Journal of Theoretical and Applied Information Technology
15th January 2019. Vol.97. No 1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

DEVELOPED CRIME LOCATION PREDICTION USING LATENT MARKOV MODEL

[1]REEM RAZZAQ ABDUL HUSSEIN,[2]Dr.MUAYAD SADIK CROOCK [3]Dr SALIH MAHDI AL- QARAAWI

[1]Informatics Institute of Postgraduate studies, Baghdad, Iraq
[2]Computer Engineering Department, University of Technology-Baghdad, Iraq
[3]Computer Engineering Department, University of Technology-Baghdad, Iraq

E-mail:[1]reemrazzak@yahoo.com [1]reem@uoitc.edu.iq, [2]120102@uotechnology.edu.iq,
[3]drsalihalqaraawi@gmail.com

ABSTRACT

Latent models, called hidden Markov models (HMMs), are types of algorithms that have been designed to detect crime activities by obtaining a sequence of observations from hidden values. The main contribution of these types of models is the fusion of coupled parameters with two types of HMM algorithms. The first algorithm is the Viterbi algorithm, which is commonly used to find the most probable path, and the accuracy of this algorithm is equal to 80%. The second algorithm is the Baum–Welch algorithm, which has been used to produce robust and accurate models. The modeling results normally focus on evaluating relative mean square errors in log likelihoods, transition matrices, and emission matrices for comparison of modeling performance based on different tolerance values. Previous reports have shown that the modified Baum–Welch algorithm can achieve good results for decreasing tolerance values. The goal of this Work is to generate a compact model that deals with ternary parameters rather than binary parameters by determining the sequential relation of past crime types and locations. Geographic locations can improve the HMM visualization in MATLAB. Moreover, crime levels and their most probable locations are predicted. The obtained results prove the goal of this work.

Keywords: *Vine Copula, Hidden Markov Models, Viterbi Algorithm, Baum Welch Algorithm, Measurement Errors*

1. INTRODUCTION

A flexible model for crime prediction is usually represented by multidimensional problems with N vectors. Thus, the simplification of the vine copula algorithm is often sought to generate an emission matrix with ternary parameters for the construction of the hidden Markov model (HMM) [1]. The HMM is used in our work to track the location of criminals in terms of time occurrence and location. Each detected location is associated with a criminal's gender, in which the proposed model assumes that a criminal may visit multiple locations. Then, the observable path is used as the inference of the hidden states. Considering that an accurate and fast algorithm is often needed, the Viterbi algorithm and the Baum–Welch algorithm can also be applied. The dataset typically consists of different types of information (e.g., numerical and categorical data) of Iraq criminals.

Past researchers in crime detection depended on data mining and developed systems that analyze patterns that have the highest probability of occurring and predict the regions where such patterns are likely to happen. Different data mining algorithms have been applied, such as naive Bayes, decision tree, and apriori [2]. the valuable connection to previously published research in this area such as ,the Baum–Welch algorithm was adopted in a past study, which collected GPS data for predicting criminal movements; this work used attributes according to their characteristics and trained an HMM for each cluster, obtaining an accuracy result of 13.85% [3]. Another work suggested two methods for crime detection and identification of criminals in India; K-means

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

clustering was used for analyzing crime detection, and KNN was adopted for classification with a visual representation via Google Maps; their approach helped agencies emphasize security in India [4]. In another past research, Bayes and mixture Gaussian distribution was utilized for predicting locations according to a synthetic dataset, not real data [5]. In another study, HMMs were implemented with the use of the Viterbi algorithm for identifying locations with a high probability of having been visited by 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 [6]. In our work, we try to overcome the problems with previous studies, we use bam Welch to get a more accurate model fusion with vine copula ,the fast handling problem, with Viterbi algorithm to get the best path of coupled parameters , using real data rather synthetic dataset .required in the first place to construct a compact model was constructed from ternary parameters, rather than binary parameters in the traditional model in HMM,

1. HMM CLASSES

An HMM model consists of process-based parameters to aid in the observance of sequences. columns.

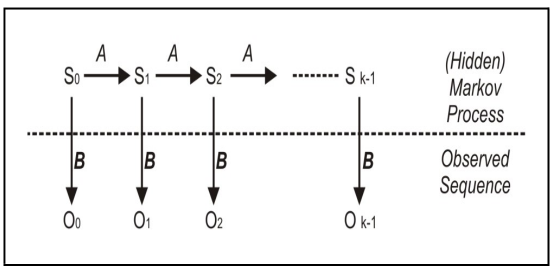

Figure 1: General Hidden Markov Model[6].

Here, the general parameters of HMM are as follows[7][9],[10],[11]:
K is the number of hidden states in the model;
S = $\{s_1, s_2, s_3 \ldots, s_T\}$ represents the hidden state sequence (location of crime);
O = $\{O_1, O_2, O_3 \ldots, O_T\}$ represents the observation sequence;
A =$\{aij\}$,where i,j=1,…,k represents the transition state;
B =$\{b_k(O_t)\}$, where t = 1,2,…T, represents the observation probability[6] [7]. All these notation of parameters to construct hidden Markov modle it may be forwared hidden Markov Model ,Viterbi algorithm ,baum Welch algorithm ,these algorithms which explained with detailed in next section (1.1) and (1.2).

