International Educational Applied Scientific Research Journal
ISSN (Online): 2456-5040
Volume: 1 | Issue: 1 | October 2016

AVL TREE AN EFFICIENT RETRIEVAL ENGINE IN CLASSIFIED FINGERPRINT DATABASE

Ahmed B. Elmadani

Department of Computer Science, Faculty of Science, Sebha University

ABSTRACT

Fingerprints are used to identify human and for crime discover. They are used to authenticate persons in order to allow them to gain access to their financial and personal resources or to identify them in big databases. This requires use of fast search engine to reduce time consumed in searching big fingerprint databases, there for choosing searching engine is an important issue to reduce searching time. This paper investigates the existing searching engine methods and presents advantages of AVL tree method over other methods. The paper will investigate searching speed and time consuming to retrieve fingerprint image. Experiment shows use of AVL tree is the best searching algorithm.

Key Words: Fingerprint Databases, fingerprint classifies, AVL tree, Access Methods Algorithms.

1. INTRODUCTION

Fingerprints have been used as the most popular biometric authentication and verification measure because of their high acceptability [15], immutability and uniqueness [1]. Fingerprint consists of a pattern of interleaved ridges and valleys [2]. There for they are used in differentiating people [10]. Fingerprint classification is a technique to assign a fingerprint into one of the several pre-specified types or classes namely, whorl, right loop, left loop, arch, and tented arch [8]. Fingerprint has two main types of features says [9]:

1. Global ridge and furrow structures which form special pattern in central region of the fingerprint, and

2. Local ridge and furrow minute details.

Fingerprint classification is based on only the first type of features and uniquely identified based on the second type of features, such features are ridge endings, bifurcation known as minutiae, orientation field and singular points in a fingerprint image [9][10].

Data whatever it is must first be sorted to facilitate searching and make it easy for any retrieval, so far different sorting methods are invented, among those methods or algorithms Bubble sort, Merge sort, Insertion sort, Selection sort, Quick sort and tree based algorithms[11].

Data when it stored in a big store, searching in it becomes a difficult task and time consume especially when it became grows [10], searching methods known also as accessing methods in computer science. Many advanced algorithms and data structures have been devised for the sole purpose of making accessing more efficient, therefore there is possibility to use them in retrieving biometrics images from its big database [12].

Access methods in literature are, sequential (Queue), binary, graph and tree based search algorithms, in practical they are vary in their suitableness, means for each type of task a particular algorithm is used, because some are slow other are fast so selection is made on the type of use [11].

Graph search it is a traversal technique visits every node exactly one in a systematic fashion. The search process in a graph can be seen as applying a set of operators to the graph's nodes until the goal node is found it initially defines the start node. Therefore, there is a general approach, applicable for any graph search, following the constraints. Two standard graph search techniques have been widely used, Depth-First Search (DFS) and Breadth-First Search (BFS) [11].

SQL search a structured query language (SQL) search, is a special-purpose designed for getting information from and updated database. It searches data in a relational database management system (RDBMS), or for stream processing in a relational data stream management system (RDSMS) [12].

Hash access method, data is stored in an extended linear hash table. The key and the data used for Hash records can be of arbitrarily complex data. Duplication of records is optionally supported in this method [16].

Queue access method it is similar to what known as sequential access method, data are stored in a queue as fixed-length records. Each record uses a logical record number as its key. This access method is designed for fast inserts at the tail of the queue, and it has a special operation that deletes and returns a record from the head of the queue. This access method is unusual in that it provides record level locking.
This can provide beneficial performance improvements in applications requiring concurrent access to the queue [12].

Tree based search algorithms consist of trees, they are used in computer to represent algebraic formula, as an efficient method for searching large databases, dynamic lists, and diverse

applications. A tree consists of a finite set of elements, called nodes. Trees can be represented generally in which each node can have an unlimited out-degree and binary tree in which no node have more than two sub-trees [4].

A binary tree is a method of placing and locating files (called records or keys) in a database, especially when all the data is known to be in random access memory (RAM) [13].

