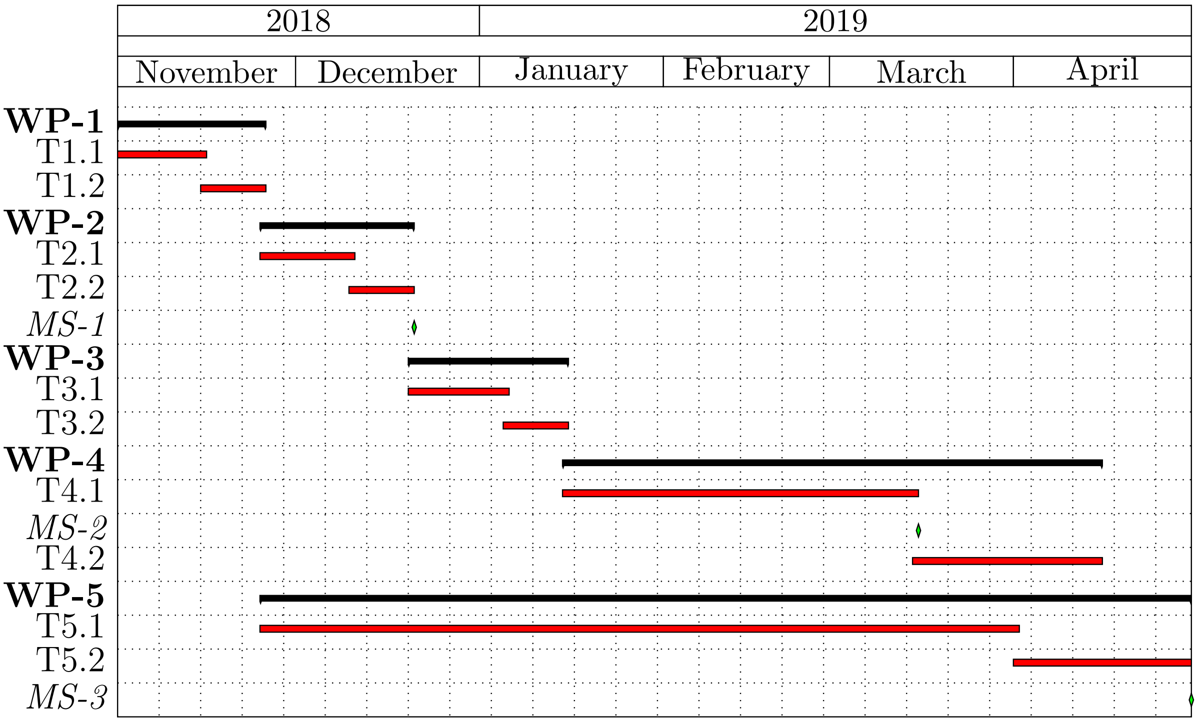
At the end of this research two main objectives will be obtained.

- Deep reinforcement learning based resource management algorithm design for C-V2V.
- Evaluation and comparison of results with classical resource management algorithms. Is the DRL algorithm can address the problem of resource management for C-V2V?

3.3 Research Plan

It is assumed that the student has a previous knowledge about 4G/5G protocols&specifications, and experiences with similar network simulators e.g. ns-3, OMNeT++ and machine learning libraries.

- Work Package 1: Thorough literature research. Identification and understanding the relevant work.
    1. Study on 5G - V2X architecture(release 14, TS 23.285).
    2. Listing of classical resource management algorithms for LTE and 5G.



- Work Package 2: Identify, test and learn about the required software libraries.
    1. Learning the protocol stack of 5G - V2X and testing possible V2X examples with the software.
    2. Developing a strategy(e.g. flowchart) to implement resource management algorithms into the simulation(*Milestone 1*).
- Work Package 3: Testing of mostly used 2-3 algorithms in the simulation environment and evaluation of their strengths and weaknesses.
    1. Implementation of the determined algorithms into the simulation.
    2. Creating plots: probability of V2V links that satisfy latency requirements vs number of V2V links. Throughput vs number of V2V links. Impact of various allocation time intervals(e.g. 20ms, 100ms etc.).
- Work Package 4: Design, specification and implementation of a DRL algorithm.
    1. Design and implementation of an algorithm, testing for different scenarios.(*Milestone 2)*.
    2. Utilization of the algorithm for better results.
- Work Package 5: Writing final report.
    1. Writing a literature review, algorithm specifications.

