The following publication Lu, L., & Wang, S. (2019, October). Literature review of analytical models on emergency vehicle service: location, dispatching, routing and preemption control. In 2019 IEEE Intelligent Transportation Systems Conference (ITSC) (pp. 3031-3036). IEEE is available at https://doi.org/10.1109/ITSC.2019.8916885

Literature Review of Analytical Models on Emergency Vehicle Service: Location, Dispatching, Routing and Preemption Control

Lili Lu, Shuaian Wang

Abstract— **Emergency vehicle service is one of the most important public services. It plays a vital role in saving people's lives and reducing the rate of mortality and morbidity. In the past a few decades, intensive efforts have been devoted by researchers to design effective and efficient strategies to ensure the quality of the service of emergency vehicles under different environments. This study seeks to provide a holistic review of the models and strategies designed in literature to improve the emergency vehicle service. According to the steps of an emergency vehicle to response to a call demand, this study identifies four problems in the field of the emergency vehicle service, i.e., the location problem, the dispatching problem, the routing problem, as well as the preemption control problem. We will review the past studies devoted to addressing each problem explicitly. This review can help the interested reader to trace the evolution of the models and current keen research interests in the field of emergency vehicle service.**

I. Introduction

Emergency vehicles (EVs) such as the ambulances, fire trucks, and police car play an important role on saving lives and reducing property damage. They are expected to reach the incident location as fast as possible. Relevant studies demonstrated that if the serious wounded people can be rescued in 30 minutes, the survival rate is 80%, and this number reduce to 40% and 10%, respectively, when the rescued time is 60 minutes and 80 minutes[1]. However, the stochasticity of call demand in space and time, the limited supply resources of and the traffic congestion are all disruptive factors that bring tremendous challenges for quality emergency service.

In general, the whole process for the emergency vehicles responding to the incident can be divided into the following steps: receive the emergency call, dispatch the emergency vehicles, determine the travel route, and transport the patient or injured people to the hospital or aid station. Correspondingly, the base location, the dispatch strategy, the route planning and preemption policy and control for emergency vehicles should be carefully designed to enable EVs to respond to the emergency call more quickly, reliably and safely. Over the past 40 years, considerable amount of research has been devoted to address each of these issues. Moreover, with the emerging of connected vehicle, autonomous vehicle and big data technology in recent years, several new trends have become evident in the field of emergency vehicle operation research. This paper will focus on the important research work in the past, conclude the state of the art and point out the promising future direction in this research field.

II. Location problems for emergency vehicles

Generally, the emergency vehicles have limited coverage range to ensure the timeliness to response to call (9 minutes in general). Thereby, the decisions of EV locations are crucial to improve the service level to the population. The location problem for emergency vehicle seeks to find the optimal standby sites and the number of emergency vehicles allocated for each site to maximize a predetermined performance indicator. This section will provide a detailed review of the models proposed in literature to address the location problem. To better illustrate the interrelationships between these

*Resrach supported by Hong Kong Scholars Program.

Lili Lu is with the Department of Logistics and Maritime Studies of Hong Kong Polytechnic University, (corresponding author to provide phone: +8618758362393; e-mail: lulili_seu@163.com).

Shuaian Wang is with the Department of Logistics and Maritime Studies of Hong Kong Polytechnic University. e-mail:hans.wang@polyu.edu.hk

models, we grouped these models into three categories according to the characteristics of the variables and assumptions made in these models. They are deterministic models, the probabilistic models and the dynamic models. Figure 1 shows the categories of the models for location problems which will be discussed in this section.

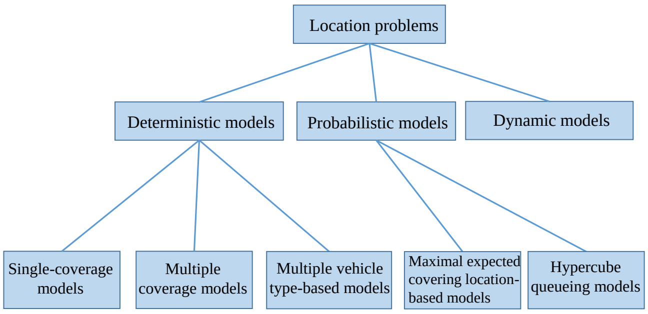

Figure 1. The categories of the models for location problems for emergency vehicles

It should be noted that most of the deterministic models and the probabilistic models address the location problems at the tactical level, aiming at deploying the emergency vehicles optimally to satisfy the call demand during a period of time, while the dynamic models address this problem in the operation level to position the vehicle in real time to serve the call adequately.