Adelson Velskii and Landis
(AVL) tree is one of this implementation[6]. It is implementation of binary tree, for each node in the tree, the height of its left sub-tree differs from the height of its right sub-tree by no more than 1 [6]. The AVL Tree ensures that al operations can be performed in time logarithmic in the size of the tree [7]. The AVL Tree must meet strict balance requirements to maintain is O (log *n*) search capabilities [14], which is resulted in search time reduction [5].

As we seen there are different searching algorithms, in this research an experiment of retrieving fingerprint image based on its class will applied using different searching algorithms to select a suitable and fast search method as continuation of an algorithm written by [10], the following paragraphs will discuss used algorithm.

To choose one of available access method for small working datasets that fit entirely in memory, there is no difference between Binary Tree and Hash. Both will perform just as well as the other says [16]. Binary Tree can be used if keys have some locality of reference where Hash can be used when a dataset is extremely large. Moreover, if your dataset becomes so large that database will almost certainly have to perform disk I/O to satisfy a random request, then Hash will definitely outperform Binary Tree because it has fewer internal records to search through than does Binary Tree [11].

Next paragraphs will propose an algorithm uses different access method to retrieve information from a data base consists of information collected from fingerprint images.

2. METHODOLOGY

The following paragraphs will describe the proposed algorithm, experiment, obtained results and their analysis. The proposed algorithm can be summarized in two points as following:

2.1 Fingerprint database construction

a. Fingerprint image collection and image preparation
b. Fingerprint image class assigning
c. Fingerprint feature extraction
d. Fingerprint key (FPkey) calculation from features
e. Store FPkey in index file and Fingerprint image in database.

2.2 Searching algorithms

a. Follow the points from section 2.1 (a to d).
b. Search from image using FPkey Apply different algorithms.
c. Write obtained results for different search algorithms
d. Analysis results.

2.3 Fingerprint database construction

A description of a fingerprint database image collection and preparation for farther use will be given in the following paragraphs.

2. 3.1 Data collection

Used Database contains about 10000 fingerprint images known as fdb10k is used. The fdb10k is a collection of fingerprint images that collected from internet FVC2000 combined with other captured locally from student and from staff.

2.3.2 Image preparation

Fingerprint image as any image must be filtered to remove noise and resized. Winner filter was used to remove noise then each image was resized to 200x200 pixels. Sorting is beyond our research but a quick sort algorithm is used because it fast compared with other method and also suitable for this type of data.

2.4 Fingerprint image class assigning

Each fingerprint image before storing it in database, it has to assign one of the five classes by following an algorithm described in [10] which is based on finding the singular points of the fingerprint image. An image of the fingerprint will be stored to its database class and this will facilitate fatherly searching process. Nine thousand five hundred and eighty one (**9581**) fingerprint images are used in classification process. Result of applying mentioned algorithm is shown in Table 1. 1996 as **A**rch, 989 as **T**ented arch, 2663 as **L**eft **L**oop, 685 as **R**ight **L**oop, and 3248 as **W**horl.

Table 1: Classification result of the collected fingerprint images

Total	A	T	LL	RL	W
9581	1996	989	2663	685	3248

2.5 Fingerprint feature extraction

A procedure was written in C++ language to extract feature from fingerprint images. Calculated features are mean denoted by Mean equation (1), standard division denoted by Std-div equation (2) and variance denoted by Var equation (3), used formulas are shown below**:**

International Educational Applied Scientific Research Journal
ISSN (Online): 2456-5040
Volume: 1 | Issue: 1 | October 2016

$$Mean = \sum_{i,j}^{N,M} I(i,j) \quad (1)$$

where $i = 0,1,3....n$, $\phi = 0,45,90,135$, $I_{i\phi}$ = sector in image ϕ, and $M_{i\phi}$ = Mean of pixel value in $I_{i\phi}$.

$$Std\text{-}div = \sqrt{(I_{i\phi}(x,y) - Mean_{i\phi})^2} \quad (2)$$

where $i = 0,1,3....n$, $\phi = 0,45,90,135$, $I_{i\phi}$ = sector in image ϕ, and $M_{i\phi}$ = Mean of pixel value in $I_{i\phi}$.

