TELKOMNIKA Telecommunication, Computing, Electronics and Control
Vol. 18, No. 5, October 2020, pp. 2378~2384
ISSN: 1693-6930, accredited First Grade by Kemenristekdikti, Decree No: 21/E/KPT/2018
DOI: 10.12928/TELKOMNIKA.v18i5.13398

Viterbi optimization for crime detection and identification

Reem Razzaq Abdul Hussein[1], Salih Mahdi Al-Qaraawi[2], Muayad Sadik Croock[3]
[1]College of Business Informatics, University of Information Technology and Communications, Iraq
[2,3]Computer Engineering Department, University of Technology, Iraq

Article Info

Article history:

Received Jun 24, 2019
Revised Mar 3, 2020
Accepted Mar 13, 2020

Keywords:

Baum Welch algorithm
Decision tree
Hybridization
Viterbi algorithm

ABSTRACT

In this paper, we introduce two types of hybridization. The first contribution is the hybridization between the Viterbi algorithm and Baum Welch in order to predict crime locations. While the second contribution considers the optimization based on decision tree (DT) in combination with the Viterbi algorithm for criminal identification using Iraq and India crime dataset. This work is based on our previous work [1]. The main goal is to enhance the results of the model in both consuming times and to get a more accurate model. The obtained results proved the achievement of both goals in an efficient way.

Corresponding Author:

Reem Razzaq Abdul Hussein,
College of Business Informatics,
University of Information Technology and Communications,
Baghdad, Iraq.
Email: reemrazzak@yahoo.com

1. INTRODUCTION

The data mining techniques are widely used in e-government, especially in the criminology field, wherein they help police departments predicate criminals and detect information about crime locations. The study of crime is expected not only to control present crime but also to analyze the criminal activities so that future occurrences of the same incidents can be overcome. Government is trying to improve the effectiveness of prevention of crimes that can be happened in clustering criminal features depending on spatial and temporal criteria which are different in each country [1-11]. It also tries to discover the correlation between different attributes, such as crime and demographics. Analyzing historical data is useful to discover the crime pattern [12]. Past researchers in crime detection depended on data mining and machine learning to develop systems that analyze patterns to have the highest probability of occurring and predict the regions where such patterns are likely to happen. Some researchers utilized the algorithm of Baum-Welch in past studies, which collected GPS data for predicting criminal movements. The work of [13] used attributes according to their characteristics and trained an HMM for each cluster, obtaining an accuracy result of 13.85%. The authors suggested data mining methods for crime detection and identification of criminal for Indian country during the period of 2000 to 2012. They depended on clustering, used k-means, are based on crime attributes, and made visualizations using Google Maps. They also used ANNs and forest decision trees (DTs) for classification. The outcomes were computed by the Weka tool. The accuracy of the ANN was 90.02. They applied 10-fold cross validation for the forest DTs, the accuracy in RMSE was 0.0751, and the time taken to build the model was 4.24 s [14]. In [15], the proposed approach helped agencies to emphasize the security in Indian, their work needs to improve with more data mining method. Other researchers considered the use of the Viterbi algorithm for identifying locations with a high probability of having a criminal depending on the relationship between the location and types of crimes, such as murder, thief, and assault, they handle idling time rather than accuracy problem. The Latent hidden Markov models are types of algorithms that have been designed to detect crime activities by obtaining a sequence of observations from hidden values. They developed an algorithm that fits regular vine copula to generate tree structures. These trees had been employed to produce an emission matrix for hidden Markov model (HMM) algorithm, the fusion of coupled parameters with two types of HMM algorithms, this work needs to improve accuracy and reduce execution time. [1]. In this work, we attempt to overcome the problems in previous studies and develop the work of past researcher [1, 3, 14], by proposing a hybridization between Baum Welch algorithm and Viterbi algorithms in a side, and Viterbi algorithms for optimization of data set then applied DT, at the other side. The objective of proposed work to handles an important topic of monitoring criminal activities and movements and indicates the level of danger in locations, the preposd hybridization method reduces the consumed time and increases accuracy, respectively.

