Cross Entropy Approach for Patrol Route Planning in Dynamic Environments

Xu Chen*

Email: cx008@ie.cuhk.edu.hk
*Department of Information Engineering
The Chinese University of Hong Kong

Tak-Shing Peter Yum*†

Email: tsyum@ie.cuhk.edu.hk
†Hong Kong Applied Science and Technology
Research Institute Company Limited

Abstract—**Proper patrol route planning increases the effectiveness of police patrolling and improves public security. In this paper we present a new approach for the real-time patrol route planning in a dynamic environment. We first build a mathematic framework, and then propose a fast algorithm developed from the Cross Entropy method to meet the real-time computation requirement needed for many applications. In addition, as the randomness is an important factor for practices, the entropy concept is used for designing the randomized patrol routes schedule strategy. Numerical studies demonstrate that the approach has fast convergence property and is efficient in dynamic patrol environment.**

I. Introduction

Police patrolling is an important public service. It helps to deter and prevent crimes, and creates a sense of public security. Police resources are limited, and there is an obvious public interest in developing low-cost approaches for improving the efficiency of police patrols. More efficient patrol route planning is one such approach. An effective solution for the patrol route planning problem has other potential applications. It can be used in a military context, for example for information gathering by Unmanned Aerial Vehicles (UVAs), and in forestry, for detecting forest fires.

In policing operations, a city is often divided into several police precincts. A precinct consists of multiple patrol beats. In a patrol beat, a patrol unit (often one or two policemen) is dispatched to conduct surveillance via patrolling. The patrols tour different locations in the area, particularly crime hot-spots, in order to deter crimes by their presence, and to intervene if a crime is in progress. Patrol route planning must therefore take into account crime patterns. Fig. 1 shows the daily criminal data statistics of a policing precinct in Vancouver City [1]. It reveals the dynamic behavior of crime arrivals. An efficient patrol route plan should ensure that the patrol unit covers the critical places at the critical times.

In the light of recent advances in security surveillance instruments and information processing techniques, a preventive approach based on crime tendency prediction is now possible. With new information gathered and the crime prediction updated, a more updated patrol route is required. This raises the real-time computation requirement for the patrol route planning problem.

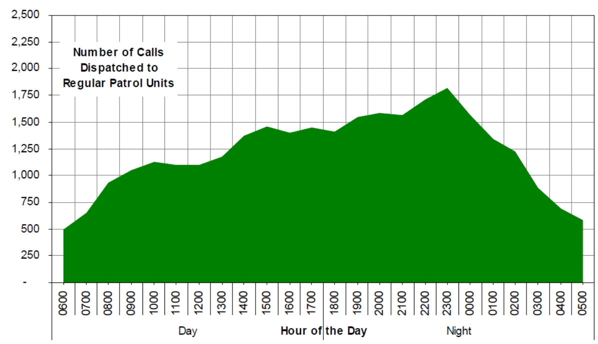

Fig. 1. Daily Criminal Data Statistics of Vancouver Police Precinct

Several approaches have been proposed for the patrol route planning problem. Chevaleyre et al. [2] modeled the patrol area as a graph, and showed that the shortest Hamiltonian cycle is the optimal strategy for maximizing the frequency of visiting each location, on the assumption that all the locations are equally important. This solution needs be generalized when there exist crime hot-spots. Praveen et al. [3] modeled the interactions between the police and terrorists as a two-stage Stackelberg game, and derived an equilibrium strategy. This approach assumes that a player always predicts how his opponent will behave and chooses the best response. Sui et al. [4] modeled the patrolling process as a continuous-time Markov decision process, and proposed a learning algorithm to maximize the long-term decision rewards. The crime arrival rate of a location is assumed to be a constant.

Previous works assumed that the patrol environment is static. They did not cope with the dynamic crime arrivals of individual locations. In this paper, we consider the patrol route planning problem in a dynamic environment. We show that the resultant problem is NP-hard, and propose a Cross Entropy algorithm for fast near-optimal solutions. For the randomized patrol strategy design, we introduce the entropy concept to evaluate the randomness and propose the maximum entropy patrol routes scheduling strategy. Numerical results show that the strategy has over 24% patrol rewards improvement over the random walk based patrolling when approaching the maximum entropy.

