International Journal of Research
Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 04 Issue-17
December 2017

Optimizing crime hotspots and cold spots using Hidden Markov Model

Mekathoti Rajesh Babu & Dr. Kondapalli Venkata Ramana
Pursing M.Tech, Department of CSSE, AU College of Engineering, Visakhapatnam
Assistant Professor, Department of CSSE, AU College of Engineering, Visakhapatnam

Abstract- Public security is a key concern around the world. Efficient patrol strategy can help to increase the effectiveness of police patrolling and improve public security. In this paper, patrol pocess is formulated as Markov process, By taking the all the nearest locations and probability of types of crimes happened in each location as Hidden Markov Model. Using Viterbi algorithm we can find most likely sequence of hidden states here hidden states are the areas need to concentrate. After finding the path for patrolling we can find the crime hotspots and cold spots in the current area for better surveillance.

Keywords: ***Police Patrolling, Crime hotspots, cold spots, Markov decision process, Repeat addressing, Hidden states***

.

I.INTRODUCTION

POLICE patrol is an important public service. It helps to prevent crime rate and creates a sense of public security. As police resources are limited, it is obviously in the public interest to develop low-cost approaches for achieving efficient police patrols [1]. One such approach is to optimize the patrol strategy by setting an effective route or path. Patrolling officers limits criminal activities and maximize safety in particular areas, protecting both people and property. Patrols involve police officers visiting an area not to respond to a specific call but to travel through the area and observe it. Patrolling officers may travel on foot or by bicycle, motorcycle, car or even boat.

A key aspect of police patrolling is to provide a high visibility for the police force in the areas where crimes occurred frequently and the areas known for the crime occurrence. Visible patrolling officers can strengthen sense of safety and security. That's one reason officers who are serving on certain types of patrol, particularly in neighbourhoods, will appear conspicuously in uniforms and marked patrol cars. They will visit both prominent locations (Hot spots) and locations where past crimes have occurred to improve a sense of safety in those areas.
Police patrol is often done in a hierarchical fashion. A city can be divided into several police divisions, and division consists of multiple patrol beats in the sub localities. In a patrol beat, patrol units are dispatched to conduct surveillance via patrolling. The patrol units walk different locations in the area, particularly crime hot spots (i.e., areas of concentrated crimes), to deter and prevent crimes by their presence, and they intervene if crimes are in progress. To shape an efficient patrolling plan to reduce the crimes and strengthen the public security.

Efficient patrol strategy depends on setting an efficient route to patrol, several approaches have been proposed for the patrolling problem. Chevaleyre [2] modeled the patrol area as a graph and showed that the shortest Hamilton cycle is the optimal patrol strategy for minimizing the maximum idling time among the locations, assuming that all the locations are equally important. This solution needs to be generalized when there exist crime hot spots. Chawathe [3] model the road network as a weighted graph with weights representing the importance of the roads and applied the shortest densest path algorithm to determine the most important patrol route.

Assuming that the crime arrival rate of a location to be a constant, Make the patrol process as a Markov model[4]. Consider the locations as states can form Markov chain. Patrol route is set through the hidden Markov model approach, A hidden Markov model is a Markov model in which the state is only partially observableIn other words, observations are related to the state of the system, but they are typically insufficient to precisely determine the state.Because we only know the crimes occurred in the area we do not know the kind of people and other circumstances of the location. Several well-known algorithms for hidden Markov models exists, one such algorithm is Viterbi algorithm. The Viterbi algorithm is a dynamic programming algorithm for finding the most

International Journal of Research
IJR
International JournalofResearch
Available at https://edupediapublications.org/journals

likely sequence of hidden states –called the Viterbi path- that results in a sequence of observed events here the observed events are crimes happened the area.After setting the route we need to map the crime hotspots using repeat addressing techniquefor better surveillance.The key results and contributions of this paper are summarized as follows

• *Patrol path setting by using hidden Markov model approach* : By taking the locations as states and the previously occurred crimes as observations we set the patrol path.

• *Pointing the crime hotspots in the map for current area through repeat addressing* : For the current area we can map the crime hotspots by using repeat addressing technique by setting some crime threshold.

The rest of this paper is organized as follows. Introduction of the patrol model by the hidden Markov approach is explained in Section II. In Section III, Crime hotspots using the repeated addressing for current patrolling area. Finally, Conclusion of this paper and discuss topics for future investigation in section IV.

