A Decentralized Reinforcement Learning System for Patrol Routing

Avijit Roy, Nisheeth Srivastava

avijit@iitk.ac.in, nsrivast@iitk.ac.in

Indian Institute of Technology, Kanpur

Abstract

We model the placement of first-responder patrol vehicles on a city map as a multi-agent reinforcement learning problem, where individual agents learn desirable locations for parking based on dynamically updated geo-localized emergency call records. The model is able to outline reasonable patrol locations and routes, adapting to changes in the geographical pattern of call locations, and permits optimization of routes accommodating fuel economy and other cost-based concerns into account in a principled way. We also present an actual patrolling system we have developed around this model, and present simulated results comparing it's performance vis-a-vis centroid-based patrol location prediction and judgments made by humans.

Video Presentation: https://youtu.be/nsMXzXfeEmM

Introduction

It is a common adage that, when you need the police immediately, they tend to be just five minutes away. Emergency response services (henceforth EMS), therefore, are tasked with responding as quickly to emergencies as possible given existing resources. Since existing resources are increasingly constrained, EMS have to optimize first response resources intelligently.

The problem of optimizing first responder resources is most commonly studied in the context of police responses. In the policing context, first responders are usually patrol officers. Researchers commonly address this problem as one concerning the optimal design of patrol routes, being in turn a subset of the more general multi-agent patrol routing problem (Almeida et al. 2004).

It is commonly recognized that deterministic strategies are not good fits for the underlying problem in policing contexts. This is for three reasons. One, deterministic fixing of patrolling schedules permit criminals to reflexively schedule crimes coordinating with the patrol schedules. Two, these approaches are not adaptive to situational changes on the ground affecting the number and distribution of potential first-responders. Three, the solutions proposed by such methods are frequently considered excessive and wasteful by practitioners. In any realistic setting involving motorized patrols, planners want to balance fast response times with managing fleet fuel consumption. No approaches, to the best of our knowledge, currently explicitly tackle the problem of managing the trade-off between response times and patrolling resource consumption.

Reinforcement learning appears to be a natural fit for the multi-agent patrol routing problem. In this paper, we present a decentralized reinforcement learning solution for the multi-agent patrolling problem particularized for the context of policing EMS. Our solution is simpler than previous reinforcement learning solutions for this task, highly scalable with respect to the number of patrolling units, and addresses the three concerns identified above that have bedeviled the adoption of previous proposals in real-world deployments. Notably, we use call records from a real police EMS for designing and testing our system, and present a final solution that is currently being field-tested for adoption within the same real-world EMS.

System Description

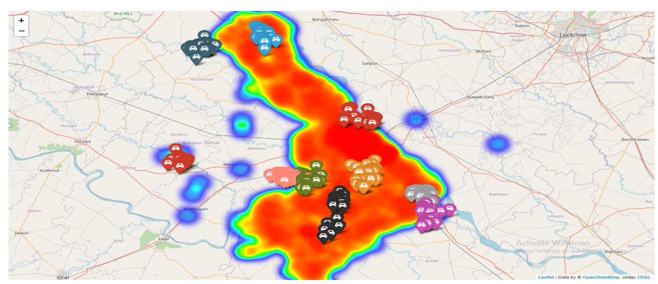

Figure 1: This figure show an overview of our Application. The markers shows locations of vehicles at different points in time. The heatmap reports density of call records at different locations in the map.

The final version of this project will link to a live deployment of the system we have created. We describe it using text and screenshots. Figure 1 provides an overview of the application.

Under the hood, we divide the entire area covered by the EMS into district-wise zones, approximately identified by rectangular bounding boxes. We create routes for different zones separately by running the algorithm once for each zone. A web-based application allows officers to monitor and update vehicle routes. We built the web-application on the MERN stack to interact with the model. We use human interaction to improve the routes over time, in the sense that our algorithm replaces its own selected patrol point with a human-generated patrol point stored in the database if the human suggestion is within a certain radius of the algorithmically proposed point.

We initially generate a route in Python using the algorithm mentioned above. We save the generated route in the database. Then we show the route and take user feedback using a web-application. We also incorporate the user feedback while generating routes for the next day.