2. HMM

The HMM is a robust model which can be used when states in a process are not observable, but observed data depends on these hidden states. HMM depends on two main properties, which are:

- The observation at time t is produced by a process whose state H_t is hidden from the observer.
- The state of the hidden process represents the Markov chain [16-23].

At the other hand, the Viterbi is a dynamic programming algorithm that depends on transition and emission matrices. In particular, the Viterbi algorithm can obtain path (state sequences) to generate output sequences. It works by finding the maximum overall possible state sequence by considering a forwarding algorithm. The Baum (1970) proposed the use of the Baum–Welch algorithm based on the computation of the probabilistic methods of the Markov model. It also called the forward-backward method, wherein the backward part represented the probability of partial observation sequences from the time (t+1) to end, can be computed iteratively [24].

3. OVERVIEW OF ADOPTED DATASETS

In this work, two datasets have been used in the criminal field. The first one is the Iraqi dataset, while the second dataset is India dataset [25]. The Iraq dataset consists of features such as as{age, gender, ID, crime types, locations, gang, longitude, latitude} the type of features is categorization whereas the features of Indian data set is included {states, murder, attempted to murder ,…, thief}, making a partitioning based on clustering needs more investigation and analysis of the contexts in a discrete sequence dataset. Here, the needed to predict the locations of the crimes, and the HMM to obtain the hidden pattern is required. As mentioned earlier, the dataset is collected from a website and social media, where Baghdad city can be divided into different main locations as shown in proposed Table 1.

While the representation of the second dataset is shown in propsed Table 2. The proposed system is also applied to India dataset. As mentioned above, the HMM must generate ten locations: Loc1, Loc2, …., Loc_{32}, where these locations represent the workflow of the proposed system in the clustering. Here, the needed to predict the locations of the crimes, and the HMM to obtain the hidden pattern in this work as well as the two hidden states with one observe the state.

Table 1. Location of Iraq representation

Seq	Location	Representation
1	Adamya	Location 1(Loc_1)
2	Albayaa	Location 2(Loc_2)
3	Hayalamel	Location 3(Loc_3)
4	kadmiya	Location 4(Loc_4)
5	Madaan	Location 5(Loc_5)
6	Mansour	Location 6(Loc_6)
7	Baghdadgedeeda	Location 7(Loc_7)
8	Hasania	Location 8(Loc_8)
9	Madeena	Location 9(Loc_9)
10	Abogreeb	Location 10(Loc_10)

Table 2. Location of Indian states representation

state		Representation
1	ANDHRA PRADESH	Location 1(Loc_1)
2	ARUNACHAL PRADESH	Location 2(Loc_2)
3	ASSAM	Location 3(Loc_3)
4	BIHAR	Location 4(Loc_4)
5	CHHATTISGARH	Location 5(Loc_5)
.		
.		
.	.	
32	PUDUCHERRY	Location 32(Loc_32)

4. PROPOSED METHOD

The Viterbi algorithm as an optimizer algorithm in two strategies is proposed. The first one is built with Baum Welch, while the second one has been created in combination with the DT algorithm. Viterbi is a supervised clustering algorithm that makes partitioning-based clustering for learning contexts in a discreet sequence dataset. And it works as an outlier detector that removes low-probability data (abnormal data) that explains in detail the following section.

4.1. Hybridization Viterbi algorithm with Baum Welch

We explain the elements of the proposed workflow that include the hybridization of the Viterbi algorithm with Baum Welch to appear in three stages as shown in Figure 1. The Viterbi plays a major role in enhancing the clustering model; in addition, the dataset formulates the structure of the dynamic model using vine while the movement of criminals is determined by the generation of dynamic transition matrix from dataset.

