

Changing and Unchanging Domination in Fuzzy Graph	M. Nithya kalyani Assistant Professor	Mathematics	International Journal of Mathematics Trends and Technology	February 2018	ISSN: 2231-5373	UGC APPROVED JOURNAL IN 2017 https://www.ijmtjournal.org/2018/Volume-54/number-4/IJMTT-V54P538.pdf
---	--	-------------	--	---------------	-----------------	---

International Journal of Mathematics Trends and Technology (IJMTT) – Volume 54 Number 4- February 2018

Changing and Unchanging Domination in Fuzzy Graph

M. Nithya kalyani^{#1}, K. Varunika^{*2}
P.G Head and Assistant Professor[#], M.Phil Scholar^{},
 Department Of Mathematics^{#*},
 Sakthi College of Arts and Science for Women^{**},
 Oddanchatram-624619^{**}, Tamilnadu, India.*

Abstract - In this paper the concepts of Changing and unchanging domination in Fuzzy Graph and for some standard theorems and examples are discussed.

Keywords - domination in fuzzy graph, changing domination in graph, unchanging domination in graph, changing and unchanging domination in fuzzy graphs.

I. INTRODUCTION

The study of domination set in graphs was begun by [5] V.R. Kulli and Bidarhali Janakiram (2004). The domination in fuzzy graph discussed by the [1] A. Nagoor Gani, and V.T. Chandrasekaran, (2006). [2] A. Somasundaram, S. Somasundaram, "Domination in fuzzy Graphs I" (1998), Domination alternation sets in graphs discussed by the [3] D. Bauer, F. Harary, J. Nieminen, S.L. Suffer, changing and unchanging domination in graph [4] J.R. Carrington, F. Harary, T.W. Haynes, changing and unchanging domination in fuzzy graphs $G(V, \rho, \mu)$ of $\gamma(G)$.

II. PRELIMINARIES

A **fuzzy graph** $G = \langle \sigma, \mu \rangle$ is a set with two function $\sigma: V \rightarrow [0, 1]$ and $\mu: E \rightarrow [0, 1]$ such that $\mu(x, y) \leq \sigma(x) \wedge \sigma(y)$ for all $x, y \in V$, hereafter we write $\mu(x, y)$ or $\mu(xy)$. A fuzzy graph $H = \langle \tau, \rho \rangle$ is called a **fuzzy subgraph** of G if $\tau(v_i) \leq \sigma(v_i)$ for all $v_i \in V$ and $\rho(v_i, v_j) \leq \mu(v_i, v_j)$ for all $v_i, v_j \in V$.

A subset S of V is called a **dominating set** in G if for every $v \notin S$, there exists $u \in S$ such that u dominates v . The minimum fuzzy cardinality of a dominating set in G is called the **domination number** of G and is denoted by $\gamma(G)$ or γ .

A dominating set S of a fuzzy graph G is said to be a minimal dominating set if no proper subset of S is a dominating set of G .

The removal of an edge from a graph G can increase by the domination number by at most one and cannot decrease the domination number. $\gamma(G - e) = \gamma(G) + 1$.

The removal of a vertex from a graph G can increase by the domination number by at most one and cannot decrease the domination number. $\gamma(G - v) = \gamma(G) + 1$

Domination color transversal **bondage number** of a graph G denoted by bb_{st} is defined to be the minimum cardinality of collection of sets $E' \subseteq E$ such that $\gamma_{st}(G) = \gamma(G - E')$. If bb_{st} is not defined for K_1 and K_2 .

A graph for which the domination number changes when an **vertex is removal (CVR)** has $V = V^- \cup V^+$ Observed V^0 is never empty for a tree, hence, no tree is in CVR.

$V^0 = \{ v \in V ; \gamma(G - v) = \gamma(G) \}, V^+ = \{ v \in V ; \gamma(G - v) > \gamma(G) \}, V^- = \{ v \in V ; \gamma(G - v) < \gamma(G) \}$

The domination number is unchanged when an arbitrary **vertex is removed class UVR**, then $V = V^0$.

III. CHANGING and UNCHANGING DOMINATION in FUZZY GRAPH

Definition

A Fuzzy graph for which the domination number changes when an arbitrary vertex is removed (CVR) has $V = V^- \cup V^+$.

$$\gamma_f(G - v) \neq \gamma_f(G) \text{ for all } v \in V$$

A Fuzzy graph for which the domination number is unchanged when an arbitrary vertex is removed, (UVR) has $V = V^0$

$$\gamma_f(G - v) = \gamma_f(G) \text{ for all } v \in V$$

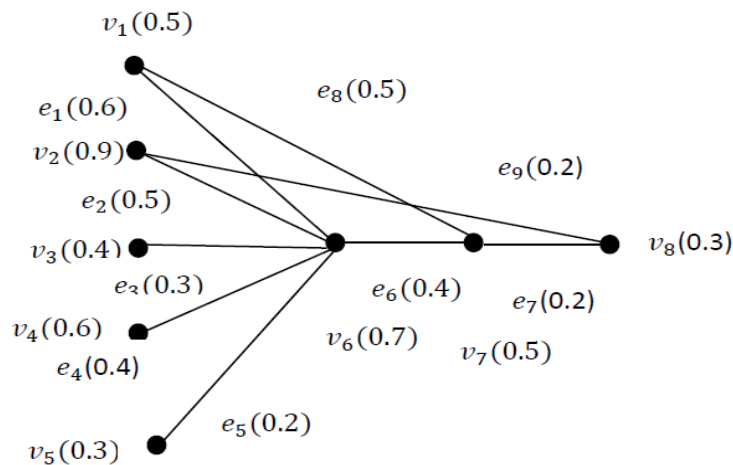


Figure1

Changing and unchanging in fuzzy graph

Theorem

For any fuzzy tree T with $n \geq 2$ there exists a vertex $v \in V$ such that $\gamma(T - V) = \gamma(T)$

Proof

Given that T^f is a fuzzy tree with $n \geq 2$.

Since T^f is a fuzzy tree if G is a acyclic with $n \geq 2$. Let $\gamma(T)$ be a minimum dominating set in a fuzzy tree T^f .

For any vertex $v \in V$. Now we remove the vertex v in fuzzy tree T^f .

Therefore $\gamma(T^f - V)$ is a minimum dominating set of $G(T^f - V)$.

Hence G is a acyclic with $n \geq 2$, then $\gamma(T^f) = \gamma(T^f - V)$.

Theorem

For any graph G , $\gamma_f(G) \leq n - \Delta_f(G) - r_f(G) + 1$

Proof

For any fuzzy graph G with $\gamma(G) \geq 2$. Let $\mu = n - \max\{|N[S]|\}$ for all $S \subset V$ having $|S| = \gamma(G) - 1$. Let $r_f(G)$ to be smallest number of edges which must be added to G to decrease the domination number.

$$\text{Hence } \gamma_f(G) \leq n - \Delta_f(G) - r_f(G) + 1.$$

Theorem

If a graph $G \in CVR$ has order $n = (\Delta(G) + 1)(i(G) - 1) + 1$. Then G is regular fuzzy graph.

Proof

Suppose $G \in CVR$ with $n = (\Delta(G) + 1)(i(G) - 1) + 1$

By theorem

Suppose $i(G) \neq i(G - V)$ for all $v \in V$ then $V^0 = \emptyset$ and so $v \in V^+$ or $v \in V^-$. If $v \in V^+$, then for each $v \in V^+$ $P_{fn}[v, S]$ contains two nonadjacent vertices. These vertices are in $v - S$ and so again if $P_{fn}[v, S] = \{v\}$ then $S - \{v\}$ is an independent fuzzy dominating set of $G - V$ which is a contradiction of $v \in V^+$

Hence V has a private neighbor in $V - S$.

Suppose $\langle P_{fn}[v, S] \rangle$ is complete. Then $S - \{v\} \cup \{u\}$ for any $u \in P_{fn}[v, S]$ is an $i(G)$ -set of G not containing v again a contradiction.

Hence, $\langle P_{fn}[v, S] \rangle$ contains at least two nonadjacent vertices they are not in V^+ . Since $G \in CVR$ we have $V = V^-$. Let S_u denote an $i(G)$ -set of $G - u$. So that $|S_u| = i(G) - 1$.

In order to dominate the $(\Delta(G) + 1)(i(G) - 1)$ vertices of $G - u$ each element of S_u must dominate exactly $(\Delta(G) + 1)$ vertices and has degree $\Delta(G)$. Thus no two vertices in S_u have a common neighbor.

To prove u is regular. It is enough to prove that for any arbitrary vertex, $x \in S_u$ for some u .

Let $r \in S_x$ we prove that $r \in S_{rf}$. Suppose $x \notin S_{rf}$

Since $G \in CVR$, $S_{rf} \cap N[r] = \emptyset$ each vertex $S_x - \{r\}$ dominates fuzzy unique vertex of S_r . So remaining vertex in S_{rf} which is not x must be dominated by S_x and so must be dominated by rf which is a contradiction as $S_{rf} \cap N(rf) = \emptyset$.

Hence $x \in S_{rf}$ and so G is regular fuzzy graph.

Theorem

Let G be a unchanging fuzzy graph. If P_p is a path on P vertices. Then $P_p \in UEA$ if and only if $P = 3k + 2 (k \geq 1)$.

Proof

Suppose $P_p \in UEA$. Let $P = 3k$.

Let $P_p = (1, 2, \dots, k)$ that $k > 1$. Consider the edge $e = (2, 5)$. Then, $i(P_p + e) = k + 1$ where $P_p = k$ and so $P_p \notin UEA$ which is a contradiction.

Hence $P \neq 3k (k > 2)$

Similarly,

If $P = 3k + 1$ with $e = (1, 3)$. Then, $i(P_p + e) = k$ where as, $i(P_p) = k + 1$ and so $P_p \notin UEA$ which is a contradiction. Hence $n \neq 3k + 1$ and so either $P = 3$ or $P = 3k + 2$.

Conversely,

Assume $P = 3k + 2$ ($k \geq 1$). Then $i(P_p) = k + 1$

We prove that $P_p \in UEA$ by induction on k . When $k = 1, P = 5$ and one can verify that $i(P_5 + e) = i(P_5) = 2$ for all $e \in E(\bar{G})$. Assume that the result is true for k .

Now we prove the result for $P = 3(k + 1) + 2 = 3k + 5$. Let $P_{3k+5} = (1, 2, \dots, 3k + 5)$ by induction hypothesis any edges joining two vertices among $(1, 2, \dots, 3k + 2)$ will not change $i(G)$. An adding $e = (3k + 3, 3k + 5)$ we observe that $i(G + e) = k + 2$. Also it is easy to verify that any edge joining a vertex of $\{1, 2, \dots, 3k + 2\}$ and a vertex of $\{3k + 3, 3k + 4, 3k + 5\}$ does not change $i(G)$.

Hence $P_p \in UEA$

IV. CONCLUSIONS

The concept of changing and unchanging domination are analyzed. In this paper changing and unchanging domination in fuzzy graph are introduced and the some standard theorems are discussed.

REFERENCES

- [1] A.Nagoor Gani, and V.T.Chandrasekaran, "Domination in Fuzzy Graph", Advances in Fuzzy Sets and Systems, 1(1) (2006), 17-26.
- [2] A.Somasundaram, S.Somasundaram, "Domination in Fuzzy Graphs –I", Elsevier sciences, 19 (1998) 787-791. Discrete Mathematics 33(1981) 249 - 258.
- [3] D.Bauer, F.Harary, J.Nieminen, S.L.Suffel, Domination alteration sets in graphs, Discrete math.47(1983) 153-161.
- [4] J.R.Carrington, F.Harary, T.W.Haynes, Changing and unchanging domination, J.Combin.Math.Combin.Comput.9(1991) 57-63.
- [5] V. R. Kulli and Bidarhalli Janakiram, "The dominating graph," Graph Theory Notes of New York XLVI, 5-8 (2004).

Changing and Unchanging Domination in Fuzzy Graph	M. Nithya kalyani, Assistant Professor	Mathematics	International Journal of Mathematics Trends and Technology	2018	2456-3307	UGC APPROVED JOURNAL IN 2017 https://www.ijmttjournal.org/2018/Volume-54/number-4/IJMTT-V54P539.pdf
---	--	-------------	--	------	-----------	---

International Journal of Mathematics Trends and Technology (IJMTT) – Volume 54 Number 4- February 2018

Changing and Unchanging Domination in Fuzzy Graph

M. Nithya kalyani^{#1}, K. Varunika^{*2}
P.G Head and Assistant Professor[#], M.Phil Scholar^{},
 Department Of Mathematics[#],
 Sakthi College of Arts and Science for Women^{**},
 Oddanchatram-624619^{**}, Tamilnadu, India.*

Abstract - In this paper the concepts of Changing and unchanging domination in Fuzzy Graph and for some standard theorems and examples are discussed.

Keywords - domination in fuzzy graph, changing domination in graph, unchanging domination in graph, changing and unchanging domination in fuzzy graphs.