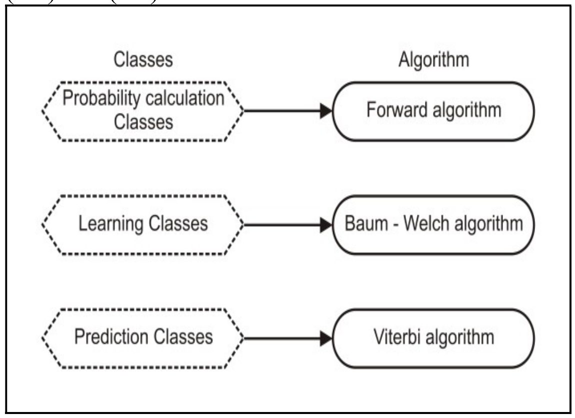

Figure 2: Correspondence between problems and class of HMM

Three classes are shown in Figure (2). Each class has a specific problem that can be solved by an HMM. The problems can be classified as follows[6]:

1) Probability calculation class. It calculates the probability $(O \mid \gamma)$, where HMM $\gamma = (A, \beta, \pi)$. Then, it calculate the observation sequence, where T is the length of time and $O = \{o_{1}, {}_{2}, \ldots, o_{T}\}$.

(2) Learning class. In estimating the parameters of $\gamma = (A, \beta, \pi)$, the observation sequence is $O = \{o_{1}, {}_{2}, \ldots, o_{T}\}$ and the probability $p(O \mid \gamma)$ of the class must be maximized.

(3) Prediction of the path. It computes the path by using the sequence $\Gamma = \{j1, j\ 2, \ldots, j\ T\}$ under maximum probability. When HMM $\gamma = (A, \beta, \pi)$, the observation sequence is presented as T and the length of time is $O = \{o_{1}, {}_{2}, \ldots, o_{T}\}$[10].

1.1 VITERBI ALGORITHM

The Viterbi algorithm, which was established by Viterbi (1976), is a dynamic programming algorithm used to construct transition and emission matrices. In particular, the algorithm can obtain the most likely path (state sequences) to generate output sequences. It works by obtaining the maximum overall possible state sequence. Figure (3) illustrates the sequence analysis to obtain observation $b_k(o_i)$ with maximum probability. This method can be used with gender sequences to predict the actions of criminals relative to their locations. We compute the probability of state sequences with $P(S_t, S_{t+1})$ and $\{o_t, o_{t+1}\}$ for each observation as follows [10]:

Initialization:

Journal of Theoretical and Applied Information Technology
15th January 2019. Vol.97. No 1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$\delta t(i) = \pi_i \times b_i(o_t)$, where $1 \leq i \leq 10$ … (1)

Induction:
$\delta_{t+1}(j) = \max [\delta_t(i) \times a_{ij}] \times b_j(o_{t+1})$, where $1 \leq j \leq 10$ …(2)

Termination:
$P(S_t,S_{t+1}) = \max [\delta_{t+1}(i)]$, where $1 \leq j \leq 10$ …(3)

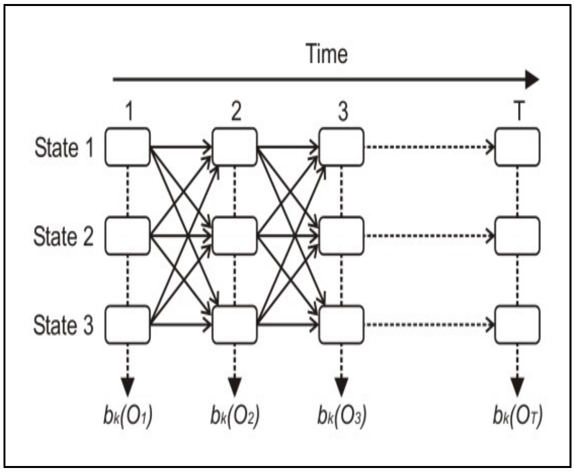

Figure 3: Viterbi scheme[6]

Viterbi algorithm
Function Viterbi (observations of length T, a state with length N**)** **returns** best sequence state to construct a path probability matrix viterbi[N+2, T] **for** each state s = 1: N do; initialization Viterbi[s,1]←a_0,s × bs(o_1) backpointer[s,1]←0 **for** each time step t =2t: T do; recursion for each state s 1: N do viterbi[s,t]← N max s_0=1 viterbi[s0,t−1] × as0,s × bs(ot) back_pointer[s,t]←N argmax s_0=1viterbi[s_0,t−1] × as$_0$,s **return** the backtrack path by following back_pointers to states back in

Many state sequences often result in the same output (sequence), but the probabilities are different. Calculating the probabilities through HMM may generate the output sequence together with the computation of the summation of all possible paths. However, the Viterbi algorithm cannot accurately solve the abovementioned problems. By contrast, the Baum–Welch algorithm can efficiently derive the local maximum likelihood. The forward-backward algorithm and a special case of expectation–maximization (EM) algorithm are discussed over this work.

2.1 Baum–Welch algorithm

Baum (1970) proposed the use of the Baum–Welch algorithm based on the statistical analysis of the probabilistic methods of the Markov model. The forward-backward method, wherein the backward part represents the probability of partial observation sequences from time t+1 to end, can be computed iteratively [3][6].

