Deciphering Crime and Forensics through the lens of Graph Theory.

A Dissertation for

MAT-651 Discipline Specific Dissertation

Credits: 16

Submitted in partial fulfilment of Masters Degree

M.Sc. in Mathematics

by

Ms. ARYAA ALIAS PALLAVI ASHISH S REGE

22P0410005

ABC ID : 961-493-798-641

201906523

Under the Supervisor of

Dr. JESSICA FERNANDES E PEREIRA

School of Physical & Applied Sciences

Mathematics Discipline

GOA UNIVERSITY

APRIL 2024

Examined by:

Seal of the School

Declaration

I hereby declare that the data presented in this Dissertation report entitled, "Deciphering Crime and Forensics through the lens of Graph Theory" is based on the results of investigations carried out by me in the Mathematics Discipline at the School of Physical & Applied Sciences, Goa University under the Supervision of Dr. Jessica Fernandes e Pereira and the same has not been submitted elsewhere for the award of a degree or diploma by me. Further, I understand that Goa University will not be responsible for the correctness of observations / experimental or other findings given the dissertation.

I hereby authorize the University authorities to upload this dissertation on the dissertation repository or anywhere else as the UGC regulations demand and make it available to any one as needed.

Signature:

Student Name: Aryaa A S Rege

Seat no: 22P0410005

Date: MAY 08, 2024

Place: GOA UNIVERSITY

Certificate

This is to certify that the dissertation report "Deciphering Crime and Forensics through the lens of Graph Theory" is a bonafide work carried out by Ms. Aryaa Ashish S Rege under my supervision in partial fulfilment of the requirements for the award of the degree of Master of Science in Mathematics in the Discipline Mathematics at the School of Physical & Applied Sciences , Goa University.

Signarure :

Supervisor : Dr. Jessica Fernandes e Pereira

Date: MAY 08, 2024

Signature of HoD of the Dept

Date: 10/5/2024

Place: Goa University

School Stamp

PREFACE

This report summarizes the information and data of the research carried out in my dissertation work, as a part of my M.Sc. programme at Goa University.

Initial part of my report includes the materials collected by reviewing the research papers available on the selected topic while the latter part summarizes the conclusions obtained through research work and analysis conducted.

Chapter 1 provides an overview of some basic terms and concepts used during the analysis along with the scope and motivation of the dissertation while the results, methods, algorithms and data obtained from the research papers related to the topic forms the 2nd chapter.

Chapter 3 and 4 summarizes the outcomes of the research work conducted. This includes the robbery case analysis using a specific methodology and fingerprint analysis using few algorithms. Chapter 3 includes analysis of two robbery cases satisfying the required constraints while chapter 4 includes the results obtained through analysis of fingerprints over a sample space of 80 participants. Chapter 5 forms the conclusion of my dissertation work.

Acknowledgements

First and foremost, I would like to express my gratitude towards my guide, Dr. Jessica Fernandes e Pereira, for constantly motivating, supporting and guiding me throughout the period of my dissertation work. She inspired me to think differently. This endeavour would not have been possible without her help and supervision. I would also like to thank all the participants who provided me with their fingerprint for conducting my analysis.

A special thanks to the IPS Arvind Gauns and PSI Sanket Pokhre, for spending more than 4 hours to provide me with the detailed information of the 2 robbery cases for conducting the required analysis. Their time, readiness and efforts made the journey of analysis smoother. I would also like to express my gratitude to Dr. Waghmare, the Head of Forensic Lab, Verna Goa for guiding me through the current procedures of fingerprint matching which helped me with a clear vision for my dissertation.

Lastly, I would also like to thank my parents and friends for their constant help and support.

Abstract

This report summarizes the application of Graph Theory in Forensic science and Criminology based on the Research papers referred.
Graph theory forms a helpful tool in analysing relations between all possible suspects and objects under investigation, to find the culprit of the crime. Evidences found at crime scene form the basis for finding the culprit where one of the main evidence forms a fingerprint. Various concepts such as Graph colouring, Matching, Planarity, etc can be applied in solving crime and analysing fingerprints found at crime scene.
Two of the most asked question in this field are:

1. Can Graph Theory be used to provide a solution to reduce crime ?

2. In the case where fingerprint is not found in the data base, can Graph Theory be used to eliminate the suspects or the criminal ?

This paper tries to attempt these questions in two ways. Firstly, few of the research papers dealing with this topic are reviewed. Secondly,Fingerprint analysis is conducted on a smaller sample size to check the inheritance of the patterns and differences between the two genders. The positive result obtained from the latter will help to eliminate suspects.

Table of contents

List of figures

List of tables

Chapter 1

INTRODUCTION

1.1 Introduction

Graph Theory:

Graph theory is the study of graphs, which are mathematical structures used to model pairwise relations between objects. It forms a helpful tool to quantify as well as simplify various interconnections. It was discovered by Leonard Euler.

Forensics and Criminology

The study of application of scientific principles and techniques to objects under investigation by court of law is defined as Forensics. Fingerprint analysis, network analysis, DNA related analysis, etc are dealt under forensics. To do such analysis, various subjects like Chemistry, Biochemistry, Physics etc can be applied.

The scientific study of crime and criminals is called Criminology.

Motivation:

Importance of Graph and Networks!

Many advance methods are required to be developed in criminology and forensics to fasten and improve the process of identifying the culprit through evidences found at crime scene.

Graph theory forms a helpful tool for the same. By definition, Graph theory is the study of graphs to model relation between two points. So, here it acts as a tool to find relation between two entities like objects found at crime scene, suspects etc. This helps us to model a graph where these entities form vertices and their relations as edges which then can be solved using various concepts of graph theory. The advantage of graph theory over other methods is that the former is an effective tool i.e. cost effective, faster and a simpler tool.

Fingerprint Analysis

The fingerprint analysis was first advocated by Sir William Herschel. He suggested that it could be used to identify the suspects.

Fingerprint patterns are present inside and on the tips of the finger. The ridges present on the skin are called friction ridges which together with the valley forms a unique pattern. The uniqueness and immutability of fingerprint forms the basis for fingerprint analysis. These patterns are developed in the womb due to the embryonic pattern in the fetal development stage. Accidents and diseases can alter the fingerprint patterns.

The ridges and valleys were first identified in 1686 by Marcello Malpighi but fingerprints were recognized as means of personal identification in 1880 by Henry Faulds.

There are two methods of identifying the fingerprints. First is by comparing the lifted prints and second by live scanning. The former is used in forensics while latter is used for authentication purposes.

The most important feature of the fingerprint which must be used in identification is the ridge characteristics, more specifically the Ridge Ending and Bifurcation as shown in the Figure 1.1.

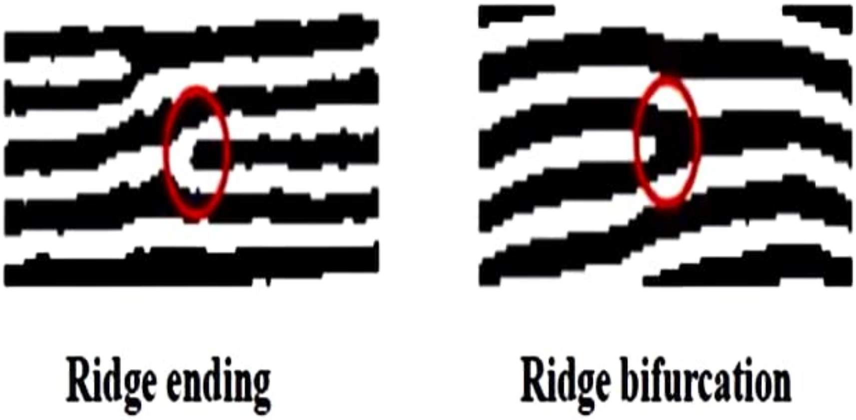

Fig. 1.1 Characteristic

1.2 Basic Results/Definitions

Graph : A graph G is a finite set of elements called vertices and a possibly empty set of pair of distinct elements from the vertex set. Sum of edges incident on the vertex or total number of out-degree and in-degree of a vertex is defined as the degree of a vertex where the total number of vertices adjacent from u is the out-degree of u and the total number of vertices adjacent to u is the in-degree of u. Similarly, a $u-v$ path can be defined as a $u-v$ walk in which no vertex is repeated.

Connected graph : A graph is connected if for every two vertices u and v of G there exists a $u-v$ path in G.

Vertex colouring : Vertex colouring of graph G on n vertices is a function $f : V(G) \rightarrow \{c_1, c_2, \ldots, c_n\}$ such that if uv belong to $E(G)$ then $f(u)$≠f(v) .

