Proceedings of the 9th Annual World Conference of the
Society for Industrial and Systems Engineering,
2020 SISE Virtual Conference
September 17-18, 2020

Evaluating Response Time Performance of Multiple Patrol Dispatching Using Stochastic Simulation

J Holguin-de la Cruz

Department of Industrial and Manufacturing Engineering
Universidad Autónoma de Ciudad Juárez
Juárez, Chihuahua, México

Corresponding author's Email: jholguin@uacj.mx

Abstract: Our research considers one police quadrant that integrates four patrolling zones, and evaluates the impact of response time affected by reducing the number of servers from one to ¼ th of a patrol per patrolling zone. This serving capacity reduction is based on a recently implemented police patrolling strategy, where four patrols are assigned for regular preventive patrolling and to serve calls for service of a 911 Emergency Response System. The previous patrolling and serving strategy had assigned one police patrol per patrolling zone in a police quadrant with four patrolling zones. Given the multiple vehicle strategies of delinquent operations, police departments also have been forced to adopt similar combat strategies. Based on an international reference of 3 minutes maximum response time, our focus is to analyze the waiting time in the queue of calls for service and their total response time of the new multiple vehicle patrolling and dispatching strategy by using stochastic simulation, and to compare these results with the previous deployment strategy of one patrol per zone. Data was obtained from a 911 Emergency Response System in Mexico.

Keywords: Police patrols, Emergency Response System, Multiple patrol dispatching, Response time.

1. Introduction

The new federal government initiated in December1st of 2018 has been slowly implementing the National Guard, and replacing the army to combat crime under a recently modified juridical frame to validate its intervention. However, the perceived historic and highest, national insecurity levels, according to the Perception and Victimization National Poll (ENVIPE), are 79.4% and 78.9% for 2018 and 2019 respectively (INEGI, 2019; 2020). These values represent an average increment of 12.55% from its lowest but still high value of 66.6% in 2012. These are alarming statistics that translate to have almost 8 persons out of 10 that feel they live in an insecure environment.

Due to the increased combat power of delinquency, using multiple vehicles and assault weapons, police patrolling strategies have adopted as well, multiple patrol dispatching approaches to protect themselves and to enhance the probability of contention and success. This relatively new police strategy, concentrates patrol resources of a police quadrant by single call for service event, leaving the rest of the police quadrant temporally unprotected if new calls for service are received while attending a given request for service. These incoming calls are placed in a queuing waiting system for the complete police quadrant. Transportation time is as well incremented since now the service area is the complete quadrant instead of one patrolling zone.

Being response time the most important performance parameter of a Security Emergency Response System (Piyadasun et al., 2017; van Barneveld et al., 2018), it is of fundamental importance to optimize all involved processes participating in this activity. Response time is measured from a time a call is received to the time the serving unit arrives to the location of the event (Piyadasun et al., 2017; van Barneveld et al., 2018). After a call is answered by the emergency response system, its level of urgency is identified. In our ERS of interest, which follows national operational procedures, the level of urgency is categorized using three levels: One or Red, which denote the highest level, Two or Yellow represents less urgent events, and lastly Three or Blue utilized for events with a low level of urgency. Subsequently, the patrol dispatching area contacts the patrol allocated to the particular patrolling zone where the incident is reported, or the nearest patrol available, considering that patrolling zones may not have a patrol allocated to serve the call. In the event that all potential serving patrols within a given proximity were busy, the call is placed in a waiting queue until a patrol is available. This waiting process in the queue is part of the response time parameter, and it could represent a very sensible parameter directly impacting the quality of the service provided and the degree of success in protecting the community of the law enforcement function.

Under the new patrolling and dispatching operating strategies utilizing multiple vehicles coming from consolidating all available allocated patrols to the police quadrant, the serving system passes from an average of four patrols per quadrant, to only one patrol per quadrant, and the waiting time for one server formed by a group of four patrols would be drastically incremented. Our research considers the ideal reference response time of three minutes maximum, which was provided by the National Advisory Commission on Criminal Justice (NACCJ, 1973) in the United States of America for a service level of 90%, which was also utilized by Guo et al. (2010).