2. Comparison of all the results and evaluation of the DRL algorithm(*Milestone 3*).

3.4 Contengency plan

Cellular V2X is a very novel topic, therefore software libraries and tools are still at the developing stage. One possible problem might be lackness or underdevelopment of some required libraries for the latest C-V2X architecture. In this situation, one should develop short part of the library to be able to implement and test resource management algorithms.

References

[1] M. Peden, R. Scurfield, D. Sleet, D. Mohan, and et.al., "World report on road traffic injury prevention," tech. rep., World Health Organization, April 2016.

[2] American Automobile Association, "Crashes vs. congestion - what's the cost to society?," tech. rep., april 2016.

[3] S. Eichler, "Performance evaluation of the ieee 802.11p wave communication standard," in *Proceedings of Vehicular Technology Conference*, pp. 2199 – 2203, 2007.

[4] G. Dimitrakopoulos, *Vehicular Communication Standards*, ch. 2, pp. 13–28. Springer International Publishing, 2017.

[5] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures." 3GPP, Tech. Rep. 36.213 (v14.3.0, Release 14), June 2017.

[6] C. Leung, R. Kwan, S. Hamalainen, and W. Wang, "LTE/LTEAdvanced Cellular Communication Networks," *Journal of Electronic and Computer Engineering*, pp. 1–10, 2010.

[7] H. Ye, L. Liang, G. Y. Li, J. Kim, L. Lu, and M. Wu, "Machine learning for vehicular networks: Recent advances and application examples," *IEEE Vehicular Technology Magazine*, vol. 13, pp. 94–101, June 2018.

[8] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka, H. Tullberg, M. A. Uusitalo, B. Timus, and M. Fallgren, "Scenarios for 5g mobile and wireless communications: The vision of the metis project," vol. 52, pp. 26–35, 05 2014.

[9] H. Mao, M. Alizadeh, I. Menache, and S. Kandula, "Resource management with deep reinforcement learning," in *Proceedings of the 15th ACM Workshop on Hot Topics in Networks*, HotNets '16, (New York, NY, USA), pp. 50–56, ACM, 2016.

[10] Z. Xu, Y. Wang, J. Tang, J. Wang, and M. C. Gursoy, "A deep reinforcement learning based framework for power-efficient resource allocation in cloud rans," in *2017 IEEE International Conference on Communications (ICC)*, pp. 1–6, May 2017.

[11] P. V. R. Ferreira, R. Paffenroth, A. M. Wyglinski, T. M. Hackett, S. G. Bilen, R. C. Reinhart, and D. J. Mortensen, "Multi-objective reinforcement learning-based deep neural networks for cognitive space communications," in *2017 Cognitive Communications for Aerospace Applications Workshop (CCAA)*, pp. 1–8, June 2017.

[12] I. S. Comsa, S. Zhang, M. Aydin, P. Kuonen, Y. lu, R. Trestian, and G. Ghinea, "Towards 5g: A reinforcement learning-based scheduling solution for data traffic management," pp. 1–14, Aug 2018.

[13] H. Ye and G. Y. Li, "Deep reinforcement learning for resource allocation in v2v communications," in *2018 IEEE International Conference on Communications (ICC)*, pp. 1–6, May 2018.

[14] 3GPP, "V2X Control Function to Home Subscriber Server (HSS) aspects (V4); Stage 3." 3GPP TS 29.388 version 14.0.0 Release 14, July 2017.

[15] C. Sommer, R. German, and F. Dressler, "Bidirectionally coupled network and road traffic simulation for improved ivc analysis," *IEEE Transactions on Mobile Computing*, vol. 10, pp. 3–15, Jan 2011.

[16] A. Virdis, G. Stea, and G. Nardini, "Simulte - a modular system-level simulator for lte/lte-a networks based on omnet++," in *2014 4th International Conference On Simulation And Modeling Methodologies, Technologies And Applications (SIMULTECH)*, pp. 59–70, Aug 2014.

[17] F. Hagenauer, F. Dressler, and C. Sommer, "Poster: A simulator for heterogeneous vehicular networks," in *2014 IEEE Vehicular Networking Conference (VNC)*, pp. 185–186, Dec 2014.

[18] G. Nardini, A. Virdis, and G. Stea, "Simulating cellular communications in vehicular networks: Making simulte interoperable with veins," 09 2017.