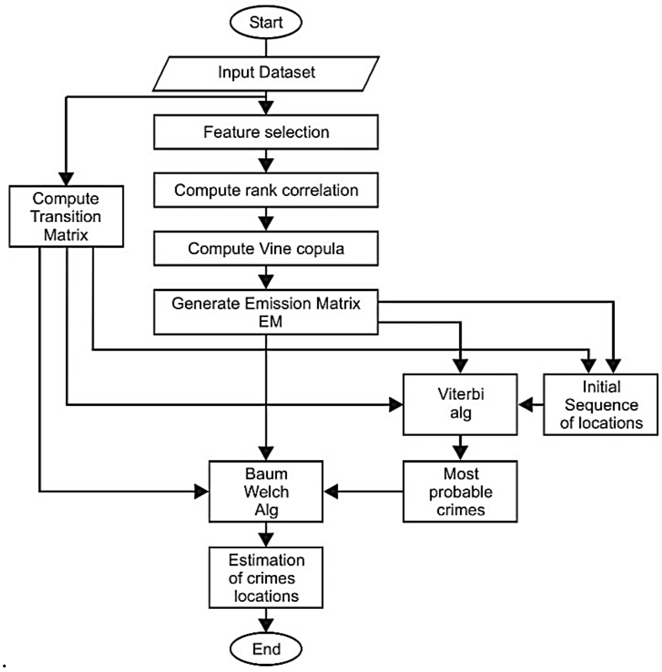

Figure 1. Hybridization Viterbi algorithm with Baum Welch

4.1.1. Emission matrix generation stage

The structure of matrix is generated using a simple vine couple to produce a simple tree structure strategy for selecting a tree model based on vine algorithm and generating a compact model. Conditional probability (Bayes theorem) is applied to the tree structure and used for assuming that the couple parameter {month, location} is combined with crime types to produce three dimensions rather than two dimensions. The Emission matrix can be embedded into HMM (Baum-Welch and Viterbi algorithms) and evaluated using in (1),

$$\textbf{Emi} = \begin{bmatrix} P(loc1/\ murder, Sep) & \cdots & P(locN/\ murder, Sep) \\ \vdots & \ddots & \vdots \\ P(loc1/\ Others, Oct) & \cdots & P(locN/\ Others, Oct) \end{bmatrix} \quad (1)$$

4.1.2. Transition matrix generation stage

The transition matrix is used to represent the Markov chain, and it is defined as a set of (location) states (S = {s1_, s2..., sn}), where Locations are represented by criminal movements in a different location. The transition is represented as the square array $T_{N\times N}$, where N= 10, for Iraq dataset, and N=32 for India dataset. The transition matrix can be construction as the following algorithm, where the summation of the probability of each row in the matrix must be an equal one.

```
An algorithm of Transition Matrix
Function Transition (x), where x is location of crime
Returns Probability matrix of criminal's movements
m = maximum value(x);
y = zeros(m,1) \\ vertice
construct probability matrix p m×m
for k=1to n-1
    y(x(k)) = y(x(k)) + 1;
    p(x(k), x(k+1)) = p(x(k), x(k+1)) + 1; where k is the number of movements
end loop
if y ==0 then
     p=0
else
     p = (p div y);
end if
```

4.1.3. Sequence generation stage

The sequence is the third parameter of HMM, in this step *is frist contribution* which defines the sequence of crime that happens, to obtain the most probable location of (crimes)states. In traditional and proposal work, the switch has been made between states and sequences (crimes) in two important phases:

- Replacing the sequence value with a state value, (i.e. the sequence as coupled parameters {crime types} is replaced with {locations}.
- Inserting states (locations) as a sequence into the Viterbi algorithm to generate the most probable sequence of crimes as an output.

4.2. DT algorithms with Viterbi algorithms

Here, the second strategy of improving decision tree algorithm is explained. The adopted steps of the proposed algorithm are illustrated in Figure 2 following the steps of:

Step 1: Start.
Step 2: Generate initialize the value of the EIMS matrix as mention in section 2 that computes the probability between {crimes type, month} and {locations}.
Step 3: Generate initialize the value of the TRNS matrix, in the same way, that mention in section, to determine criminal's movements.
Step 4: A sequence of couple parameter initialization.