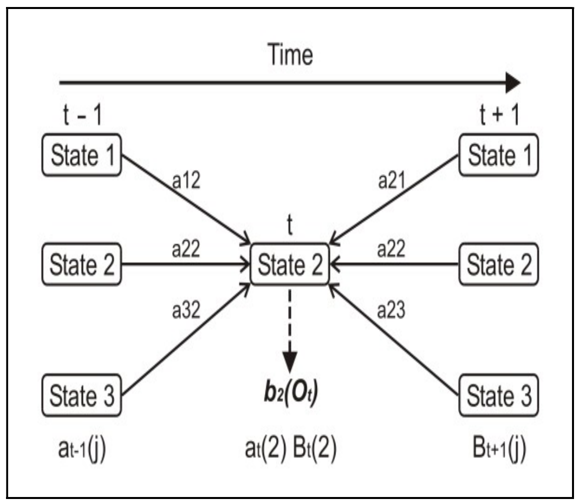

Figure 4 Forward-backward schema[6].

Figure (4) illustrates the forward-backward algorithm. The variables [$a_t(2)$ and $B_t(2)$**)** for State 2 are computed when the time is equal to t. Depending on the steps of the forward-backward variables, the algorithm computes the probability of state $a_{t-1}(j)$ and state $B_{t+1}(j)$ .the forward-backward algorithm shown as following:

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Forward-backward algorithm
Function forward_backward (observations of len T, output location W, latent state set is crime type Q) returns HMM=(A trans,B emmision) initialization the variables Ąand B continue iterate until convergence Expectation: $E_t(i)=a_t(i)\times B_t(i)/P(O/\lambda)$. Max_paramter : $M_t(i,j)=\frac{\alpha t(i)\times aij\times bj(Ot+1)\times Bt+1(j)}{P(O/\lambda)}$ $\bar{a}ij=\sum_{t=1}^{T-1} Mt(i,j)/\sum_{t=1}^{T-1} Et(i)$ $b\grave{}ij=\sum_{\substack{t=1\\ ot=wk}}^{T-1} E1(i)/\sum_{t=1}^{T-1} Et(i)$ return (Ątran,B emission)

E(i) is the probability of being in state Si at time t. Given the observation sequence O, the variables M and E satisfy the relationship.

$Et(i)=Mt(i,j),\ 1\leq j\leq 10$ (7)

When Et(i) is summed for all locations (except instant T), the expected number of time relative to location Si can be obtained. When Mt(i,j) is summed for all locations (except instant T), the number of times that a criminal has moved from location i to j can be obtained. Et(i) and Mt(i,j) are computed to re-estimate the parameters, as shown in algorithm[9][10],[11].

2.3 Vine Copula

The first step in our work includes generating tree T as a noncyclic graph, where T={V, E}, and V refer to the vertices of Tree, and they can be represented as V={crime, gender, location,…..,v_N} for N dimensions. Moreover, E is denoted as a set of edges of Tree, and it can be represented as E={e_1,e_2,….., the e_N} to link the vertices. Edge values represent a correlation coefficient in the first level, a partial correlation with one constant coefficient in the second level, and two constant coefficients in the third level of the partial correlation. These values are computed to provide a regular vine copula with different tree structures. We use two methods to select the tree structures. The first one depends on the estimation of the summation path of a tree, which is criteria for selected parameters. A straightforward correlation coefficient is computed for each edge by using the following equation [1,13,14]:

$$r=\frac{1-6\sum D2}{N\,(N2-1)}\ldots(7)[1]$$

where n= number of observations ,
di= rank(xi)-rank (y) …(8)
Then, the highest summation is computed to generate the data ranks, as shown in Figure (5).

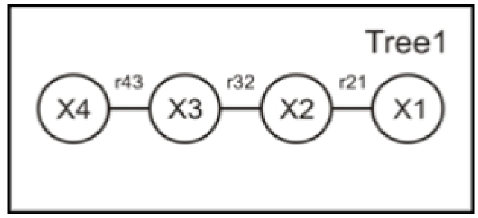

Figure 5 Level 1 of Tree ***[1][13]***

The second method uses a Partial Correlation Coefficient (PCC) with one constant, which is second criteria for selection parameters which are computed with the following formula:

Where r represents a rank correlation, for example, r12 represents a relation between x1 and x2

$$r12.3=\frac{r12-r13\,r23}{\sqrt{(1-r13^2)}\sqrt{(1-r23^2)}}\quad\ldots(9)[13]$$

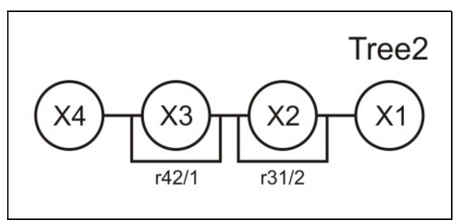

Figure (6) Level 2 of Tree [1][13]

PCC is used to test the constant conditional correlation between two variables. For example, PCC tests the variables x_4, x_2 with the constant x_3, which is represented as r42_3, as shown in Figure (6). Moreover, the partial rank correlation is computed on the basis of the standard rank.