2. Literature Review

There are several approaches documented in the literature for the allocation of scarce resources to Emergency Response Systems (ERS) as well as for improving performance parameters such as response time of calls for service (Green and Kolesar, 2004; Holguin-de la Cruz, 2017). Zaki et al. (1997) identify the problem of allocating a required level of resources to an ERS that minimizes the response time to or under a given time value. According to Green (1984), software has been created to determine the number of patrol units that are required to each precinct of a city during different times of the day and days of the week. The author establishes that the computer program Patrol Car Allocation Model (PCAM) represents a prominent example designed by Chaiken and Dormont (1978a, b), which has been used by at least 40 police departments. Green states as well that Markovian multiple server models were used, such as the M/M/c model applied by PCAM with non-preemptive priority classes previously created by Cobham (1954). One of the limitations of all these queueing models used was that they allocated a single patrol car per call (Chaiken and Dormont 1978a). Examples of other M/M/c models applied to police patrols are presented by Kolesar (1975), and Larson (1972; 1974). Additionally, Green (1984), states that one of the main disadvantages of queueing models is their lack of ability to consider multiple dispatches precisely.

According to Green and Kolesar (1984), a significant proportion of calls for police service in many cities, demand more than one patrol car. Likewise, Green (1984) establishes that in New York City, 30% of all calls for service required a multiple car response. Based on this challenging situation, Green (1984) developed a model that represents de multiple car allocation per call, which was based on a simpler model created by Green (1980). The newer model, identified as multiple car dispatch (MCD), is a multi-server, multi-priority queueing model where the number of servers that is assigned to each customer is a random variable, which is a function of the type of customer and the availability of servers (Green, 1984). Green and Kolesar (1984) compared the accuracy and performance of the MCD model compared with the Cobham approximation methods. In general terms, the authors state that the MCD model produced closer to observed delays of calls when all servers are busy, and later the authors extend this attribute to also use its performance to compare the patrols allocations of the MCD model with the Cobham approximations methods. However, Green and Kolesar (1984) do not offer a direct quantitative performance comparison of the MCD model with observed or real behavior of the system.

We would like to comment that we did not find recent methodologies associated with the multiple dispatch of police patrols to one call for service. However, there are other methodologies to assist in the analysis and improvement of security emergency response systems which include mathematical programming (Yang et al., 2015), agent-based simulation (Chen et al., 2017), and discrete event simulation (Wu et al., 2014).

3. Methodology

Our research approach is based on stochastic discrete simulation of the processes involved in receiving and serving calls for service in a 911 Emergency Response System in Mexico. Data includes 552 hours of continuous operation and main processing times were probabilistically characterized to firstly develop an actual scenario to reproduce actual operations identified as BP-Actual. Our second scenario integrates the international reference a response time of 3 minutes maximum probabilistically distributed for 90 % of the calls for service, which is identified as RT3M. Scenarios BP-Actual and RT3M are based on a simulation model that allocates only one patrol per call or customer. Both scenarios, BP-Actual and RT3M, model one police quadrant with four patrolling zones and consider one dedicated patrol P_{dj} per zone, where sub index d identifies the patrol number from 1 to 4, and sub index j denotes the patrol zone from 1 to 4. In addition, the police quadrant has four backup patrols B_{ij}, where sub indexes i and j denote back up patrol number and zone number respectively, which are used in a priority basis if all dedicated patrols are busy.

Subsequently, scenario 4D+4B Actual is based on scenario BP-Actual except that the simulation model allocates four patrols per call or customer where there are two groups of four patrols identified as 4D and 4B, which are dispatched as a group to serve single calls for service and are freed likewise as a group when the service has been completed. This scenario allocates the group of servers based on a priority basis, where group 4D is the group that has been assigned as fixed to one

ISBN: 97819384961-9-6

police quadrant and serves all calls for service in the quadrant if the serving group is not serving a call. If serving group 4D is busy, then serving group 4B operates as a backup resource to serve incoming calls. This backup serving group is assumed to serve the four quadrants in a police district. Due to this function of the serving group 4B, transportation times are adjusted to reflect a longer transportation distance. For simplification purposes, backup serving group 4B does not consider demand for serving the other three police quadrants as a common backup resource for the complete police district. Lastly, scenario 4D Actual is based on scenario 4D+4B Actual, except that this scenario does not have a backup serving group 4B to evaluate the quality of the service if the serving group 4B was not financially feasible.

Scenarios BP-Actual and RT3M were run for 552 hours and 10 replicates and their results were imported from previously published results (Holguin-De La Cruz, 2017). On the other hand, scenarios 4D+4B Actual and 4D Actual were run for 552 hours and 15 replicates, where averages are reported by replicate from R1 to R15. Additionally, these two scenarios were run for 552 hours, 100 replicates and results are reported as an average in R100. All scenarios were validated comparing produced outputs with observed data. Our research selected Police District 5 and Police Quadrant 1 considering that their demand represents districts with high demand for service.