$$Var = \sum_{i,j}^{N,M} (I(i,j) - Mean) \quad (3)$$

2.6 Fingerprint key (FPkey)

The FPkey is a fingerprint key that calculated using the formula (4) to be use as a unique key for retrieving the requested fingerprint image

$$fpr = abs(\sum Mean*10^0 + Std - div*10^1 + Var*10^2)$$
$$FPkey = INT(fpr) \quad (4)$$

2.7 FPkey field as index of Fingerprint image

As mentioned above The FPkey field to be a unique key as shown in Table 2. It is used in searching an indexed file, where the class is used to indicate the database in which the fingerprint image can be find. By using the index file's FPkey it will be easy to find the image name as shown in Table 3. Fingerprint databases consist of five separate file each contains one of mentioned is class, 1 for **A**rch database, 2 for **T**ented arch database, 3 for **L**eft **L**oop database, 4 for **R**ight **L**oop database, and 5 for **W**horl database.

Table 2: FPkey index file class retrieval

No.	FPkey	Class
1	113530	1
2	410118	4
3	410207	4

Table 3: FPkey index file image retrieval

No.	FPkey	Image name
1	113530	A20
2	410118	R6000
3	410207	R6020

2.8 Fingerprint image searching in classified Fingerprint database

Designing Databases with separate classes will minimize searching time, because each will contain less number of images and only a signal class compared to whole databases.

When a fingerprint image is searched, it switched to different search engines each engine uses a search method different to the other based on that a fast searching method will be discover. An experiment was taken place by selecting several access methods Queue, Graph, SQL, Hash, Binary, and AVL tree as will be described in next sections.

2.9 Results for different Access algorithms

An experiment of retrieving fingerprint image was performed using Dual-Core CPU, HDD 500GB, and 4 GB RAM. First only 100 fingerprints for each class were used, it shows delay in queue method where AVL tree is the fastest specially in retrieving a left loop (LL) fingerprint class. Chart – 1, Shows results of the performed experiment.

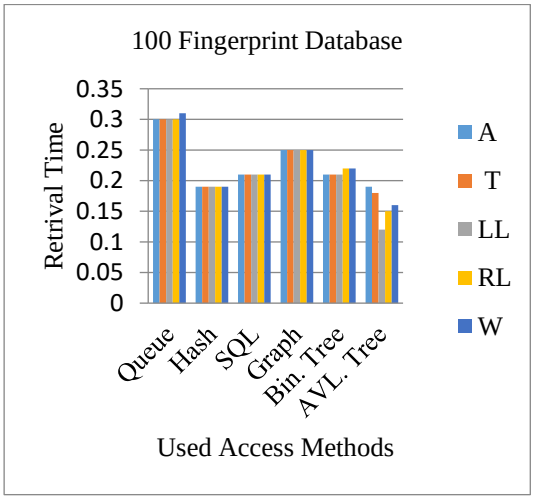

Chart – 1: Time retrieval from 100 Fingerprint Database

Another experiment was performed to retrieve fingerprint images from databases based on using big amount of fingerprints starting from 100 to ≈10000 using a single access methods to discover which the fastest method, used method are Queue, Graph, SQL, Hash, Binary Tree, and AVL tree

Time retrieve for different Class based on single access method, experiment using a binary tree as retrieved algorithm, it shows that the arch class is fastest in retrieving were whorl is the slowest. Char – 2, shows results of using binary tree in retrieving fingerprint class.

International Educational Applied Scientific Research Journal
ISSN (Online): 2456-5040
Volume: 1 | Issue: 1 | October 2016

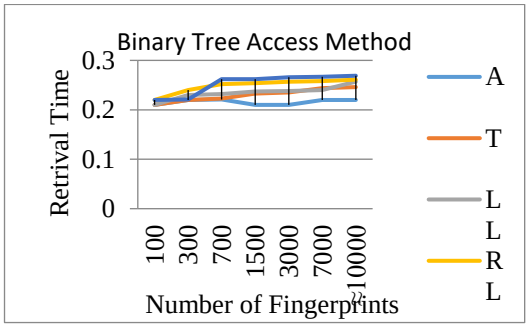

Chart – 2: Results of testing binary tree

Then another experiment of retrieving fingerprint from database using graph method shows that Tarch is the fastest were arch is the slowest. Chart – 3, shows results of using graph in retrieving fingerprint class.