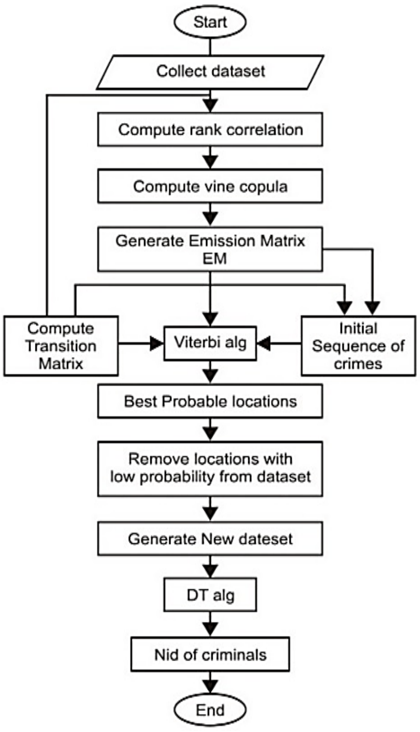

Figure 2. DT algorithms with Viterbi algorithms

Step 5: apply the Viterbi algorithm which can be used as an optimization algorithm by the elimination of the location with low probability as an outlier to maximize to increase the efficiency of time execution,
Step 6: Remove locations with low probability from the dataset.
Step 7: Generate new dataset, with improved quality of the dataset.
Step 8: Apply Decision Tree algorithm with hyperparameter, obtaining Nid as a label for first dataset, and kidnapping as a label for the second dataset.
Step 9: End

5. CONTRIBUTIONS

The Vine Copula Baum-Welch algorithm can achieve good results with the hybrid Vine Copula Viterbi algorithm for determining the sequential relation of past crime types, date, and locations. The speed of execution and accuracy of the Baum-Welch algorithm are enhancedNid of Iraq can help to facilitate the identification of criminals by the police and checkpoints. The contribution of this work, including a dataset, is optimized using a Viterbi algorithm for outlier detection and generates a new dataset.which is used to generate DT.

6. RESULT AND IMPLEMENTATION

The implementation of the proposed algorithm in both strategies can be explained in numerous stages: Stage 1: Generate the Emission matrix, Each value of the probability matrix related to crime type in a specific location. The main locations can then be determined from the column of crime type after transposition, as mention Emission matrix initially It was built by the implementation of the theory of Bayes, then was trained using latent Markov method ,the Implementation of generated matrix of the first dataset is shown in Table 3. While the Implementation of generated Emission matrix of the second dataset is shown in Table 4. The two features selected {thief, other IPC crimes}, the porobabilty crime occring appering in (32) states in indain country, as mention Emission matrix initially It was built by Bayes thory, then was trained using latent Markov method. Finally, The Emission matrix had been constructed with three parameters which had the main role HMM algorithms.

Table 3. The implementation emission matrix of the frist dataset

	Loc1	Loc2	Loc3		Loc10
Theft	0.1250	0	0	…	0
Murder	0.3333	0	0	…	0
Others	0.5000	0	0	…	0
Theft	0.4286	0.2857	0	…	0
Murder	0.0213	0.4894	0.1277	…	0.2340
Others	0	1.0000	0	…	0

Table 4. Emission matrix of the second dataset

	loc1	loc2	loc3	…….	loc32
Thief	0.0538	0.0010	0.0137	………..	0.0076
Other IPC Crimes	0.0019	0	0.0259	………….	0.0019

Stage 2: Generating a transition matrix for both datasets. The locations representation as a state diagram of Baghdad city is shown in Figure 3. It is important to note that the transition stages express the movement of criminals among location. The locations are represented in digraph of the Markov chain. Where the circle shape represents the nodes (cities) in which the crimes occurred and the edges connecting the nodes between them, to represent the possibility of transition between nodes.
Stage 3: Entering the sequence of crimes for both datasets.
Stage 4: Table 5 to Table 6 illustrates the comparison of the proposed algorithm with the traditional methods in terms of accuracy, relative mean square error (RMSE) and consumed time. The RMSE is considered for accuracy expression for HMM methods that includes Baum Welch and the proposed Baum Welch+Vetribi as shown in Table 5. It is well shown that the proposed algorithm for both datasets decreases the RMSE, clearly by (0.13).

At the other hand, Table 6 explains the comparison between the proposed (DT with Viterbi) and traditional DT. In this table, the accuracy of a confusion matrix is considered as the main indicator. The DT based on Matlab library for implementation.The accuracy of DT before the optimization is approximately

(98%) for the first dataset, whereas it is nearly (93.8%) for the second dataset. The accuracy of DT after optimization is improved by approximately (98.8%) for the first dataset. For the second dataset, the accuracy is also improved by around (95.9%) after optimization. The obtained results show that the proposed DT+Viterbi increases the accuracy for both datasets.