In Section II, we formulate the patrol route planning problem. In Section III, we first introduce the generic Cross Entropy method, and then propose the Cross Entropy algorithm for the patrol route planning problem. In Section IV, we describe the randomized patrol routes schedule strategy. In Section V, we present the numerical results. In Section VI, we present our conclusion and discuss topics for future investigation.

II. Problem Formulation

When a patrol unit is dispatched, an efficient patrol plan is needed. Fig. 2 depicts the life cycle of the patrol planning procedure. The first step is to identify, from the historical crime data, the strategic patrol locations in the area e.g. notable crime hot-spots and infrastructures. After the crime tendency prediction is made, the expected crime distribution over both spatial and temporal dimensions is specified. Based on these, a patrol route that maximizes the patrol effectiveness over a desired planning horizon (e.g. an 8-hour patrol shift) is generated. When important information is received and the prediction needed to be updated, the procedure is repeated to generate an updated patrol route.

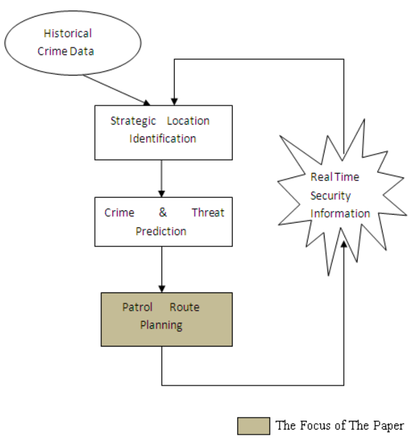

Fig. 2. Life Cycle of Patrol Planning

A. The Patrol Model

Many techniques for strategic location identification and crime tendency prediction have been developed, such as crime mapping and prediction [5]. The focus of this paper is the patrol route planning problem. We first introduce the resultant patrol model after the stages of strategic location identification and crime prediction as followings:

- $\{1, 2, ..., N\}$ denotes a finite set of nodes, with a node n representing a strategic patrol location in the area, such as a crime hot-spot or infrastructure. The unit will patrol around after it reaches the node.
- w_n denotes the importance of node n in the patrol area.
- s_n denotes the prescribed patrol dwell time in node n. Based on previous justice research, it is found that the optimal police dwell time for each visit to a crime hot spot is 11 to 15 minutes [6].
- t_{nm} denotes the average travel time for the patrol unit moving from node n to node m.
- $[T_0, T_K]$ denotes the planning time horizon, and $T_0 < T_1 < T_2 < \cdots < T_{K-1} < T_K$ denote a K-time interval partition of $[T_0, T_K]$. Typically, a time interval is taken as 30 or 60 minutes. Motivations for such a partition are two-folds: first we can approximate the dynamic crime arrival rate by a piecewise constant function. This is a common approach for dealing with difficult dynamic problems [7]. Second, since patrol coverage is regarded as an important index in justice research [8], we also utilize the partition to set the patrol time separation for a node, which rules out greedy solutions that focus the patrol's effort on only a small set of locations at all times. In this study, we take a time interval as the separation time. Our method also applies well to the case that the separation time consists of multiple time intervals.
- λ_n^k denotes the expected crime arrival rate of node n in time interval $[T_{k-1}, T_k]$. Let $\lambda_n(t)$ be the dynamic crime arrival rate function of node n, for $t \in [T_0, T_K]$. Then, we can obtain λ_n^k by

$$\lambda_n^k = \frac{\int_{T_{k-1}}^{T_k} \lambda_n(t)dt}{T_k - T_{k-1}} \tag{1}$$

B. Patrol Route Planning

Based on this patrol model, if there is no patrol deployment in the area, the risk can be measured by the weighted expected number of crime arrivals in $[T_0, T_K]$, which is $\sum_{n=1}^{N} w_n \int_{T_0}^{T_K} \lambda_n(t)dt$. Once the patrol unit arrives in a location, crime in that location are either discouraged or intercepted. The patrol reward therefore is just the risk being reduced. Hence keeping the area as safe as possible is equivalent to maximizing the reward through patrolling.