I. INTRODUCTION

The study of domination set in graphs was begun by [5] V.R. Kulli and Bidarhali Janakiram (2004). The domination in fuzzy graph discussed by the [1] A. Nagoor Gani, and V.T. Chandrasekaran, (2006). [2] A. Somasundaram, S. Somasundaram, "Domination in fuzzy Graphs I" (1998), Domination alternation sets in graphs discussed by the [3] D. Bauer, F. Harary, J. Nieminen, S.L. Suffer, changing and unchanging domination in graph [4] J.R. Carrington, F. Harary, T.W. Haynes, changing and unchanging domination in fuzzy graphs $G(V, \rho, \mu)$ of $\gamma(G)$.

II. PRELIMINARIES

A **fuzzy graph** $G = \langle \sigma, \mu \rangle$ is a set with two function $\sigma: V \rightarrow [0, 1]$ and $\mu: E \rightarrow [0, 1]$ such that $\mu(x, y) \leq \sigma(x) \wedge \sigma(y)$ for all $x, y \in V$. hereafter we write $\mu(x, y)$ or $\mu(xy)$. A fuzzy graph $H = \langle \tau, \rho \rangle$ is called a **fuzzy subgraph** of G if $\tau(v_i) \leq \sigma(v_i)$ for all $v_i \in V$ and $\rho(v_i, v_j) \leq \mu(v_i, v_j)$ for all $v_i, v_j \in V$.

A subset S of V is called a **dominating set** in G if for every $v \notin S$, there exists $u \in S$ such that u dominates v . The minimum fuzzy cardinality of a dominating set in G is called the **domination number** of G and is denoted by $\gamma(G)$ or γ .

A dominating set S of a fuzzy graph G is said to be a minimal dominating set if no proper subset of S is a dominating set of G .

The removal of an edge from a graph G can increase by the domination number by at most one and cannot decrease the domination number. $\gamma(G - e) = \gamma(G) + 1$.

The removal of a vertex from a graph G can increase by the domination number by at most one and cannot decrease the domination number. $\gamma(G - v) = \gamma(G) + 1$

Domination color transversal **bondage number** of a graph G denoted by bb_{st} is defined to be the minimum cardinality of collection of sets $E' \subseteq E$ such that $\gamma_{st}(G) = \gamma(G - E')$. If bb_{st} is not defined for K_1 and K_2 .

A graph for which the domination number changes when an **vertex is removal (CVR)** has $V = V^- \cup V^+$ Observed V^0 is never empty for a tree, hence, no tree is in CVR.

$V^0 = \{ v \in V ; \gamma(G - v) = \gamma(G) \}$, $V^+ = \{ v \in V ; \gamma(G - v) > \gamma(G) \}$, $V^- = \{ v \in V ; \gamma(G - v) < \gamma(G) \}$

The domination number is unchanged when an arbitrary **vertex is removed class UVR**, then $V = V^0$.

III. CHANGING and UNCHANGING DOMINATION in FUZZY GRAPH

Definition

A Fuzzy graph for which the domination number changes when an arbitrary vertex is removed (CVR) has $V = V^- \cup V^+$.

$$\gamma_f(G - v) \neq \gamma_f(G) \text{ for all } v \in V$$

A Fuzzy graph for which the domination number is unchanged when an arbitrary vertex is removed, (UVR) has $V = V^0$

$$\gamma_f(G - v) = \gamma_f(G) \text{ for all } v \in V$$

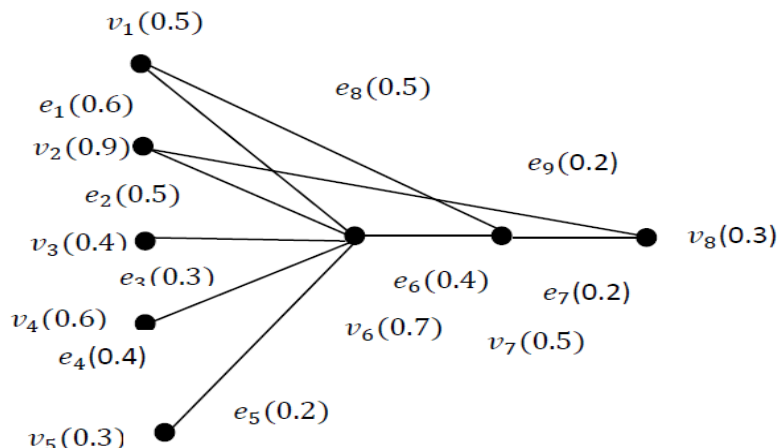


Figure1

Changing and unchanging in fuzzy graph

Theorem

For any fuzzy tree T with $n \geq 2$ there exists a vertex $v \in V$ such that $\gamma(T - V) = \gamma(T)$

Proof

Given that T^f is a fuzzy tree with $n \geq 2$.

Since T^f is a fuzzy tree if G is a acyclic with $n \geq 2$. Let $\gamma(T)$ be a minimum dominating set in a fuzzy tree T^f .

For any vertex $v \in V$. Now we remove the vertex v in fuzzy tree T^f .

Therefore $\gamma(T^f - V)$ is a minimum dominating set of $G(T^f - V)$.

Hence G is a acyclic with $n \geq 2$, then $\gamma(T^f) = \gamma(T^f - V)$.

Theorem

For any graph G , $\gamma_f(G) \leq n - \Delta_f(G) - r_f(G) + 1$

Proof

For any fuzzy graph G with $\gamma(G) \geq 2$. Let $\mu = n - \max_{S \subset V} |N[S]|$ for all $S \subset V$ having $|S| = \gamma(G) - 1$. Let $r_f(G)$ to be smallest number of edges which must be added to G to decrease the domination number.

$$\text{Hence } \gamma_f(G) \leq n - \Delta_f(G) - r_f(G) + 1.$$

Theorem

If a graph $G \in CVR$ has order $n = (\Delta(G) + 1)(i(G) - 1) + 1$. Then G is regular fuzzy graph.

Proof

Suppose $G \in CVR$ with $n = (\Delta(G) + 1)(i(G) - 1) + 1$

By theorem

Suppose $i(G) \neq i(G - V)$ for all $v \in V$ then $V^0 = \emptyset$ and so $v \in V^+$ or $v \in V^-$. If $v \in V^+$, then for each $v \in V^+$ $P_{fn}[v, S]$ contains two nonadjacent vertices. These vertices are in $v - S$ and so again if $P_{fn}[v, S] = \{v\}$ then $S - \{v\}$ is an independent fuzzy dominating set of $G - V$ which is a contradiction of $v \in V^+$

Hence V has a private neighbor in $V - S$.

Suppose $\langle P_{fn}[v, S] \rangle$ is complete. Then $S - \{v\} \cup \{u\}$ for any $u \in P_{fn}[v, S]$ is an $i(G)$ -set of G not containing v again a contradiction.

Hence, $\langle P_{fn}[v, S] \rangle$ contains at least two nonadjacent vertices they are not in V^+ . Since $G \in CVR$ we have $V = V^-$. Let S_u denote an $i(G)$ -set of $G - u$. So that $|S_u| = i(G) - 1$.

In order to dominate the $(\Delta(G) + 1)(i(G) - 1)$ vertices of $G - u$ each element of S_u must dominate exactly $(\Delta(G) + 1)$ vertices and has degree $\Delta(G)$. Thus no two vertices in S_u have a common neighbor.

To prove u is regular. It is enough to prove that for any arbitrary vertex, $x \in S_u$ for some u .

Let $r \in S_x$ we prove that $r \in S_{rf}$. Suppose $x \notin S_{rf}$

Since $G \in CVR$, $S_{rf} \cap N[rf] = \emptyset$ each vertex $S_x - \{r\}$ dominates fuzzy unique vertex of S_r . So remaining vertex in S_{rf} which is not x must be dominated by S_x and so must be dominated by rf which is a contradiction as $S_{rf} \cap N(rf) = \emptyset$.

Hence $x \in S_{rf}$ and so G is regular fuzzy graph.

Theorem

Let G be a unchanging fuzzy graph. If P_p is a path on P vertices. Then $P_p \in UEA$ if and only if $P = 3k + 2 (k \geq 1)$.

Proof

Suppose $P_p \in UEA$. Let $P = 3k$.

Let $P_p = (1, 2, \dots, k)$ that $k > 1$. Consider the edge $e = (2, 5)$. Then, $i(P_p + e) = k + 1$ where $P_p = k$ and so $P_p \notin UEA$ which is a contradiction.

Hence $P \neq 3k (k > 2)$

Similarly,

If $P = 3k + 1$ with $e = (1, 3)$. Then, $i(P_p + e) = k$ where as, $i(P_p) = k + 1$ and so $P_p \notin UEA$ which is a contradiction. Hence $n \neq 3k + 1$ and so either $P = 3$ or $P = 3k + 2$.

Conversely,

Assume $P = 3k + 2$ ($k \geq 1$). Then $i(P_p) = k + 1$

We prove that $P_p \in UEA$ by induction on k . When $k = 1$, $P = 5$ and one can verify that $i(P_5 + e) = i(P_5) = 2$ for all $e \in E(\bar{G})$. Assume that the result is true for k .

Now we prove the result for $P = 3(k + 1) + 2 = 3k + 5$. Let $P_{3k+5} = (1, 2, \dots, 3k + 5)$ by induction hypothesis any edges joining two vertices among $(1, 2, \dots, 3k + 2)$ will not change $i(G)$. An adding $e = (3k + 3, 3k + 5)$ we observe that $i(G + e) = k + 2$. Also it is easy to verify that any edge joining a vertex of $\{1, 2, \dots, 3k + 2\}$ and a vertex of $\{3k + 3, 3k + 4, 3k + 5\}$ does not change $i(G)$.

Hence $P_p \in UEA$

IV. CONCLUSIONS

The concept of changing and unchanging domination are analyzed. In this paper changing and unchanging domination in fuzzy graph are introduced and the some standard theorems are discussed.

REFERENCES

- [1] A.Nagoor Gani, and V.T.Chandrasekaran, "Domination in Fuzzy Graph", Advances in Fuzzy Sets and Systems, 1(1) (2006), 17-26.
- [2] A.Somasundaram, S.Somasundaram, "Domination in Fuzzy Graphs –I", Elsevier sciences, 19 (1998) 787-791. Discrete Mathematics 33(1981) 249 - 258.
- [3] D.Bauer, F.Harary, J.Nieminen, S.L.Suffel, Domination alteration sets in graphs, Discrete math.47(1983) 153-161.
- [4] J.R.Carrington, F.Harary, T.W.Haynes, Changing and unchanging domination, J.Combin.Math.Combin.Comput.9(1991) 57-63.
- [5] V. R. Kulli and Bidarhalli Janakiram, "The dominating graph," Graph Theory Notes of New York XLVI, 5-8 (2004).

Extraction of Top k Itemsets From High Utility Itemsets Using Faster High-Utility Itemset Miner	S.Kavitha, Assistant Professor	Computer Science	International Journal of Scientific Research in Computer Science, Engineering and Information Technology	March 2018	2456-3307	https://ijsrcseit.com/CSEIT1833194 Google scholar UGC CARE Listed Journal No : 64718
---	--------------------------------	------------------	--	------------	-----------	--



International Journal of Scientific Research in Computer Science, Engineering and Information Technology

© 2018 IJSRCSEIT | Volume 3 | Issue 3 | ISSN : 2456-3307

Extraction of Top K Itemsets From High Utility Itemsets Using Faster High-Utility Itemset Miner

M. Geetha¹, S. Kavitha²

¹Research Scholar, Department of Computer Science, Sakthi college of Arts and Science For Women, oddanchatram, Tamil Nadu, India

²Head and Associate Professor, Department of Computer Science, Sakthi college of Arts and Science For Women, oddanchatram, Tamil Nadu, India

ABSTRACT

Frequent itemset mining is the recent research topic in the data mining systems. It generally composes of tremendous volume of frequently searched/retrieved item with low/ high itemset values. This dilemma doesn't satisfy the user's requirements. The utility itemsets is an important topic and it can be measure in terms of weight, value, quantity and all other information's depending on the user's requirements. If the utility itemset is no less than user specified min utility, so this itemset is called a utility of high itemset. It contains a many applications like biomedicine, mobile computing, market analysis, etc. In database, the HUI is a difficult, because in FIM used the downward closer property is does not hold the utility of itemsets. Superset the low utility itemset can be a high utility so the3 HUI pruning search space is also difficult. To overcome this issue, we discovered fittest threshold for mining the relevant itemsets from set of itemsets. Setting of min-util value to the user is a daunting task. In order to find an efficient threshold value for the users, the behaviors of the users are studied. In this work, we proposed two mechanisms, namely, mining top k utility itemsets and mining top k utility itemsets in single phase in which k is the number of covered HUI mining. Initially, we give an auxiliary examination of the two calculations with talks on their preferences and restrictions. Exact assessments on both genuine and manufactured datasets demonstrate that the execution of the proposed calculations is near that of the ideal instance of best in class utility mining calculations.