3. PROPOSED METHOD

Here, we need to predict the locations of the crimes, and we used HMM to obtain the hidden pattern.

The data were collected from a website and social media. Baghdad city is divided into two parts as shown in Table 1:

- kurahk
- Resafa

Table1: Main Locations in Baghdad City

seq	Location	Representation
1	kurahk	Location A(Loc_A)
2	Resafa	Location B (Loc_B)

HMM must generate two main locations, locA and LocB, these locations represent the workflow of the proposed system is shown in figure (7).

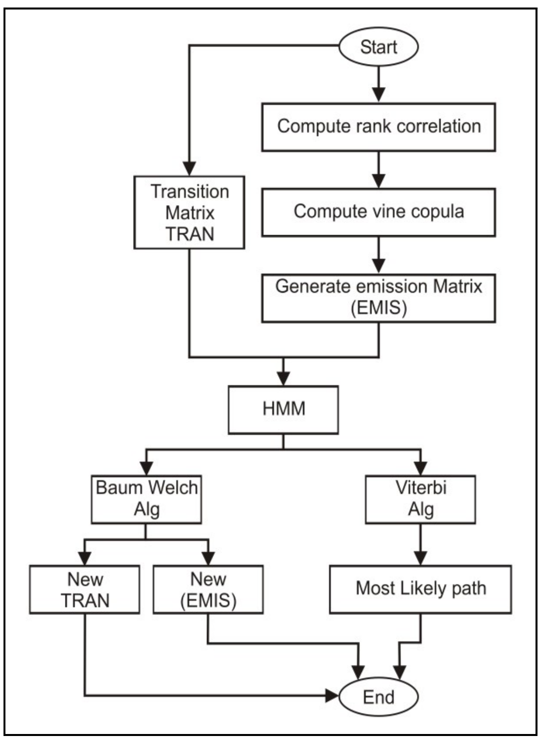

Figure (7) Workflow of the proposed method

We explain the elements of Workflow in the next sections.

3.1 Emission Matrix Generation

The proposed method involves the computation of a correlation, and thus, we applied a vine copula to obtain the necessary parameters. Our work is A modified Dißmann's algorithm was used for the proposed method to increase the dependency at each level of the tree, and a rank-correlation coefficient was adopted instead of a simple correlation coefficient (Kendall's measure). The hybridization of the method depended on the estimation of the summation path of the tree. The PCC method was adopted on the basis of the type of criminal attribute with response value {x1,x2,…., N}. Table (1) presents the categories of the variables.

Table 2: Categories of variables

Seq	Names	Classes
1	Gender	{Male, Female}
2	Crime Type	{Murder, Theft, Other Crimes}
3	Location	{locA, locB}
4	Age	{Young, Old}

Conditional probability (Bayes theorem) is applied to the tree structure and used for calculation. Assuming that criminals are murders, with couple parameters {gender, location}, to produce three dimensions rather than two dimension

$P(murder/locA,male)=\frac{P(murder \cap locA,male)}{P(locA,male)}$

, where $P(locA,male) \neq 0$…(10)

$P(murder/locB,male)=\frac{P(murder \cap locB,male)}{P(locB,male)}$,

where $P(locB, male) \neq 0$…(11)

$(murder/locA,female)=\frac{P(murder \cap locA,female)}{P(locA,female)}$,

where $P(locA, female) \neq 0$…(12)

$P(murder/locB,female)=\frac{P(murder \cap locB,male)}{P(locB,male)}$,

where $P(locB, female) \neq 0$…(13)

When criminals are thieves, with couple parameters {gender,location} , the Bayes theorem adopts the following equations:

$P(thief/locA,male)=\frac{P(thief \cap locthe\ the\ A,male)}{P(locA,male)}$,

where $P(locA, male) \neq 0$…(14)

$P(thief/locB,male)=\frac{P(thief \cap locB,male)}{P(locB,male)}$,

where $P(locB, male) \neq 0$…(15)

$P(thief/locA,female)=\frac{P(thief \cap locA,female)}{P(locA,female)}$, where $P(locA, female) \neq 0$…(16)

$P(thief/locB,female)=\frac{P(thief \cap locB,female)}{P(locB,female)}$, where $P(locB, female) \neq 0$…(17)

When the criminals commit other crimes, with couple parameters the corresponding equations are as follows:

$P(others/locA,male) = \frac{P(others \cap locA,male)}{P(locA,male)}$, where $P(locA, male) \neq 0$…(18)

$P(others/locB,male)=\frac{P(others \cap locB,male)}{P(locB,male)}$, where $P(locB, male) \neq 0$…(19)

$P(others/locA,female)=\frac{P(others \cap locA,female)}{P(locA,female)}$, where $P(locA, female) \neq 0$…(20)

$P(others/locB,female)=\frac{P(others \cap locB,female)}{P(locB,female)}$, where $P(locB, female) \neq 0$…(20)

The emission matrix can be embedded into HMM (Baum–Welch and Viterbi algorithms). An observation matrix is obtained from Equations (9)–(20). The emission matrix is obtained and represented as follows:

$$Emi = \begin{bmatrix} P(murder/locA,male) & \cdots & P(murder/locC,female) \\ \vdots & \ddots & \vdots \\ P(others/locA,male) & \cdots & P(others/locC,female) \end{bmatrix}$$

Finally, The emission matrix had been constructed with three parameters which had main role HMM algorithms.