To quantify the effectiveness of the patrol, we define the patrol reward r_n^k as the weighted expected number of potential crime intercepted by a unit patrolling at a node n in time interval k, i.e.

$$r_n^k = w_n \lambda_n^k s_n \tag{2}$$

The patrol route planning problem can be described formally as: *Given a planning time horizon* $[T_0, T_K]$*, plan a patrol route that maximizes the total patrol rewards.*

We have built a mixed integer programming formulation for the problem. Due to the space limit, we omit the lengthy formulation and write a high-level maximization framework instead. First, we denote the segment on a route that node n is patrolled in time interval k as g_n^k. Let

$$I_{\{g_n^k \in \ell\}} = \begin{cases} 1 & \text{Patrol route } \ell \text{ contains segment } g_n^k \\ 0 & \text{Otherwise} \end{cases} \tag{3}$$

The patrol route planning problem is then formulated as:

$$\begin{aligned} \max \quad & \textstyle\sum_{k=1}^{K} \sum_{n=1}^{N} r_n^k I_{\{g_n^k \in \ell\}} \\ \text{s.t.} \quad & I_{\{g_n^k \in \ell\}} \in \{0, 1\} \\ & \ell \in \Omega \end{aligned} \tag{4}$$

where Ω is the set of feasible patrol routes satisfying the required constraints, e.g. flow balance constraint, time constraint and sub-tour elimination constraint. For the complete formulation, please refer to [9].

Note that for the special case where there is only one time interval, i.e. $K = 1$, the problem degenerates to the orienteering problem. It is shown in [10] that the orienteering problem is NP-hard. Therefore, the patrol route planing problem is also NP-hard. In our computational experiences, it typically takes more than 10 hours to obtain an acceptable solution using popular solvers such as MOSEK for the number of nodes $N \geq 12$ on an ordinary PC. A fast algorithm is therefore essential for real-time solutions.

III. Cross Entropy Algorithm

A. The Cross Entropy Method

The Cross Entropy (CE) method was first used to estimate the probability of rare events. It was adapted for use in optimization observing that sampling around the optimum of a function is a rare event [11]. Due to its fast convergence and simplicity, the CE method has been successfully applied to a number of difficult optimization problems, including the maximal cut problem, the traveling salesman problem, and various kinds of scheduling problem [11].

To apply the CE method, the deterministic optimization problem must first be converted into a stochastic estimate problem. Then the following three steps are performed repeatedly until convergence:

1) Generate random data samples according to a parameterized random mechanism.
2) Select the elite samples according to a performance criterion.
3) Update the parameters of the random mechanism based on the selected elite samples for producing better new samples.

Formally, let $U(x)$ denotes the objective function. Suppose we wish to maximize $U(x)$ over the set $\mathcal{X}$ by the CE method. In step 1, M random samples $X_1, ..., X_M$ are generated by the parameterized random mechanism $f(x, v)$, a probability density function with parameter v. In step 2, the elite samples are selected according to the performance criterion, i.e. samples for which $U(x) \geq \gamma$ are selected. In step 3, the CE method uses minimization of the Cross Entropy (also known as Kullback-Leibler distance) between updated random mechanism and the probability distribution of the selected elite samples as the update criterion. According to [11], it is equivalent to solving:

$$\max_{v} \frac{1}{M} \sum_{i=1}^{M} I_{\{U(X_i) \geq \gamma\}} \ln f(X_i, v) \tag{5}$$

where,

$$I_{\{U(X_i) \geq \gamma\}} = \begin{cases} 1 & \text{If } U(X_i) \geq \gamma \\ 0 & \text{Otherwise} \end{cases} \tag{6}$$

The CE method uses an iterative approach. In each iteration, the $\lfloor \rho M \rfloor$ best samples are selected as the elite samples, typically with $0.01 \leq \rho \leq 0.03$. As shown in [12], the CE method guarantees convergence to the global optimum with proper parameters setting.

