Abstract—Several challenges in traffic control in route guidance system causes increasing number of vehicles to transport goods and people in our society. The concept of autonomous agents fits most actors in transportation systems: the traffic, the expert, the driver. More so, traffic signals and intersection can also be regarded as an autonomous agent. Though, there are increased number of agents, typical agents make response to changes in their environment and are highly self-adaptive, but create an unpredictable collective pattern, and response in a highly coupled environment, most challenges for standard techniques are created by this domain in route guidance system from multi-agent systems such as reinforcement learning and self-adaptive. This research has two main goals in route guidance system: first, to present problems, methods, new approaches; and second, open problems and challenges are highlighted so that future research in route guidance system using multi-agent systems will be able to address them.

Keywords-self-adaptive multi-agent system (SAMAS), route guidance system (RGS), multi-agent reinforcement learning (MARL), shortest path problem (SPP)

I. INTRODUCTION

Route Guidance System (RGS) is one of the several type of traffic information system [1] to offer routes that solves the traffic problem, which provides an optimum route solution and certain traffic information prior to the trip to help drivers to get to their station as fast as they can. Nowadays, most route guidance systems compute the best and shortest time or distance for the driver [2]. RGS integrates the some technologies like information technology, networks, communication, and computer science etc. All the drivers can be received the dynamic information and optimal path; this can result in to congestion transfer with the application of on-board unit with the view of GPS and GIS [3]. Khaled [4] also analyse the result of traffic information and services due to information dissemination strategy dynamic traffic components and behaviours. In this paper, emphasis was made on user behaviour in RGS and the significance of traffic information. A new feedback strategy that instils the available routing strategy advantages was proposed by Wang [5]. Also, demonstration of decision support system which was applied in the free flow of traffic for motorways and urban roads was also proposed by Schutter [6]. The system takes a view of the multi-agents approach. Wang Fei-Yue and LIU Xiao-Ming [7] make an extensive study on coordination control and traffic flow with a view on the agent based technology. An entity that allow for a sensory input from an environment is known as an agent; and makes decision autonomously, while executing actions that have effect on that same. This paper focuses on two kinds of multi-agent based in RGS, which are MARL and self-adaptive. Using this idea is meant to give the driver optimum direction on their trip. The remainder of this paper is organized as follows. First, we start with the overview of self-adaptive and MARL in route guidance system in Section II. Next, in Section III, we propose a self-adaptive and multi-agent reinforcement learning model and a formal problem definition. Then in Section IV, empirical study is presented, followed by Section V where simulation results and experimental comparison are analyzed with several samples. Finally, discussion was made on the conclusion and future works in Section VI.

II. RELATED WORK

The modelling of urban traffic systems based on Multi-agent system was studied by Chabrol et al. [8], [9]. Reserving time and space is a new approach in the urban road network for vehicles at intersection and junction which was studied in [10]. Moreover, multi-agent approach is must better for traffic detection and management has analyzed by most literature [11] and [12] and most agent based application was concentrated on modelling and simulation. However, in recent decade studies centralized on theories and implementation technologies of RGS. Due to the variety of road network environment area applications, the corresponding route guidance systems are more complex, and present many new characteristics. Therefore the vehicle driver needs to use a RGS based on self-adaptive and multi-agent systems. Any changes in the environment could perceive by the self-adaptive and multi-agent system because the self-adaptive and multi-agent system is a complex system as it has ability to change structure and behaviour itself [13], [14]. However, in real world applications, there must be efficiency and research work to be done. The goals of this research work is to make multiple agents to learn exact behaviors in a dynamic environment using reinforcement learning in route guidance system. Multi-agent reinforcement learning is entirely a new area of study in artificial intelligence, and there is a need to know if it is most likely for MARL systems to learn having cooperation between agents.