4. Results

The characterization of the main processes for receiving and serving the calls for service were obtained for police quadrant one integrated in police district five. Table 1 presents the probability distributions for Police District Five for the inter arrival time estimated from patrolling zones, and for response time and patrol busy time obtained from police quadrants. From this table it can be observed that *inter arrival times* are primarily represented by Gamma (43.7%), Lognormal (29.1%), and by Weibull (25%) probability distributions respectively. Likewise, the *response time* parameter is represented mainly by Lognormal (91.6%) probability distributions. Similarly, the parameter *patrol busy time* is described by Lognormal (58.3%), and Exponential (33.3%) probability distributions.

Table 1. Characterization of Arrivals and Service of the City's ERS

Parameter	D_k[1]	Probability Distributions (95% C.I.)							Total
		Exponential	Gamma	Loglogistic	Lognormal	Normal J-T	Normal	Weibull	
Interarrival Time	D_5	1	21		14			12	48
Response Time	D_5		1		11				12
Patrol Busy Time at L[2]	D_5	4			7			1	12
Total		5	22		32			13	72

[1] = (D_k) Police District *k*
[2] = (L) Location of Event

Table 2 illustrates the simulation results by scenario and performance parameters, which include: (1) Average Number of Times Used (ANTU), (2) Average Time per Usage, and (3) Percent Average Utilization (%AU).

Table 2. Simulation Results by Scenario: District and Quadrant Averages by Performance Parameters

D_k[1]	Scenario	Parameter	Dedicated Patrols				Inventory Back up Patrols			
			P_{11}	**P_{22}**	**P_{33}**	**P_{44}**	**B_{1j}**	**B_{2j}**	**B_{3j}**	**B_{4j}**
D_5	**BP-Actual**	**ANTU**	127.38	122.23	160.10	141.78	74.78	12.95	2.80	1.75
		ATPU	24.75	25.49	24.88	24.71	25.03	24.14	22.73	27.98
		% AU	9.54	9.47	12.04	10.62	5.67	0.93	0.18	0.13
Single Patrol Dispatch	**RT3M**	**ANTU**	127.03	124.28	170.15	153.15	56.70	6.45	1.98	1.60
		ATPU	16.61	16.32	16.51	16.19	16.91	15.80	16.68	21.84
		% AU	6.34	6.15	8.45	7.45	2.93	0.33	0.09	0.09
$D_{5\ Q1}$	**4D+4B Actual**	**ANTU**	413.70				239.50			
		ATPU	37.69				37.28			
		% AU	47.05				26.91			
Group Dispatch	**4D Actual**	**ANTU**	655.80							
		ATPU	37.36							
		% AU	73.93							

[1] = (D_k) Police District *k*
[2] = (ANTU) Average Number of Times Used (From 4 Quadrants)
[3] = (ATPU) Average Time per Usage (From 4 Quadrants): Transportation Time + Service Time at Location
[4] = (%AU) Average Percent Utilization of Patrol Time (From 4 Quadrants)

As it can be observed, the single patrol dispatch scenarios have eight police patrols that primarily use the four dedicated patrols and have the valuable assistance of the four backup patrols to serve calls when dedicated patrols are busy. However, group dispatch scenario 4D+4B Actual scenario, which only has two groups or servers, drastically presents a demand concentration of 63.3% in group 4D, and of 36.7% in group 4B. In scenario 4D Actual, 100% of the demand for service in the police quadrant is absorbed by only serving group 4D or one server, which presents a significantly high serving percent average utilization of 73.93%.

According to Green and Kolesar (1984), one of the main objectives of developing the MCD model was to obtain more accurate waiting times of calls in the queue when including multiple patrol dispatch, since the Cobham approximation models significantly underestimated waiting times in the queue. In our research, we estimated these waiting times in the queue for our group dispatching scenarios 4D+4B Actual, and 4D Actual. Figure 1 illustrates the call waiting times in the queue obtained by the stochastic discrete simulation model for scenario 4D+4B Actual, where we can observe a potential waiting time of 3.3 to 10 minutes for Priority 3/Blue calls, 7 to 13.3 minutes for Priority 2/Yellow calls, and 7.8 to 14.7 minutes for Priority 1/Red calls. In this figure we can also observe values for 99.9% Confidence Intervals and average values by call priority based on 100 replicates. Similarly, Figure 2, presents the call waiting times in the queue obtained by the simulation model for scenario 4D Actual, where we can observe a potential waiting time of 52.5 to 163.8 minutes for Priority 3/Blue calls, 64.4 to 161.1 minutes for Priority 2/Yellow calls, and 70.9 to 180.4 minutes for Priority 1/Red calls. Likewise, in this figure we can also observe values for 99.9% Confidence Intervals and average values by call priority based on 100 replicates. A comparison between scenarios 4D+4B Actual and 4D Actual demonstrates how waiting time is drastically incremented when the number of servers is only one patrolling group.