3.2 Transition Matrix Generation

The transition matrix is used to represent the Markov chain, and it is defined as a set of states (S ={s1_,s2,...,sn}), where states are represented by criminal movements from one location to another. Each location movement is called a step. If the chain is in state location A, then it moves to state location B within the next step and a probability denoted by $P_{loctation\_AB}$. We used this chain in the study, wherein each state is denoted by coupled parameters represented as fully connected (ergodic) HMM in the following:

S_1=male,locA…{location A, when the criminal is male}

S_2= male,locB…{location B, when the criminal is male}

S_3=female,locA…{location A, when the criminal is female}

S_4= female,locB…{location B, when the criminal is female}

Probabilities of $P_{(loc\_AB)}$ represent the transition probabilities. The criminal may remain in the same state if this situation occurs with probabilities P_{loc_AA} and P_{loc_BB}. An initial probability is defined as the starting state. The transition is represented as the square array $T_{n\times n}$, where n= 4×4.

$$T= \begin{bmatrix} P(locA,male) & \cdots & P(locB,female) \\ \vdots & \ddots & \vdots \\ P(locA,male) & \cdots & P(locB,female) \end{bmatrix}$$

By determining a specific state as the starting state in the transition matrix, the entries in the first row of matrix T can then be represented by probabilities of various types of crime occurrence.

4. MODIFICATION HIDDEN MARKOV MODELS

HMMs can represent a label with multiple states. This work assumes that the number of the chain is C with N states. Thus, HMM has N^C states, which is an inefficient system because large amounts of data are needed to train such a system.

Copula with HMM requires (N × C) states, and HMM is impractical even for a few classes and states. To solve this problem, the coupling is embedded in HMM. The coupling model was explained in Section 3, and the result of each state obtained from vine copula summarizes one or more class.

The new findings of our work show as follows :

1. we modified Diβmann's algorithm Spearman correlation coefficient, and partial correlation for validation data and applied on Iraq dataset.
2. use both AIC and BIC for finding the truth of value.
3. validation data to select best-fit tree structure.
4. Generate an emission matrix from the simplified tree structure.
5. vine copula algorithms are embedded with Viterbi and Baum–Welch algorithms. This modification allows for an N-dimensional space rather than only two dimensions in standard HMM, wherein each state transition is a compact state.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

5. RESULTS AND DISCUSSION

MATLAB software is used to obtain the analysis and outcome of our work. Common setting issues are solved in the following stages:
The rank correlation matrix whenever the between gender and crime type scores (p = 0.6245), whereas no significant correlation is found between
the row is equal column the value of the cell is equal to one, for example, the correlation between (Gender, Gender), (Age, Age),…is equal one age and gender scores (p = 0.3700). The rank correlation matrix whenever

- Initialization stage: This stage computes the rank correlation matrix shown in Table (3). For example, a significant correlation is found

Table 3: Ranked correlation

	Gender	Age	crime type	location
Gender	1.000	0.370	0.624	0.524
Age	0.370	1.000	0.462	0.494
Crime Type	0.624	0.462	1.000	0.664
Location	0.524	0.494	0.664	1.000

Table 4: Average estimation and PCC

No	Tree	Sum of path	Ave _Estimation	Conditional_copula	PCC
T1	−2>−1>3	r12+r13	−0.2123	C23_1	−0.6812
T2	−1>−2>3	r12+r23	−1.1057	C13_2	0.0176
T3	−1>−3>2	r13+r32	0.1491	C12_3	0.3936

.

Table 5: AIC and BIC evaluation

No	AIC	BIC
Tree 1	−1.2208	−1.2235
Tree 2	−2.0347	−2.0373
Tree 3	−1.0548	−1.0574

In Table (4) the sum of the path is further computed. PCC is the value of the conditional (partial) correlation. Three structures can be generated, and the suitable model can choose either the sum of paths or conditional copula (Dißmann's algorithm). The results of implementing both strategies have the highest value in T3. If each strategy obtains different tree structures, then AIC and BIC evaluation must be performed. Table (5) shows the selection of the best structure. The initialization of the two matrices is conducted in this stage. The first matrix is called the transition matrix, as shown in Table (6). The second matrix is called an emission matrix, which applies Bayes theorem, as shown in Table (7).

Table 6:initial transition martrix

	locA, m	locB, m	locA,f	locB, f
locA,m	0.5	0.3	0.1	0.1
locB,m	0.2	0.6	0.1	0.1
locA,f	0.2	0.1	0.5	0.2
locB,f	0.1	0.2	0.1	0.6

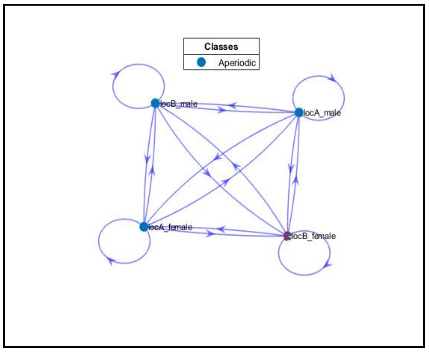

Figure (8) Aperiodic Markov

Table (7) initial emission matrix

	locA, male	locB, male	locA, female	locB, female
theft	0.8302	0.0566	0	0.1132
murder	0.5882	0.1373	0.0784	0.1961
other crimes	0.8276	0.0345	0	0.1379
Non criminal	0	0	0	0