A. *Deterministic models*

In deterministic models, the inputs (e.g., demand of zones, the travel time and the service time) and outputs (e.g., locations of vehicles, the fleet size) are all deterministic. These models are extensively studied in literature due to convenience in modeling and designing solution algorithm compared with other models (e.g., probabilistic models and dynamic models). In this section, we further categorize the deterministic models into three types, i.e., the single-coverage models, the multiple-coverage models and multiple vehicle type-based models. We will discuss each of them in detail.

The single-coverage models seek to deploy the emergency vehicles optimally so that most of the demand zones or the calls can be covered by at least one emergency vehicle. In this line, Toregas et al. first proposed a set covering model to minimize the number of vehicles needed to cover all zones [2]. They assume all demand points are of equal importance, and propose a static covering model. The objective is to minimize the total cost. Church and ReVelle [3] proposed a new model that aims to locate a fixed number of vehicles to cover the maximum number of calls.

The single-coverage models assume that a vehicle is always available upon arrival of an emergency call, which does not allow busy periods (e.g., the period when a vehicle assigned for a call). To reduce the possibility of failure to serve a call due to busy periods, the multiple-coverage models are proposed to ensure the call and regions can be covered by multiple vehicles simultaneously. Daskin and Stern [4] first proposed a two-layer model to ensure multiple coverage. The first-layer model seeks to minimize the number of vehicles needed to cover all demand zones. Based on the solution set of the first-layer model, the second-layer model seeks to choose a particular solution which maximizes the number of zones with multiple vehicles. Eaton et al. [5] and Hogan and ReVelle [6] extended this research by introducing a constraint to indicate the number of demand in each zone and a constraint to indicate the total number of vehicles that can be allocated, respectively. Revelle et al. [7] extended the maximal covering model to the case where facility sites may not be used to cover their own zones. They developed a multiple objective programming problem to simultaneously maximize the covered demand and the number of vehicle sites that are at least double covered. This problem was also addressed based on the concept of double coverage, where a proportion of the demand can be coved with a prescribed time frame (S) and all the demand can be covered by another prescribed time frame S' ($S' > S$). The representative research can be found in Gendreau et al. [8], Doerner et al. [9] and Laporte et al. [10].

The above mentioned studies ignores the fact that vehicles of several types are needed to dispatch to the scene for some incidents (e.g., ambulances and fire engines) simultaneously. In North-American cities, the emergency medical services typically include two type of units, the basic life support (BLS) units and advanced life support (ALS). The BLS is assured by firemen trained as paramedics while the ALS is coved by ambulance. To deploy different type of emergency vehicles optimally, Schilling et al. [11] extend the maximal coverage model to deploy both ALS and BLS vehicles to maximize the amount of demand that is covered by both types of vehicles. ReVelle and Snyder [3] proposed a multi-objective programming problem to simultaneously maximize the call coverage for fire and call coverage for ambulances. A solution algorithm is developed to find the solutions that are "efficient" in the sense that there is no other solution which is better on both objectives. Other research seeks to address the multi vehicle-type deployment problem by considering coherent covering location, in which the ALS vehicles can provide ALS and BLS service while BLS vehicles provide only BLS service (see e.g., [12, 13]).

B. *Probabilistic models*

In deterministic models, the demand of zones, the travel time and the service time are all deterministic. Thereby, they only deal with idealistic situations where the uncertainties of the system induced by the variations of these factors are ignored. These models may not be robust to ensure the system performance when applied in reality. The probabilistic models address this problem by considering the uncertainties in traffic demand, travel time and/or the service time. According to how these uncertainties are captured in the models, the probabilistic models can be divided into maximal expected covered location-based (MECL) models and hypercube queueing models. The MECL models seek to find the set of vehicle locations that maximizes the expected coverage considering vehicle's availability. The hypercube queueing models characterize the busy fraction of vehicles and other system performance measures by leveraging queueing models. They are able to model the stochasticity of travel time and the service time.