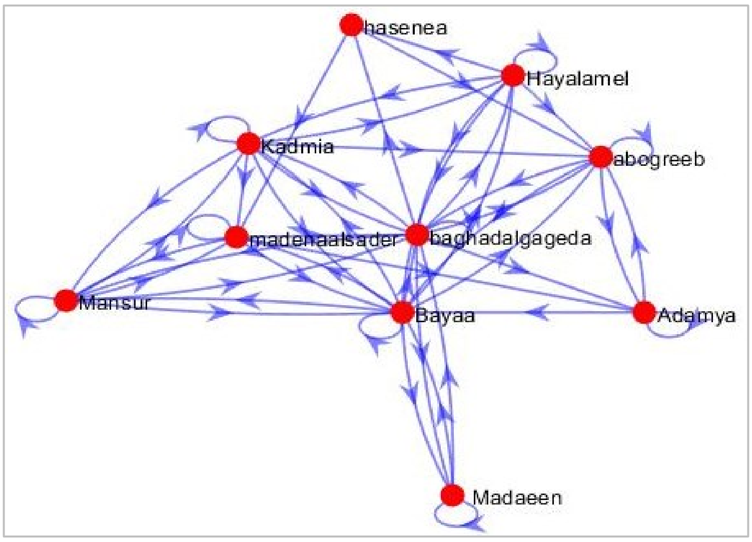

Figure 3. State transition before optimization of Iraq dataset

Table 5. Comparison of results for RMSE

	Baum Welch [2]	Baum Welch +Viterbi (preposed work)
Dataset1	0.17	0.12
Dataset2	0.04	0.03

Table 6. Comparison of results for accuracy in the confusion matrix

	Accuracy in the confusion matrix	
	DT [3]	Prepoded DT+ viterbi (preposed work)
Dataset1	~%98	~%98.8
Dataset2	~%93.8	~%95.9

Table 7. Shows the time-consuming comparison of the proposed methods (Baum Welch + Viterbi) with the traditional (Baum Welch) algorithms. Applied in two datasets, both improved to 0.05 and 0.08 respectively. It is proved that the proposed methods including Baun Welch outperform the traditional approaches which had been improved by (~0.15 sec) while Deepika K.K and at [14], achieves 4.24 s with 0.0751 as RMSE.

Table 7. Comparison of results for the consumed time

	Consumed Time in (second)	
	Baum Welch [1]	Baum Welch +Viterbi (preposed work)
Dataset1	0.2	0.05
Dataset2	0.1	0.08

7. CONCLUSION

In this paper, a proposed HMM has produced to predicate the level of danger in a specific region. The combination was done using Baum Welch with Viterbi from a side and DT with Viterbi at the other side to get a more accurate model. The proposed methods can achieve promising results for considering the Viterbi algorithm to determine the sequential relation of past crime types, with two latent parameters {month and locations}. The optimization of using DT method helped the proposed algorithm in mining the optimal solution. In addition, the consumed time is reduced in efficient performance.

REFERENCES

[1] Reem R. A., Muayad S. C., Salih M. A., "Developed Crime Location Predication Using Latten Markov Model," *Journal of Theoretical and Applied Information Technology,* vol. 96, no. 1, pp. 290-301, 2019.

[2] Reem R. A., Salih M. A., Muayad S. C., "Role of Data Mining in E-Government Framework," *Iraqi Journal for Computers and Informatics*, vol. 44, no. 1, pp. 14, 2018.

[3] Reem R. A., Muayad S. C., Salih M. A., "Improvement of Criminal Identification by Smart Optimization Method," *MATEC Web of Conferences* EDP Sciences, 2019.

[4] Chen H., "Special Issue Digital Government: technologies and practices," *Decision Support Systems,* vol. 34, no. 3, pp. 223-227, 2003.

[5] Tongwei Y., Peng C., "Data Mining Applications in government Information Security," *Procedia Engineering,* vol. 29, pp. 235-240, 2012.

[6] Qayyum, Shamaila, Hafsa S. D., Tehmina A., "A Survey of Data Mining Techniques for Crime Detection," *University of Sindh Journal of Information and Communication Technology,* vol. 2, no. 1, pp. 1-6, 2018.