Keywords: Cloud computing, Cloud security, Peer to Peer, Resource Description Framework.

I. INTRODUCTION

Data mining is the field of our study. The applications of data mining are tremendously growing due to the growth of information technologies. In general context, data mining is explained as follows:

- (i) Extracting the relevant knowledge from the set of unidentified or identified set of resources.
- (ii) The formation of meaningful pattern by exploring the data in a hyperplane system.

The real world data may be in structured or unstructured form. The main objective of the work is to find the relations or similarity between the data for deriving useful knowledge. The behavior of data implies lot of information from its elementary form. It also plays a vital role in the data analysing process. It authorizes users to analyze data from several diverse dimensions or angles, categorize it, and sum up the relationships acknowledged.

Several users make use of data mining for discovering the knowledge from variant aspects. In some cases, knowledge discovery is a developmental step in the

data mining process. It comprises of several steps as follows:

- ✓ Cleaning the data which removes the noise or inaccurate data.
- ✓ Merging the data for achieving better data availability.
- ✓ Selecting the data for achieving better data retrievability.
- ✓ Transformation of data that completely migrates the data into better interpretability form.
- ✓ Mining the data that extracting the relevant data patterns.
- ✓ Evaluating the patterns that derive the data from the similar patterns.
- ✓ Presenting the knowledge from the derived databases.

1.1 Working procedure of data mining systems

The data analysis may be carried out in large or small scale data. Based on their data association, the data are classified or clustered. The queries are transacted over those datasets with unrestricted no. of queries. It has been applied to various fields like statistical, machine learning and neural networks. The data relationship is constructed from four models:

Classes: Initially, the received data is stored in an indexed form. The similar data is grouped based on their representative classes. By doing so, it reduces the effects of data traffic.

Clusters: Based on their logical relationships, the data are clustered.

Association: Based on the consecutive purchasing behavior of the users, the data are associated.

Sequential mining: In order to predict the behavior patterns, the past knowledge are used for finding the pattern similarity score. Depends on those score, the relevant data is achieved.

1.2 Characteristics of Data Mining

The main characteristics of the data mining are listed as follows:

Higher volume of data: It is enormous in real time process. Each data has to be analyzed effectively for

constructing the relationships.

Incomplete of the data: It differentiates the quality of data from its original resources.

Data structure: It is complex in nature which predicts the statistical analysis process. The stored data may be in heterogeneous form

1.3 Benefits of Data Mining

The advantages of the data mining system are the:

- ✓ It is one of the best data rendering services.
- ✓ It depicts the better data retrieving services.
- ✓ It also helps to store and retrieve the data based on their behavior.
- ✓ It also genuinely derives the valuable information.
- ✓ It loads the data with the association of data system.
- ✓ It develops the better relationship with each other.
- ✓ It helps to develop extraordinary data promotion systems.

The stored data is flexible in nature.

1.4 Application of Data Mining

The main applications of the data mining are:

Marketing systems: In the marketing field, the data is constructed based on their historical data. It helps for promoting their brands via direct mail, online marketing, campaign etc. In order to maintain their retaining strategy, data mining techniques are widely adopted. As a result of market basket analysis, a store can include an appropriate production collection in a way that customers can buy frequent buying products in concert with satisfying. In addition it facilitates the retail companies to offer positive discounts for particular products by that it will pull towards a lot of customers.

Banking systems : To derive the knowledge from the financial data, data mining techniques are widely studied. Based on the historical transactional data, the loan prediction system is formed. Data mining aids banks identify fraudulent credit card

transactions to save from harm credit card's owner.

Manufacturing applications: In order to find any errors in the equipments, the parameter optimization developed using data mining process are employed. Still, some imperfection may occur due to the ranges of control parameters. Then, those optimal control parameters are utilized to manufacture wafers with preferred eminence.

Governments scenario: Data mining facilitates government agency by means of excavating and analyzing records of financial transaction to build patterns that can identify money decontaminate or criminal activities.

Law oriented application :Data mining can assists law enforcers in recognizing illegal suspects as well as arresting these criminals by investigating inclinations in location, crime type, habit, and additional patterns of behaviors.

Researching field : Data mining can aids researchers by speeding up their process of analyzing the data; therefore, permitting those more time to work on other projects.

II. FREQUENT ITEMSET MINING (FIM)

Frequent Itemset mining is studied for the market basket analysis. It is a kind of data analysis technique that finds the certainties and uncertainties in the data systems. It is mainly predictable for predicting the purchasing behavior. It deals with the products recognition and mail delivery subsystem. In certain cases, multitude of the data assigning tasks to be followed.

Already from the start, Frequent Itemset Mining (FIM) has been an essential part of data analysis and data mining. FIM tries to extract information from databases based on frequently occurring events, i.e., an event, or a set of events, is interesting if it occurs frequently in the data, according to a user given minimum frequency threshold. Many techniques

have been invented to mine databases for frequent events. These techniques work well in practice on typical datasets, but they are not suitable for truly Big Data. Applying frequent itemset mining to large databases is problematic. First of all, very large databases do not fit into main memory. In such cases, one solution is to use level-wise breadth first search based algorithms, such as the well known Apriori algorithm, where frequency counting is achieved by reading the dataset over and over again for each size of candidate itemsets. Unfortunately, the memory requirements for handling the complete set of candidate itemsets blows up fast and renders Apriori based schemes very inefficient to use on single machines. Secondly, current approaches tend to keep the output and runtime under control by increasing the minimum frequency threshold, automatically reducing the number of candidate and frequent itemsets. However, studies in recommendation systems have shown that itemsets with lower frequencies are more interesting.

Item Set Enumeration derives the general top-down search scheme for item set enumeration from the fundamental properties of the support measure, resulting in breadth-first and depth-first search, with the sub-problem and item order providing further distinctions. Database Representations reviews different data structures by which the initial as well as conditional transaction databases can be represented and how these are processed in the search. Advanced Techniques collects several advanced techniques that have been developed to make the search maximally efficient, including perfect extension pruning, conditional item reordering, the k-items machine, and special output schemes. Intersecting Transactions briefly surveys intersecting transactions as an alternative to item set enumeration for finding closed (and maximal) item sets, which can be preferable in the presence of (very) many items. Extensions discusses selected extensions of the basic approaches, such as association rule induction, alternatives to item set support, association rule and item set ranking and

filtering methods, and fault-tolerant item sets.

2.1 Working of Frequent Itemset Mining

First, we build the header table which consists of item name and link field corresponding to each item. All link entry of header table is initially set to null. Whenever an item first time added into the tree, the corresponding entry of header table is updated. The root, labeled as “null”, is created. Children are added by scanning the database.

TID	ITEMS	Frequent Items
T1	A, B	A, B
T2	C, D, E	D, E, C
T3	A, B, C, F	A, B, C, F
T4	A, C, D	A, D, C
T5	A, D, F	A, D, F
T6	B, E, I	B, E
T7	A, B, D, E	A, D, B, E
T8	D, E, F, I	D, E, F

Figure 1. Transaction details

First, a path that shares same prefix is require to be search. If there exist a path that is same as any prefix of current items of transaction (in ist order) then the count of prefix portion is incremented by one in the tree, remaining items of same transaction (which do not share the path), are added from the last node of sharing portion in ist order and their count value is set to 1. If items of a transaction do not share any path of tree then they are added from the root in ist order. Each path of prefix tree (FP-tree) represents a set of transactions.

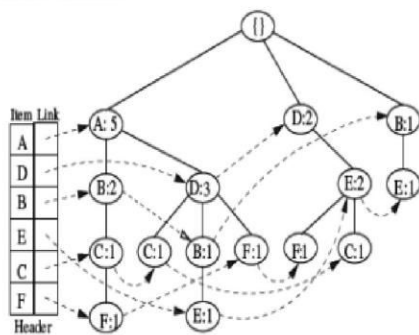


Figure 2. FP tree

The limitations of these existing methods are the ones inherited from the original methods. The size of the data for the level-wise generate-and-test techniques affects their scalability and the pattern-growth techniques require a lot of memory for

accommodating the dataset in the data structures, such as the FP-tree, especially when the transactions do not share many items. In the case of uncertain data, not only the items have to be shared for a better compression but also the existence probabilities, which is often not the case.

III. MOTIVATION

The mining tasks are performed on ordered/unordered database. It is very sensation and significant research system in the data mining process. Several kinds of applications are available for predicting the futuristic scope with the help of past knowledge. The applications such as telecommunication, user activities, crime detection, illness recovery etc are determined from the frequent based analysis. The objective of the itemset mining is to discover the association among those frequently used items. Relied upon the threshold, the high frequencies are sorted and the lists are updated. In some cases, the threshold level is less in number.

In real-life scenarios, market based analysis is used widely studied. Based on the customer's history, the similar profiles are collected and stored. In order to discover the purchasing behavior, these products similarity are studied. Market basket analyses gives retailer proper information regarding related sales on collection of goods basis Customers who buy s bread frequently moreover buy several products associated to bread like milk, butter or jam. It makes intellect that these groups are situated alongside in a retail center in order that customers can contact them rapidly. This sort of connected customer behavior analysis helps to aware about the customers via logical systems.

IV. PROBLEM DEFINITION

In general context, pattern mining is the field of study that discovers the relationships between each item. The frequency may be captured on the basis of similar itemsets, subsequences and the substructures

from homogeneous and heterogeneous data. The frequency should not lead the defined thresholds. The apriori algorithm is purely depends on the association rule mining systems. Based on the candidates and its substructures, the apriori algorithm is defined. It is further divided into two patterns, namely, apriori frequent pattern growth and the Equivalence CLASS transformation (ECLAT). Most of the transactional databases offer variant knowledge discovery process. Though the transposing of the databases is an easier model, the procedure to be carried out is not easy. In this study, the extraction of similar patterns makes use of transposed database method. A tremendous amount of data has been generated by the transposed database model. It composes of large set of resources and objects. Each object contains a set of attributes. In order to minimize the search space, the data are stored onto vertical format.

V. THE HUN-MAX ALGORITHM

Input: DB: a transaction database, min_util: a user-specified threshold

Output: the set of high-utility itemsets

Step-1: First Scan DB

Step-2: Then calculate the TWU for single items;

Step-3: Identify the set ITWU which contains each item i

Step-4: Each item i such that $TWU(i) < min_util$;

Step-5: Let the total order of TWU ascending values on ITWU be α ;

Step-6: Scan DB to build the utility-list of each item $i \in ITWU$

Step-7: Build the Utility Matrix structure;

Step-8: Find Extensions such as (NULL, ITWU, min_util, Utility Matrix);

VI. EXPERIMENTAL RESULTS

The shopping analysis is taken from the public repository as an application scenario to implement. The comparison of experimental analysis is shown on synthetic and real datasets. It describes the result and

performance analysis of the proposed algorithm. Java is the programming tool used for validating the proposed algorithm. The Java programming language is a high-level language that can be characterized by all of the following buzzwords:

- ✓ Simple
- ✓ Architecture neutral
- ✓ Object oriented
- ✓ Portable
- ✓ Distributed
- ✓ High performance
- ✓ Interpreted
- ✓ Multithreaded
- ✓ Robust
- ✓ Dynamic
- ✓ Secure

product_itemID	product_item_names	shopping_items_id	shopping_items_names	id	Products	Total
1	Recreation Centre...	1	Department Store...	1	REACT...	8494
2	Gym Trainers	2	American Truck...	2	HOME_S...	2212
3	Trainers Gym	3	Art Supplies Toy...	3	HOTEL...	5882
4	Gym Hoods	4	Coffee & Tea Ac...	4	NIGHT...	21306
5	Trainers Gym V...	5	Hardware Stores...	5	RESTAUR...	24057
6	Parks Landmark...	6	Auto & Crafts Bo...	6	SHOPPL...	16153
7	Museum Aquari...	7	Bookstores Vinyl...	7	PETS	1617
8	Parks Lakes	8	Accessories Lamp...	8	Total_P...	84622
9	Gym Trainers	9	Adults Lampre...			
10	Lounges Decore...	10	Draperys Cent...			
11	Gym Trainers	11	Toy Stores Card...			
12	Zooz Vennor & ...	12	Adults Entertain...			
13	Gym Trainers P...	13	Food Vennor & ...			

Figure 3. Product Items Separation

Figure 3 represents a loading of the items on real and synthetic datasets in the terms of product items separation. The selected dataset is shopping. There are three tables which contain the overall product, shopping items and category of the product respectively. The first table consists of the entities: product item id and product item name, by which all the products can be viewed like the items are recreation centre, Gym trainers, Gym hotels etc. and the second table is containing two entities: shopping items id and shopping item names by which items can be viewed like Department stores, art supplies toy, hardware stores etc. and the third table contains the entities: id, products, Total. This consequences the total products.