Daskin [14] first proposed a MECL model to locate a given number of vehicles in order to maximize the expected coverage. It assumes that each ambulance has the same and independent probability (called the busy fraction) of being busy and is unavailable to answer a call. To relax the assumption of uniform busy probability, ReVelle and Hogan [15] modified the MECL model assuming the local busy fractions depending on both the number of available EMS vehicles and the aggregated level of demand within each local area. Sorensen and Church [16] present a probabilistic location model for EMS planning which integrates the area-specific busy fraction.

MECL model assumes that the busy/idle statuses of different vehicles are independent. This may not be true in practice as if a vehicle in one location is busy, it is also likely that another vehicle in the same location is busy. The hypercube queueing models are able to relax the assumption of independent busy probability. They assume the service time and the travel time follow certain distributional function, and uses the queue theory to track the status of each vehicle (e.g., busy or idle). Larson [17] introduced a hypercube queueing model in which the call arrival process is assumed to be a Poisson process and that the service time for each call is exponentially distributed. He studied the steady-state busy fractions, loss probabilities, average response time, and expected coverage, for any fixed configuration of facilities. To reduce the computational complexity of the hypercube queueing model developed before, Larson [18] developed a hypercube approximation model. It incorporates a so called "Q-factors" to relax the independency assumption of vehicle busy probabilities. Burwell et al. [19] extend this model to capture the preferences for EMS vehicles for an incident.

C. *Dynamic models*

Most of the above mentioned models belong to the class of the static strategy which only concerns the determination of the base stations and the fleet size assigned to an incident. They assume that after a service is finished, the vehicle will return to its designated base. However, the call demand may vary significantly in space and time. The method to reposition the vehicles according to the variation of demand to ensure adequate service in the busy area is of core interest to decision-makers. The dynamic models seek to find effective strategies to address this problem. Gendreau et al. [20] proposed the first ambulance relocation model that considers the dynamic nature of the demand and the relocation costs. A tabu search metaheuristic is proposed to solve the model in parallel to reduce the computational time. Moeini et al. [21] extended this model considering that only some demand zones really require double coverage. Andersson and Värbrand [22] introduced a different ambulance relocation model that maximizes the system's capability of fulfilling future demands. The authors developed a solution algorithm (labeled DYNAROC) to determine the best relocation plans. The main idea of this solution algorithm is to minimize the maximal travel time required to perform a relocation. Naoum-Sawaya and Elhedhli [23] proposed a two-stage stochastic programming approach that minimizes the cost related to vehicles' relocation, as well as the cost associated to demands that cannot be served within the prescribed delay.

III. Dispatching problems for emergency vehicles

Dispatching decision determines which emergency vehicle will be sent to respond to a call based on the nature and location of calls. An emergency vehicle can be dispatched after a call is received or after the vehicle becomes idle after serving a call. Dispatching strategies significantly affect the system's capability to adequately serve future demands.

The most simple and straight dispatching strategy is to send the closest idle vehicle to serve a call. However, as indicated in many literatures (see e.g., [24, 25]), such strategy may result in suboptimal solution where the performances such as average response time, coverage can be deteriorated. Indeed, Schmid [24] showed with a dynamic programming model that more flexible dispatching policies coupled with good repositioning strategies can improve the performance achieved by the nearest vehicle one. Gendreau et al.[20] proposed an integer programming problem to find the optimal dispatching strategy provided that several vehicles are available to reach the incident scene. The objective of this problem is to minimize subsequent relocation costs. Andersson and Värbrand [22] proposed a simple heuristic method to dispatch the vehicle whose dispatch will cause the smallest degradation of system capability to serve future demand. McLay and Mayorga [26] and Bandara et al. [27] seek to find optimal dispatching policies to maximize the coverage level and the patient survival using an approach based on Markov decision process (MDP). Bandara et al. [27] proposed a priority-based heuristic dispatching rule in which the closest available vehicle is dispatched for high priority calls while the less busy vehicle is dispatched to serve the less priority calls. Application show that this dispatching rule can significantly improve the patient survivability, the coverage, and reduce the average response time for high priority calls with just slight degradation of the service quality of the less priority calls. Recently, Toro-Diaz et al. [28] [29] proposed a programming problem which combined the location and dispatching decisions. A queuing approach based on the hypercube model is used to capture the dynamics. They designed a fixed preference list according to which the first available vehicle on the list will be dispatched to answer a call. Application shows that the nearest vehicle policy can improve performances of both response time and coverage for location problems. However, this policy can lead to suboptimal solution for joint location and dispatching decisions.