B. CE Algorithm for the Patrol Route Planning Problem

To use the CE method for patrol route planning, we first need to define a random route generation mechanism. Observing that the patrol environment is static in a time interval, we define K auxiliary $N \times N$ transition matrices $\{P^k : [p^k_{nm}]\}_{k=1}^K$, where p^k_{nm} is the probability of moving from node n to node m in time interval k. In the followings, we present the algorithm for sample route generation, and the method for transition matrices update.

1) Sample Route Generation: In each iteration of the CE method, we need to generate M sample routes over the planning time horizon $[T_0, T_K]$. With the transition matrices, a route can be generated in the time-interval by time-interval manner. Algorithm 1 below gives the details.

Algorithm 1 Sample Route Generation Algorithm

/* In the followings, a route is specified by sequence of node visits. Each node visit is specified by 3-tuple (n_c, r_c, t_c), where c is the sequence index, n_c is the c-th patrol note on the route, r_c is the patrol reward obtained at c-th node, and t_c is the patrol arrival time on the c-th node. */

Generate M sample routes, for each route:

1) Start from a random starting node n_0, set $c = 0, t_0 = T_0$.
2) Repeat until $t_c \geq T_K$
   (a) Obtain the patrol reward $r_c = r^k_{n_c}$ when $t_c \in [T_{k-1}, T_k]$;
   (b) Generate the next node n_{c+1} according to the transition matrix P^k and the separation time requirement;
   (c) Update the patrol arrival time on the $c+1$-th node $t_{c+1} = t_c + s_{n_c} + t_{n_c n_{c+1}}$, and set $c = c + 1$;
3) Given a route $\ell = \{(n_0, r_0, t_0), ..., (n_{c-1}, r_{c-1}, t_{c-1})\}$, calculate the total patrol reward on route as:

$$R(\ell) = \sum_{i=0}^{c-1} r_i \tag{7}$$

2) Transition Matrices Update: After the sample routes are generated, the critical issue is the updating of the transition matrices, or solving the problem in (5). Theorem 1 gives the optimal update formula.

Theorem 1. *Given M sample routes $\ell_1, ..., \ell_M$, and the elite sample selection criterion $I_{\{R(\ell_i) \geq \gamma\}}$, the optimal transition matrices update that minimizes the Cross Entropy between the updated transition matrices and the probability distribution of selected elite routes is*

$$p^k_{nm} = \frac{\sum_{i=1}^{M} I_{\{R(\ell_i) \geq \gamma\}} I_{\{\mathcal{L}^k_{nm} \in \ell_i\}}}{\sum_{i=1}^{M} I_{\{R(\ell_i) \geq \gamma\}} \sum_{m=1}^{N} I_{\{\mathcal{L}^k_{nm} \in \ell_i\}}} \tag{8}$$

where $I_{\{\mathcal{L}^k_{nm} \in \ell_i\}}$ an indicator whether the route ℓ_i contains a visit from node n to node m in time interval k.

Proof: For simplicity of notation, we denote the set of the transition matrices as $\mathbf{P} \triangleq \{P^k\}_{k=1}^K$. Using (5), the optimal update problem becomes:

$$\begin{aligned} \max_{\mathbf{P}} \quad & \tfrac{1}{M}\textstyle\sum_{i=1}^M I_{\{R(\ell_i)\geq\gamma\}} \ln f(\ell_i, \mathbf{P}) && (9) \\ \text{s.t.} \quad & \textstyle\sum_m p_{nm}^k = 1 && \forall n = 1, ..., N; \\ & && k = 1, ..., K. \end{aligned}$$

Since $f(\ell_i, \mathbf{P})$ is the probability distribution function that the route ℓ_i is generated by the transition matrices $\mathbf{P}$, we have

$$\ln f(\ell_i, \mathbf{P}) = \sum_{k=1}^K \sum_{n=1}^N \sum_{m=1}^N I_{\{\mathcal{L}_{nm}^k \in \ell_i\}} \ln p_{nm}^k \qquad (10)$$

This function is concave as it is a sum of logarithm functions. Therefore, the objective function in (9) is concave, and constraint set is convex. This guarantees that any local maximum is also the global maximum.