Figure 4. User Transaction (Bought) Items

Figure 4 represents a loading of the dataset and the selected dataset is shopping. There are three tables; one is containing the user id, Product id and the user rating by which it is used to view the ratings given by the user, second table is containing the entities shopping users, shopping items, shopping ratings, and the last table contains the entities id, names like Beauty & S..., Home_ser, shopping etc. the outcome of this data set is total bought items.

Figure 5. Top most User frequency

Figure 5 represents a loading of the dataset and the selected dataset is shopping. There are two tables; one is for user frequency by including the entities shopping user id and user utility items count by which it consequences the Reputation of the user frequency. Another table is for top most user frequency includes entities shopping items top most user frequency. The outcome of this figure is to show the frequency of top most users.

Figure 6. Find out the top k Utility Itemsets

Figure 6 represents a loading of the dataset of the Utility Items. There are two tables; one is for utility items by including the entities shopping items id and utility items and as well as it used to check the frequent items in a dataset. Another table is for top most utility items including the entities shopping items and top most utility. It determines the top k utility items from all the utility items.

Figure 7. Find out the top k utility itemsets in one phase.

Figure 8. Shopping

Figure 7 represents a loading of the dataset of the Utility Items. There are two tables; one is for utility

items by including the entities shopping items id and shopping items name and items utility count. Another table is for top most utility items including the entities shopping items, shopping items name and top most utility. It determines the top k utility items in one phase from all the utility items.

Figure 8 represents a loading of the dataset of the Utility Items. There are two tables; one is for negative utility items by including the entities shopping items id and shopping items name and topmost utility, the shopping items, in this table are art supplies, drug stores, toy stores, adult entertainment. Another table is for positive utility items including the entities shopping items, shopping items name and items utility. It finds out both the negative and positive utility items.

shopping_item_id	shopping_item_name	Topmost_utility
61	Home Decor F...	18
147	Women's Cloth...	17
236	Women's Cloth...	18
544	Men's Clothing ...	13
765	Home Decor F...	12
1163	Men's Clothing ...	15
1888	Men's Clothing ...	15

Figure 9. Negative profit Utility Itemset

Figure 9 represents a loading of the dataset of the Utility Items. There is a negative utility profit items dataset. This dataset contains the shopping items id, shopping items name and topmost utility items. The outcome of this is to find the negative profit utility items.

shopping_item_id	shopping_item_name	items_utility_count
18	Museum Art Gall...	63
29	Kitchen & Bath Fi...	62
56	Department Stores...	62
93	Electronics Phone...	69
236	Shopping Centers...	64
544	Men's Clothing We...	123
751	Electronics Comp...	61
892	Farmers Market Sh...	67

Figure 10. Dataset of the Utility Items

Figure 10 represents a loading of the dataset of the Utility Items. There is a positive utility profit items dataset. This dataset contains the shopping items id, shopping items name and topmost utility items the shopping items are such as museum art gallery, department stores, shopping centers etc. The outcome of this is to find the positive profit utility items.

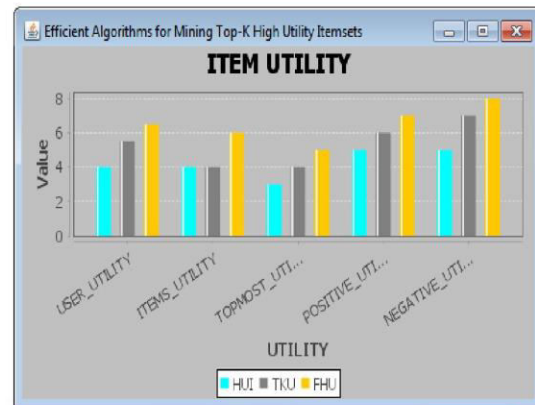


Figure 11. Performance analysis of utility items.

Figure 11 represents the performance analysis in the terms of user utility, items utility, topmost utility, positive utility and negative utility between the proposed HUI (High Utility Items), and existing TKU (Top K Utility) and FHU (Frequently High Utility). It is inferred that the proposed HUI consumes time range of 0-4s than the existing techniques.

VII. CONCLUSION

Frequent itemset mining is one of the recent research study focused by the researchers. In order to limit the size of the output, the itemsets with efficient utilities has to be selected from the pool of resources. Thus, mining of top k itemsets is a tedious task in which k is the required number of itemsets. Based on the characteristics of the users, the threshold value is defined. This study concentrates on developing an efficient top k itemsets with defined min_util thresholds. Two algorithms, namely, mining Top k utility itemsets (TKU) and mining top k utility itemsets in one phase (TKO) is developed for defining min_util threshold value. The mining

performance is enhanced significantly since both the search space and the number of candidates are effectively reduced by the proposed strategies. In the experiments, different types of real datasets are used to evaluate the performance of our algorithm. The experimental results show that TKU outperforms the baseline algorithms substantially and the performance of TKU is close to the optimal case of the state-of-the-art utility mining algorithm.

As a future work, prior techniques have some challenging issues such as, large itemset database required more and more scan iterations which is time consuming task and degrades the efficiency and system performance. Scalability is the major issue as large number of itemsets has been generated during processing. Thus, an efficient solution is required for overcoming dynamic data challenges.

VIII. REFERENCES

- [1]. Vincent S. Tseng, Cheng-Wei Wu, Philippe Fournier-Viger, and Philip S. Yu, "Efficient Algorithms for Mining the Concise and Lossless Representation of High Utility Itemsets", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 27, No. 3, 2015.
- [2]. Chowdhury Farhan Ahmed, Syed Khairuzzaman Tanbeer, ByeongSoo Jeong, and Young-Koo Lee, "Efficient Tree Structures for High Utility Pattern Mining in Incremental Databases ", *IEEE Transactions on Knowledge and Data Engineering*, Vol.21, No 12, December 2009, pp 1708-1721.
- [3]. Vincent S. Tseng, Bai-En Shie, Cheng-Wei Wu, and Philip S. Yu, "Efficient Algorithms for Mining High Utility Itemsets from Transactional Databases", *IEEE Transactions on Knowledge and Data Engineering*, Vol.25, No. 8, AUGUST 2013, pp 1772-1786.
- [4]. Chun-Jung Chu, Vincent S. Tseng, Tyne Liang, "An efficient algorithm for mining high utility itemsets with negative item values in large databases", Elsevier, 2009. doi:10.1016/j.amc.2009.05.066.
- [5]. Hua-Fu Li, Hsin-Yun Huang, Suh-Yin Lee, "Fast and memory efficient mining of high-utility itemsets from data streams: with and without negative item profits", Springer, 2010. DOI 10.1007.
- [6]. Sen Su, Shengzhi Xu, Xiang Cheng, Zhengyi Li, and Fangchun Ya, "Differentially Private Frequent Itemset Mining via Transaction Splitting", *IEEE Transactions on Knowledge and Data Engineering*, Vol.27, No 7, July 2015
- [7]. Vincent S. Tseng, Cheng-Wei Wu, Viger, Philip S. Yu, "Efficient Algorithms for Mining Top-K High Utility Itemsets", *IEEE Transactions on Knowledge and Data Engineering*, DOI 10.1109/TKDE.2015.
- [8]. Alva Erwin, Raj P. Gopalan, and N. R. Achuthan, "Efficient Mining of High Utility Itemsets from Large Datasets", In *Proc. of PAKDD 2008*.
- [9]. Shankar, S.; Purusothaman, T.; Jayanthi, S. "Novel algorithm for mining high utility itemsets" *International Conference on Computing, Communication and Networking*, Dec. 2008.
- [10]. Raymond Chan; Qiang Yang; Yi-Dong Shen, "Mining high utility itemsets" In *Proc. of Third IEEE Int'l Conf. on Data Mining*, November 2003.
- [11]. Ramaraju, C., Savarimuthu N. "A conditional tree based novel algorithm for high utility itemset mining", *International Conference on Data mining*, June 2011.
- [12]. Ying Liu, Wei-keng Liao, Alok Choudhary "A Fast High Utility Itemsets Mining Algorithm" In *Proc. of the Utility-Based Data Mining Workshop*, 2005.
- [13]. Adinarayanareddy B ,O Srinivasa Rao, MHM Krishna Prasad, "An Improved UP-GrowthHigh Utility Itemset Mining" *International Journal of Computer Applications* (0975-8887) Volume 58-No.2, November 2012.

- [14]. P. Asha, Dr. T. Jebarajan, G. Saranya, "A Survey on Efficient Incremental Algorithm for Mining High Utility Itemsets in Distributed and Dynamic Database" *IJETAE Journal*, Vol.4, Issue 1, January 2014.
- [15]. L. Sweeney, "k-anonymity: A model for protecting privacy," *Int. J. Uncertainty Fuzziness Knows.-Base Syst.*, vol. 10, no. 5, pp. 557–570, 2002.
- [16]. Y. Bastide, R. Taouil, N. Pasquier, G. Stumme, and L. Lakhal. Mining frequent patterns with counting inference. *SIGKDD Explorations Newsletter*, 2(2):66–75, December 2000.
- [17]. J. Pei, J. Han, and R. Mao. Closet: An efficient algorithm for mining frequent closed itemsets. In *DMKD 00: ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery*, pages 21–30, May 2000.
- [18]. M. J. Zaki and C.-J. Hsiao. Charm: An efficient algorithm for closed itemset mining. In *SDM 02: Proceedings of the second SIAM International Conference on Data Mining*, April 2002.
- [19]. K. Gouda and M. J. Zaki. Genmax: An efficient algorithm for mining maximal frequent itemsets. *Data Mining and Knowledge Discovery*, 11(3):223–242, 2005.
- [20]. G. Grahne and J. Zhu. Efficiently using prefix-trees in mining frequent itemsets. In *FIMI 03: Proceedings of the ICDM 2003 Workshop on Frequent Itemset Mining Implementations*, November 2003.
- [21]. R. Agrawal and R. Srikant, "Fast algorithms for mining association rules," in *Proc. Int. Conf. Very Large Data Bases*, 1994, pp. 487–499.
- [22]. C. Ahmed, S. Tanbeer, B. Jeong, and Y. Lee, "Efficient tree structures for high-utility pattern mining in incremental databases," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 12, pp. 1708–1721, Dec. 2009.
- [23]. K. Chuang, J. Huang, and M. Chen, "Mining top-k frequent patterns in the presence of the memory constraint," *VLDB J.*, vol. 17, pp. 1321–1344, 2008.
- [24]. R. Chan, Q. Yang, and Y. Shen, "Mining high-utility itemsets," in *Proc. IEEE Int. Conf. Data Mining*, 2003, pp. 19–26.
- [25]. P. Fournier-Viger and V. S. Tseng, "Mining top-k sequential rules," in *Proc. Int. Conf. Adv. Data Mining Appl.*, 2011, pp. 180–194.

Distributed Intrusion Detection System For Cognitive Radio Networks Based on Weighted Fair Queuing Algorithm	S.Kavitha, Assistant Professor	Computer Science	International Journal of Scientific Research in Computer Science, Engineering and Information	March 2018	2456- 3307	http://ijsrcseit.com/CSEIT1833232 Google scholar UGC CARE Listed Journal No : 64718
--	--------------------------------------	---------------------	---	---------------	---------------	---



International Journal of Scientific Research in Computer Science, Engineering and Information Technology

© 2018 IJSRCSEIT | Volume 3 | Issue 4 | ISSN : 2456-3307

Distributed Intrusion Detection System for Cognitive Radio Networks Based on Weighted Fair Queuing Algorithm

¹M. Indhumathi, ²S. Kavitha

¹Research Scholar, Department of Computer Science, Sakthi College of Arts and Science for Women, Oddanchatram, India

²Head & Associate Professor, Department of Computer Science, Sakthi College of Arts and Science For Women, Oddanchatram, India

ABSTRACT

Reliable detection of intrusion is the basis of safety in cognitive radio networks (CRNs). So far, few scholars applied intrusion detection systems (IDS) to combat intrusion against CRNs. In order to improve the performance of intrusion detection in CRNs, a distributed intrusion detection scheme has been proposed. In this paper, a method base on Dempster-Shafer's (DS) evidence theory to detect intrusion in CRNs is put forward, in which the detection data and credibility of different local IDS Agent is combined by D-S in the cooperative detection center, so that different local detection decisions are taken into consideration in the final decision. The effectiveness of the proposed scheme is verified by simulation, and the results reflect a noticeable performance improvement between the proposed scheme and the traditional method.

Keywords : Safety, cognitive radio networks, intrusion detection, IDS Agent, cooperative detection center, Dempster-Shafer's evidence theory

I. INTRODUCTION

1.1 INTRODUCTION ABOUT TO IDS

1.1.1 IDS Defined

Intrusion detection is the process of identifying computing or network activity that is malicious or unauthorized. Most all Intrusion Detection Systems (IDS) have a similar structure and component set. This consists of a sensor (or agent) that monitors one or more data sources, applies some type of detection algorithm, and then initiates zero or more responses. Usually there is a management system that provides for monitoring, configuration and analysis of intrusion data.