IV. ROUTING

After dispatching decision is made, a route need to be designated for each dispatched vehicle to reach to the accident scene or a patient. The route travel time is significantly contribute to the reaction time. Thereby, the strategy for optimal routing is important to improve the emergency vehicle service level. Panahi and Delavar [30] designed a dynamic routing strategy that enables emergency vehicles to select the routes with small travel time by leveraging the real-time information system. Jotshi et al. [31] developed an integrated dispatching and routing strategies for real-time operation. It calculates the routes to and from the accident scene, respectively, based on the real-world road networks, existing road damage, and congestion. Ardekani et al. [32] discussed two vehicle routing heuristics for inter-facility patient transfer to travel times and the violation of working hours simultaneously. Talarico et al. [33] developed two ambulance routing models considering the priority of the demand.

V. PREEMPTION CONTROL

Emergency vehicle preemption is an important and effective approach to reduce the travel time of emergency vehicle. A typical preemption strategy is to provide the additional green time for emergency vehicles when they arrive at the signalized intersection. The related studies have been demonstrated that the emergency vehicle preemption can reduce the travel time by 10%. And it is able to save up to 45s for one intersection when the traffic volume is high at the intersection [34]. The critical issues related to emergency vehicle preemption include the emergency vehicle detection, the control strategies, and the impact of emergency vehicle preemption service (EVPS) on other vehicles.

The emergency vehicle preemption is similar to the transit priority but has a higher priority. It is a rule in most places that the emergency vehicles does not need to obey the traffic rules (e.g., traffic signals) when they are on the way to the incident location. Nowadays, multiple vehicle priority systems exist in USA, Canada, Australia, Japan, etc. For instance, the FAST (fast emergency vehicle preemption system) in Japan provides the emergency vehicle the optimum routes and the priority at the signalized intersection [35]. Its main component is the signal priority system which enable the signal to turn to green when the emergency vehicle equipped with an infrared beacon pass the detector. Expect for the infrared based detector technology, the sound-based, the light-based, the radio-based and loop-detector-based equipment is also used to identify the emergency vehicle. Nowadays, with the dedicated short range communication (DSRC) technologies, the emergency vehicles are capable to connect with the infrastructure which improves the accuracy of emergency vehicle detection. Moreover, since the emergency vehicle equipped with GPS that send out the real-time location information, the arrival time of emergency vehicle can be predicted by taking the traffic state into account.

The emergency vehicle preemption system can be classified into two categories: pre-preemption and real-time preemption. Previously, the emergency vehicle preemption system is activated by pushing the button when the emergency service control center receives the request and will stay in effect with pre-designed control strategy for a certain period of time until the switch is turned off. Nowadays, the emergency vehicles are detected or tracked with the DSRC and GPS technology. When the system receives the request from the emergency vehicles, the signal scheme is adopted to allocate the additional green time to the vehicles in the same direction with the emergency vehicles. A number of control strategies have been developed with the purpose to optimize the delay of the emergency vehicles at the intersection. However, it has been realized that it is also important to reduce the adverse impact of the emergency vehicle on normal traffic to avoid the grid-lock, traffic accidents. Qin et al. [36] considered that there are two transition periods for the emergency vehicle signal control. The transition 1 period begins from the time that emergency vehicle is detected to the emergency vehicle clears the intersection, the transition 2 period the emergency vehicle signal preemption back to the normal operation. Two control strategies are proposed to optimize the control scheme and minimize the negative impact of emergency vehicles on the normal traffic. Similarly, Wang [37] proposed a multilayer fuzzy model is used to determine the degree-of -priority based on emergency vehicle preemption demand intensity and preemption influence intensity. Huang et al. [38] used the timed petri nets to design the traffic control policy for emergency vehicle preemption. Meanwhile, with the emerging of connected vehicle, the emergency vehicle signal preemption system are designed and implemented based on the Cooperative Vehicle-Infrastructure Technology [39]. Nowadays, with the emerging of the connected vehicle technology, the traffic control preemption for EVs has become more practical and effective.

VI. Conclusion

Although considerable efforts have been devoted to address the problems in both planning and operation for emergency vehicle service, potentially fruitful areas for future research remain.