- The implementation of Markov chain is shown in Figure (8). It is derived from a specified state transition matrix. The class of MarkovMarkov chain is aperio
- dic, wherein each state can be obtained in another state with one step. The items in the transition matrix are non-zeroes.
- Recurrent State: ["locA_male" "locB_male" "locA_female" "locB_female"].
- The initial emission matrix is shown in Table (7)
-
- Forward Stage: Viterbi algorithms are calculated, and the inputs are included in the transition matrix, emission matrix, and sequence of states. that the length of the sequence be equal to 10. The most probable path of states is as follows:
  [{'male,location A'} {'male,location B'} {'male,location B'} {'male,location B'} {'male,location B'} {'male,location B'} {male,location B'} {'male,location B'} {'male,location B'} {'male,location B'}….].

Table 9: RMSE of Baum–Welch Algorithm

		RMSE_Baum Welch		
Tolerance	Iteration	log	Tran	Emi
0.1	3	0.00164548	0.00198764	0.00656947
0.01	3	.00164548	0.00198764	0.00656947
0.001	88	7.1291e−06	9.9268e−04	9.1968e−04
0.0001	163	1.78e−05	3.91e−08	3.91e−08
0.00001	163	1.78e−05	3.91e−08	3.91e−08
0.000001	164	4.32802e−07	8.73161e−12	8.73162e−12
0.00000001	165	3.61403e−12	4.31962e−13	4.36401e−13

Tolerance values help in determining the number of iterations that the training algorithm executes. Baum–Welch algorithm is terminated when the log-likelihood, length of the input sequence, and transition and emission matrices are lower than the value that needs to be satisfied. We further compute the Relative Mean Square Error (RMSE) of the transition and emission matrices at different tolerance levels, as shown in Figure (9). When the algorithms fail to reach the accepted tolerance, there is another way to control a number of iterations, by indicating the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

max number of iterations, finally, the limitation and assumption had been undertaken in both Viterbi algorithms and Baum in parameters, there are three parameters that Viterbi need :

- Transition Matrix.
- Emission Matrix.
- Sequence
- The limitation of accuracy depends on a sequence.

While Baum Welch the iteration of learning to depend on the value of tolerance Require parameters show as follows :

