Abstract — The concept of vehicular networks requires quick and reliable services due to the nature of the road environment. Vehicle to Infrastructure (V2I) communication should provide consistent real-time response to be useful for the driver. The majority of the services provided by vehicular networks are delay intolerant. Road accidents or increased traffic can only worsen the situation and increase the overall delay. Many services of such kind have no value once the deadline has passed. In our research, we propose a robust recourse allocation model which decreases the number of service requests that miss the deadline without compromising the Quality of Service (QoS). The model includes a Load Balancer component that allocates the arriving tasks between neighboring Roadside Units prior the internal allocation to the virtual machines (VMs). To decrease a deadline missing rate and, therefore, increase the QoS we propose a probabilistic allocation model based on historical data of the tasks completion times. Historical data stored at the edge and updated through the cloud. Task allocation directly depends on the calculated probability of the task to meet its deadline in all of the neighboring Roadside Units. Additionally, different scheduling policies within the Roadside Unit have been considered to increase robustness. Simulation results demonstrate that proposed allocation model provides 40 % decrease in task deadline missing rate compare to conventional Roadside Unit architecture.

Index Terms—Roadside Unit, Vehicular Network, Edge Computing, V2I.

Introduction.

Recent advancements in communications and networking stimulated a rapid development of vehicular networks technology. The field of Vehicle to Everything (V2X) communications got the attention of the Federal Communications Commission which already reserved the 5.850 to 5.925 GHz frequency band specifically for V2X. Two types of vehicular communications can be distinguished. The first one is V2V or Vehicle to Vehicle communication. The second type is defined as Vehicle to Infrastructure communication (V2I). V2V and V2I supposed to complement each other to provide reliable vehicular communication. In reality, considering the nature of the vehicular movement, V2V is pretty hard to implement, and V2I draws the majority of work to itself. The United States Department of Transportation (DoT) names V2I communications the next generation of Intelligent Transportation Systems [1].

Infrastructure in V2I is represented by a special Roadside Unit or, as we are going to call it in this article, a Base Station. A Base Station is a stationary device that is installed on the side of the road and possesses some computational power as well as communication capability [2]. Therefore, every vehicle in the network is able to access additional recourses provided by the Base Stations. Wrong Way Driver warning, Cooperative Forward Collision Warning, Lane Change warning, Weather reports, and other services now can be provided to moving vehicles in a real-time fashion. According to such real-time nature, the majority of the services in V2I systems are delay sensitive and require a very short response time. Moreover, after a certain period of time, the execution of a delay sensitive task loses its value.

Deadline is the main execution constraint, and it is directly dependent on the End-to-End Delay. End-to-End Delay in V2I communication is an accumulation of three components: the uplink delay, processing delay, and the downlink delay. Unfortunately, conventional cloud architecture increases the latency problem. Communication with distant servers only contributes to the overall delay time. One approach currently used is edge computing concept. The computational power of the edge eliminates the delay produced by communication with the cloud. However, network delay is not the only source problematic point within a vehicular network. In V2I communications Base Stations can have different computational power and communication medium (e. g., wireless, fiber, optic). Such heterogeneity notably contributes to the overall delay. Another complication can occur during the emergency situations like a natural disaster or road accidents as well as during the rush hours. A rapid increase of service requests can affect the response time significantly. Therefore, heterogeneity and oversubscription can notably decrease Quality of Service (QoS) in V2I systems.

The major challenge for the vehicular Ad-Hoc networks is an efficient resource allocation which will decrease the deadline missing rate and increase the robustness of the network. There are several approaches have been proposed. [3] proposes to remove unpromising service requests from the waiting queue before scheduling as well as to balance upload and download requests. [4] suggests transferring delay tolerant requests to the neighboring Base Stations. It also proposes a scheduling algorithm which prioritizes delay-sensitive, small, popular tasks over the delay tolerant, large, unpopular tasks. [5] suggests special models of local resource reservation and reallocation based on the priority of services.

The specific question our research addresses in this article is how to allocate the tasks arriving at an oversubscribed Base Station in a V2I network so that the number of tasks missing their deadline is minimized. We propose a probabilistic resource allocation model that assures robustness of a V2I service to satisfy QoS for the end-users in presence of heterogeneity and fluctuating request arrival in the system. Previous works mostly concentrated on prioritizing the service requests to define allocation algorithm. Our model uses historical data to calculate probabilities of certain task types to meet the deadline in respective Base Stations and allocates according to the calculated probabilities.

In summary, the main contributions of this article are:

• Defining robustness for the proposed model with a presence of heterogeneity along the Base Stations and stochastic task arrival rate.

• Defining robustness of a task type as its probability to meet a required deadline in respective Base Station.

• Proposing an efficient task allocation algorithm that considers robustness the task type and the level of the Base Station subscription.

• Providing heuristics to define task type robustness across the Base Stations.

The rest of the paper is organized as follows. Section 2 introduces the scenario and assumptions made to describe the model. Section 3 includes the problem statement and the formulation of the proposed model. Section 4 discusses probabilistic allocation approach. Sections 4 and 5 present heuristics and simulation. Experimental stage and performance evaluations particularly described in section 6. Finally, section 7 wraps up the article with a conclusion.