Using the Lagrange multipliers $\{\mu_n^k\}_{n=1,...,N}^{k=1,...,K}$ for solving the maximization problem in (9), we have

$$\max_{\mathbf{P}} \frac{1}{M}\sum_{i=1}^M I_{\{R(\ell_i)\geq\gamma\}} \ln f(\ell_i, \mathbf{P}) + \sum_{k=1}^K \sum_{n=1}^N \mu_n^k (\sum_{m=1}^N p_{nm}^k - 1) \qquad (11)$$

Differentiating (11) with respect to p_{nm}^k yields

$$\frac{1}{M}\sum_{i=1}^M I_{\{R(\ell_i)\geq\gamma\}} I_{\{\mathcal{L}_{nm}^k \in \ell_i\}} + \mu_n^k p_{nm}^k = 0 \qquad (12)$$

Summing over $m = 1, ..., N$ gives us

$$\mu_n^k = -\frac{1}{M}\sum_{i=1}^M I_{\{R(\ell_i)\geq\gamma\}} \sum_{m=1}^N I_{\{\mathcal{L}_{nm}^k \in \ell_i\}} \qquad (13)$$

which substituting μ_n^k in (12) gives us the optimal solution (8).

■

We note that instead of updating the transition matrices with (8) directly, a smoothing procedure combining the updates between two successive iterations is more efficient, i.e.

$$\tilde{p}_{nm}^k(i) = \alpha p_{nm}^k(i) + (1-\alpha)\tilde{p}_{nm}^k(i-1) \qquad (14)$$

where $p_{nm}^k(i)$ is from (8), and $\tilde{p}_{nm}^k(i-1)$ denotes the smoothing update in the last iteration. This step is necessary to prevent the algorithm from getting stuck in a local optimum when $p_{nm}^k(i)$ is zero. Values of $0.7 \leq \alpha \leq 0.9$ are found to give the best results [11]. For initialization, we set each transition matrix as a uniform distribution among the nodes, i.e. $\tilde{p}_{nm}^k(0) = \frac{1}{N}$. This guarantees that all feasible routes can be generated.

To sum up, we describes the CE algorithm for the patrol route planning problem in Algorithm 2.

Numerical studies in Section V show that the CE algorithm always converges to a solution that is optimal, or very close to the optimum in less than one minute in an ordinary PC.

Algorithm 2 The CE Algorithm for Patrol Route Planning

1) Set the initial transition matrices as the uniform distributions.
2) Generate M sample routes $\ell_1, ..., \ell_M$ by the sample route generation algorithm.
3) Determine $\lfloor \rho M \rfloor$ elite samples.
4) Use the selected elite samples to update the transition matrices according to formula (8).
5) Apply the equation (14) to smooth the transition matrices.
6) Repeat from step 2 until convergence.

IV. Strategy for Randomized Patrol Routes Schedule

The aim of police patrol is to provide the public with a sense of security and discourage potential crimes. One good way of discouraging crime is to adopt a randomized patrol strategy, so that criminals can never be sure when a patrol might arrive on the scene. The use of randomized police patrols has been one of the key factors in the significant fall in the crime rate in New York city in recent years [13].

A. Randomness of Patrol Strategy

To evaluate the randomness of a patrol strategy, we use the "entropy" measure in information theory. For a discrete probability distribution $p_1, ..., p_n$, the entropy H is defined as

$$H = -\sum_{i=1}^n p_i \log p_i \qquad (15)$$

where the uniform distribution gives the maximum entropy. Applying to patrolling problem, the entropy of a patrol strategy π is:

$$H_\pi = -\sum_{k=1}^K \sum_{n=1}^N p_n^k \log p_n^k \qquad (16)$$

where p_n^k denotes the probability of visiting node n in time interval k under the patrol strategy π. It captures the randomness in both spatial and temporal dimensions.