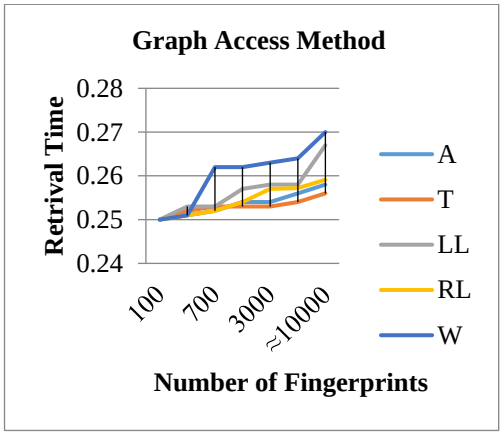

Chart – 3: Results of testing graph method

Same experiment was carried using SQL and hash which shows that retrieving arch is the fastest, were in using AVL tree retrieving lift loop is the fastest. Chart – 4, shows experiment of using SQL, Chart – 5, shows experiment of using Hash and Chart – 6, shows experiment AVL tree.

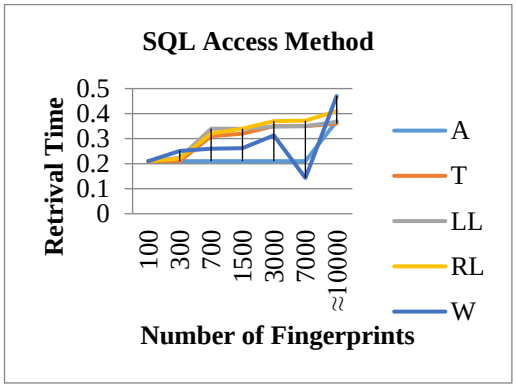

Chart – 4: Results of testing SQL method

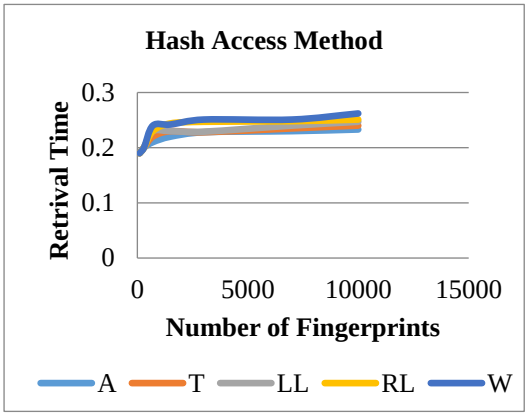

Chart – 5: Results of testing Hash method

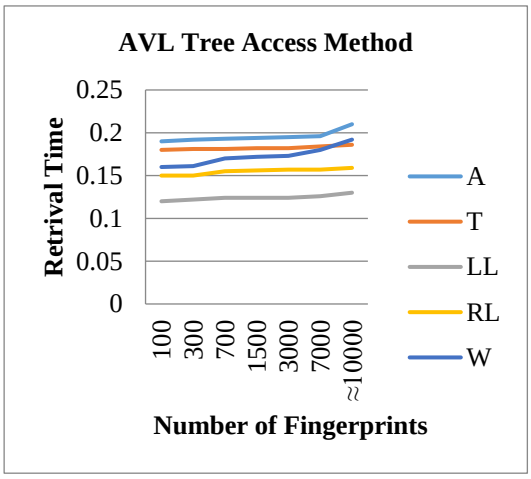

Chart – 6: Results of testing AVL tree method

An experiment also carried with use of different access method to retrieve a signal fingerprint class. In the experiment six access methods are used they are queue, sql, graph, binary tree, hash and AVL tree. In this experiment as shown in Chart – 7, AVL tree scores fastest in retrieving a Tarch class were queue is the slowest.