- Transition Matrix.
- Emission Matrix.
- Sequence.
- States

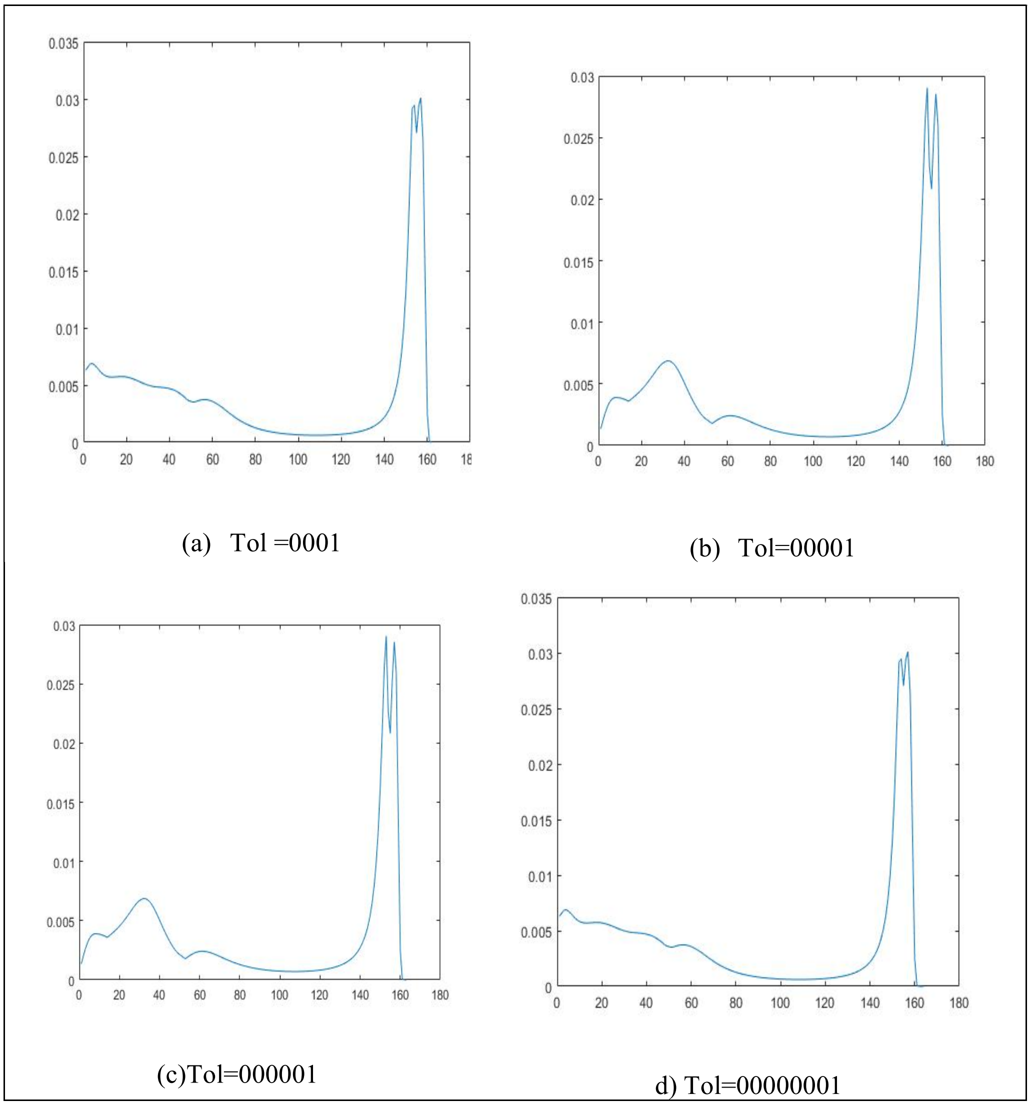

Figure: 9 RMSE

According to Figure (9), the results of the decreasing error is related to the tolerance value. Here, the lowest error was obtained when tolerance was equal to 1e−7, the RMSE of the transition matrix was equal to 4.31962e−13, and the emission matrix was equal to 4.36401e−13. The sequence of the path is very important in reflecting the effect of the accuracy of the algorithm .solved the high dimensionality problem. Assuming that criminals are murders, theft, and other types

Journal of Theoretical and Applied Information Technology
15th January 2019. Vol.97. No 1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

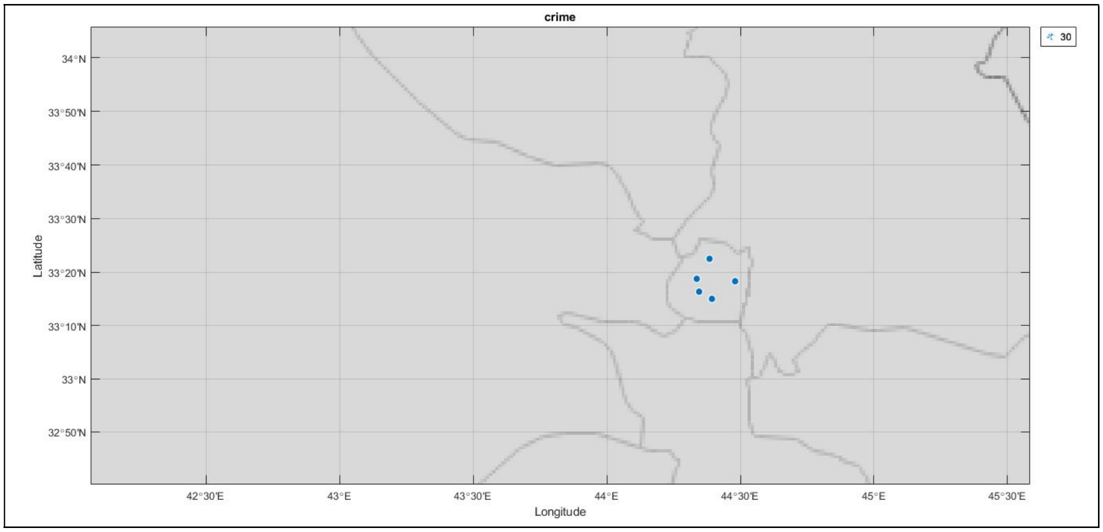

Figure 10: Crime location

,that coupled with gender parameter that.Figure (10) visually represents crimes on a map of Baghdad City in accordance with the regions where they occurred. These regions (sub-locations) belong to locA and locB, as mentioned in Section (3). The high-dimensionality problem is solved with the assumption that that criminals are murderers, thieves, or other types, coupled with the gender parameter.

6. CONCLUSION

In this paper, we used a coupled latent Markov model to predict crime locations. The proposed the model used the correlation between the crime location and criminal information, such as gender in detecting the expected next crime location. It applied copulas to the Viterbi algorithm and Baum–Welch algorithm. Both algorithms were implemented in many stages, and then each stage was evaluated. The rank correlation was used to evaluate the initial stage, whereas AIC and BIC were used to evaluate the choice for the best tree structure. The forward stage represented the most probable path with an accuracy of 80%. RMSE was used to evaluate the performance of the learning stage at the different tolerance level

REFERENCES

[1]: Kraus, Daniel, and Claudia Czado. "Growing simplified vine copula trees: improving Di {\ss} Mann's algorithm." *arXiv preprint arXiv:1703.05203* (2017).

[2]: Sathyadevan, Shiju, and Surya Gangadharan. "Crime analysis and prediction using data mining." *Networks & Soft Computing (ICNSC), 2014 First International Conference on*. IEEE, 2014.

[3]: Mathew, Wesley, Ruben Raposo, and Bruno Martins. "Predicting future locations with hidden Markov models." *Proceedings of the 2012 ACM conference on ubiquitous computing*. ACM, 2012.

[4]: Tayal, Devendra Kumar, et al. "Crime detection and criminal identification in India using data mining techniques." *AI & Society* 30.1 (2015): 117-127.