B. Maximum Entropy Patrol Routes

On the basis of the entropy concept, we propose the maximum entropy patrol routes scheduling strategy: a daily patrol route is randomly selected from a designated route set with the goal of maximizing the entropy H_π. In this case, the entropy H_π depends on the structure characteristics of the candidate routes. We denote the set of Z candidate routes as $\mathcal{L} \triangleq \{\ell_1, ..., \ell_Z\}$. Let routes be randomly selected from $\mathcal{L}$ with equal probability. Then the probability that node n is visited in time interval k is

$$p_n^k = \frac{\sum_{i=1}^Z I_{\{g_n^k \in \ell_i\}}}{\sum_{i=1}^Z \sum_{k=1}^K \sum_{n=1}^N I_{\{g_n^k \in \ell_i\}}} \qquad (17)$$

where $I_{\{g_n^k \in \ell_i\}}$ is defined in (3). We can then obtain the corresponding entropy $H_\pi(\mathcal{L})$ by formula (16).

Since the entropy is also a quantitative measurement of structure diversity [14], a larger entropy H_π means the routes are more dissimilar, i.e. they share fewer common segments. This property is desirable since more diverse structures provide less similarity information for prediction. However, if a route is very dissimilar to the optimal one, the reward is likely to be significantly lower. Therefore, a careful trade-off between the randomness and the rewards of the candidate routes needs be established. This leads to the constrained maximum entropy candidate routes generation problem as follows: *Given the optimal patrol route ℓ_1 and the minimum acceptable patrol reward $R_{min} < R_{opt}$, generate $Z-1$ other patrol routes that maximizes the entropy H_π of the Z candidate routes.*

To solve the constrained maximum entropy candidate routes generation problem, we first need to generate a large number of routes satisfying the minimum acceptable reward requirement. Given that the CE algorithm employs an iterative approach, we can run the algorithm until all its generated sample routes satisfy the minimum acceptable reward requirement R_{min}. We denote the satisfied routes as the feasible set Φ. From Φ, we choose $Z-1$ routes together with the optimal route ℓ_1 as the candidate routes set $\mathcal{L}$, such that the entropy $H_\pi(\mathcal{L})$ is maximized.

When the size of Φ is large, it is difficult to enumerate all possible solutions. We now use the CE method again to solve this problem. Similar to Algorithm 2, the three steps are as follows:

- Generate M random solutions $\{\mathcal{L}_j\}_{j=1}^{M}$ according to the random mechanism $\{q_i\}_{i=0}^{|\Phi|}$, where a solution $\mathcal{L}_j$ consists of Z routes, q_i is the probability that route ℓ_i from the feasible set Φ is chosen.
- Calculate the corresponding entropies of the solutions $\{H_\pi(\mathcal{L}_j)\}_{j=1}^{M}$, and select $\lfloor \rho M \rfloor$ elite solutions.
- Update the random mechanism $\{q_i\}_{i=0}^{|\Phi|}$ as following:

$$q_i = \frac{\sum_{j=1}^{M} I_{\{H_\pi(\mathcal{L}_j)\geq\gamma\}} I_{\{\ell_i\in\mathcal{L}_j\}}}{\sum_{j=1}^{M} I_{\{H_\pi(\mathcal{L}_j)\geq\gamma\}}} \tag{18}$$

  where $I_{\{H_\pi(\mathcal{L}_j)\geq\gamma\}}$ is an indicator of the selected elite solutions, and $I_{\{\ell_i\in\mathcal{L}_j\}}$ denotes whether the sample solution $\mathcal{L}_j$ contains the route ℓ_i. The formula (18) is derived with the same method in Section III.

V. Numerical Results

A. Simulation Setup

We run the simulation experiments on a PC with a 3.2GHz CPU and 1G memory. The program is written in C language. The experiment involved an 8-hour police patrol shift in a square patrol area. The patrol area is divided into 16 atom areas, each contains a patrol node in the center. We set the dwell time in a node equal to one time unit, which corresponds to 6 minutes in real time. We assume it takes 4 time units to cover one side of the square area. The travel time from one node to another node is proportional to their physical distance. The planning time horizon is 80 time units, and this horizon is partitioned into 8 time intervals, each 10 time units long. The patrol reward for a node in a time interval is randomly generated from the set $\{0.1, 0.2, ..., 1.0\}$.