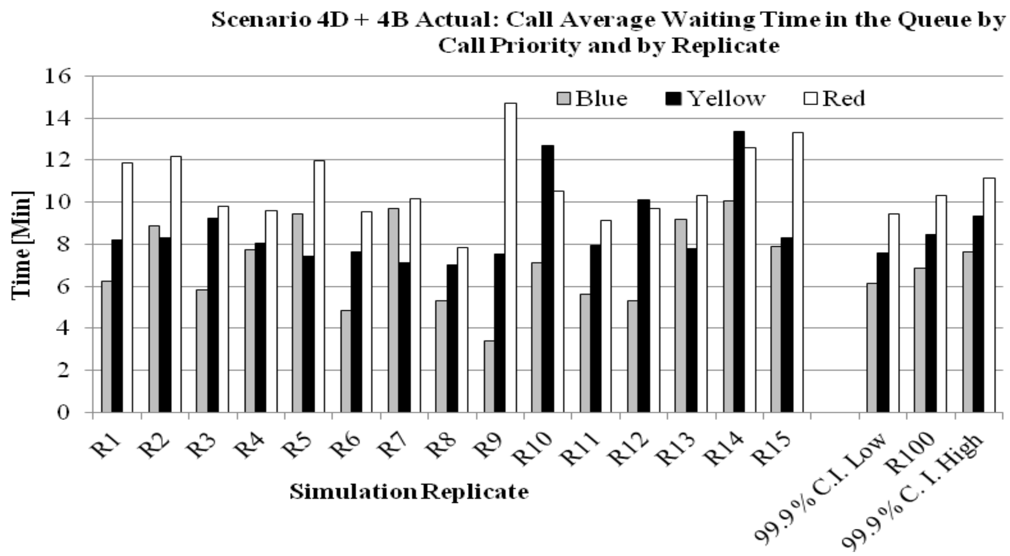

Figure 1. Scenario 4D+4B Actual: Average Waiting Time in the Queue by Call Priority and by Replicate

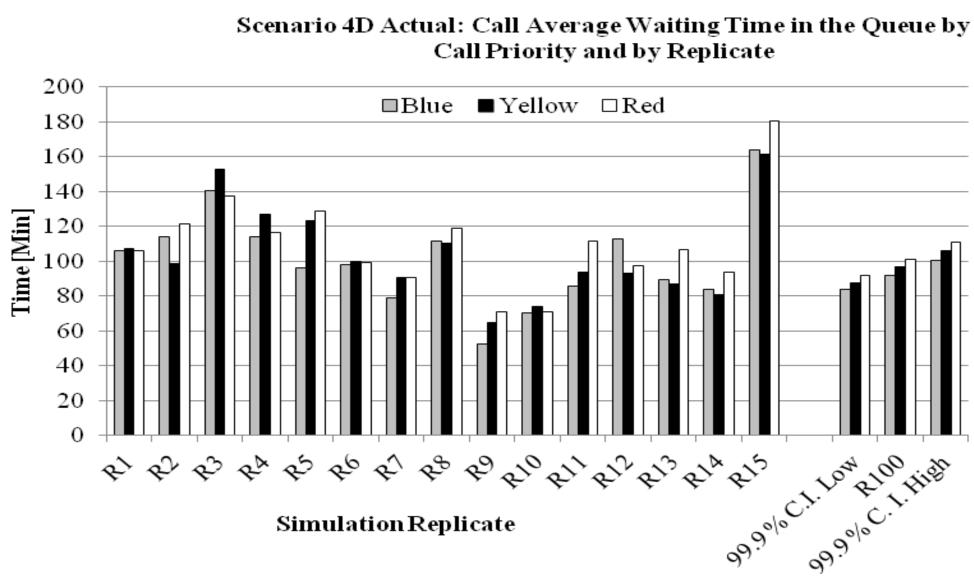

Figure 2. Scenario 4D Actual: Average Waiting Time in the Queue by Call Priority and by Replicate

5. Conclusions

We conclude that stochastic discrete simulation models represent an appropriate modeling strategy which considers actual probability characterizations of the 911 Emergency Response System. Similarly, we observe that current queueing modeling strategies of multiple patrol dispatching for one call or customer, assume Markovian behavior that may not reflect the real operating conditions of the system. Additionally, it was identified how multiple vehicle patrolling and dispatching could significantly increment the waiting time in the queue of incoming calls when servers are busy making it even more difficult to comply with the international response time reference of 3 minutes maximum.