II.PATROL MODEL

When patrol unit is dispatched in a beat to conduct patrolling, an efficient patrol strategy is needed. The patrol problem is essentially a sequential Markov model problem which is Markov chain gives rise to HMM (Hidden Markov model) formulation.

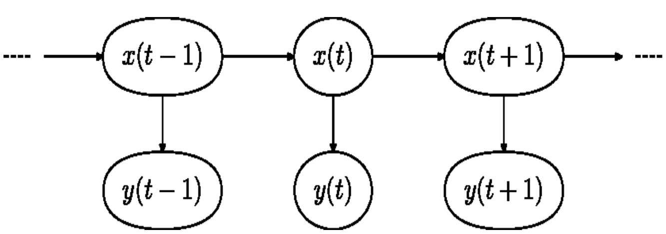

Fig1. General architecture of the HMM

The diagram above shows the general architecture of an instantiated HMM. Each elliptical shape represents a random variable. The random variable $x(t)$ is the hidden state at time t . The random variable $y(t)$ is the observation at time t (with $y(t) \in \{ y_1, y_2, y_3, y_4 \}$). The arrows in the diagram denote conditional dependencies.The conditional probability distribution of the hidden variable $x(t)$ at time t, given the values of the hidden variable x at all times, depends *only* on the value of the hidden variable $x(t-1)$; the values at time $t-2$ and before have no influence. This is called the **Markov property**. Similarly, the value of the observed variable $y(t)$ only depends on the value of the hidden variable $x(t)$ both at the time t .In the standard type of HMM the state space of hidden variable is discrete and the observations are either discrete or continues. The parameters of HMM are two types transmission probabilities and emission probabilities. The transmission probability controls the way the hidden state at time t is chosen given the hidden state at time t

The state transition and the emission probabilities on each state is visualized in the fig2. Fig2 depicts all the probabilistic parameters in the HMM. Circles represents the states or the areas in the patrol zone, The rectangular boxes are the observations, The arrows with pointing head are the transition probabilities from one state to the other and the lines between the observations and the states are the probabilities of the observation on particular state.

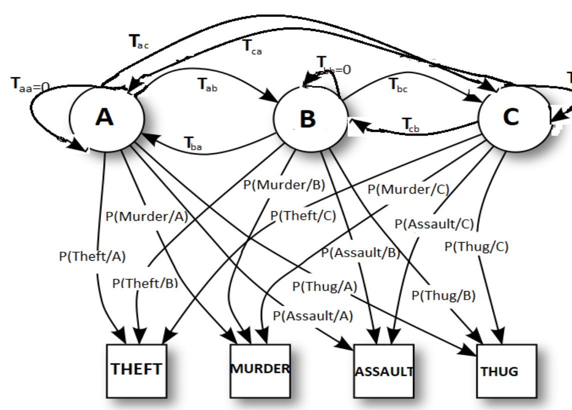

Fig2. Probabilistic parameters of HMM

Patrol zone is divided into N sublocalities and by considering the crimes as the observations with their probabilities in each state among all the observations recorded in sub locations. If we consider the 3 sub locations(A,B,C) as states and observations $(O_1,O_2,\ldots,O_n)$ suppose K crimes recorded in the patrol zone then the startingprobability of each sub locality is number of crimes recorded in the sub location divide by total crimes recorded in patrol zone.

International JournalofResearch
Available at https://edupediapublications.org/journals

$$SP_A = \frac{\sum_{k=0}^{n} O_k\ (\mathrm{A})}{\sum_{k=0}^{n} O_k}$$

For remaining all the states starting probabilities can be find by using above stated formula. An N*N transition matrix can be formed with the transition probabilities to transit from one state to the other.

$$T_{[A][B]} = \frac{\sum_{k=0}^{n} O_k\ (\mathrm{B})}{\sum_{k=0}^{n} O_k - \sum_{k=0}^{n} O_k(A)}$$

For a given parameters Viterbi algorithm iterate through all the states for each observation and takes the maximum probable state in each step, It iterates K number of times here K is the number of observations. As earlier said N*K emission matrix contains the emission probabilities for the observations on states.

$$P(S_i\ |Y_i) = \frac{P(Y_i|S_i)\ P(S_i)}{P(Y_i)}$$

For example, Fig3 shows a table that contains the number of crimes observed in each state and the total observations of each type given. By substituting the values in above formulas we can get the probabilities. From the table the start probability of A is 19/50 = 0.38. The transition probability from A to B is16/50-19 = 0.41 and the emission probability of "THEFT" on A is4/15=0.26 Like wise we can find all the possible combinations.