$Acolouringinwhich$k$coloursareusedfor$V(G)$iscalled$k$-colouringof$GandG$issaidtobe$k$-colourable.$

Edge colouring: Proper edge colouring is colouring of edges of graph in such a way that two adjacent edges receive different colours. A minimum number of colours that can be used to colour the edges of G is defined as edge chromatic number.

Planar graph: Graphs which can be drawn on a plane without edge crossing are called planar graphs and a graph in which no edges cross is a plane graph.

Crossing number of G is the minimum number of edges that cross among all the drawings of G in a plane.

G and H are called **Isomorphic graphs** if there exists a bijection ϕ from vertex set of G to vertex set of H such that uv is an edge in G iff $\phi(u)\phi(v)$ is an edge in H. These graphs have same number of edges, same number of vertices and same degree of vertex.

Let G be a graph and U be subset of $V(G)$ then the subgraph induced by the set U is the graph consisting of vertices in set U and for u, v belonging in U, uv is edge in U if uv is an edge in G.

G is called a **k-Partite Graph** if the vertex set of G can be partitioned into k-subsets such that every edge of G joins a vertex of V_i to a vertex of V_j where $i \neq j$. Now, If $k = 2$, G is said to be bipartite graph. A bipartite graph G with partition V_1 and V_2 is called **Complete Bipartite graph** if for every pair of vertices $v_i \in V_1, v_j \in V_2$ $v_i v_j \in E(G)$.

Digraphs or directed graphs consist of a finite non empty set of elements called vertices and a set (possibly empty) of ordered pairs of distinct vertices called arcs or directed edges. For digraphs, sum of in-degree of each vertex together is equal to sum of out-degree.

Independent edges: Two edges are said to be independent if they are not adjacent in graph G. A set of pairwise independent edges in G is called a matching in G.

Ridge is a path of raised skin on the pad of a finger while **Core** is the approximate center of a fingerprint. The triangular reference point on a fingerprint showing orientation is defined as **Delta**.

Minutia are the points which can be identified other than Core and Delta, which is unique in each individual. The two **Minutia Types** are Bifurcation and Ridge Ending.

Valley is the gap between two ridges. Minutia and other characteristics like short ridge,lake, trifurcation are called as **Ridge characteristics**.

Adjacency Matrix is a square matrix used to represent a finite graph where if the pair of vertices are adjacent then it takes value 1 or else value 0. This type of matrix is symmetric and diagonal elements is 0.

Incidence Matrix is a matrix is the relation between vertices and edges of a graph in which if the edge is incident on a vertex then it takes value 1 or else 0. The row wise sum of this type of matrix gives the degree of each vertex.

Labelled Graphs is a graph that has labels on the vertices or edges.

Numbers can be assigned as labels to the edges of a labelled graph in order to enable them to be used in application problems. Such an assignment is called weight of edges.**Weighted Graph** is a graph whose edges are given labels.

Chapter 2

Literature Review

In this chapter some research papers on the topic are reviewed.

2.1 Paper: Utilising Graph Theory to Model Forensic

Utilising Graph Theory to Model Forensic [1]

This paper is a basic paper from which the thought of applying Graph Theory to Forensics initiated. The paper highlights the application of Graph Theory in Forensics. Identifying suspects, victims, evidence items etc as vertices and the connection between the vertices represent the edges.

The graph obtained from forensic evidences are useful when the graph is weighted and directed. The direction of arcs should be from initiator to the target. For example, if a person downloads files from a website then edge is directed from person to the website.

A case of a particular company was considered which was infected by a spyware in order to steal data. A former employee is suspected and suppose at least one current employee has visited the site. So, the first step is identifying vertices and their connections. So, the suspect, infected website and victim form vertices. If there is a connection between former and the victim then consider drawing an edge giving a simple graph. Below graph indicates this situation where suspect is vertex A, infected is B and victim is C.

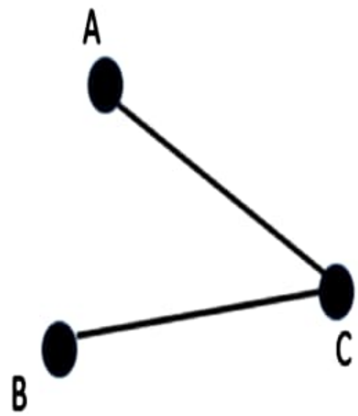

Fig. 2.1 Graph

The same can be represented as an adjacency matrix as shown below.

	A	B	C
A			1
B			1
C	1	1	

Table 2.1 Adjacency matrix

The adjacency matrix is preferred because it can be stored in the data consuming lesser space as compared to the graph. The data can also be stored as an incidence matrix which is also preferred for the same reason. For example, consider the graph shown in Figure 2.2

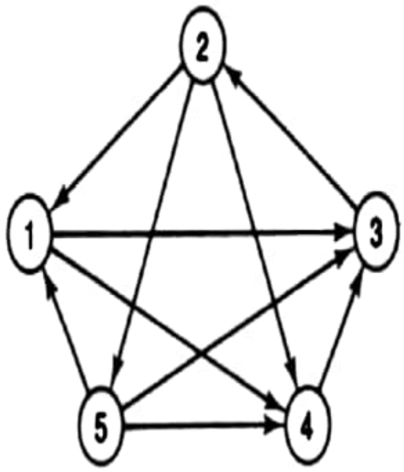

Fig. 2.2 Directed Graph

The figure 2.2 can be represented as an incidence matrix as shown below.

0	0	1	1	0
1	0	0	1	1
0	1	0	0	0
0	0	1	0	0
1	0	1	1	0

Table 2.2 Incidence matrix

Conclusion

The paper suggests graphical solutions that help in providing ways to control a crime in the real world scenario. The removal of isolated vertices and the loops where the isolated vertices represented dis-connectivity from the remaining vertices which can be obtained through events like imprisonment in order to stop various interactions, can be used to control crime.

It therefore can be concluded that Graph Theory can be applied effectively in solving as well as controlling crime and forensics, especially the digital forensic science.

2.2 Paper: Having Fun with Graph Theory and Fingerprint

Having Fun with Graph Theory and Fingerprint [3]

The paper focuses on identification of different patterns of fingerprints as well as classifying them into various categories. The paper also provides 3 algorithm to check for an appropriate match.

The fingerprint is first analysed to identify core, delta and the various ridge characteristics as shown in the figure below.

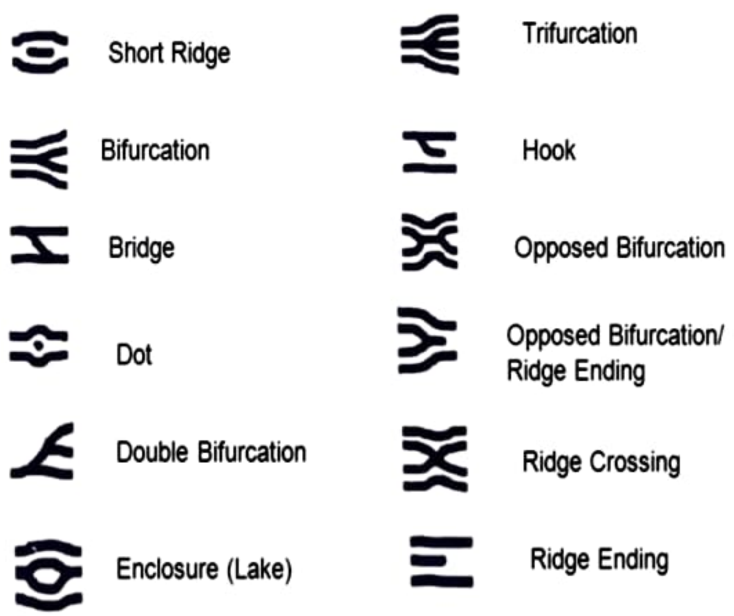

Fig. 2.3 Types of Ridges

Once these are identified, each print is classified based on the Pattern along with the Sub-pattern type mentioned below. The 3 Pattern type are Arch, Loop and Whorl.

Arch: The pattern having no delta is an Arch

Loop: The patter having one core and one delta is described as a Loop.

Whorl: The pattern having more than 1 deltas are described as a Whorl.

Fig. 2.4 Fingerprint points

Each pattern is further sub-categorised into sub-patterns:

1. The Arch is further divided into Plain Arch or Tented Arch on the basis of presence of a core. Therefore, the pattern where neither core nor delta is present is defined as a Plain Arch while the pattern with 1 core and no delta is defined as a Tented Arch as shown in the figure.