6. References

Chaiken, J.M. and Dormont, P. (1978a). A Patrol Car Allocation Model: Background. *Management Science*, 1978a; 24: 1280-1290.

Chaiken, J.D. and Dormont, P. (1978b). A Patrol Car Allocation Model: Capabilities and Algorithms. *Management Science*, 1978b; 24: 1291-1300.

Chen, H., Cheng, T., and Wise, S. (2017). Developing an Online Cooperative Police Patrol Routing Strategy. *Computers, Environment and Urban Systems,* 2017; 62: 19–29.

Corham, A. (1954). Priority Assignment in Waiting Line Problems. Oper. Res., 1954; 2: 70-76.

Green, L.V. (1984). A Multiple Dispatch Queueing Model of Police Patrol Operations. *Management Science,* 1984; 30: 653-664.

Green, L.V. and Kolesar, P.J. (2004). Improving Emergency Responsiveness with Management Science. *Management Science,* 2004; 50: 1001-1014.

Green, L.V. and Kolesar, P.J. (1984). A Comparison of the Multiple Dispatch and M/M/c Priority Queueing Models of Police Patrol. *Management Science,* 1984; 30: 665-670.

Guo, S., Fang, X., Tong, H., Rui, L. (2010). Police Cars Deployment and Patrol Models. *2010 International Conference on Computational Intelligence and Software Engineering*, 1-5. doi: 10.1109/CISE.2010.5676759.

Holguin-De La Cruz, J. (2017). Improving Response Times in a Safety Emergency Response System Using Stochastic Simulation. *Proceedings of The 6th Annual World Conference of the Society for Industrial and Systems Engineering*, 84-89.

INEGI (2018). Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública ENVIPE 2018. *Instituto Nacional de Estadística Geografía e Informática*. Aguascalientes, Ags., México.

INEGI (2019). Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública ENVIPE 2019. *Instituto Nacional de Estadística Geografía e Informática*. Aguascalientes, Ags., México.

Kolesar, P, Walker, W.E. (1975) *A Simulation Model of Police Patrol Operations: Executive Summary*, New York, NY: The New York City Rand Institute

Larson, R. C. (1972). *Urban Police Patrol Analysis*, Cambridge, MA, MIT Press.

Larson, R.C. (1974). A Hypercube Queueing Model for Facility Location and Redistricting in Urban Emergency Services. Comput. & Ops. Res., 1974; 1: 67-95

NACCJ (1973) National Advisory Commission on Criminal Justice Standards and Goals: Task Force on Police. *Police,* 1973; 153

Piyadasum, T., Kalansuriya, B., Gangananda, M., Malshan, M., Bandara, D.H.M.N., Marru, S. (2017). Rationalizing Police Patrol Beats Using Heuristic-Based Clustering, *2017 Moratuwa Engineering Research Conference (MERCon)*, 2017; 431-436.

van Barneveld, T., Jagtenberga, C., Bhulaia, S., and van der Mei, R. (2018). Real-Time Ambulance Relocation: Assessing Real-Time Redeployment Strategies for Ambulance Relocation. *Socio-Economic Planning Sciences*, 2018; 62: 129-142.

Wu, W., Shen, L., Ji, X. and Jin, W. (2014). Analysis of Freeway Service Patrol with Discrete Event-Based Simulation. *Simulation Modelling Practice & Theory*, 2014; 47: 141-151.

Yang, B., Hu, Z.H., and Zhou, J.X. (2015). Configuration, Deployment, and Scheduling Models for Management and Optimization of Patrol Services. *Mathematical Problems in Engineering*, 2015, 1-13.

Zaki, A.S., Cheng, H.K. and Parker, B.R. (1997). A Simulation Model for the Analysis and Management of an Emergency Service System. *Socio-Econ. Plann. Sci.,* 1997; 31: 173-189.