First, related to the location problem, while attention have been given to address the impacts of the uncertainty and dynamics inherent to these problems, only a limited number of contributions has been made in this area and there is still need for the development of more sophisticated dynamic and stochastic approaches. Second, related to the dispatching problem, so far few research interest have been given to balance the workload of the crews and service quality. Note that the labor scheduling policies are viable, effective strategies are called for to dispatch all emergency vehicles optimally in space and time to reduce the gaps in workload. Third, for routing problem, so far, it does not consider the stochasticity of travel time induced by non-recurrent and recurrent traffic congestion. Therefore, using the traffic big data to gain insights of the traffic state at the urban scale withe the deep learning technology will the future direction in this research area. Further, effective solution algorithms are needed to reduce the computational time for routing problem for large amount of vehicle in large-scale network. Last but not the least, there are few practical and effective algorithms for traffic control at the intersection for emergency vehicles. In the future, the connected autonomous vehicle technology should be taken into consideration to design an effective EVSP.

VII. Acknowledgement

The research work is sponsored by the Hong Kong Scholars Program and the National Natural Science Foundation of China (No. 71771050).

References

[1] J. Yao, “Review on Emergency Traffic Control in Urban Area with Probe Data,” Appl. Mech. Mater., vol. 409–410, pp. 1343–1346, 2013.

[2] C. Toregas, R. Swain, C. ReVelle, and I. Bergman, “The location of emergency service facilities,” Operations research, vol. 19(6), pp. 1363-1373, 1986.

[3] C. Revelle, and S. Snyder, “Integrated fire and ambulance siting: A deterministic model,” Socio-Economic Planning Sciences, vol. 29, no.4, pp. 261-271, 1995.

[4] M. Daskin, and E. Stern, "A hierarchical objective set covering model for emergency medical service vehicle deployment." Transportation Science, vol. 15, pp. 137-152, 1981.

[5] D. J. Eaton, D.J, R. R. Lantigua, and Morgan, J. “Determining ambulance deployment in Santo Domingo, Dominican Republic.” Journal of the Operational Research Society, vol. 37, no. 2, pp.113-126, 1986.

[6] K. Hogan, and C. ReVelle. “Concepts and applications of backup coverage,” Management science, vol. 32, no.11, pp.1434-1444, 1986.

[7] C. Revelle, J, Schweitzer, and Snyder, S., 1996. The maximal conditional covering problem. INFOR: Information Systems and Operational Research, vol. 34, no.2, pp.77-91.

[8] M. Gendreau, G. Laporte, and F. Semet, “Solving an ambulance location model by tabu search,” Location science, vol.5, no.2, pp.75-88, 1997.

[9] K. F. Doerner, W. J. Gutjahr, R. F. Hartl, M. Karall, and M. Reimann, “Heuristic solution of an extended double-coverage ambulance location problem for Austria,” Central European Journal of Operations Research, vol.13, no.4, pp.325-340, 2005.

[10] G. Laporte, F. V. Louveaux, F. Semet, and A. Thirion, 2009, “Application of the double standard model for ambulance location,” In Innovations in distribution logistics (pp. 235-249). Springer, Berlin, Heidelberg.

[11] D. Schilling, D. J. Elzinga, J. Cohon, R. Church, and C. ReVelle, “The TEAM/FLEET models for simultaneous facility and equipment siting,” Transportation Science, vol.13, no.2, pp.163-175, 1979.

[12] D, Serra, “The coherent covering location problem. Papers in Regional Science,” vol.75, no.1, pp.79-101, 1996.

[13] M. B. Mandell, “Covering models for two-tiered emergency medical services systems,” Location Science, vol.6, no. (1-4), pp.355-368, 1998.

[14] M. Daskin, “A maximal expected covering location model: Formulation, properties, and heuristic solution,” Transportation Science, vol. 17, pp. 48-69, 1983.

[15] C. ReVelle, and K. Hogan, “The maximum availability location problem,” Transportation Science, vol. 23-3, pp.192 – 200, 1989.

[16] P. Sorensen, and R. Church, "Integrating expected coverage and local reliability for emergency medical services location problems," Socio-Economic Planning Sciences, vol.44, no.1, pp.8-18, 2010.

[17] R. C. Larson, "A hypercube queuing model for facility location and redistricting in urban emergency services," Computers & Operations Research, vol. 1, no.1, pp.67-95, 1974.

[18] R.C, Larson, "Approximating the performance of urban emergency service systems," Operations research, vol.23, no.5, pp.845-868, 1975.

[19] T. H. Burwell, J.P.Jarvis, and M.A. McKnew, "Modeling co-located servers and dispatch ties in the hypercube model," Computers & Operations Research, vol.20, no.2, pp.113-119, 1993.