1.1.2 Evolution of IDS

The first IDS were host-based, and looked at system operating logs performing simple pattern matches

against a small set of signatures. This approach quickly expanded to systems that looked at network traffic, initially also for simple patterns. As IDS gained a level of protocol-awareness, they were able to look for certain single packet traffic types known to be malicious, examining the source and destination IP addresses, along with source and destination ports. Further sophistication brought an awareness of network sessions and the ability to examine dialogs between systems for multi-packet activity. More recent IDS can examine and respond to entire conversations between hosts, using knowledge of protocols and network sessions to analyze traffic for malicious activity based on how that traffic would appear at the destination-a task often requiring specialized network drivers to operate at full wire-speed (For a good discussion of the evolution and genealogy of IDS, see article by Inella). The emerging class of IDS take this one step

further by combining log analysis, along with information from other IDS and anti-virus software to correlate events in an effort to identify and respond to intrusions in real time.

1.1.3 Relation of IDS to dIDS

From the above, it is clear that as IDS grow in function and evolve in power, they also evolve in complexity. Agents of each new generation of IDS use agents of the previous generation as data sources, applying ever more sophisticated detection algorithms to determine ever more targeted responses. Often, one or more IDS and management system(s) may be deployed by an organization within its own network, with little regard to their neighbors or the global Internet. Just as all individual networks and intranets connect to form "The Internet", so can information from stand-alone

A secure computer system provides guarantees regarding the confidentiality, integrity, and availability of its objects (such as data, purpose, or services). However, systems generally contain design and implementation flaws that result in security vulnerabilities. An intrusion can take place when an attacker or a group of attackers exploits the vulnerabilities and thus damages the confidentiality, integrity or availability guarantees of a system. Intrusion Detection Systems (IDSs) detect some set of intrusions and execute some predetermined actions when an intrusion is detected. Over the last one and half decade, research in the field of intrusion detection has been heading towards a distributed framework of systems that do local detection and provide information to perform global detection of intrusions.

These distributed frameworks of intrusion detection have some advantages over single monolithic frameworks. Most of these distributed systems are hierarchical in nature. The local intrusion detection components look for local intrusions and pass their analysis results to the upper levels of the hierarchy.

The components at the upper levels analyze the refined data from multiple lower level components and attempt to establish a global view of the system state. However, such IDSs are not fully distributed systems because of the centralized data analysis performed at the higher levels of the hierarchy. An agent-based architecture is proposed for performing intrusion detection in a distributed environment. By employing a suitable communication mechanism, the resource overhead is minimized in the distributed intrusion detection process.

1.2 INTRODUCTION TO DIDS

Some of the existing distributed IDS frameworks are discussed briefly. DIDS is a distributed intrusion detection system consisting of host managers and LAN managers doing distributed data monitoring, and sending notable events to the DIDS director. These managers also do some local detection, passing the summaries to the director. The director analyzes the events to determine the security state. AAFID is distributed IDS developed in CERIAS at Purdue University. It employs agents at the lowest level of the hierarchy for data collection and analysis and transceivers and monitors at the higher levels for controlling agents and obtaining a global view of activities. It provides a subscription-based service to the agents.

A prototype called the Hummingbird System is developed at University of Idaho. It is a distributed system that employs a set of Hummer agents, each assigned to a single host or a set of hosts. Each Hummer interacts with other hummers in the system through a manager, a subordinate, and the peer relationships. It enables a system administrator to monitor security threats on multiple computers. Architecture of an intrusion detection system using a collection of autonomous agents has been proposed in. In cooperation and communication model proposed by the authors, agents request and receive information solely on the basis of their interests. They can specify new interests as a result of a new

event or alert. This avoids unnecessary data flow among the agents.

However, most of these intrusion detection systems have the following drawbacks: (i) **Analysis hierarchy**: as there is a hierarchy in data analysis these systems are very difficult to modify. Changes may have to be made at many levels if any new distributed attack is developed. (ii) **Data refinement**: when a module from a lower level sends results of analysis to a higher level, some data refinement is done. However, the knowledge of what events are important in a system-wide level is difficult to anticipate at the lower levels of the hierarchy, and thus data refinement may result in loss of important information.

1.3 Wireless Sensor Networks Applications



Figure-1: Wireless Sensor Networks Applications

- (i) These networks are used in environmental tracking, such as forest detection, animal tracking, flood detection, forecasting and weather prediction, and also in commercial applications like seismic activities prediction and monitoring.
- (ii) Military applications, such as tracking and environment monitoring surveillance applications use these networks. The sensor nodes from sensor networks are dropped to the field of interest and are remotely controlled by a user. Enemy tracking, security detections are also performed by using these networks.
- (iii) Health applications, such as Tracking and monitoring of patients and doctors use these networks.

(iv) The most frequently used wireless sensor networks applications in the field of Transport systems such as monitoring of traffic, dynamic routing management and monitoring of parking lots, etc., use these networks.

(v) Rapid emergency response, industrial process monitoring, automated building climate control, ecosystem and habitat monitoring, civil structural health monitoring, etc., use these networks.

Wireless Sensor network (WSN) is a network system comprised of spatially distributed devices using wireless sensor nodes to monitor physical or environmental conditions, such as sound, temperature, and motion. Cluster-based data transmission in WSNs has been investigated by researchers to achieve the network scalability and management, which maximizes node lifetime and reduce bandwidth consumption by using local collaboration among sensor nodes. A CH aggregates the data collected by the leaf nodes (non-CH sensor nodes) in its cluster, and sends the aggregation to the base station (BS). To prevent quick energy consumption of the set of CHs, LEACH randomly rotates CHs among all sensor nodes in the network, in rounds. LEACH achieves improvements in terms of network lifetime. In this paper, for convenience, we call this sort of cluster-based protocols as LEACH-like proto-cols. Researchers have been widely studying CWSNs in the last decade in the literature.

The multi-constrained QoS routing is NP-hard and heuristic algorithms are proposed to find solution for the problem. But these algorithms are too complex and cannot obtain best global solution. QoS may be more accurately determined by using fuzzy logic instead of static values. Fuzzy Inference System (FIS) accepts more number of uncertain and imprecise data as inputs and thereby achieves flexibility, robustness, and low cost solution. But, FIS uses human-determined membership functions (MFs) that are fixed. Therefore, they are rarely optimal in

terms of reproducing the desired outputs. Tuning membership functions of parameters is a time consuming task. Neural networks overcome most of the complex problems to adapt dynamically to the system operating conditions, and to make correct decisions, if the signals are uncertain. But the integration of neural network into the fuzzy logic system makes it possible to learn from prior obtained data sets. This paper proposes an approach which integrates both neural and fuzzy techniques to select a server from a number of group members belonging to any cast group by considering QoS constraint route and server with higher stability in MANETs. This section presents some of the related works, software agent concept and our contributions.

The set of destinations is identified by unique any cast address and provide the same services. Searching for services on networks often depends on the broadcast or multicast mechanism to acquire the information, which usually results in large overhead. It will be a serious problem in ad-hoc wireless networks, where the bandwidth is limited and each node moves arbitrarily. Any casting scheme in ad hoc wireless networks can simplify access management in distributed service, improve the robustness and performance of an ad hoc network when mobility and link disconnections are frequent, and reduces the communication overhead.

The source node does not need to know about picking a single server and is determined by routing scheme. The server in any cast routing may be chosen by minimum hops, delay or other metrics. Any casting along the minimum hops path may result in inefficient use of network resources, because it forwards packets along already congested shortest path, and also may not satisfy the Quality of Service (QoS) requirements for multimedia and real time application services. Set of mobile or semi mobile nodes with no available pre-established communications is a MANET Forming a short-term network. Each mobile node in the network acts as a

computer switching program that transfers incoming messages to outgoing links via the most efficient route possible, e.g. over the Internet i.e., a router. This kind of networks are characterize by the relationships between parts linked together in a system such as a computer network topologies, continuation of bandwidth constrain and variable capacity links, energy constrain operations and are highly intensity to security threats. Due to all these characteristics routing is a major issue in ad hoc networks. The routing protocols for ad hoc networks have been classified as: (a) Proactive or table driven for example Destination Sequenced Distance Vector (DSDV) and Optimized Link State Routing (OLSR), (b) Reactive/On-demand, e.g. Dynamic Source Routing Protocol, Ad hoc On-Demand Distance Vector routing protocol, Temporally Ordered Routing Algorithm. In table driven or proactive routing, each node has one or more tables that include the latest information of the routes to any node in the network. Each row has the subsequently hop for reaching to a node or subnet and the cost of this route. Different table-driven protocols vary in the way the information about alter in topology is spread through all nodes in the network. The two kinds of table keep informed in proactive protocols are the periodic update and the triggered update.

1.4 COGNITIVE RADIO NETWORK

Cognitive radio (CR) is a form of wireless communication in which a transceiver can intelligently detect which communication channels are in use and which are not, and instantly move into vacant channels while avoiding occupied ones. This optimizes the use of available radio-frequency (RF) spectrum while minimizing interference to other users.

In its most basic form, CR is a hybrid technology involving software defined radio (SDR) as applied to spread spectrum communications. Possible functions of cognitive radio include the ability of a transceiver to determine its geographic location,

identify and authorize its user, encrypt or decrypt signals, sense neighboring wireless devices in operation, and adjust output power and modulation characteristics.

II. LITERATURE REVIEW

Jaydip Sen, A Survey on Wireless Sensor Network Security [1] Wireless sensor networks (WSNs) have recently attracted a lot of interest in the research community due their wide range of applications. Due to distributed nature of these networks and their deployment in remote areas, these networks are vulnerable to numerous security threats that can adversely affect their proper functioning. This problem is more critical if the network is deployed for some mission-critical applications such as in a tactical battlefield. Random failure of nodes is also very likely in real-life deployment scenarios. Due to resource constraints in the sensor nodes, traditional security mechanisms with large overhead of computation and communication are infeasible in WSNs. Security in sensor networks is, therefore, a particularly challenging task. This paper discusses the current state of the art in security mechanisms for WSNs. various types of attacks are discussed and their countermeasures presented. A brief discussion on the future direction of research in WSN security is also included. In addition to traditional security issues like secure routing and secure data aggregation, security mechanisms Deployed in WSNs also should involve collaborations among the nodes due to the decentralized nature of the networks and absence of any infrastructure. In real-world WSNs, the nodes cannot be assumed to be trustworthy apriori. Researchers have therefore, focused on building a sensor trust model to solve the problems which are beyond the capabilities of traditional cryptographic mechanisms. In this chapter, we present a survey of the security issues in WSNs. First we outline the constraints of WSNs, security requirements in these networks, and various possible attacks and the corresponding countermeasures. Then a holistic view

of the security issues is presented. These issues are classified into six categories: cryptography, key management, secure routing, secure data aggregation, intrusion detection and trust management. The advantages and disadvantages of various security protocols are discussed, compared and evaluated. Some open research issues in each of these areas are also discussed.

Efficient and Secure Routing Protocol for Wireless Sensor Networks through SNR based Dynamic Clustering Mechanisms [2] Advances in Wireless Sensor Network Technology (WSN) have provided the availability of small and low-cost sensor with capability of sensing various types of physical and environmental conditions, data processing and wireless communication. In WSN, the sensor nodes have a limited transmission range, and their processing and storage capabilities as well as their energy resources are limited. Triple Umpiring System (TUS) has already been proved its better performance on Wireless Sensor Networks. Clustering technique provides an effective way to prolong the lifetime of WSN. In this paper, we modified the Ad hoc on demand Distance Vector Routing (AODV) by incorporating Signal to Noise Ratio (SNR) based dynamic clustering. The proposed scheme Efficient and Secure Routing Protocol for Wireless Sensor Networks through SNR based dynamic Clustering mechanisms (ESRPSDC) can partition the nodes into clusters and select the Cluster Head (CH) among the nodes based on the energy and Non Cluster Head (NCH) nodes join with a specific CH based on SNR Values. Error recovery has been implemented during Inter cluster routing itself in order to avoid end-to-end error recovery. Security has been achieved by isolating the malicious nodes using sink based routing pattern analysis. Extensive investigation studies using Global Mobile Simulator (GloMoSim) showed that this Hybrid ESRP significantly improves the Energy efficiency and Packet Reception Rate (PRR) compared to SNR unaware routing algorithms like Low Energy Aware Adaptive Clustering Hierarchy (LEACH) and Power-Efficient Gathering

in Sensor Information Systems (PEGASIS). Sensor Network Wireless is widely considered as one of the most important technologies for the twenty-first century. The sensing electronics measure ambient conditions related to the environment surrounding the sensors and transform them in to an electrical signal. In many WSN applications, the deployment of sensor

Node informing that the misuse IDS system is operational. The messages sent to the Central IDS Node are formatted using the extended signed IDMEF format. In addition, the upper tier process listens for commands from the Central IDS Node. It receives parameters for the rate limiting of alert messages, configuration for the Snort process and new attack signatures.