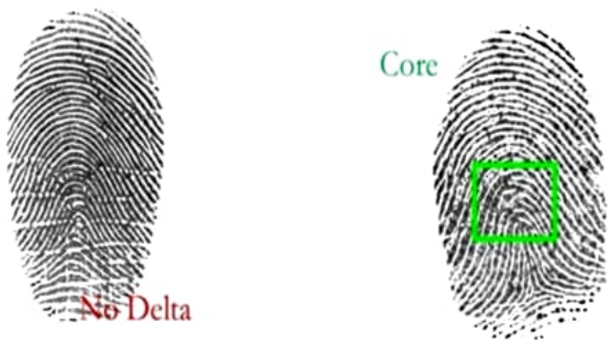

Fig. 2.5 Arch

2. Loop is further classified as Radial Loop and Ulnar Loop on the basis of the direction of the loop. So, if the pattern has 1 core, 1 delta and is directed towards the thumb,

the sub-pattern is called as a Radial Loop while if the pattern has 1 core, 1 delta and is directed towards the ring finger, the sub-pattern is defined as a Ulnar Loop.

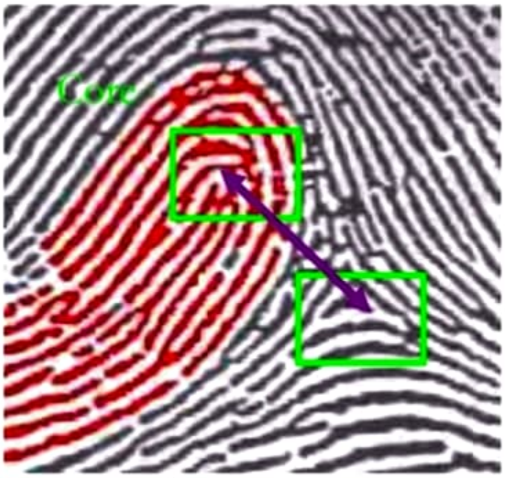

Fig. 2.6 Radial Loop

3. Whorl is further classified into 4 sub-patterns which are Plain whorl, Central Pocket whorl, Double loop and Accidental whorl. Each of these are displayed in the image below.

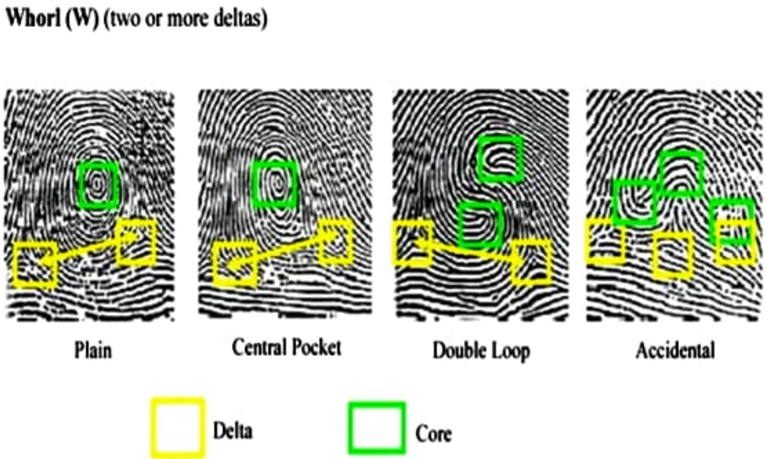

Fig. 2.7 Types of Whorls

2.2.1 Algorithm 1

Algorithm 1: Associating weighted and coloured graph with the fingerprint.

1. Identify all possible ridge characteristics
2. Associate a vertex to each of the ridge characteristic with different colours.
3. Draw an edge between any two vertices that lie on a connected ridge flow.
4. Associate weight to each edge calculated by measuring in cm.
5. Associate (m, n) to each vertex where m = abbreviation and n = degree of vertex in the graph.

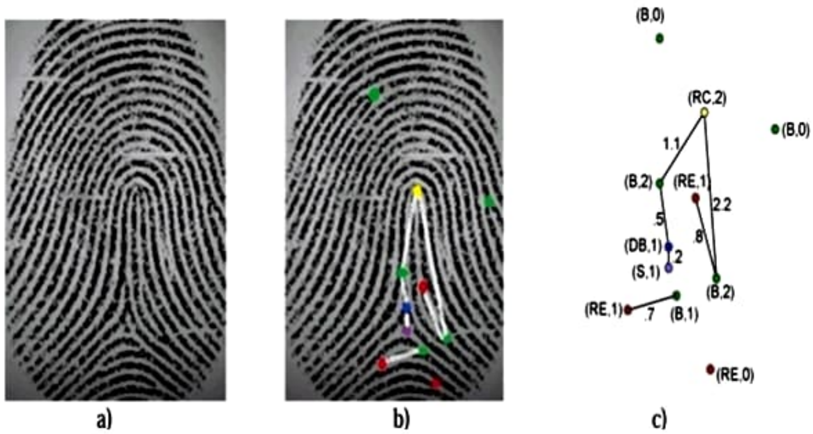

Fig. 2.8 Example

2.2.2 Algorithm 2

Algorithm 2 : Classification of fingerprint.

1. Determine if fingerprint has a core. If yes, move to 3. Otherwise to 2.
2. Classify it as Plain Arch and move to 8.
3. Determine if fingerprint has delta. If yes, move to 5. Otherwise to 4.
4. Classify it as Tented Arch and move to 8.
5. Find the number of deltas. If only 1 then move to 6. Otherwise go to step 7.
6. Classify it as Radial or Ulnar loop and move to 8.
7. Classify fingerprint as Whorls and move to 8.
8. Count the number of different ridge characteristics.
9. Level 1 classification: Classify fingerprint as $(PT, Sub-PT, No.of RC)$
10. Count the number of connected components(CC), total no. of vertices(V) and total no. of edges(E).
11. Level 2: Classify as (CC, V, E)
12. Find RC associated to each vertex [m_i=1,2,...]
13. Find degree of each vertex (n_i) and weight of edges incident to each vertex ($R_i = 1, 2, ..., n_i$) Level 3: Classify each vertex in the graph as $(m, n, R_1, R_2, ...R_n)$

Pattern Type (PT): A

Sub-Pattern Type (SPT): TA

# of different types of Ridge Characteristics (RC): 5

# of Connected Components (CC): 5

Number of vertices (V): 11

Number of Edges (E): 6

Characterize each vertex by label:

$$(m, n, R_1, R_2, \ldots, R_n)$$

where R_i=weight of the edges incident with the vertex, listed in increasing order

Associated classification:

1. (A, TA, 4)
2. (5, 10, 5)
3. (DB, 1, 0.5) (B, 2, 0.8, 2.2) (B, 1, 0.7) (B,0)
   (B, 2, 0.5, 1.1) (RE, 1, 0.8) (RE, 0) (S,1,0.2)
   (RC, 2, 1.1, 2.2) (RE, 1, 0.7) (B,0)

Fig. 2.9 Example

2.2.3 Algorithm 3

Algorithm 3: Matching a Recovered Fingerprint with Fingerprint from the Database.

This algorithm is used to match the fingerprint obtained at crime scene with the one already present in the database. When two graphs match completely, they are said to be identical and hence isomorphic.

1. Choose the set of fingerprints from database corresponding to first level of classification of the recovered fingerprint from crime scene.

2. Among those chosen set of fingerprints from 1, choose the ones that match the second level of classification of the recovered fingerprint.

3. From the set of fingerprints chosen in 2, choose the fingerprint that matches the third level of classification with recovered fingerprint.

4. The fingerprint that matches all the 3 levels is the required fingerprint output.

2.3 Paper: Minutiae Based Fingerprint Verification using Graph Model.

Minutiae Based Fingerprint Verification using Graph Model [5].

Proposed work

1. The first step is the pre-processing step in which image enhancement, noise reduction etc are performed inorder to get a clear image to analyse.

2. Step 2 deals with minutiae extraction which includes identification of vertices based on identifying the ridge characteristics and minutia followed by marking of those vertices.

3. False Minutiae Identification and elimination is performed in this step. False minutiae extraction are the points or ridge characteristics which are spotted in the fingerprint image but are not present in the original print. These may be present due to some issues such as using less ink while giving a print which would then show breaks in the ridge or extra ink which joins two or more ridges leading some other patterns.

4. Once the vertices are identified correctly, the edges are to be drawn i.e. construction of graph from the extracted minutiae points.

5. After extracting a graph, comparing the resulting graph with the main graphs in the database is the last step. This is called Matching process. Fingerprint matching is done using sub-graph isomorphism rather than graph isomorphism, since this would give a more accurate outcome.