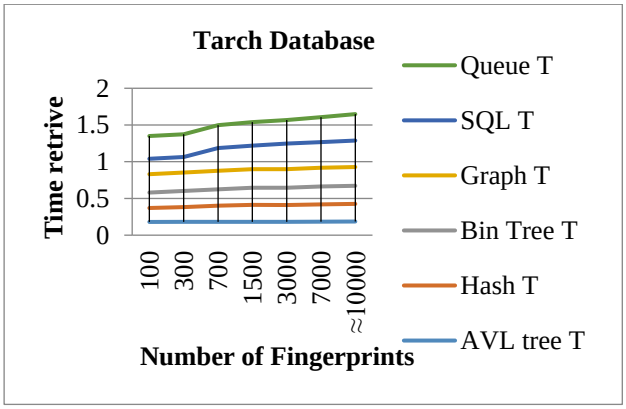

Chart – 7: Time retrieve for Tarch Class using different access methods

International Educational Applied Scientific Research Journal
ISSN (Online): 2456-5040
Volume: 1 | Issue: 1 | October 2016

Same experiment was repeated using arch class, whorl class, right loop class it shows that AVL tree is the fastest where queue the lowest. Chart - 8, Chart - 9 and Chart - 10, show that AVL tree always the fastest and queue is the slowest. Figures also shows that other methods still can be used but when a discussion is concerned to the accessing of information in database minimum time is preferred. So use of AVL tree for this type of data is best solution.

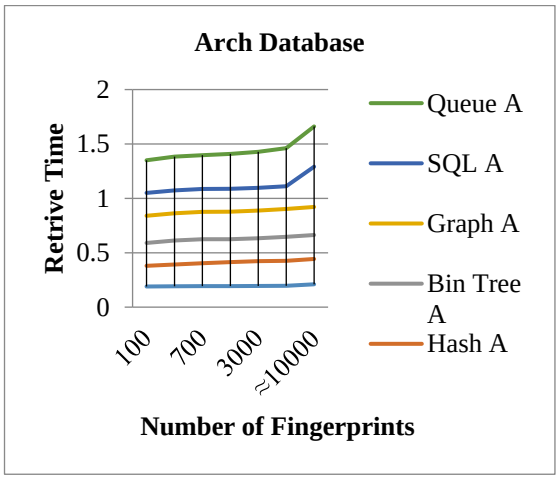

Chart – 8: Time retrieve for Arch Class using different access methods

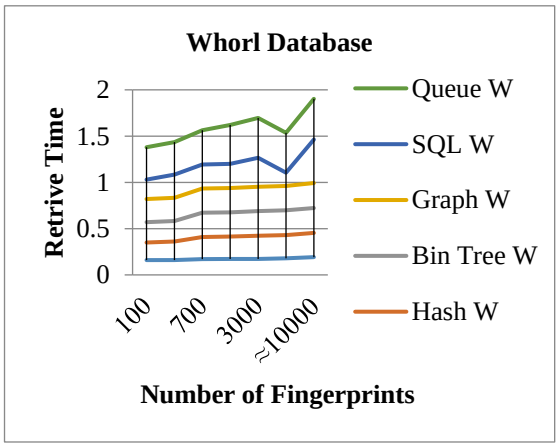

Chart – 9: Time retrieve for whorl Class using different access methods

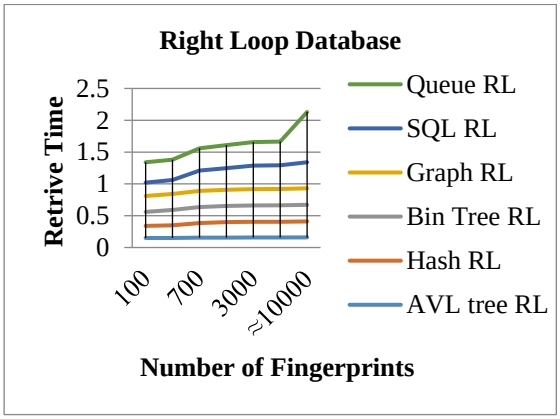

Chart 10: Time retrieve for Right Loop Class using different access methods

2.10 Results Analysis

Results show an importance of separating fingerprint class in minimizing searching databases and in time retrieval. This fact was noted at pre-tests of the proposed algorithm. The AVL tree because it adjust it salve after any insertion or deletion so it can easily access any particular information in the given database, this gives it an advantage even with binary tree that share some similarity specially when a small set of item are used as shown in charts Chart - 8, Chart - 9, and Chart - 10. The queue method is not suitable for this type of data so it scores a big retrieval time even with small set database as shown in mentioned figures. Table 4, shows access time that scored from different used access method and it shows that AVL tree is the fastest among all.