III. METHODOLOGY

(i) Security Agents

3.1. Misuse Detection Agent

As we previously mentioned, each security agent consists of two tiers. The lower tier comprises of the process that handles the misuse detection within our network. Snort [6] has been chosen as the misuse IDS software for our system. Snort is a libpcap-based [7] software that can be used as a sniffer, packet logger or network intrusion detection system. In our case, we used Snort as a misuse intrusion detection tool. The detection of malicious packets is based on known attack signatures. Snort is able to detect a variety of attacks such as DoS/DDoS attacks, Portscans, HTTP, DNS, SMTP, IMAP, POP3 attacks and Virus/Worm attacks.

Alerts generated from Snort are passed to the upper tier of our agent. The upper tier of the Misuse Detection Agent receives alert messages from the lower tier and stores them for a defined period of time in a buffer. For every different case of attack, that is, source IP address and port, target IP address and port and known attack signature, the upper tier process uses a unique alert identification. Rate limiting is achieved independently for different types of attacks, sending the alert message only once in the specified period of time.

Agent's upper tier process is also responsible for sending the heartbeat messages to the Central IDS

3.2. Anomaly Detection Agent

For the lower tier of the Anomaly Detection Agent we developed a prototype anomaly detection tool [8] that currently focuses on DoS Attacks. The prototype tool consists of two main modules: the collector and the detector. The collector module is responsible for asynchronously receiving flow data from the Netflow-enabled [9] router; information is analyzed, mean values and adaptive thresholds are calculated and stored in a local data structure.

The tool extracts and stores packet and flow counters per destination IP address, as well as total counters and mean values for each pair of input-output interfaces. The detector process is responsible for calculating the metrics for the interface pairs stored by the collector, and comparing the results to detection thresholds. It is periodically activated, implements extensive logging of detection events and generates notifications with security alerts which are sent to the upper tier.

The upper tier process receives the alerts and sends them to the Central IDS Node using the signedIDMEF Format. Moreover, the Central IDS Node adjusts Anomaly Detection Agent's parameters (metrics and thresholds for the DoS attack detection algorithm).

3.3. SNMP Query Agent

As the other two agents, the SNMP Query Agent is comprised of two tiers. The lower tier process is a

custom SNMP client that performs SNMP queries at the routers of the network. Values like CPU and memory usage, active and inactive flows are polled from routers at specific intervals. The upper tier accepts the values from the SNMP queries and forwards them to the Central IDS Node after formatting them using the signed-IDMEF data model. The upper tier process is also responsible for sending heartbeat messages to inform the Central IDS Node that the SNMP client is operational. Instructions from the Central IDS Node are sent to the SNMP Query Agent, giving information about the router and the SNMP objects to be polled.

(ii) Intrusion Detection System

Intrusion detection mechanism can detect malicious behavior on the network and identify malicious users. So Intrusion detection mechanism can protect the reliability of the network, especially it is more important in distributed cognitive radio network which absents center facilities. The traditional intrusion detection system (IDS) was proposed by Denning in 1987. It is composed of main body, object, audit record, activity profile and exception record and activity rules. A more detailed description of IDES is given as follows.

There are six main parts in the IDS model [12].

1) Subject: Active initiator in the system operation, the entity that moves on the target system, such as the process of the computer operating system, the service connection of the network and so on.

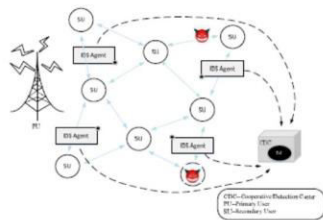


Figure-2:. System of Distribute Intrusion Detection in CRNs

2) Object: Resources that are managed by the system- files, devices, commands, for example.

3) Audit Records: when a subject operates on an object, such as user registration, command execution,

and file access, data will be produced by the target system.

4) Activity Profile: Preserve the information about the normal activity, and the specific implementation depends on the detection method.

5) Anomaly Record: Used to indicate the occurrence of an abnormal event. The format is <Event, Time-Stamp, Profile>

6) Activity rule: An action that should be taken when an audit record, update profile, an exception record, detect relate anomalies to some suspected intrusions or abnormal behavior is produced.

Actually, the Denning model is can be described as a rule based pattern matching system. After generating an audit record, it will match against the profiles. Then type information will determine the rule to report anomalies detection. It's largely system-independent for the rules and profile structures to establish profile templates. Not all of the IDS can be fully consistent with the model.

IV. CONGESTION CONTROL FOR WIRELESS SENSOR NETWORK

4.1 Congestion Control

The flow type is of high importance to guide a real congestion control. Flow types may include a single packet, few packets, a large number of packets, which require light control, medium level control, and tight control, respectively. When a large number of nodes transmit information, their flows will cross at intermediate nodes. This high number of sources increases the congestion but helps improving the reliability. For example, in tree architectures, every intermediate node can suffer from congestion causing packet loss, which in turn decreases network performance and throughput and cause energy waste. It is very difficult to predict the intersection points due to network dynamics (addition or removal of sensors or a change in the report rate), variability in radio channel quality over time. All these can

transform uncongested parts of the network to under-provisioned and congested regions. The area around the intersection will become a hotspot and there is a possibility of congestion (buffer overflow) and contention (links interference). For these reasons, a congestion control algorithm for data packet transmission is necessary.

Contention-based Congestion: when many nodes within range of one another attempt to transmit simultaneously, losses occur due to interference and packet loss is engendered. If the packet generation rate is sufficiently small, simultaneous transmission becomes independent of the rate. Rather, it depends on the exact time generation of the packet. Explicit local synchronization (or also named phase shifting) among neighbours can reduce this type of loss, but it cannot eliminate the problem as non-neighbouring nodes can still interfere (hidden nodes). The contention may happen between different flows in the same area, and between different packets of the same flow, especially in the case of high density networks. Consequently, the nodes' channel capacity becomes time-variant. **Buffer-based Congestion:** each node uses a buffer for the packets waiting to be sent. The overflow of this buffer causes congestion and packets loss. This is due to high reporting rate that varies in time due to dynamic channel conditions. The many-to-one nature (or converge cast) of WSNs causes congestion, in addition to the other causes shared with general wireless networks.

4.2 Congestion Detection Strategies

Many congestion detection mechanisms are used and tested. The most used are: packet loss, queue length, packet service time, the ratio between packet service time and packet inter-arrival time, delay. In many cases, a single parameter cannot indicate congestion accurately.

Packet loss: It can be measured at the sender if ACKs (Acknowledgements) are used; this suggests reliability to be ensured by the protocol. It can also

be measured at the receiver with sequence numbers use. Further, CTS (Clear To Send) packet loss can be used as congestion indication.. Not overhearing the parent's forwarding on the upstream link, by a child node over the downstream link, can be used as an indication for packet loss as well. The time to repair losses (if reliability ensured) can be used as a congestion indication. Loss ratio is also used in some protocols.

Queue length: As each node has a buffer; its length can serve a simple and good indication of congestion a fixed threshold is used and the congestion is signalled as soon as the buffer length exceeds this threshold, the remaining buffer length from the overall size is used. In the difference between the remaining buffer and the traffic rate is used as congestion indication. The traffic rate represents the excess rate, which is the difference between the output rate and the sum of sourced and forwarded rates. In the buffer length is used in addition to the difference of output and input time, which is quite similar to output and input rate. In buffer length and capacity of the node are used together. The number of non-empty queues can indicate congestion level. When there is a congestion, this number is larger than 0. This number increases with network load. If the link layer applies retransmissions, link contention will be reflected through buffer length.

Queue length and Channel load: In case of increase in packets collision, and after several unsuccessful MAC (Medium Access Control) retransmissions, packets are removed. Consequently, the decrease in buffer occupancy due to these drops may mean the absence of congestion when only buffer state is used for congestion detection. Therefore, for accurate congestion detection, a hybrid approach is required using queue length and channel loading as a congestion indication. Channel busyness ratio or channel load is the ratio of time intervals when the channel is busy (successful transmission or collision) to the total time. In the authors use the busyness channel ratio, similarly to channel load, but apply it

to a subset of nodes, and queue length for another set of nodes. The node activates channel monitoring only when it receives a packet to forward. Therefore, there is no overhead to measure channel loading.

4.3 Channel Busyness Ratio and Throughput Measurement

Throughput is addition to channel busyness to take into account the effects of hidden nodes problem in multi-hop environment. The throughput quantifies the number of successful transmissions.

Packet service time: The inverse of packet service rate, it is the interval between packet arrival at the MAC layer and its successful transmission. It covers packet waiting, collision resolution, and packet transmission time at the MAC layer.

The congestion control cannot be decoupled from the MAC protocol, and adequate protocol should first be used to avoid congestion. In applications where the event cannot be known a priori, random access contention based MAC protocols are necessary (CSMA "Carrier Sense Multiple Access"-based). Continuous periodic applications with high rate a TDMA "Time Division Multiple Access"-like scheme is more appropriate.

V. RESULTS AND DISCUSSION

5.1 EXPERIMENTAL SETUP

Congestion in a network may occur if the load on the network the number of packets sent to the network is greater than the capacity of the network the number of packets a network can handle. Congestion control refers to the mechanisms and techniques to control the congestion. The congestion control having a different type of models but it have some disadvantage .To overcome the drawbacks, we proposed evolutionary algorithm, ant colony algorithm to get the optimal solution for the congestion control. In order to avoid congestion delays, the ant colony optimization paradigm is

explored to find a optimize routes and to proposed routing algorithms are simple yet efficient. The routing optimization is driven by the minimization of total latency during packets transmission between the tasks.

5.2 DYNAMIC SOURCE ROUTING (DSR)

Genetic algorithms are a part of evolutionary computing. It is also an efficient search method that has been used for path selection in networks. These stochastic search algorithms are based on the principle of natural selection and recombination. GA has been an efficient search method based on principles of natural selection and genetics. They are being applied successfully to find acceptable solutions to problems in business, engineering, and science.

We can find good solution for adequate amount of data at hand, but the complexity of data increases as GA takes time to find the solution. GA works well for network model to find the optimal path. In this, the source and the destination nodes are sure to participate in every generation. Other nodes or the genes become a part of the chromosome if they find an optimal path between the source and destination. GA is composed with a set of solutions, which represents the chromosomes. This composed set is referred to population. Population consists of set of chromosome which is assumed to give solutions. From this population, we randomly choose the first generation from which solutions are obtained. These solutions become a part of the next generation. Within the population, the chromosomes are tested to see whether they give a valid solution. This testing operation is nothing but the fitness functions which are applied on the chromosome. Operations like selection, crossover and mutation are applied on the selected chromosome to obtain the progeny. Again fitness function is applied to these progeny to test for its fitness. Most fit progeny chromosome will be the participants in the next generation. The disadvantage of this protocol is that the route maintenance

mechanism does not locally repair a broken link. Stale route cache information could also result in inconsistencies during the route reconstruction phase. The connection setup delay is higher than in table-driven protocols. Even though the protocol performs well in static and low-mobility environments, the performance degrades rapidly with increasing mobility. Also, considerable routing overhead is involved due to the source-routing mechanism employed in GAs. This routing overhead is directly proportional to the path length.

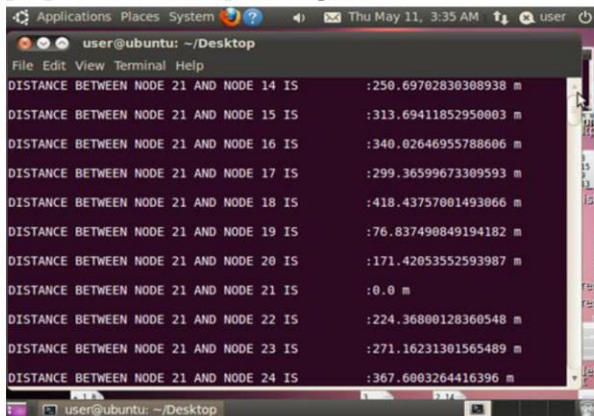


Figure-3: Calculate the Distance between all the Nodes

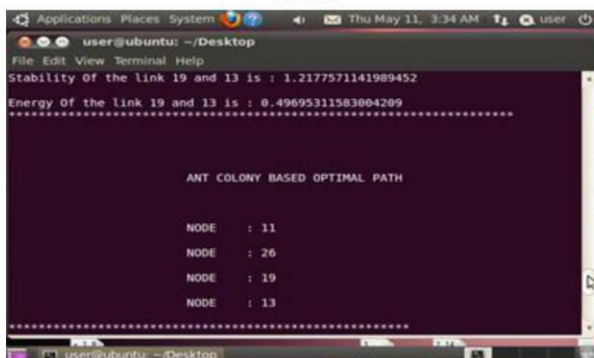


Figure-4: After the Selection of Source and Destination Calculate the Distance between their Neighbouring Nodes