Construction of graph:

Construction of graph is basically drawing edges with weights for the vertices which have been identified. The edges of the graph can be weighted in one of the following ways:

1. Define $e(i,j) = d$, where d is the shortest distance between node i and node j determined for (x_1, y_1) as position of node i and (x_2, y_2) as position of node j using euclidean shortest distance formula given by
   $d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$

2. $e(i,j)$=Number of ridge lines intersecting the straight line joining node i and node j. This method of allocating weight to edge is useful if input fingerprint to be matched is scaled or magnified. Although, the distance between the points may vary.

Example:

Suppose the set of pixel points identified are $(0.2,1)$,$(0.4,1.6)$,$(0.5,0.5)$, $(1,1)$, $(1,1.6)$, $(1.5,2)$, $(1.9,0.6)$, $(2,1)$ and $(2.5,1.6)$ then these can be written in a matrix as shown below

									1															
															2									
				2																				
									2						1									
																			1					
				1																				
														1										
															2									

Table 2.3 Pixel image matrix

Here entries are 1 or 2 denoting Ridge Ending or Bifurcation resp.

Using Euclidean distance formula, we get:

Vertices	1	2	3	4	5	6	7	8	9
1	0	1	1	1	1	2	2	2	2
2		0	1	1	1	1	2	2	2
3			0	1	1	1	1	2	2
4				0	1	1	1	1	2
5					0	1	1	1	2
6						0	1	1	2
7							0	1	1
8								0	1
9									0

Table 2.4 Edge weights

Using these edge weights, we can form a complete graph K_9

2.3.1 Properties of Graph constructed

The graph obtained using the above methods have the following properties:

1. The weight matrix is symmetric i.e $e_{ij} = e_{ji}$. This is because each edge in such a graph is independent of direction and is symmetric.

2. The graph obtained is a complete graph, since every pair of vertices are connected to each other by the stated methods of drawing an edge.

3. The graph is weighted since every edge is assigned a weight.

4. The graph is un-directed.

Now, we have two graphs i.e. the extracted graph and the database graph. The next step in the method is fingerprint matching.

2.3.2 Fingerprint Matching

Fingerprint Matching is performed using Sub graph Isomorphism:
The main fingerprint, which is usually in the database, is called Candidate Fingerprint and fingerprint which is extracted from the crime scene that needs to be analysed is called as Fragment Fingerprint. The graph of candidate fingerprint is denoted by $GC(VC,EC)$ while the graph obtained from fragment fingerprint is denoted by $GF(VF,EF)$ where it is assumed that $|VC| \geq |EC|$ to ensure that no parallel edges are included in the graphs.
For the Matching process, the two proposed algorithm used are Generalised Combinatorial Sub Graph Isomorphism and Sequential Combinatorial Sub Graph Isomorphism. Candidate graph should be of greater cardinality

The result of match between two fingerprints depends upon:

1. If $GF(VF,EF) \subseteq GC(VC,EC) \rightarrow \exists\, T | T(X) \subseteq S$
2. GF is isomorphic to a subgraph of GC.

Algorithm 1

Use of Geometric Shortest Distance Between Two Nodes To Weight The Edges Of $G(V,E)$.

1. Input minutiae extracted matrix of order rxc.
2. Number all minutiae points from 1 to n where minutiae locations are marked by 1(for ridge ending) or 2(for bifurcation)
3. Find row and column (x_1,y_1) of node i and row and column (x_2,y_2) of node j.
4. Find $d \leftarrow \sqrt{(x_2-x_1)^2+(y_2-y_1)^2}$
5. Put d in i th row and j th column of edge weight matrix where $e(i,j)=d$
   Hence the weight matrix of order nxn is obtained.

Sequential Combinatorial Sub graph Isomorphism

Here, instead of taking all possible sub graphs, only the sequential combinations are considered. Since the edge weights are consistent in a matrix, only consecutive combinations are considered. So, only $n-r+1$ number of sub graphs are checked for isomorphism.

Generalised Combinatorial Sub graph Isomorphism

Let EF be of dimension r and EC be of dimension n. In the algorithm, all the $r \times r$ sub graph combinations are extracted from the candidate print. The total number of such graphs extracted is ${}^nC_r$ and all of these are checked for isomorphism with the fragment fingerprint. If the isomorphism is obtained then it is a match. This algorithm is found to be more accurate and generalised.

Algorithm 2

:Sequential Combinatorial Sub-graph Isomorphism:

1. Take a Window counter and initialize it to 0.
2. Take number of consecutive sequences ${}^nC_r = n-r+1$
3. Construct a $2-d$ array brr containing only the consecutive sequences among all possible combinations of ${}^nC_r$. Now, brr $(n-r+1, r)$ generate all consecutive sequences.
4. Choose $drr(i, j)$ of size $r \times r$ which is a sub graph of EC such that $i, j \in brr$. Repeat steps 5,6,7 for all $n-r+1$ sub graphs of GC.
5. Initialize ctrblock to 0.
6. Choose edge $a_{ij} : i, j \in drr$ and choose edge $b_{ij} : i, j \in EF$. Choose $T|T(a) \subseteq$ b $\forall a \in EF$. Now if $T(a)$ is isomorphic to b and the sub graph of GC is isomorphic to GF then increment counter block.

7. Check if counter block $\geq r * r$-offset, where offset is determined by statistical methods. If yes, then increment Window-counter else result is fail.

8. If fail > threshold, conclude a mismatch. Here threshold is determined statistically.

9. If Window-counter $\geq \sqrt{^nC_r}$ then Graphs are isomorphic, else not.

Algorithm 3

Generalised Combinatorial Subgraph Isomorphism:

To determine whether graph GF and graph GC are isomorphic:

1. Take a Window counter and initialize it to 0.

2. ^nC_r: all possible combinations of sub-graphs $SG(r) \subseteq GC(n)$.

3. Construct a $2-d$ array brr containing all possible combinations of ^nC_r brr(^nC_r,r) $\leftarrow$ generate all possible combination of indexes of ^nC_r in recursive approach.

4. Choose $drr(i,j)$ of size $r \times r$ which is a sub graph of EC such that $i, j \in brr$

5. Initialise ctr-block to 0

6. Choose edge a_{ij} : i,j $\in drr$ and choose edge b_{ij}:i,j $\in EF$. Choose $T|T(a) \subseteq$ b $\forall a \in EF$. Now if $T(a)$ is isomorphic to b and the sub graph of GC is isomorphic to GF then increment counter block.

7. Check if counter block $\geq r * r$-offset, where offset is determined by statistical methods. If yes, then increment Window-counter else result is fail.

8. If fail > threshold, conclude a mismatch. Here threshold is determined statistically.

9. If Window-counter $\geq \sqrt{^nC_r}$ then Graphs are isomorphic, else not.

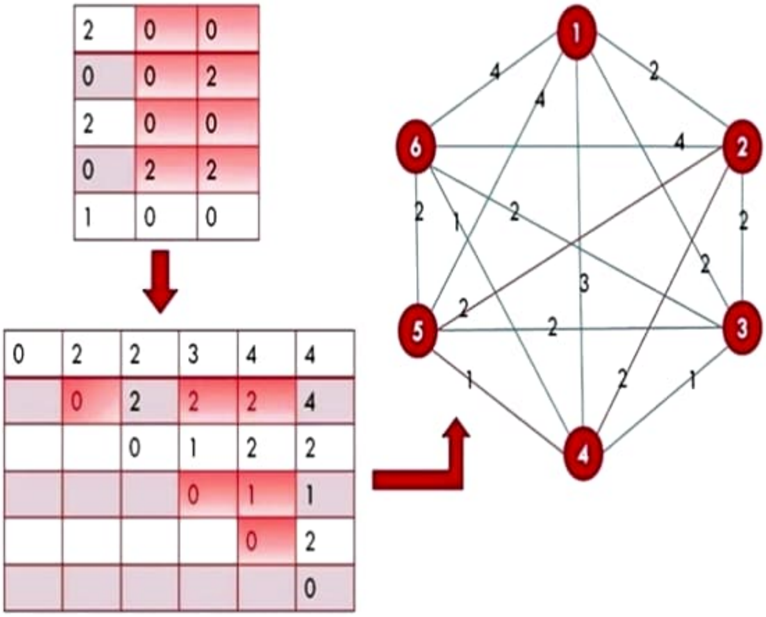

Fig. 2.10 Example 1

Threshold Determination:

According to the law and study of some countries, the minimum number of similar points of identification between two fingerprints are 12. This means that if the 2 prints are the same then at-least 12 points should be identified as a match. Hence the threshold value is taken as 13. For larger number of minutiae points identified irrespective of noises being included or not, the offset defined as the minimum number of edges that must match between 2 prints is taken as $\frac{2}{3}r$ where r is the order of Fragment fingerprint. The offset is taken as square root of $^{n}C_{r}$ since fingerprint at site may not match for all blocks.