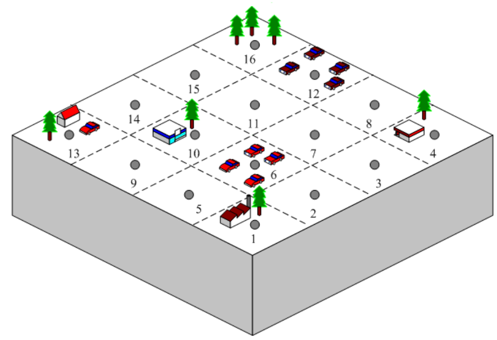

Fig. 3. Patrol Area Graph

The parameters used for the CE algorithm are as follows: fraction of samples selected for updating is $\rho = 0.02$; the smoothing factor is $\alpha = 0.7$; $M = 7000$ samples are generated in each iteration. We found that the performance of the algorithm improves with increasing M. But when $M \geq 6000$ the improvement levels off.

B. Computation Time

We ran the CE algorithm 20 times, and calculated the average computation time. The simulation results are given below:

Maximal Reward	32.2
Avg. Comp. Time (seconds)	21.3

To compare the performance of the CE algorithm, we solve optimization problem in Section II by the popular optimization solver MOSEK, and obtained the following results.

Reward	30.8	31.1	31.5	32.3
Comp. Time (hours)	1.5	5	11	24

Upon comparison, we see that at least 11 hours is needed to obtain an acceptable solution, and it takes 24 hours to obtain the optimal solution. The CE algorithm takes only 21.3 seconds to approach the optimum.

We ran the CE algorithm with $N = 20, 25$ and 30 patrol nodes respectively for the same patrol area. It was also found that the maximal reward obtained by the CE algorithm is very close to the optimal one obtained by MOSEK (See Fig. 4). The average computation times of CE algorithm are less than one minute:

Number of Nodes	20	25	30
Avg. Comp. Time (seconds)	34.1	42.9	54.2

C. Comparison of Patrol Strategies

Since we have not found any previous studies on patrol route planning in a dynamic environment, we benchmark the patrol reward by the CE algorithm with the following two strategies:

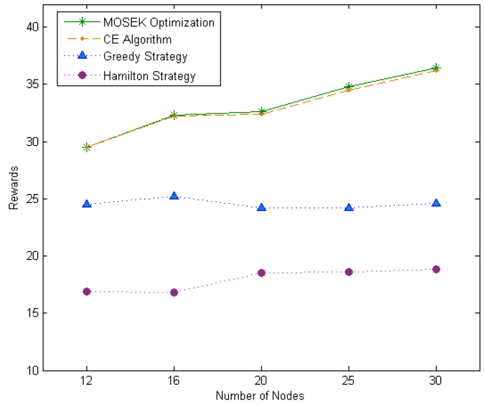

Fig. 4. Different Patrol Strategies

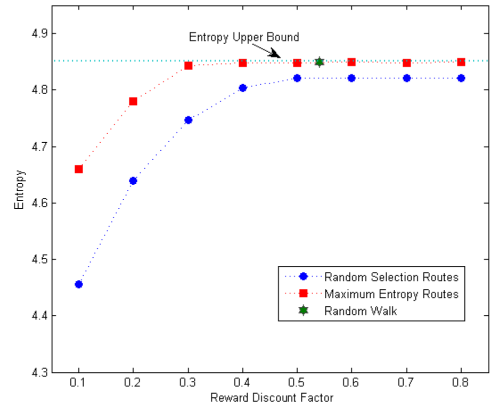

Fig. 5. Entropy and Reward Discount

- Greedy strategy: the patrol unit always choose the next node with the highest reward. This is the patrol strategy with short-term reward maximization.
- Hamilton strategy: the patrol unit visits all the nodes according to the shortest Hamilton path. This strategy was proposed in [2], with the goal to maximize the patrol frequencies in each node.

We show the rewards obtained using these patrol strategies in Fig 4. We found that the Hamilton strategy is ill adapted to the dynamic patrol environment. The CE algorithm offers $20\% \sim 40\%$ higher reward than the greedy strategy. Also, the performance gain of the CE algorithm increases with the number of nodes.