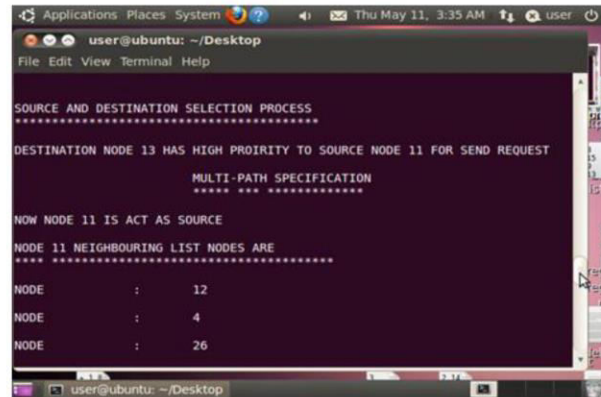


Figure-5: List out the Neighbouring Node

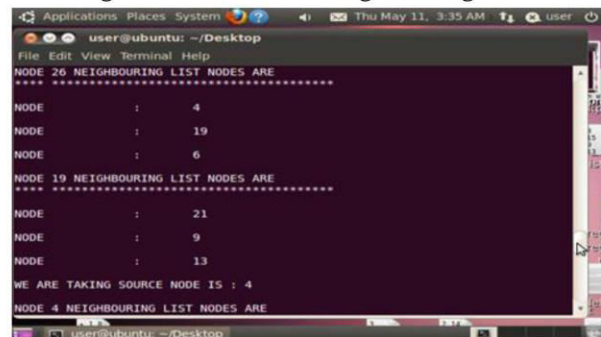


Figure-6: Finally Display the Optimal Path between Source and Destination

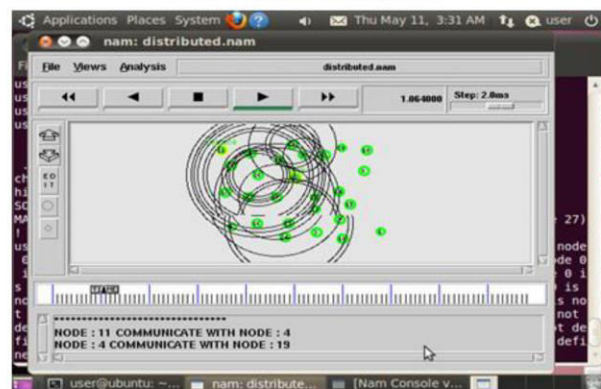


Figure-7: Neighbouring Path between the Source and Destination

VI. CONCLUSION

I propose distributed IDS for CRNs based on evidence theory. Aim to get more accurate final detection result making at CDC, we apply D-S theory of evidence to combine different detection data and credibility from every IDS Agents. Simulations presented show that the proposed system performs more excellent than the traditional Weighted Fair Queuing (WFQ) Combination algorithm.

VII. FUTURE ENHANCEMENT

In my future work, I would like to work on any cast routing protocols to check the efficiency under high throughput applications, e.g. multimedia applications by employing negotiation parameters in route request packet in finding nearest server through non congestion paths.

VIII. REFERENCES

- [1]. T. Hara, V.I. Zadorozhny, and E. Buchmann, *Wireless Sensor Network Technologies for the Information Explosion Era, Studies in Computational Intelligence*, vol. 278. Springer-Verlag, 2010.
- [2]. Y. Wang, G. Attebury, and B. Ramamurthy, "A Survey of Security Issues in Wireless Sensor Networks," *IEEE Comm. Surveys & Tutorials*, vol. 8, no. 2, pp. 2-23, Second Quarter 2006.
- [3]. A.A. Abbasi and M. Younis, "A Survey on Clustering Algorithms for Wireless Sensor Networks," *Computer Comm.*, vol. 30, nos. 14/15, pp. 2826-2841, 2007.
- [4]. W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "An Application-Specific Protocol Architecture for Wireless Micro sensor Networks," *IEEE Trans. Wireless Comm.*, vol. 1, no. 4, pp. 660-670, Oct. 2002.
- [5]. A. Manjeshwar, Q.-A. Zeng, and D.P. Agrawal, "An Analytical Model for Information Retrieval in Wireless Sensor Networks Using Enhanced APTEEN Protocol," *IEEE Trans. Parallel & Distributed Systems*, vol. 13, no. 12, pp. 1290-1302, Dec. 2002.
- [6]. S. Yi et al., "PEACH: Power-Efficient and Adaptive Clustering Hierarchy Protocol for Wireless Sensor Networks," *Computer Comm.*, vol. 30, nos. 14/15, pp. 2842-2852, 2007.
- [7]. K. Pradeepa, W.R. Anne, and S. Duraisamy, "Design and Implementation Issues of Clustering in Wireless Sensor Networks," *Int'l J. Computer Applications*, vol. 47, no. 11, pp. 23-28, 2012.
- [8]. L.B. Oliveira et al., "Sec LEACH-On the Security of Clustered Sensor Networks," *Signal Processing*, vol. 87, pp. 2882-2895, 2007.
- [9]. K. Zhang, C. Wang, and C. Wang, "A Secure Routing Protocol for Cluster-Based Wireless Sensor Networks Using Group Key Management," *Proc. Fourth Int'l Conf. Wireless Comm., Networking and Mobile Computing (WiCOM)*, pp. 1-5, 2008.
- [10]. P. Banerjee, D. Jacobson, and S. Lahiri, "Security and Performance Analysis of a Secure Clustering Protocol for Sensor Networks," *Proc. IEEE Sixth Int'l Symp. Network Computing and Applications (NCA)*, pp. 145-152, 2007.
- [11]. K. Zhang, C. Wang, and C. Wang, "A Secure Routing Protocol for Cluster-Based Wireless Sensor Networks Using Group Key Management," *Proc. Fourth Int'l Conf. Wireless Comm., Networking and Mobile Computing (WiCOM)*, pp. 1-5, 2008.

A Novel Approach for Fingerprint Liveness Detection Using Gradient and Texture Features	R. Madhumathi Assistant Professor	Computer Science	International Journal of Scientific Research in Computer Science, Engineering and Information Technology	2018	2456 - 3307	UGC Journal No : 64718 https://ijsrcseit.com/paper/CSEIT1833195.pdf
---	--------------------------------------	------------------	--	------	-------------	---



A Novel Approach for Fingerprint Liveness Detection Using Gradient and Texture Features

P. Shanthi¹, R. Madhumathi²

¹Research Scholar, Department of Computer Science, Sakthi college of Arts and Science For Women, oddanchatram, Tamil Nadu, India

²Assistant Professor, Department of Computer Science, Sakthi college of Arts and Science For Women, oddanchatram, Tamil Nadu, India

ABSTRACT

Fingerprints are good basis for individual identification by biometric authentication. Password based authentication systems are less secure than that of the fingerprint authentication where fingerprints and Iris are unique for every Individual. With the emerging use of biometric authentication systems in the past years, spoof fingerprint detection has become increasingly important. In this paper, we propose a static software approach that combines all sorts of fingerprint features. Initially, we extract the features of the fingerprint image using Gabor wavelet feature process. The extracted features are then aligned with histogram process. Each extracted features are preserved with dynamic score level integration. This dynamic approach consumes higher computational time. It has been experimented on the LivDet 2011 dataset which proves the efficiency of our proposed system. These have shown the classification rate of 9.625% with reduced error rate of 2.27%.

Keywords: Fingerprint liveness, low level features, Gabor filters, texture analysis, Biometric Security.

I. INTRODUCTION

Biometrics is burlier authentication system in the domain of security. Fingerprints are intrinsic to persons and can neither be lost nor stolen which makes it highly truthful and trustworthy. Furthermore, the accessibility of low-cost fingerprint readers united with easy integration capabilities has led to the broad spread use of fingerprint biometrics in a diversity of organizations. An organization can have unlimited benefits by appropriately deploying biometric technology. Today's economy is a developing one and technological progressions have altered the system in which organizations function and conduct businesses. Recent organizations require being adaptive, flexible and responsive to endure in the competitive business surroundings. Fingerprint technology can promote organizations in a diversity

of segments e.g. health care, government, retail enterprises, technology organizations, manufacturing industry, libraries, universities etc Employee identification and workforce management becomes faster, exact and more proficient with fingerprint technology [1]. Different magnetic strip cards or passwords, individuals constantly carry their fingerprints with them and they cannot be misplaced or elapsed. Tracking attendance of employees in industrialized organizations checks employee time thievery and diminish deceptive behavior. A biometric system facilitate automated calculation of employee hours therefore sinking paper expenditure and time exhausted in manual settlement of attendance data.

Fingerprint biometrics can give both physical access to company buildings and logical access to internal

resources such as enterprise computers and systems. Governments and private organizations, institutes everywhere in the world are opting biometric technology to contest identity fraud and security breaches, secure confidential data, and reduce costs and to develop overall user understanding. Biometrics is one of the quickly emerging eras in the information technology segment with fingerprint recognition anticipated to stay put the most leading form of biometric technology [2]. Fingerprint liveness detection has been a vigorous research area in excess of the previous several years. It has been confirmed that it is achievable to spoof standard optical and capacitive sensors. The possibility to spoof a fingerprint based authentication system generates the necessity to grow a method which can differentiate between live and fake fingerprint images.

Biometric technology presents numerous advantages over classical security methods based on moreover some information (PIN, Password, etc.) or physical devices e.g. key, card, etc. Though, providing to the sensor a fake physical biometric can be a simple mode to overhaul the system's security. Fingerprints, in particular, can be simply spoofed from ordinary resources, such as gelatin, silicone, and wood glue. Consequently, a protected fingerprint system should discriminate properly a spoof from an authentic finger. Several fingerprint liveness detection algorithms have been developed, and they can be widely divided into two approaches: Hardware and Software approach [3]. In the hardware approach a particular device is added to the sensor sequentially to detect exacting properties of living aspects such as the blood pressure, skin distortion or the odor. In the software approach, which is used in this work, fake characters are detected once the sample has been obtained with a standard sensor. Additionally, hardware based approaches are usually more costly due to the added sensors essential; next to, they need an end user to cooperate with the extra hardware. Alternatively, software based approaches do not utilize extra persistent biometric dimensions.

However, these approaches are more demanding as they need the identification of distinguishable features to discriminate between live and fake fingerprint images.

Software based approaches are additionally separated generally into dynamic and static based approaches. Dynamic software based approaches necessitate a minimum of two time series images ensuing in added computational time. So, the hardware approach would be costly but less secure and software approach is complicated to construct an algorithm which can distinguish different features between fake and live fingerprint. Software algorithm is a demanding approach and important method for fingerprint liveness detection. Because of human mistake fraud cases have turn into common and the fake fingerprint that has been prepared by the gelatin or latex will have more strength of edges in compare of live fingerprint which in some way acquires simple access to authentication systems. To decrease the limitations of the software based fingerprint authentication, we developed a static software approach in which the algorithm extracts features which are exclusive for each and every person. The extracted features from a fingerprint image can be specified as SURF, PHOG and gabor wavelet. These features are resolute to distinguish between fake and live fingerprints. The Gabor wavelets demonstrate optimal properties in both frequency and spatial domain which consecutively diminish the human based errors in the authentication systems [4].

II. STATEMENT OF THE PROBLEM

In this research, a method is proposed to overcome the restrictions faced in the static software based approaches where a single feature set unsuccessful to execute uniformly in excess of dissimilar fingerprint sensors and materials. This methodology extracts low level textural and gradient information for fingerprint liveness detection from a single image. It proposes the use of SURF features in amalgamation with PHOG to acquire gradient features that

distinguish well between fake and live fingerprint images. SURF features have a brief descriptor length which is dense and consumes less computational time in compare to LBP. Additionally, SURF is also invariant to scale and image rotation. PHOG feature descriptor extracts intensity gradient and edge directions to explain the shape and manifestation in an image. The PHOG extractor is also invariant to geometric and photometric transformation. Therefore, grouping of SURF and PHOG facilitate this method to execute likewise over a variety of sensors and materials.

Consecutively to acquire textural features, we suggest the use of Gabor wavelets as they have optimal localization properties in both the frequency and spatial domain. They extract discriminative ridge feature maps and have performed well in discerning between live and fake fingerprint images.

- ✓ To the best of our knowledge, the suggested method is one of the only some work that executes well over a large open source dataset generated using six dissimilar sensors and six dissimilar materials. In this work, we examine the use of local distinguishable feature space on live and spoof fingerprints by using PHOG, SURF, GABOR and their amalgamation.
- ✓ Experiments executed on six sensors express that the amalgamation of PHOG and SURF always works better than PHOG and SURF individually for LivDet 2011 and 2013 databases. This specifies that these descriptors accompaniment each other. Also, the amalgamation of PHOG and SURF feature vector generates a strong distinguishable feature vector which executes extremely well in the area of fingerprint liveness detection.
- ✓ Unlike, LivDet 2013 competition winner and other top four teams which do not execute well on Crossmatch sensor, this method executes extremely well on Crossmatch sensor generating an average classification error of

2.5% in compare of 31.20% obtained in LivDet 2013 fingerprint competition.

- ✓ The proposed method is entirely software based and it is computationally not expensive, rapid and flexible for future adaptations. This method can be organized in real-time applications. At last, the outcome accomplished by this method does better the state of the art appreciably.