A. Multi-agent Reinforcement Learning (MARL)

Multi-agent system (MAS) is a convenient system for the environment that in cooperate interactions between all kind of goals and proprietary knowledge by people or organizations.
A process were by an agent enhance its behavior in a domain via experience is known as Reinforcement Learning (RL). This is based on the ability of an agent performing an action, in which should be reinforced by a reward if favorable result is produced by the action, and weakened (punished) if unfavorable result is produced. By comparing RL with other learning methods, One of Its most important advantage is that; in every minute, prior knowledge is required of the domain, since we don’t need target of desirable output for each input in the cause of generating training sample. The system can be increasingly adjusted by RL algorithms like Q-learning [19, 23] in the domain of many favourable outcomes when a feedback judgment (good or bad) is provided on the system’s output for a given input. Due to this reason, RL has been identified as on among some of the extensively used learning method in MAS [17, 18, 19]. Robot soccer games [16, 20] domain are said to be some of the known application of RL in which the outcome of the game could be you win or you lose with a feed back to train the soccer team. More so, application of RL to other problems are, learning agent coordination mechanism [19], learning to schedule multiple goals [15], setting right prices in competitive marketplaces [21] and in dealing with malicious agent in a market based MAS [22].

B. Self-adaptive and multi-agent system

Gowri et al. [25] studied to support self-adaptation in multi-agent system by using an ADELFE methodology. Also, Goldsby and Cheng [26] proposed to adopt some self-adaptive systems functions. Controlling the real-time system using self-adaptive multi-agent system was proposed by Gleizes et al. [27]. Self-adaptive multi-agent system has been studied as a control tool used method in dynamic process [28].

III. PROPOSED SOLUTION

We shall propose the scheme of multi-agent system in route guidance system communication model in this section. Then, the problem definition and self-adaptive and multi-agent reinforcement learning have been proposed.

A. Proposed multi-agent model

Figure 1 shows the scheme of MAS and RGS communication. The shortest path from the current position of vehicle to destination zone shall be calculated by the RGS in other to minimize the traveling distance or the time for vehicle drivers. Destinations, with the help of installed agent in the intersection and junction were used by Dijkstra algorithm to calculate the nearest distance. The system we use the updated information of the vehicle to find the shortest path between the current vehicle positions to the destination location for vehicle drivers based on their conditions.

B. Problem definition

A presentation of the road direction graph is by G=(V, E) is a directed route guidance system which is based on a set of ’v’ nodes (V) and ’e’ directed edges (E) and an electronic map. Each R(o, t) edge represents the cost and ‘o’ is a start node and ‘t’ is a last node connected to the R(o, t). Consider a directed graph, G, which is made up of a set of s nodes and a set of e directed edges. Set a number of edges as cost (C) file table. Also, G file is filled in data file.

If O = {o1, o2, …, on}, T = {t1, t2,….,tn} then

\[ R_{(o, t)} = \{ R_{(o1,t1)}+R_{(o2,t2)}+…+R_{(o,t)} \} \]

R(o,t) consists of the sum of all edge costs that is involve in the road network path. Therefore, due to the trip origin (o) node and last node (t), solving this issue can be through optimization problem that is based on real road network which is defined bellow:

\[ R_{(o,t)} = \min \sum_{i=1}^{t} \sum_{j=1}^{t} a \ R_{(i,j)} \]  \[ a = 1 \hspace{1cm} A \text{ link between } i, j \]  \[ a = 0 \hspace{1cm} \text{Otherwise} \]

Therefore:

- i and j are represented as the movement of the current state and directions.
- \( R_{(o,t)} \) is represented from first node, ‘o’ to the last node, ‘t’ as the shortest path
- Links are independently from each other.

C. Proposed self-adaptive and multi-agent reinforcement learning (SAMARL) Model

The self-adaptive and MARL (SAMARL) approach is model based on Figure 2 and equation (1). The self-adaptive and MARL performance elements are agent, environment, learning, policy, action, and reward. More so, input state was distinguished by an agent. Policy function instill inside the multi-agent model that will be used for decision-making process will determine the action of the agent. MARL must learn on a stated policy for a given state in other to uncover the best action. The Policy will plan the current state and directions.