[4]: Vural, Mehmet Sait, and Mustafa Gök. "Criminal prediction using Naive Bayes theory." *Neural Computing and Applications* 28.9 (2017): 2581-2592.

[5]: Mekathoti Rajesh Babu m Dr. Kondapalli Venkata Ramana, "Optimizing crime hotspots and cold spots using Hidden Markov Model", International Journal of Research, Vol 04 No 17, December 2017,2348-6848.

[6]: Zhang, Yanxue, Dongmei Zhao, and Jinxing Liu. "The application of Baum-Welch

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

algorithm in the multistep attack." *The Scientific World Journal* 2014 (2014).

[7]: Won, Kyoung-Jae, Adam Prugel-Bennett, and Anders Krogh. "Evolving the structure of hidden Markov models." *IEEE Transactions on Evolutionary Computation* 10.1 (2006): 39-49.

[8]: Ghosh, Tusharkanti, Hierarchical hidden Markov models with applications to BiSulfite-sequencing data. Ph.D. thesis,2018.

[9]: Wiig, Trygve Bertelsen. Wiig, Trygve Bertelsen. "*Approximate marginal inference in binary Markov random fields using a mean squared error energy approximation and the junction tree algorithm*". MS thesis. NTNU, 2016.

[10]: Nguyen, Nam Thanh, et al. "Learning and detecting activities from movement trajectories using the hierarchical hidden Markov model." *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*. Vol. 2. IEEE, 2005.

[11]: Ismael, Bilal, et al. "Improving time series classification using Hidden Markov Models." *Hybrid Intelligent Systems (HIS), 2012 12th International Conference on*. IEEE, 2012.

[12 Pereira, Guilherme, et al. "Spatial R-vine copula for streamflow scenario simulation." *Power Systems Computation Conference (PSCC), 2016*. IEEE, 2016.

[13]: Morales, O., Dorota Kurowicka, and A. Roelen. "Eliciting conditional and unconditional rank correlations from conditional probabilities." *Reliability Engineering & System Safety* 93.5 (2008): 699-710.

[14]: Aas, Kjersti. "Pair-copula constructions for financial applications: A review." *Econometrics* 4.4 (2016): 43.

Journal of Theoretical and Applied Information Technology
15th January 2019. Vol.97. No 1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

APPENDIX

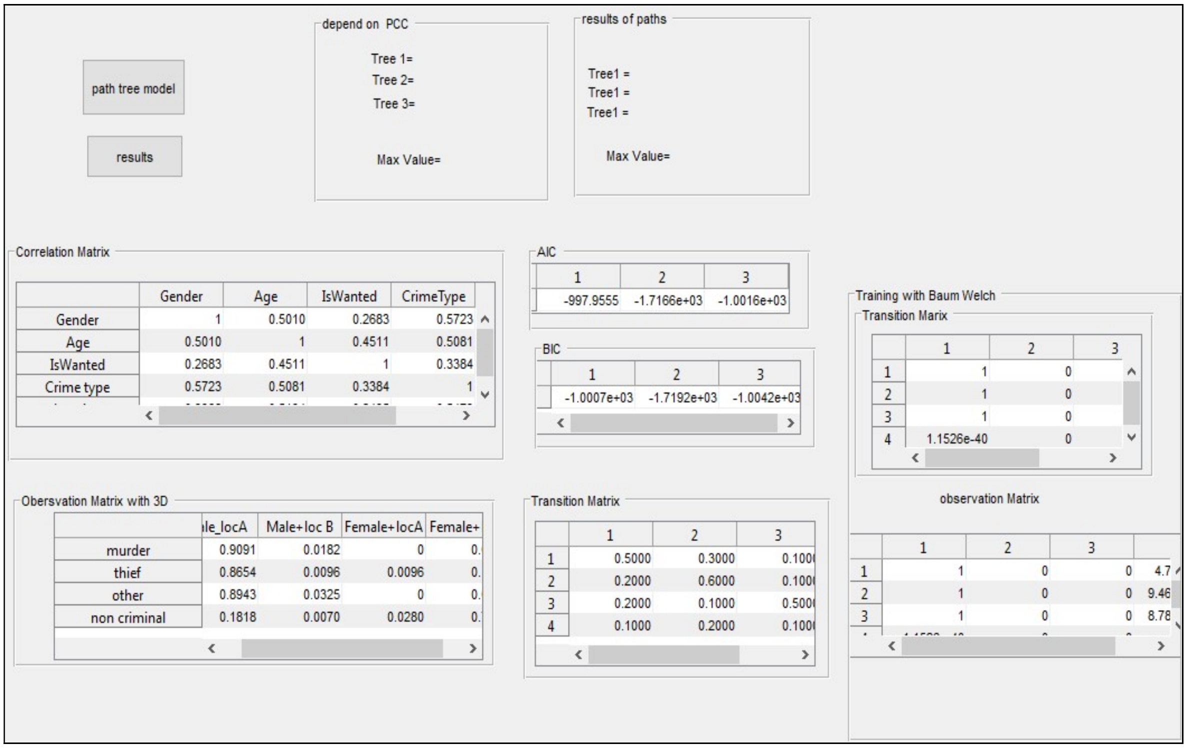

Figure 11: The Interface of Implementation