Locations / Crimes	THEFT	MURDER	ASSAULT	THUG	Total observations in each location(Σobs At A)
LOC A	4	5	5	5	19
LOC B	6	5	3	2	16
LOC B	5	4	3	3	15
Total observations of each type	15	14	11	10	50

Fig3. Example crime rate scenario

By taking the above mentioned inputs Viterbi algorithm evaluate all the observations finished, It produces the probability for the given sequence of observations on the states. Collection of state sequence from the final step to the start decode the hidden states and gives the state sequence for the observations.

1. Viterbi algorithm :

It is a dynamic programming algorithm for to find the most likely sequence of hidden states. The sequential states called as the Viterbi path –that results in sequence of observed events especially in the HMM and Markov information sources. The Viterbi path is essentially the shortest path and this can be visualized using trellis diagram.

Input for the Viterbi algorithm is (S,O,IP,TP,EP) where

- S is the state space which consists of set of states, These states are either discrete or continues.
  S = {s1,s2,…….,sn}
- O is the set observations at time t.
  O = {o1,o2,….,on}
- IP is an array of initial probabilities
  IP = {ip1,ip2,…..,ipn} such that sp1 stores the probability of sate s1
- TP is a matrix which contains the transmission probability of transmission from si to sj, This is also called as the Markov matrix. The size of TP is N*N where N is the no.of states
- EP is a K*N matrix such that EPij is the probability of observing oj from the state si, Here K is the no.of observed variables

Algorithm 1 The Viterbi algorithm for finding the sequence of states

1. Viterbi(S,O,SP,TP,EP)
2. For each state i ∈{1,2,….,n} do
3. T1[i,1]=IPi* EPiy1
4. T2[i,1]=0
5. End for
6. For each observation i=2,3,….,N do
7. For each state j∈{1,2,….,k} do
8. T1[j,i]=EPjyi * max k(T1[k,i-1].TPkj)
9. T2[j,i]=argmax k(T1[k,i-1].TPkj)
10. End for
11. End for
12. ZT=argmax k(T1[k,T])
13. ZT=Szt
14. for i=T,T-1,…,2 do
15. Zi-1=T2[zi,i]
16. Xi-1=Szi-1
17. End for
18. Return X
19. End

Evolution of states based on the observations can be done using the above algorithm.

III.MAPPING CRIME HOTSPOTS

Offences that are reported during the short period of time may have the common attributes like type of crime, the way crime happened, and the common geographical area. Relatively crimes happened in small areas are called as the hotspots or clusters [5]. A hotspot is a small area in which crime rate is above the threshold and that is highly predictable, at least over one year period. Hotspots are found by cluster of events or the locations. The high concentration of reported cases in particular are and greater probability of future crime occurring in the same geographical area make it sense for the crime prevention strategy. For to find the clusters or hotspots in large geographical area we need to use the clustering methods but here we are interested in the divisions or the sub localities with in the city. A Patrolling officer may walk through highly risk areas for to prevent the crime happening, That strategic planning is given by mapping the hotspots using the repeated addressing method. Here in this we loop through the all reported crimes occurred in particular area for to find the highly concentrated areas like streets, a small junctions, buildings.

We are not interested in the tactical crime analysis which can be use the collected data over several years. we are interested in the data that is recorded at least over a year. In
the dataset we have the GIS information like latitude-longitude, street names, area names. By iterating through the data that is recorded in the current area we can find the concentrated areas. Based on the situation we can set some threshold value for the addresses if an address count crosses the threshold value then we can mark it as the hotspot.

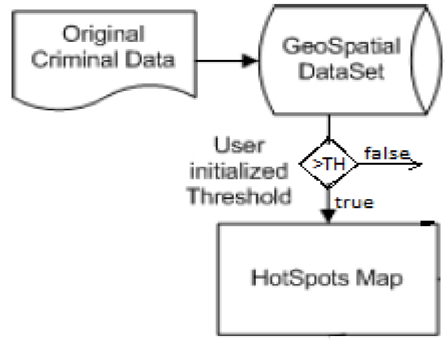

Fig4. Finding hotspots through repeat addressing

1. From the criminal database take the spatial data
2. Staring from the first row iterate through the database
3. Make the count of an addresses such(Here we can count the addresses based on the localities, sub localities, streets)
4. If any area or the address crosses the threshold save the address and place them in the map.