Table 4: Average access time using different fingerprint databases

Data Base / Access method	Avg. 100	Avg. 1500	Avg. ≈10000
Graph	0.25	0.255	0.262
Hash	0.19	0.233	0.247
Bin Tree	0.214	0.240	0.251
AVL tree	0.16	0.166	0.176
SQL	0.21	0.32	0.39
Queue	0.32	0.35	0.46

3 CONCLUSION

This work shows a model of how fingerprint can be classified in database based on class, and then uses several methods to retrieve them. The work proves that an AVL tree is the fastest access and retrieval method not only for fingerprint class but it works effectively for any type of data.

REFERENCES

[1] M. H. Bhuyan, S. Saharia, and D. K. Bhattacharyya, "An Effective Method for Fingerprint Classification" International Arab Journal of e-Technology, Vol. 1, No. 3. 2010.

[2] R. J. Tom1 and T. Arulkumaran, " Fingerprint Based Gender Classification Using 2D Discrete Wavelet Transforms and Principal Component Analysis" International Journal of Engineering Trends and Technology- Vol.4 Issue 2. 2013.

[3] *S. Gagandeep and O. Jasuja "Latent Fingerprint Development" Dye formulations LAP LAMBERT Academic Publishing. 2012.*

[4] R. F. Giberg and B. A. Forouzan, "Data Structures A Pseudocode approach with C++", Brooks/Cole, USA. 2001.

[5] Blieberger, "*Discrete Loops and Worst Case Performance*, Computer Languages", Vol. 20, No. 3, (pp. 193-212).1994.

[6] M. Medidi and D. Deo "**Parallel dictionaries on AVL trees",** Eighth International Parallel Processing Symposium Proceedings. (pp. 878 –882). 1994.

[7] K. S. Larsen "**AVL trees with relaxed balance** Eighth International Parallel Processing Symposium Proceedings. (pp. 888 –893). 1994.

[8] D. Maio and D. Maltoni "A Structural Approach to Fingerprint Classification", *in* proceedings 13th ICPR, Vienna. Vol. 3. (pp. 578 –585). 1996.

[9] K. Jain, S. Prabhakar and L. Hong "**A multichannel approach to fingerprint classification"** IEEE Transactions on Pattern Analysis and Machine Intelligence.Vol.21 Issue: 4. (pp. 348 – 359). 1999.

[10] B. Elmadani "Classification of Fingerprints based on Singular Points" Applied Science journal Sebha University. Vol. 8. No. 1. 2015.

[11] N. Thomas "Sorting and Searching Algorithms: A Cookbook". At http://members.xoom.com/thomasn/s_man.htm. Accessed 1.1. 2016.

[12] H. Feng, D. John and Z. S. Piotr " A Fast Search Algorithm for a Large Fuzzy Database" IEEE Transactions on Information Forensics and Security, Vol. 3, No. 2, June 2008.

[13] R. Chadha, R. Misal and T. Mokashi "Modified Binary Search Algorithm" International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868 Foundation of Computer Science FCS, New York, USA Vol. 7. No. 2. 2014.

[14] B. Elmadani and V. Prakash "New Technique in Fingerprint Matching based on An-AVL tree Search algorithm" Journal of Technology and commerce Brunei Darussalam. Vol 4. No. 1. 2006.

[15] R. J. Tom, T. Arulkumaran and M. E. Scholar "Fingerprint Based Gender Classification Using 2D Discrete Wavelet Transforms and Principal Component Analysis" International Journal of Engineering Trends and Technology- Vol.4 No. 2. 2013.

[16] Introduction To Electrical Engineering and computer Science http://ocw.mit.edu/courses/electrical-engineering-and computer-science/6-01sc-introduction-to-electrical-engineering-and-computer-science-i-spring-2011/unit-4-probability-and-planning/optimizing-a-search/ 2011. Accessed: 1.6.2016.