Comparison of the methods stated in the papers reviewed:

1. The Intersecting ridge line method is less effective than Euclidean distance formula method.

2. Disconnected graph may be obtained using the method mentioned in paper [3] but graph obtained using the method mentioned in paper [1] always produces a connected graph.

3. Paper [3] considers more than 2 types of ridge characteristics but paper [1] considers only 2 types of ridge characteristics.

2.4 Paper: Fingerprint Identification using Graph Matching

Fingerprint Identification using Graph Matching [2]

The paper produces a method for fingerprint matching using using partition, refining and scoring methods.

A digitised fingerprint is considered for analysis. These prints contain a light and dark bands where ridges are taken as dark and valley (space between the ridges) are taken as light pixels. From the print, minutiae (Bifurcation, Ridge Ending) are identified after which each ridge is given a unique numbering. Next, each ridge is given an order such as side 1, side 2 where initial end is selected arbitrarily and then end neighbours and side neighbours are identified and marked. Lastly, an overlap length is measured.

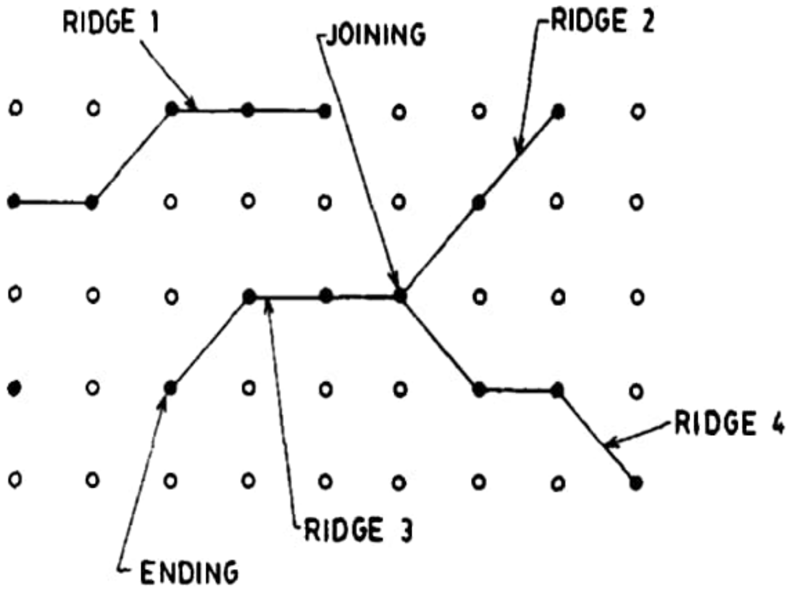

Fig. 2.11 Ridges

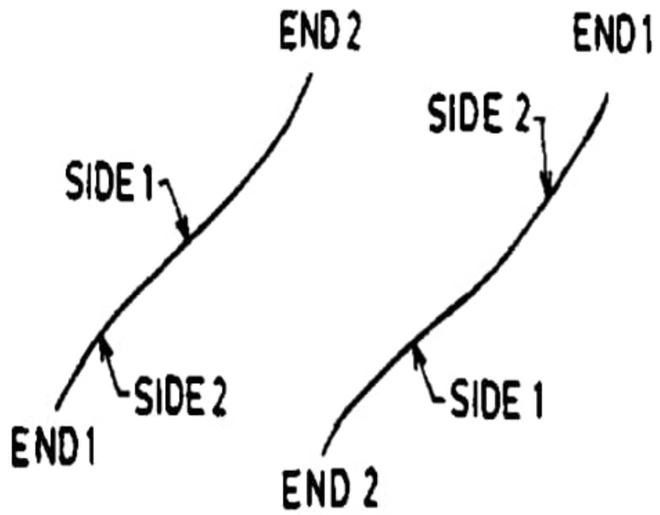

Fig. 2.12 Side neighbours

Every ridge should be properly oriented for which the ridge R should have following conditions checked and satisfied:

1. All side 2 neighbours of ridge *R* that have *R* as a neighbour only on one side are flipped, so that that side is side 1. Hence properly oriented.

2. Similarly for side 1 neighbours of ridge.

3. The 1st clockwise neighbour on end 2 of *R* and the 1st counter-clockwise neighbour on end 1 of *R*, are forced to meet with *R*'s ends such that end 1 of *R* meets end 2 of the neighbour, and end 2 of *R* meets end 1 of the neighbour. (none are side)

Figure 2.13 depicts the above case.

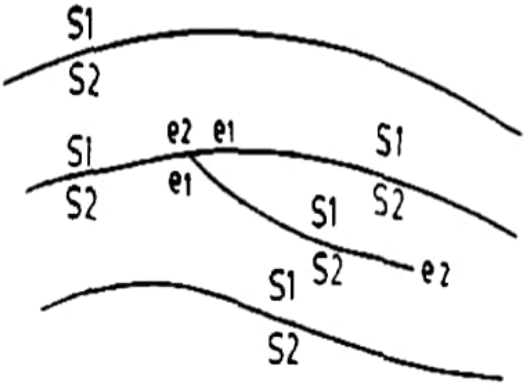

Fig. 2.13 Ridges satisfying the 3 conditions

Once all ridges have been properly oriented, level numbers can be assigned. Main criteria is that if ridge is given L then above ridges are assigned $L-1$ or less and below ridges are assigned $L+1$ or more. Figure 2.14 is an example of the level numbering.

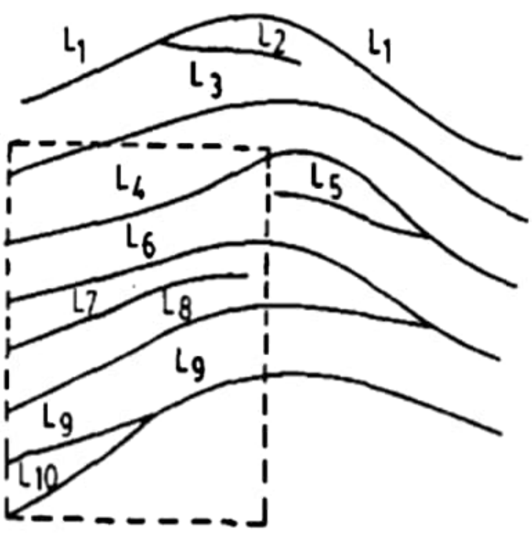

Fig. 2.14 Level Numbering

Also, even if some ridges are missing, the level numbering is not affected. An example of this is shown in figure 2.15 where L_3 of figure 2.14 is L_1 in figure 2.15.

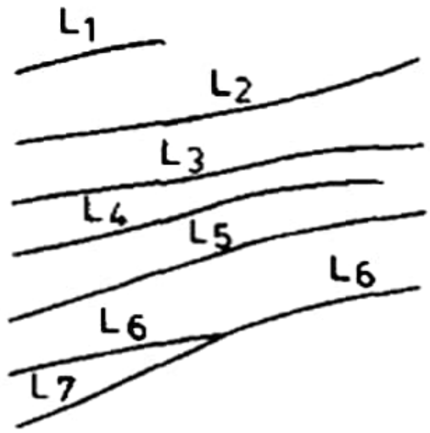

Fig. 2.15 Missing Ridge: Level Numbering

Once the level numbering as well as orientation is correctly assigned, the graph is extracted as shown in the figure below.

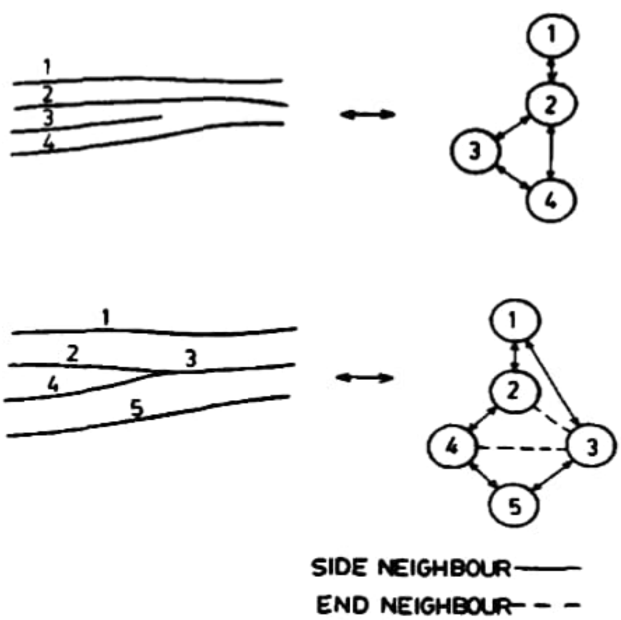

Fig. 2.16 Graph Formation

The case where the print may have some noises identified is depicted below. If the noise is a ridge break or ridge joining then it can be eliminated as shown in the figure 2.17. Figure 2.18 is another example of extraction of graph.