Experimental Evaluation

We present succinct evaluations of our system on two different dimensions - its ability to generate patrol points that domain experts would find reasonable, and its ability to accommodate fuel consumption constraints into this process. These together attempt to discern whether the output of the system is realistically usable by the EMS.

User study: patrol point quality

To compare the system-generated patrol points with human-generated points, we conducted a small user study. We have taken 4 different zones with different densities of crime. Then we generate 5 types of patrolling points for each of the zones.

G1: Random points, G2: System generated points with the assumption that the fuel prices are normal, G3: Human-generated points with the assumption that the fuel prices are normal, G4: System generated points with the assumption that the fuel prices are high, G5: Human-generated points with the assumption that the fuel prices are high.

Now for each patrol point, a user rates the patrolling points out of 100 based on the factors that how much area it covers total distance covered, and response time. The users involved in this experiment were oblivious to the type of patrolling points they are rating.

Figure2 shows the comparison between the mean score obtained by different types of patrolling points from 10 separate domain experts. We can see that our system-generated points are significantly better than random patrolling points and approaching the human-generated points level. Also, when there is a fuel constraint, the system generated even better points keeping in mind the fuel efficiency.

Qualitative analysis: fuel cost optimization

In this exercise, we compare routes for a single patrol vehicle generated for three separate areas within the district by three separate mechanisms - by a human expert, by a density estimation algorithm, and by our RL-based method. Our observations are visually summarized in Figure 3.

The first column in Figure 3 documents patrol points marked by a human expert. We can see that for the sparsely populated rural area examined in the top row, the human is actually rather inefficient in placing patrol points - they believe in patrolling along the highway for the most part, while most of the crime is reported from the rural cluster a little removed from the highway. The human's judgments are better in placing patrol points for the dense urban cluster seen in the middle row, and the sub-urban area shown in the bottom row.

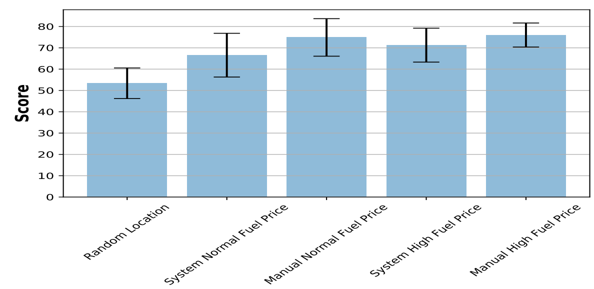

Figure 2: User evaluation of patrol points generated from different sources. Error bars represent ± 1 SD.

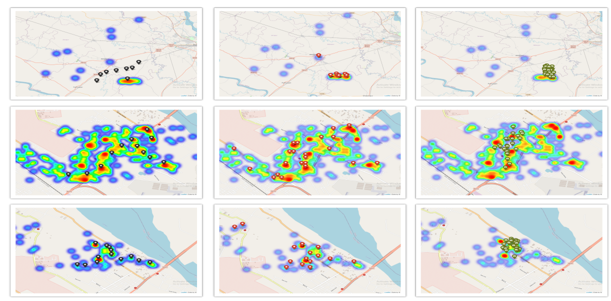

Figure 3: This shows the three Methods of Patrolling point generation. Left: Manual, Middle: DBSCAN, Right: Our Algorithm

The second column shows patrol points generated using DBSCAN, an extremely robust and powerful density estimation algorithm for spatial data (Ester et al. 1996). DBSCAN does well on the top and middle rows, but spreads the patrol points too much in the bottom row, expecting the agent to travel a long distance to patrol two points in the extreme north of the area for little profit.

The third column shows patrol points generated using our algorithm. Notice that the algorithm places patrol points really close together for the rural area, widely dispersed for the urban area, and at an intermediate level of density for the sub-urban area. Adaptive variation in plotting routes is one of the key advantages of RL-based approaches over more graph-theoretic approaches, and is evident in our results.

References

Almeida, A.; Ramalho, G.; Santana, H.; Tedesco, P.; Menezes, T.; Corruble, V.; and Chevaleyre, Y. 2004. Recent advances on multi-agent patrolling. In *Brazilian Symposium on Artificial Intelligence*, 474–483. Springer.

Ester, M.; Kriegel, H.-P.; Sander, J.; Xu, X.; et al. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, volume 96, 226–231.