[7] Keyvanpour, Mohammad R., Mostafa J., and Mohammad R. E., "Detecting and investigating crime by means of data mining: a general crime matching framework," *Procedia Computer Science,* vol. 3, pp. 872-880, 2011.

[8] V. Gupta and G. S. Lehal, "A survey of text mining techniques and applications," *Journal of Emerging Technologies in Web Intelligence,* vol. 1, no. 1, pp. 60-76, 2009.

[9] Kadhim B. A., Swadi, "A proposed framework for analyzing crime data set using decision tree and simple k-means mining algorithmsm" *Journal of Kufa for Mathematics and Computer,* vol. 1, no. 3, pp. 8-24, 2011.

[10] J. Hosseinkhani, S. Ibrahim, S. Chuprat, and J. H. Naniz, "Web Crime Mining by Means of Data Mining Techniques," *Research Journal of Applied Sciences, Engineering and Technology*, vol. 7, pp. 2027-2032, 2014.

[11] Li Xingan, "Application of data mining methods in the study of crime based on international data sources," *Tampere University Press*, 2014.

[12] Reem R. A., Muayad S. C., Salih Mahdi Al-Qaraawi, "Multistage Tree Model for Crime Dataset in Iraq," *Iraqi Journal of Computers, Communication, and Control & Systems Engineering,* vol. 19, no. 2, pp. 1-8, 2019.

[13] Mathew, Wesley, Ruben R., and Bruno M., "Predicting future locations with hidden Markov models," *Proceedings of the 2012 ACM conference on ubiquitous computing*, 2012.

[14] K. Deepika, Vinod, Smitha, "Crime analysis in India using data mining techniques," *International Journal of Engineering & Technology*, vol. 7, no. 2.6, pp. 253-258, 2018

[15] Mekathoti R. B., Kondapalli, "Optimizing crime hotspots and cold spots using Hidden Markov Model," *International Journal of Research*, vol. 4, no. 17, pp. 2348-6848, 2017.

[16] Joshi J. C., Tankeshwar Kumar, Sunita S., and Divya Sachdeva, "Optimisation of Hidden Markov Model using Baum–Welch algorithm for prediction of maximum and minimum temperature over Indian Himalaya," *Journal of Earth System Science,* vol, 126, no. 3, 2017.

[17] Görnitz, Nico, Mikio B., and Marius K., "Hidden Markov anomaly detection," *International Conference on Machine Learning*, pp. 1833-1842, 2015.

[18] Lehéricy Luc, "State-by-state minimax adaptive estimation for nonparametric hidden Markov models," *The Journal of Machine Learning Research,* vol. 19, no. 1, pp. 1432-1477, 2018.

[19] Zheng, Yuhui B. J., Le S., Jianwei Z., and Hui Z, "Student's t-hidden Markov model for unsupervised learning using localized feature selection," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 28, no. 10, pp. 2586-2598, 2018.

[20] Won, Kyoung-Jae, A. Prugel-Bennett, and Anders K., "Evolving the structure of hidden Markov models," *IEEE Transactions on Evolutionary Computation,* vol. 10, no. 1, pp. 39-49, 2006.

[21] Ghosh, Soumya K., and Shreya G., "Modeling Individual's Movement Patterns to Infer Next Location from Sparse Trajectory Traces," *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 693-698, 2018.

[22] Emdadi, Akram, Fatemeh A. M., Fatemeh Y. M., and Changiz E., "A novel algorithm for parameter estimation of Hidden Markov Model inspired by Ant Colony Optimization," *Heliyon,* vol. 5, no. 3,2019.

[23] Robinson, William N., and Andrea A., "Sequential fraud detection for prepaid cards using hidden Markov model divergence," *Expert Systems with Applications,* vol. 91, pp. 235-251, 2018.

[24] Zhang, Yanxue, Dongmei Z., and Jinxing Liu, "The application of Baum-Welch algorithm in the multistep attack," *The Scientific World Journal,* vol. 2014, pp. 1-7, 2014.

[25] Kaggle, "Daraset: Crime in India," 2017. [Online]. Available https://www.kaggle.com/rajanand/crime-in-india/.