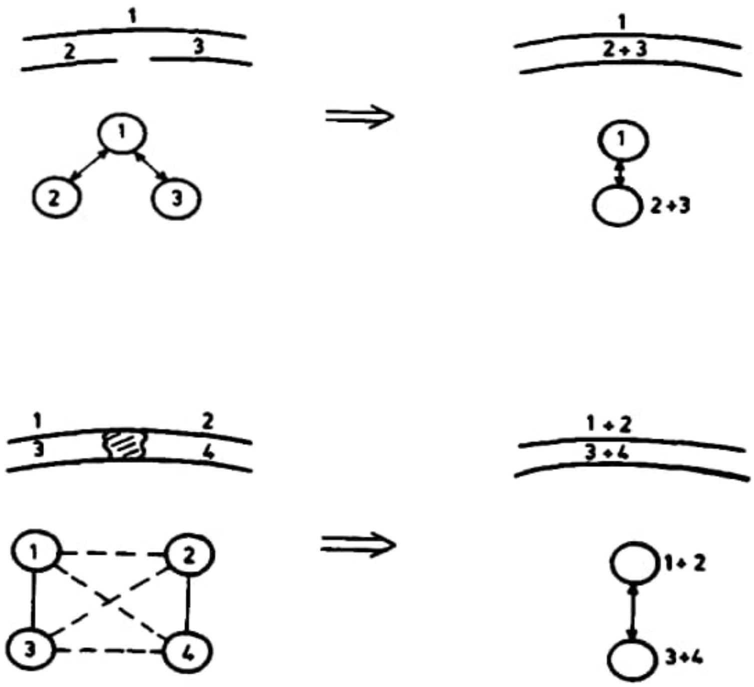

Fig. 2.17 Noise Removal

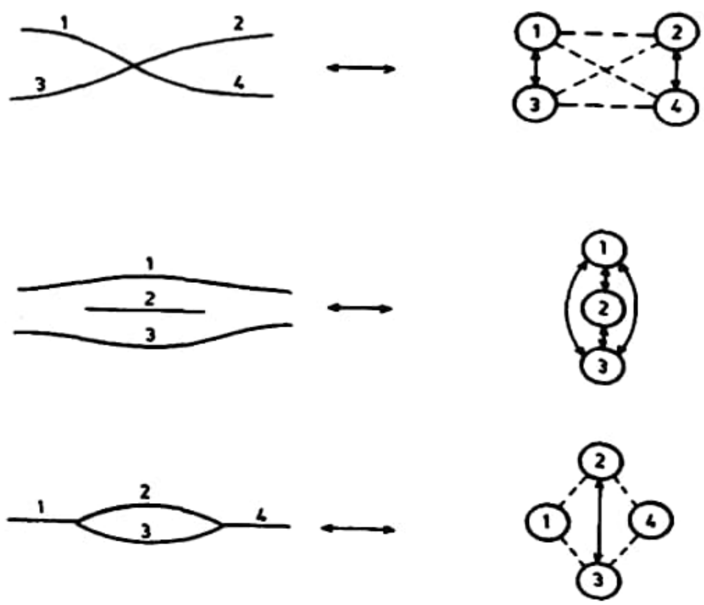

Fig. 2.18 Graph Formation

2.4.1 Matching

Matching includes 3 steps : Partitioning, Refining and Scoring.

1. Partitioning:

   Two graphs (to be matched) G_A having vertices $a_1, a_2, \cdots, a_M$ and G_B having vertices $b_1, \cdots, b_N$ are considered where a partitioning array, P, having M rows and N columns is prepared such that P_{ij} is an element if vertex b_j is possible match for vertex a_i. Each row of P forms a partition for G_B. Also, P_{ij} becomes false if corresponding ridges differ in length by more than 30% and number of neighbours on corresponding sides of the ridges differ by more than one.

2. Refining:

   At the start of this step, every vertex in G_A will have one of 0 to N vertices in G_B as potential match. Since there may be more than one vertex in each partition where chances of repetition are there, the number of vertices are reduced to 1 vertex in each partition in this step. A mapping between a_i and b_j is retained only if side neighbours of a_i map onto corresponding side neighbours of b_j.

3. Scoring:

   The notation (a_i, b_j) indicates the mapping. Using a_i and b_j as roots, a tree is traced in each graph. While tracing each tree, node l may be included as a son of node k only if node k has node l as a neighbor, and if the level number of node l is higher than that of node k. The score for the two trees grown parallel are calculated once they have reached the maximum size where the score is a value depending on the number and minutia type. The tree with the maximum score is considered as measure of match between the two prints.

Experimental conclusion:

- Small scores for good matching are due to:
  - Insufficient image size.
  - Small overlap
  - Poor image enhancement.
- Despite of these, this method identifies the match correctly.
- The ability of proposed encoding scheme to identify and repair different types of noise can be used to improve score profile, is currently being investigated.
- The whole process is conducted on $VAX-780$ computer where the encoding time for each window was 16.2CPU seconds with minimum of 12.7s and maximum of 19.5s.

2.5 Paper: Implementation of Graph Colouring Technique on crime

Implementation of Graph Colouring Technique on Crime [4]
The paper describes methodology for identification of suspects through a graphical approach based on analysis of vertex set, edge set and incidence function relating the vertices with the edges. In this paper, an arbitrary jewellery shop robbery is considered.
In the first scenario, 3 suspects were identified and named U, V and W. These are taken as vertices. Each vertex is given a different colour say blue, red and green respectively. Next the statements of these suspects were recorded and are as follows:

1. U: I am not the robber
2. V: U is the robber
3. W: I am not the robber

Now, the statement of U leads to an edge directed from U to V and U to W, the statement of V results in an edge from V to U and the statement of W leads to edge directed from W to U and W to V wherein each edge gets the colour of the vertex from which it has initiated.

Algorithm for Construction of Graph G is as follows:

1. Consider the suspects as vertices U, V and W.
2. Assign distinct colour to the vertices.
3. Represent the relation between these vertices as edges and give respective colours to the edges.
4. Draw directed edges from the initiator to the other vertices based on the statements of the suspects.

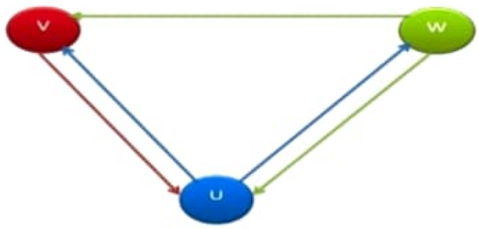

Fig. 2.19 Crime graph

Therefore, Figure 2.19 is obtained in this scenario.

Another point to be noted is that the police has a tip which says that only 1 person is truthful. In order to identify the criminal following cases are considered.

Case 1: Suppose U is the robber.

If U is the robber then V or W will not be the robber which implies that V is truthful as his statement states that U is the criminal. Since only one suspect is truthful, this implies that W is also a liar which would then mean that W is the robber. Hence a contradiction to U being robber. Graphically, we would get the following.

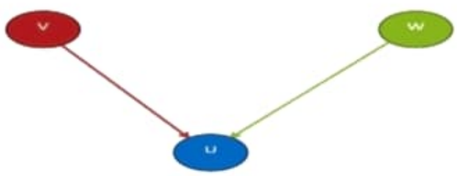

Fig. 2.20 Graph depiction

This graph is obtained because assuming U as the robber will lead to removal of edges coming out from U and since V is truthful, the edge coming into V is also removed. We have two edges into U from 2 different vertices giving contradiction to only one being truthful. Hence U is not he robber.

Case 2: Suppose V is the robber.

If V is the robber then neither U nor W is. Clearly, U is truthful which further implies that W is lying which would further imply that W is the robber giving a contradiction. Graphically, we obtain the below. Hence V is not the robber.

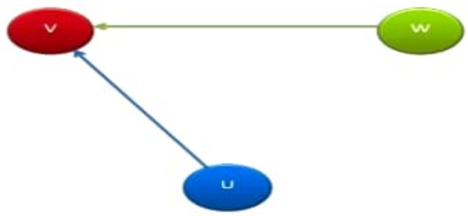

Fig. 2.21 Graph Depiction

Case 3: Suppose W is the robber.