D. Maximum Entropy Routes

Using (16), we obtained the entropy upper bound $H^*_\pi = 4.85$ for $\{p^k_n\}^{k=1,\ldots,8}_{n=1,\ldots,16}$ being uniformly distributed.

We set the minimum acceptable reward $R_{mim} = (1-\delta)R_{opt}$, where δ is the reward discount. A total of $Z = 50$ maximum entropy candidate routes are found by the CE method for δ ranging from 0.1 to 0.8. The CE method is compared with the random selection method whereby $Z = 50$ routes are randomly chosen from the feasible set Φ. Fig 5 shows the entropy values of the solutions. It is seen that both methods allow the tradeoff between the patrol rewards and the randomness. Between two methods, CE method offers a higher entropy solution. It converges to the entropy upper bound H^*_π when the reward discount is 0.3. Also, with 4% lower entropy, the patrol reward is reduced by only 10%. The random selection method is incapable of approaching the upper bound H^*_π.

We have also evaluated the random walk based patrolling where the patrol unit randomly chooses the next node to visit. It is found that this approach can achieve the upper bound H^*_π with at least 54% patrol reward reduction. The CE method needs only 30% reduction as mentioned above.

VI. Conclusion

In this paper, we presented a new approach for the real-time patrol route planning in a dynamic environment. We used an algorithm developed from the Cross Entropy method, and demonstrated its effectiveness for real-time solutions. As the randomness has a key role in patrols, we proposed the maximum entropy patrol routes scheduling strategy. Numerical results showed that significant performance improvement over the greedy, Hamilton and random walk strategies.

A challenging extension of our approach would be the multiple patrol unit planning in a single patrol area. The optimal cooperation among the patrol units would be a fascinating topic for future study.

References

[1] S. Demers, and C. T. Griffiths, "Vancouver Police Department Patrol Deployment Study," *http://vancouver.ca/police/Planning/PatrolDeploymentStudy.pdf*, 2007.
[2] Chevaleyre Y., Sempe, and Ramalho, "A Theoretical Analysis of Multi-agent Patrolling Strategy," *Proc. of AAMAS'04*, 2004.
[3] P. Paruchuri, J.P. Pearce, J. Marecki, M. Tambe, "Playing games for security: An efficient exact algorithm for solving bayesian stackelberg games", *P*roc. of the 7th Intl. Joint Conference, 2008.
[4] S Ruan, C Meirina, and F. Yu, "Patrolling in a Stochastic Environment," *10th Intl. Command and Control Research and Tech,* 2005.
[5] S. Chainey, and J. Ratcliffe, "Mapping Crime: Principle and Practice," *http://www.ncjrs.gov/pdffiles1/nij/178919.pdf*.
[6] C. Koper, "Just Enough Police Presence: Reducing Crime and Disorderly Behavior by Optimizing Patrol Time in Crime Hot Spots," *Justice Quarterly,* 1995.
[7] S. Chick, and P.J. Sanchez, "Modelling and Simulaton of A Telephone Call Center", *Proce. of 2003 Winter Simulation Coference*, 2003.
[8] S. Walker, "The Police in America: An Introduction", *McGram-Hill Gollege*, 1999.
[9] Xu Chen, and Tak-Shing Peter Yum, "Cross Entropy Approach for Patrol Route Planning in Dynamic Environments", *Technical Report, http://personal.ie.cuhk.edu.hk/∼cx008/CEPRP.pdf*
[10] B.L. Golden, L. Levy, and R. Vohra, "The Orieentering Problem," *Naval Reserach Logistics*, Vol. 34, pp. 307-318, 1987.
[11] R. Rubinstein, and D. P. Kroese, "The Cross Entropy method: An unified approach to Combinatorial Optimization, Monte-Carlo Simulation, and Machine Learning", *Information Science and Statistics*, Springer, 2004.
[12] L. Margolin, "On the Convergence of the Cross Entropy method", *Annals of Operations Research*, Springer, 2005.
[13] N. Billante, "The Beat Goes On: Policing for Crime Prevention," *http://www.cis.org.au/IssueAnalysis/ia38/ia38.htm*, 2003.
[14] E.C. Pielou, "Ecological Diversity", *John Wiley*, 1975.