![Figure 1. The scheme of MAS and RGS communication](image-url)
required so as to receive rewards and send actions among equipped vehicles and the environment. Installation of the agent takes place in each intersection and junction which will communicate with its environment by collecting new information about the status of the road in the intersection and forward the status of the road to vehicles. RL is a problem of learning interaction with the agent’s environment. A policy search is executed by the learner, the use of critic to aid the reward inputs as guides to improve the policy in some solutions (Figure 1).

Multi-agent reinforcement learning in RGS

We will address Q-learning within reinforcement learning, a variant in which an agent increasingly computes from its interaction with its domain, a table of expected aggregate future rewards, with values discounted as they extend into the future. An environment is made up of several states. When an action takes place a new state is observed, this is because an action has an effect on the state. Reinforcement is provided when an action is completely taking. It’s not necessary that the action is a deterministic one, i.e. we let this action to have probabilistic outcomes. Actions that are to be taken in any given state are detected by Policy. Optimal policy of an agent to actions that maximize the long term reinforcement signal is known as mapping of states. The optimal value mapping can also be defined. Also, the action that is predicted by the value mapping is performed as being the optimal one in a given state.

IV. EXPERIMENT AND RESULT

Simulation has been performed on different road networks in this Section. It is made up of 5 to 20 nodes with various edges from 10 to 50 edges. The new approach results have been proposed in Section III and are compared with dijkstra algorithm results having three cases network graphs (the graph of 16 nodes and 24 edges as Case 1, 12 nodes and 23 edges as Case 2, and 15 nodes and 30 edges as Case 3).

A. Case 1

1) Using Dijkstra algorithm

The main objective of Figure 3 is to calculate the shortest path from node 1 to node 11. This is drawn (shortest path) with the use of Matlab software. These Matlab commands uses dijkstra algorithm to find the shortest and optimal path. The result will find the minimum distance between node 1 and node 11 that is shown in Figure 3A. Therefore, the new optimal path is Node 1 → Node 4 → Node 5 → Node 11 in Figure 3A.

2) Using Proposed model

The route from node 4 to node 5 is shown in figure 3b, that is blocked (or cost = 99) due to traffic congestion or accident; in this case, some of the connected routes are changed automatically in the new shortest and optimum route path. Therefore, the new optimal path in Figure 3B is Node 1 → Node 2 → Node 5 → Node 11. However, the distance between Node 4 and Node 5 tends to ∞ and the previous shortest path in Figure 3A, 15 is changed to 19 in Figure 4A.

B. Case 2

1) Using Dijkstra algorithm

Also, the objective of Figure 4 is to calculate the shortest path from node 1 to node 23. This is drawn (shortest path) with the use of Matlab software. These Matlab commands uses dijkstra algorithm to find the shortest and optimal path. The result will find the minimum distance between node 1 and node 23 which is shown in Figure 4A. Therefore, the new optimal path in Figure 4A is Node 1 → Node 2 → Node 6 → Node 10 → Node 11 → Node 12.

2) Using Proposed model

Figure 4B shows the edge among node 2 and node 6 that are congested or blocked (or cost = 99) because of traffic congestion or accident; in this case, some connected edges are changed automatically in the new shortest and optimum route path. Therefore, the new optimal path in Figure 4B is Node 1 → Node 5 → Node 6 → Node 10 → Node 11 → Node 12.
A. Case 3

1) Using Dijkstra algorithm

The main objective of Figure 5 is to calculate the shortest path from node 1 to node 15. This is drawn (shortest path) with the use of Matlab software. These Matlab commands uses dijkstra algorithm to find the shortest and optimal path. The result will find the minimum distance between node 1 and node 15 that is shown in Figure 5A. Therefore, the new optimal path in Figure 5A is Node 1 → Node 2 → Node 6 → Node 10 → Node 13 → Node 15.

However, the distance between Node 2 and Node 6 tends to ∞ and the previous shortest path in Figure 5A, 27 is changed to 28 in Figure 5A.