[20] M. Gendreau, G. Laporte, and F, Semet, "A dynamic model and parallel tabu search heuristic for real-time ambulance relocation," Parallel Computing, vol.27, no.12, pp.1641-1653, 2001.

[21] M. Moeini, Z. Jemai, and E. Sahin, "Location and relocation problems in the context of the emergency medical service systems: a case study," Central European Journal of Operations Research, vol.23, no.3, pp.641-658, 2015.

[22] T. Andersson, and P. Värbrand, "Decision support tools for ambulance dispatch and relocation," Journal of the Operational Research Society, vol. 58, no.2, pp.195-201, 2007.

[23] J. Naoum-Sawaya, and S. Elhedhli, "A stochastic optimization model for real-time ambulance redeployment," Computers & Operations Research, 40(8), pp.1972-1978, 2013.

[24] V. Schmid. "Solving the dynamic ambulance relocation and dispatching problem using approximate dynamic programming," European Journal of Operational Research, vol. 219, no.3, pp.611-621, 2012.

[25] D. Bandara, M.E. Mayorga, and L.A. McLay, 2014. Priority dispatching strategies for EMS systems. Journal of the Operational Research Society, 65(4), pp.572-587.

[26] L. A. McLay, and M. E. Mayorga, "A model for optimally dispatching ambulances to emergency calls with classification errors in patient priorities," IIE Transactions, vol. 45, no.1, pp.1-24, 2013.

[27] D. Bandara, M.E. Mayorga, and L.A. McLay, "Optimal dispatching strategies for emergency vehicles to increase patient survivability," International Journal of Operational Research, vol. 15, no.2, pp.195-214, 2012.

[28] H. Toro-DíAz, M.E. Mayorga, S. Chanta, and L.A. Mclay, "Joint location and dispatching decisions for emergency medical services," Computers & Industrial Engineering, vol. 64, no.4, pp.917-928, 2013.

[29] H. Toro-Díaz, M.E. Mayorga, L. A.McLay, H.K. Rajagopalan, and C, Saydam. "Reducing disparities in large-scale emergency medical service systems," Journal of the Operational Research Society, vol. 66, no.7, pp.1169-1181, 2015.

[30] S, Panahi, and M. R. Delavar, "Dynamic shortest path in ambulance routing based on GIS," Int J Geoinform vol. 5, no.1, 13–9, 2009.

[31] A. Jotshi, Q, Gong, R. Batta. "Dispatching and routing of emergency vehicles in disaster mitigation using data fusion," Socio-Econ Plann Sci, vol. 43, no.1:1-24, 2009.

[32] L. H. Ardekani, D. Haight, A. Ingolfsson, M. Salama, M. Stanton, "Scheduling and routing ambulances that provide inter-facility patient transfers Technical Report," Working paper, 2014.

[33] L. Talarico, F. Meisel, K Sorensen. "Ambulance routing for disaster response with patient groups." Computer Operation Research, vol.56, 120-33, 2015.

[34] Y. Wang, Z. Wu, X. Yang, and L. Huang, "Design and implementation of an emergency vehicle signal preemption system based on cooperative vehicle-infrastructure technology," Adv. Mech. Eng., vol. 2013, 2013.

[35] N. Koike and T. Fukai, "Study on Signal Priority. Systems for Vehicles in Emergency Dispatch," IFAC Proc. Vol., vol. 36, no. 14, pp. 197–202, 2017.

[36] X. Qin and A. M. Khan, "Control strategies of traffic signal timing transition for emergency vehicle preemption," Transp. Res. Part C Emerg. Technol., vol. 25, pp. 1–17, 2012.

[37] J. Wang, W. Ma, and X. Yang, "Development of Degree-of-Priority Based Control Strategy for Emergency Vehicle Preemption Operation," Discret. Dyn. Nat. Soc., vol. 2013, pp. 1–10, 2013.

[38] Y. S. Huang, J. Y. Shiue, and J. Luo, "A traffic signal control policy for emergency vehicles preemption using Timed Petri nets," IFAC-PapersOnLine, vol. 28, no. 3. pp. 2183–2188, 2015.

[39] Y. Wang, Z. Wu, X. Yang, and L. Huang, "Design and implementation of an emergency vehicle signal preemption system based on cooperative vehicle-infrastructure technology," Adv. Mech. Eng., vol. 2013, 2013.