W is the robber implies that U is truthful as he states he is not while V is a liar since U is truthful, which is logically the required result. Graphically, we get the below figure.

Hence W is the robber.

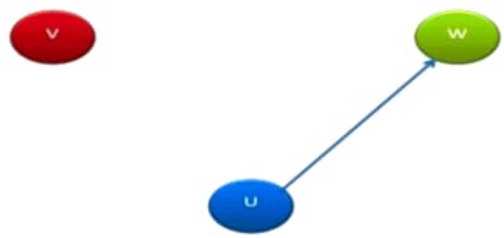

Fig. 2.22 Graph Depiction

Thus, we obtain a solution for the given crime problem using the Graph Colouring Technique.

The observation here is that the in-degree of U, V and W is given by (2,2,1) where in-degree implies the number of people being truthful for that particular vertex to be the criminal.

Below is another algorithm stated in the paper

Algorithm to Construct a Graphical Sequence

1. Consider vertex in graph G
2. Count the number of lines entering in that vertex.
3. Repeat the iteration for each vertex in the graph G.
4. Write the sequence $(in-deg(a), in-deg(b), ..., in-deg(n))$ where $a, b, ..., n$ are vertices.

In the second scenario, 4 suspects were identified by the police.

Construction: Let P, Q, R, S be suspects. Assign blue, red, green and yellow to vertices.

The statements of the suspects were as follows:

- P: I am not the robber.
- R: I am not the robber
- Q: P is the robber
- S: Q is the robber

From these statements, a graph is drawn with P, Q, R and S as vertices where the statement forms the relation between two vertices i.e. the edges. These are directed edges. Explanation: P states that he is not the robber which implies that the rest could be the criminal and so, the edges will direct from P to Q, R and S. Similarly for R. Next, Q states that P is the robber which implies that the edge should direct from Q to P while according to S, the edge should direct from S to Q. Hence a graph is obtained consisting of these edges mentioned.

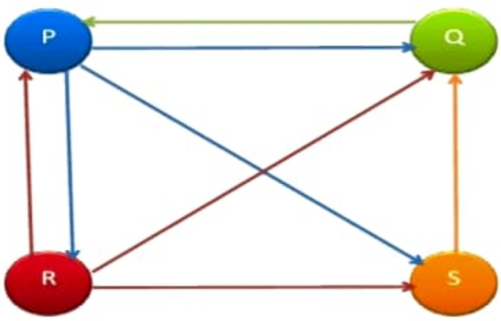

Fig. 2.23 Graph obtained

From the graph obtained, we calculate the in-degree of each vertex as $(P, Q, R, S) = (2, 3, 1, 2)$. The in-degree here denotes the number of individuals giving a true statement for that vertex to be the criminal. This implies that The vertex, say P, is the criminal if 2 suspects are truthful.

Therefore, it is observed that:

- R is the robber if only 1 suspect is truthful
- P or S is the robber if 2 suspects are truthful
- Q is the robber if 3 suspects are truthful.

Hence, we can also conclude that vertex with maximum in-deg is the robber if most are truthful and vertex of smallest in-deg is the robber if least individuals are truthful. This method can be extended to large number of suspects, which will then help to eliminate the suspects.

Chapter 3

Application of Graph Colouring Technique on real life Robbery Cases

3.1 Case-study on Graph Colouring Technique

3.1.1 Robbery Case 1

A robbery case was registered in the police station by an owner X of a clothing store consisting of 5 employees (1 cleaner, 2 sales people, 1 cashier and 1 helper).

Scenario: X had brought a bag consisting of cash which was then placed in a drawer in the store. The cleaner was on a break on the day of robbery due to which only 4 employees were present. At 12 pm, X realised that the bag had gone missing. According to X, the robbery happened between 10.30 am to 12 pm and during this period only 4 people visited the store (3 : frequent customers and 1 for direction). X suspected one of the 4 employees to be the robber.

Suspects: The police verified the background and records of the customers. The employees where then asked basic questions from which the suspects were identified as the

4 employees. The owner also mentioned few details which were cross checked. Out of those, two details are used while investigating the case. These are:

- There was a mistake in the billing.
- A customer was searching for some clothing for quite some time.

The 4 suspects, P, Q, R and S, were interrogated. Their statements were as follows:

- P: I was displaying dresses to customers around that time.
- Q: I was with a customer displaying salwar-suit.
- R: I was at the billing counter
- S: I was helping R due to some fault in the billing.

Solution using Graph Theory:

Consider P, Q, R and S as the vertices. The above statements are considered to draw the edges as seen in the paper [4]. From the graph below, we get the in-degree of each vertex as $(P, Q, R, S) = (3, 3, 2, 3)$

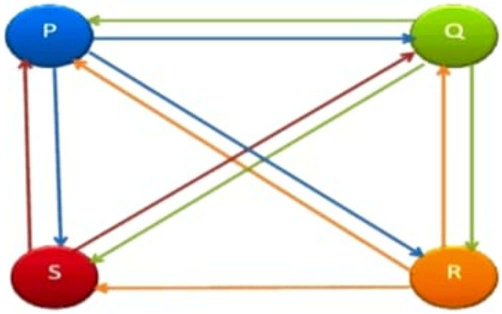

Fig. 3.1 Graph Obtained

Conclusion

- Combining the information and the graph, we can conclude that the robber is either P or Q.
- On further investigation, Police found that the robber was Q.
- Hence, we can conclude that the result obtained for such scenario helps to eliminate the suspects.

3.1.2 Robbery Case 2

A robbery of cash and jewellery from a house at Dona Paula was reported after few days from the robbery. The owner stated that while looking for some cash, it was noticed that the cupboard was open and the key was in the lock as well. On hearing the detailed incident, it was known that there were 4 members who visited the house which include driver, maid and the couple. The owner also mentioned that the ex-driver had done robbery of some cash 4 years back due to which the prime suspect according to the couple was the ex-driver. While reporting the incident, the following key points were obtained:

- House had CCTV which showed blur video or images which revealed body structure and a peculiar wrist watch.
- There were no marks of forceful entry.
- The footage showed black and white image from which the gender was identified as male.
- The area only has hotels and bungalows.

The suspects were then identified as Driver (D), Ex-driver (E), Maid (M) and some outsider (S)

The statements of the suspects were as follows:

- M: I worked here for past 10 years. I have no idea.
- D : I never enter the house.
  M also will not betray (On asking about her)
- E: No, I wasn't in town.

From the key points it can be concluded that *M* is truthful since robber is a male and *D* is truthful since his body structure is completely different from the robber.

The below graph depicts the relation between these vertices which helps to find the in-degree of each vertex.

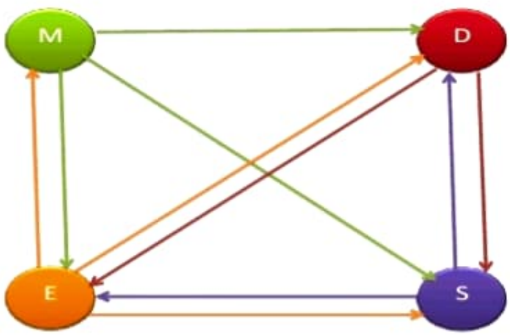

Fig. 3.2 Graph

From the graph, the in-degree of vertices are found to be $(M, E, D, S) = (1,3,3,3)$. It is then possible to conclude that either *E* or *S* is likely to be the robber since we have already seen that *M* and *D* are truthful. Hence, there has to be one more suspect who is truthful. The edges directing from *S* to *E* and *S* to *D* is obtained from the point which stated that the robber is a male and also that there was no forceful entry marks found.

Hence from this method we can say that the suspect elimination is found to be effective. Although further investigation to find the criminal need to be held. From the wrist watch mentioned in key point and on verifying the statement of E, S can be suspected to be the robber. Now, S being a stranger requires further enquiry.

Conclusion

From the two robbery cases analysed, it can be concluded that the methodology mentioned in the Graph Colouring Technique paper can be used to eliminate the suspects and in some cases, it can be used to identify the criminal.

The procedure for analysis was as follows:
At first, fingerprints were collected on paper slips by applying ink on the finger and placing the fingerprint with minimum pressure such that every ridge was visible containing no noise. Each fingerprint was then analysed to identify the various ridge characteristics and marked as vertices giving each type of ridge characteristics a specific colour. Edges are drawn between two vertices if they are connected on the same plane. So, after edges were drawn, the edge weights were assigned to these edges by finding distance in cm between every two vertices which were connected by an edge. Hence, we obtain a graph for each fingerprint for each family. Later, each print was classified into 3 levels mentioned in paper [3]. With this, we have obtained extracted graphs with levels for each fingerprint which then needs to be compared with the other fingerprints.