2) Using Proposed model

The route from node 2 to node 6 is shown in Figure 5B, that is blocked (cost = 99) due to traffic congestion or accident; in this case, some of the connected routes are changed automatically in the new shortest and optimum route path. Therefore, the new optimal path in Figure 5B is Node 1 → Node 5 → Node 6 → Node 10 → Node 13 → Node 15.

However, the distance between Node 2 and Node 6 tends to ∞ and the previous shortest path in Figure 5A, 27 is changed to 28 in Figure 5A.

However, consider the Figure 3, 4, in the proposed model (SAMARL), in Section III and Table I show the results obtain by using proposed approach in experimental cases. For simulation result, we assumed that there are 50 equipped vehicles in two groups and that 25 of them were guided by the SAMARL model and dijkstra algorithm as group 1 while the remaining 25 were not guided by SAMARL model but used only dijkstra algorithm as group 2. Also, the traveling distance (km) and the average speed (km/h) of two groups in Case 1 are 150 km and 75 km/h, the traveling distance (km) and the average speed (km/h) of two groups in Case 2 are 310 km and 75 km/h and the traveling distance (km) and the average speed (km/h) of two groups in Case 3 are 410 km and 75 km/h. Considered in Table I is the average travel time (AveTT1) of group 1 and the average travel time (AveTT2) of group 2 in Case 1 are 610 and 700 seconds respectively. Also, in Case 2, the average travel time (AveTT1) of group 1 and the average travel time (AveTT2) of group 2 are 1100 and 1400 seconds respectively. In Case 3, the average travel time (AveTT1) of group 1 and the average travel time (AveTT2) of group 2 are 1600 and 2100 seconds respectively. Therefore, with regard to the saving time formulation i.e. SveTime = (AveTT2 - AveTT1 / AveTT2)*100% the saving time in group 1 and group 2 are 12%, 21% and 24% respectively.

### Table I - Using SAMARL evaluation

<table>
<thead>
<tr>
<th>Case</th>
<th>Traffic congestion</th>
<th>Group 1 AveTT1</th>
<th>Group 2 AveTT2</th>
<th>SveTime %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4-5</td>
<td>610</td>
<td>700</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>2-6</td>
<td>1100</td>
<td>1400</td>
<td>21</td>
</tr>
<tr>
<td>3</td>
<td>2-6</td>
<td>1600</td>
<td>2100</td>
<td>24</td>
</tr>
</tbody>
</table>

The following are the application of the roles of multi-agents learning in RGS:

- Learning to cooperate with other agents;
- Finding the actions to be given to the agent;
- Learning to schedule multiple goal;
- Route request are been received from the driver vehicles;
- The optimal path are been calculated and forwarded to drivers;
- Collect or forward information to other agents.

V. CONCLUSION AND FUTURE WORK
This study includes a new approach (SAMARL) for finding the optimal path in route guidance system. We highlight the importance of using features of the self-adaptive and multi-agent reinforcement learning to improve route guidance system. In all experiment cases, average saving time of dijkstra algorithm and has better performance results. The shortest road graph are been find by the simulation of the SAMARL model cases from 5 to 50 nodes. Consequently, the future work of this paper shall be looking at some investigation that could be put in to consideration in the areas mentioned below: 1- Employing different learning techniques for different aspects by developing some hybrid learning architecture for e-commerce activities. 2- Improving of negotiation decision making by studying the task of the way to put together the learned model of your opponent. 3- An Experimental decision tree learning with some other preferable real-world training data. (4) Improving agent performance by Investigate the possible application of improved decision tree learning by making it adaptive to the changing domain. Subsequently, in our future work we shall optimize the architecture, and the evaluation of the system reliability and effectiveness. Therefore, when the installed agent (in intersection) reports the road traffic, it is then reported in a specific route of the past shortest path to vehicle drivers that they direct to the next priority of shortest path.

VI. ACKNOWLEDGMENT

The authors wish to thank Ministry of Higher Education Malaysia (MOHE) under Fundamental Research Grant Scheme (FRGS) and Universiti Teknologi Malaysia under the Malaysia (MOHE) under Fundamental Research Grant.

REFERENCES