We can place the data points or the address points in the map in several ways. some of the map pointing techniques are.

a) Graduate symbols: when looking for high concentrated places dot maps are superior to other forms of mapping. For to identify the high crime locations one can choose the graduated dots or the symbols, So that dot size is proportional to the number of crimes at particular location. This technique is best used on the large scale mapping.

b) Colour gradient dots: Two approaches are used to put colour gradient on maps is useful for small scale areas.one is to use colour gradient sky blue to red, for example to depict the number of crimes at the location sky blue indicates the low crime rate, a thick blue colour that indicates the moderate crime rate and the red colour indicates the places with high concentration of crimes usually these are the crime hotspots.

As there are several categories of crimes we only concentrate on the personal crimes. The other crimes like property crimes and cybercrimes are not at all interest to the patrolling officers. As the other types of crimes are not interrupted by the patrolling officials. Through strategic patrol planning patrol officials can find the latest trends and patterns in the crimes happening in the particular area.

In this The city can be divided into several patrolling divisions, And each divisions consists of several sub localities. For a given division this will give you the strategic plan to patrol over the sub localities. Sub locality is the current location of the patrolling officer then he/she can view the map for to find the crime hotspots, Usually the crime hotspots in the sub locality are the streets.

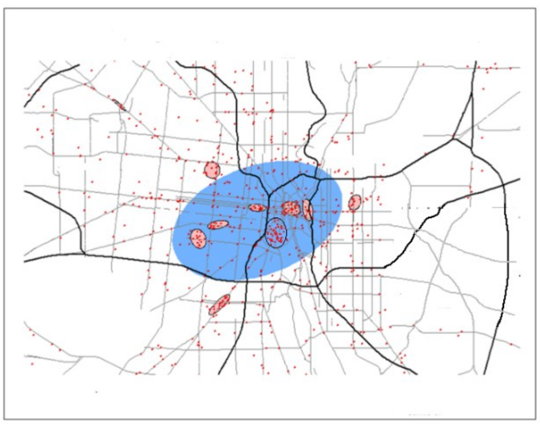

Fig2.Generalised view of Hotspots on map

Above diagram that shows the hotspots for the personal crimes. This is a generalised view of the crime hotspots this can be differentiated to several divisions and each division

IV. EXPERIMENTAL SETUP

For to achieve this in real time requirements need to have is Intel processor with support for Intel VT-x,, and Execute Disable (XD) Bit functionality, or AMD processor with support for AMD Virtualization, Android studio is used as developing tool with java language. HTML/CSS is used to design map for to plot hotspots.

V. CONCLUSION

In this paper we have formulated the patrolling problem as Markov process. Using Viterbi algorithm we found the patrolling path that contains sequence of patrolling areas as a strategic patrolling plan. Repeat address mapping for to find the crime hotspots and the cold spots so that patrolling officer may roam around the highly concentrated areas to prevent the crimes.

An interesting and challenging extension of this paper is to consider the idling times for a particular area. If we consider the idle times at the particular area and the time for the patrolling officers wait to take on call. Consideration of idling times among the patrol locations would be the fascinating topic for the future study.

References

[1] Ahmed and A. R. Tripathi, “Security policies in distributed CSCW and workflow systems,” IEEE Trans. Syst., Man, Cybern. A, Syst., Humans, vol. 40, no. 6, pp. 1220–1231, Nov. 2010.

[2] Y. Chevaleyre, F. Semp, and G. Ramalho, “A theoretical analysis of multiagent patrolling strategy,” in Proc. Int. Joint Conf. Autonom. Agents Multiagent Syst., 2004, pp. 1524–1525

[3] S. S. Chawathe, “Organizing hot-spot police patrol routes,” in Proc. IEEE Conf. Intell. Security Informatics, 2007, pp. 79–86. Discovery Data, vol. 2, no. 4, pp. 1–40, 2009.

[4] S. Ruan, C. Meirina, and F. Yu, “Patrolling in a stochastic environment,” in Proc. 10th Int. Command Control Res. Technol., McLean, VA, 2005

[5] L. Sherman and D.Weisburd, “General deterrent effects of police patrol in crime hot spots: A randomized, controlled trial,” Justice Quart., vol. 12, no. 4, pp. 625–648, Dec. 1995.