Figure 4.1 is an example of the graph extracted from the fingerprints of a family consisting of 3 members, using the procedure mentioned above. The figure is divided into 4 parts depicting the 4 prints, since these are the required fingerprints to understand the analysis. Beginning with the left is the left fingerprint of parent 1 followed by right fingerprint of parent 2, to its right. The below half consists of left and right fingerprint of the child respectively.

The vertices coloured red represent the bifurcation while the vertices coloured green represent the ridge ending identified in the fingerprint. The node coloured blue is the core which helps in identifying the pattern type.

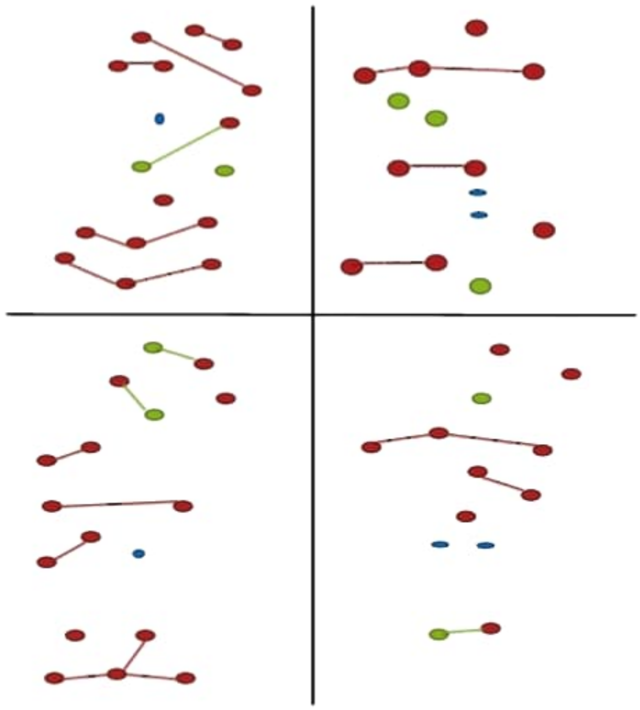

Fig. 4.1 Graph Extracted

On assigning the 3 level of classification as shown in paper [3], the following results were obtained.

Relation		No. of RC	PT	SPT	CC	V	E
Parent 1	Left	2	L	RL	8	16	8
	Right	3	L	RL	13	16	3
Parent 2	Left	3	W	DL	10	12	2
	Right	2	W	DL	8	12	4
Child	Left	2	L	RL	8	16	8
	Right	2	W	DL	8	12	4

In this particular example, the child has inherited patterns and sub patterns from both parents. It was therefore observed that the left fingerprint of parent 1 and child matches the first and second level of classification while the right fingerprint of parent 2 and child matches the first and second level of classification.

In general, the 3rd level shows variations in degree, edge weights, etc but 4 of the nodes match which implies that the fingerprint is not exactly same but has common nodes of inheritance.

Analysis of Fingerprints Collected

On analysing all the 80 fingerprints using the algorithms mentioned in the paper [3], the following results were obtained:

1. Level 1 and 2 of classification matches between the father and the child more frequently. It is also observed that one of the hand of the child matches with the father while other one with the mother.
2. Special characteristics always is inherited from the father/parent to the child.
3. Hence, pattern and sub-pattern type is inherited from the parent to the offspring.

The results obtained were verified using the procedure used in forensic lab wherein once the graph is extracted, orientation of the 4 points/vertices common are checked.

Similarly, in the next analysis technique, the minutiae points of fingerprints were identified as vertices after which the edges were drawn between every 2 vertices irrespective of whether they are connected on the plane or not. Edges were assigned edge weights using the Euclidean distance formula. We, therefore, get a complete graph.

The observations were as follows:

1. For a fingerprint to completely match at least 5 points should match and if they do not match then maximum 1 point would match. Using this, it was observed that if there 3 points that match then the fingerprint should be related in some way.
2. It was noticed that the edge weight set obtained for the parent graph and child graph was the same but since the graph obtained is always complete. a generalised similarity could not be concluded. The reason being that the paper mentions technique for verification of the graph being the exact same rather than comparing between two different prints.

3. It was observed that the male fingerprints differ from the female ones based on the distance between the ridges as well as the number of ridges found. The males have broader valley then females while the females have more ridges then the males.

Hence, the papers reviewed can be used effectively in identifying the fragment fingerprints as the candidate fingerprint and the analysis conducted gives a positive result for the inheritance of fingerprint patterns on this population size.

Chapter 5

Conclusion

5.1 Dissertation Outcomes

Objective of Dissertation

The objectives of the Dissertation are as follows:

- Review 4-5 papers on application of Graph Theory in Forensics and Criminology.
- Collect fingerprint samples having sample size of 50-60 approximately for analysis.
- Analyse the inheritance of the fingerprints collected between 2 or 3 generations using Graph Theory.
- Analyse fingerprints to find difference between male and female.
- Analyse at least 1 robbery case to check applicability of graph colouring technique on real life scenario.

Implemented Work

5 Research papers were reviewed in which 4 were based on fingerprint algorithms and techniques while 1 was based on criminal identification.

Conclusion

On reviewing and comparing the papers mentioned in chapter 2, following points were noticed:

1. In the paper [1], identification of vertices is based on identification of only 2 ridge characteristics unlike paper [3]. where the identification of vertices is based on identification of various (more than 2) types of ridge characteristics as mentioned in figure 2.3. Hence the result obtained on matching a recovered fingerprint with the fingerprint in the data base is found to be more accurate in paper [3] than in paper [1].
2. On visiting the forensic lab, it was noticed that the manual procedure of identification of ridge characteristics relates to the procedure used in paper [3].
3. In paper [1], the intersecting line method to obtain edge weights is found to be less effective than the Euclidean distance formula method.
4. The graph obtained in the paper [1] is always connected, complete, weighted and symmetric while the graph obtained in the paper [3] may be disconnected and weighted.

Data Collection and Outcomes:

Due to the time restraint, 80 fingerprints of people, consisting of 42 male fingerprints and 38 female fingerprints, were collected from different places in Goa. The sample included 3 families in which the 3 generations contributed for the fingerprint analysis. All the 80 fingerprints were then analysed using the research papers on fingerprint where the 2 hypothesis of my fingerprint analysis were that "Fingerprints are likely to be inherited" and "The male and female prints differ".

5.2 Observations

On analysing the fingerprints, positive result is obtained for first hypothesis. The following is the outcome depicting a positive result.

1. On analysing the paper [3] and paper [2] for the fingerprints collected, it was observed that the patterns (Loop, Arch, Whorl), sub-pattern and number of ridge characteristics present are inherited where the paternal print has higher weight-age for inheritance.

2. Special characteristics (Lake, Dot etc) if present in the father is then inherited in the offspring.

3. Peculiar and generalised differences between male and female prints need further investigation.

Data on the 2 robbery cases were collected with the help and contribution of PSI Sanket Pokhre and IPS Arvind Gauns, which were then analysed based on the review paper. The conclusions were as follows:

1. The Graph Colouring Technique mentioned in paper [4] can be applied to identify suspects for the real life robbery cases (having the similar conditions) but is not found to be effective in identifying the criminal.

2. Among the 2 methods mentioned in the paper, the method to eliminate suspects using in-degree of vertices is found to be faster and more efficient.

Conclusion

It can therefore be concluded that "**Graph Theory** is an effective tool in solving and controlling **Forensic Science and Criminology** cases". This can also be concluded from the reviewing of various papers on this topic.

Further fingerprint analysis on a larger sample size ($\geq$ 1000) should be conducted to check applicability and effectiveness of the conclusions obtained as well as develop an algorithm for fingerprint inheritance analysis is the requirement.

References

[1] Chuck Easttom. Utilizing graph theory to model forensic examination. *International Journal of Innovative Research in Information Security (IJIRIS)*, 4, 03 2017.

[2] D.K. Isenor and S.G. Zaky. Fingerprint identification using graph matching. pattern recognition, 19, 111-112. *Pattern Recognition*, 19:113–122, 12 1986.

[3] M A Karim and Violeta Vasilevska. Having fun with graph theory and forensics: Csi fingerprint analysis, whose fingerprint is this one?, 11 2015.

[4] B. Logeshwary. Implementation of graph colouring technique in crime science. *International Journal of Recent Technology and Engineering (IJRTE)*, 8:4234–4237, 03 2020.

[5] Sonali Sen. Minutiae based fingerprint verification using graph model. 11 2021.