III. SCOPE OF THE RESEARCH

Biometric sensors are broadly employed to distinguish between particular ones that are allowed to involve in an activity and individuals that are not allowed to involve in that activity. For e.g., fingerprint sensors are generally utilized to find out whether a fingerprint provided by an individual matches information in a database, and if a match is find out, then the individual may be allowed to involve in an activity. For e.g. the individual may be permitted to go into a building or room or permissible to utilize an electronic device like as a mobile phone or an application running on a mobile device. Biometric sensors can be mislead and thus authorize an illegal individual to employ in an activity that is kept back for legal individuals. Spoofing a fingerprint sensor may be achieved in dissimilar ways. These consist of via a fake fingerprint, with body parts other than a finger, and by means of a dead finger from a person. While it is improbable that the exacting type of spoofing to be utilized on a fingerprint sensor will be recognized in proceed, it is significant to guard alongside all types of spoofs. As increasingly biometrics is utilized for user identification and/or verification, liveness detection becomes gradually more vital in turn to make sure admission security and correctness. Liveness detection is significant since a lot of methods of misleading an identification system and/or verification system make use of spoofs that are not alive. For e.g. a latex finger may be made to have ridges and valleys like a fingerprint of a legal user.

IV. METHODOLOGY

Image Acquisition: Image acquisition in image processing can be widely defined as the action of retrieving an image from a few sources, generally a hardware-based source, thus it can be accepted during whatever processes require to come about later. Performing image acquisition in image processing is all the time, the primary step in the workflow sequence because, exclusive of an image, no processing is achievable. The image that is attained is entirely unprocessed and is the result of whatever hardware was used to produce it, which can be very significant in some areas to have a reliable baseline from which to work.

Preprocessing: The objective of pre-processing is an enhancement of the image data that contains unnecessary distortions or improves some image features significant for additional processing. We improved the quality of the image by first cropping the fingerprint region in the image and median filtering is afterward applied on the cropped images devoid of diminishing the sharpness of the input image. To end with, histogram equalization is carried out to advance the compare of the image by expanding the intensity range over the entire cropped image. The output achieved after this stage is an image with a condensed noise and enhanced description of the ridge structure.

Feature Extraction: In fingerprint authentication systems, the image is generally captured from various subjects by using the dissimilar scanners. Hence, fingerprint images are usually obtained to be of dissimilar scales and rotations. In definite circumstances, the fingerprint images are partly captured caused by human errors. Sequentially to acquire features that are invariant to these troubles, various features use which capture properties of live fingerprint images. In this work, we decide to employ SURF as it is invariant to enlightenment, scale and rotation. SURF is also utilized because of its brief descriptor length. Although SURF is invariant

to object orientation and scale transformation, it is not invariant to geometric transformations. Therefore, sequentially to recompense the restrictions of SURF, PHOG descriptors are used to extract local shape information to achieve more distinguishable features. Additionally, Gabor wavelet features are also integrated for texture analysis.

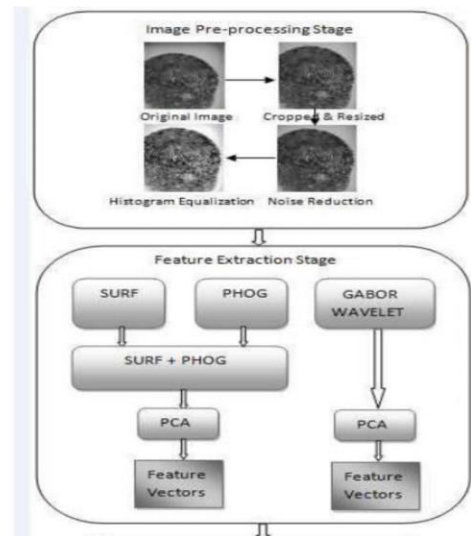


Figure – 1: : System architecture of the proposed Method

Feature Reduction using PCA: Extreme features increase computation times and storage memory. Moreover, from time to time they make classification more difficult that is called the curse of dimensionality. It is necessary to decrease the number of features. PCA is an efficient tool to diminish the measurement of a data set comprising of a large number of consistent variables although keeping most of the variations. It is accomplished by transforming the data set to a novel set of prearranged variables according to their variances or importance.

Classification: The classification procedure is done over the extracted features. Here, main innovation is the acceptance of SVM and Random Forest. RF and SVM classifier is applied over the features and the classification is done.

V. EXPERIMENTAL RESULTS

5.1 Importing and Exporting Images

Image Processing Toolbox chains images produced by a broad range of devices, containing digital cameras, satellite and airborne sensors, medical imaging devices, microscopes, telescopes, and other scientific instruments. It may visualize, analyze, and process these images in various data types, containing single- and double-precision floating-point and signed and unsigned 8-, 16-, and 32-bit integers. There are numerous modes to import and export images into and out of the MATLAB background for processing. Image Acquisition Toolbox can be used to obtain live images from Web cameras, frame grabbers, DCAM-compatible cameras, and other devices. Using Database Toolbox, images can be accessed which are stored in ODBC/JDBC-compliant databases.

5.2 Displaying and Exploring Images

Image Processing Toolbox expands MATLAB graphics to offer image display capabilities which are extremely customizable. It can construct displays with multiple images in a single window, interpret displays with text and graphics, and create specialized displays e.g. histograms, profiles, and contour plots. Additionally, to display functions, the toolbox provides a suite of interactive tools for exploring images and building GUIs.

5.3 Preprocessing and Post Processing Images

Image Processing Toolbox supports reference-standard algorithms for preprocessing and post-processing responsibilities that resolve frequent system problems, e.g. interfering noise, low dynamic range, out-of-focus optics, and the dissimilarity in color demonstration between input and output devices.

Image enhancement techniques in Image Processing Toolbox assist to improve the signal-to-noise ratio and accentuate image features by altering the colors or intensities of an image. It can:

- ✓ Perform histogram equalization

- ✓ Perform decorrelation stretching
- ✓ Remap the dynamic range
- ✓ Adjust the gamma value
- ✓ Perform linear, median, or adaptive filtering

5.4 Analyzing Images

Image Processing Toolbox gives a widespread collection of reference-standard algorithms and graphical tools for image analysis tasks e.g. statistical analysis, feature extraction, and property measurement.

Statistical functions analyze the common characteristics of an image by:

- ✓ Computing the mean or standard deviation
- ✓ Determining the intensity values along a line segment
- ✓ Displaying an image histogram
- ✓ Plotting a profile of intensity value

Edge-detection algorithms identify object boundaries in an image. These algorithms contain the Sobel, Prewitt, Roberts, Canny, and Laplacian of Gaussian methods. The dominant Canny method can detect true weak edges without being "fooled" by noise.

5.5 Working with Large Images

Few images are outsized that they are complicated to process and display with standard methods. Image Processing Toolbox offers exact workflows for working with larger images than or else possible. Devoid of loading a large image completely into memory, can create a reduced-resolution data set (R-Set) that partitioned an image into spatial tiles and resample the image at dissimilar resolution levels. This workflow develops performance in image display and navigation. A block processing workflow can be used to apply a function to each distinct block of a large image that considerably reduces use of memory.

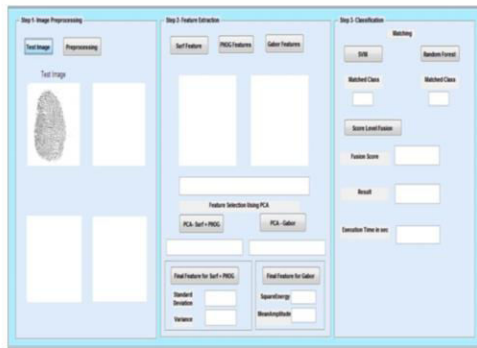


Figure 2. Input image

Figure 2 shows the collection of Test Image. This image will be used for preprocessing.

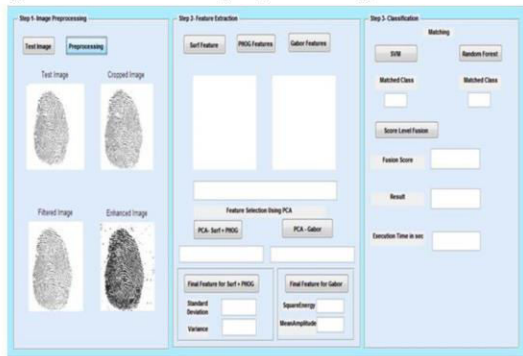


Figure 3. Image Preprocessing

Figure 3 depicts the preprocessing of the images. By Image Preprocessing we found the preprocessed Images e.g. Test Image, filtered image, cropped image, Enhanced image.



Figure 4. Extraction of Gradient Features from SURF

The above figure 4.3 depicts the Extraction of gradient Features from SURF. After Preprocessing of the Input image, SURF features are extracted from preprocessed image.



Figure 5. Extraction of gradient features from PHOG

Figure 5 depicts the Extraction of gradient features from PHOG. After extracting the SURF features the features are extracted from PHOG.

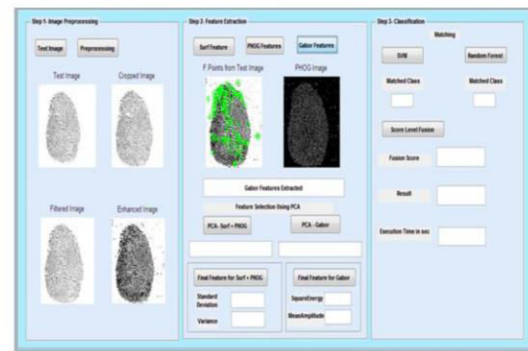


Figure 6. Extraction of texture features from Gabor

Figure-6 depicts the Extraction of texture features from Gabor wavelet. After extracting the PHOG features the Gabor features are extracted.



Figure 7. Feature selection process using PCA method by combining SURF and PHOG features.

Figure 7 depicts the feature selection process. To perform this, we have used PCA method for selecting optimal features.



Figure 8. Feature selection process using PCA – Gabor

Figure 8 depicts the feature selection process using PCA- Gabor method. After selecting the optimal features using PCA with SURF and PHOG, to perform this, we have used PCA method with Gabor features.



Figure 10. Final feature extraction for Gabor

Figure 10 depicts the final feature for Gabor. After extracting the final feature for SURF and PHOG we found the final feature for Gabor method.

VI. CONCLUSION

A new technique for fingerprint liveness detection by combining low level features is developed which comprises gradient features from SURF, PHOG, and texture features from Gabor wavelet. Additionally, an efficient dynamic score level integration module is developed to unite the outcome from the two individual classifiers. Experiments are carried out on two most commonly used databases from LivDet competition 2011 and 2013. In detail comparison is done with the current state of the art, and the winner of LivDet 2011 and 2013 fingerprint liveness detection competition. ACE rate of 2.27% in comparison to the 12.87% of the 2013 LivDet competition winner is an important concert gain. The proposed method scored constantly low EER on the whole six sensors which were not experiential in the state of the art techniques.

VII. REFERENCES

- [1]. Manju Kulkarni, Harishchandra Patil "Liveness detection in fingerprint recognition technique using first order texture features" IJAET/Vol.II/ Issue IV/October-December, 2011.
- [2]. Ana F. Sequeira and Jaime S. Cardoso "Fingerprint Liveness Detection in the

- Presence of Capable Intruders" *Sensors* 2015, 15, 14615-14638; doi:10.3390/s150614615.
- [3]. Sajida Parveen et. al. "Face anti-spoofing methods" *current science*, vol. 108, no. 8, 25 April 2015.
- [4]. Emanuela Marasco and Arun Ross "A Survey on Anti-Spoofing Schemes for Fingerprint Recognition Systems" *ACM Comput. Surv.* 47, 2, Article A, September 2014. DOI:<http://dx.doi.org/10.1145/0000000.0000000>
- [5]. Y. Chung and M. Yung "Fingerprint Liveness Detection Based on Multiple Image Quality Features" *LNCS 6513*, pp. 281-291, Springer-Verlag Berlin Heidelberg 2011
- [6]. Yujia Jiang and Xin Liu "Spoof Fingerprint Detection based on Co-occurrence Matrix" *International Journal of Signal Processing, Image Processing and Pattern Recognition* Vol.8, No.8 (2015), pp.373-384 <http://dx.doi.org/10.14257/ijcip.2015.8.8.38>
- [7]. Athos Antonelli et. al. "Fake Finger Detection by Skin Distortion Analysis" *Ieee Transactions on Information Forensics and Security*, Vol. 1, no. 3, September 2006.
- [8]. Qinghai Gao "A Preliminary Study of Fake Fingerprints" *I.J. Computer Network and Information Security*, 2014, 12, 1-8 Published Online November 2014 in MECS (<http://www.mecs-press.org/>) DOI: 10.5815/ijcnis.2014.12.01
- [9]. Arunalatha G. and M. Ezhilarasan "Spoof Detection of Fingerprint Biometrics using PHOG Descriptor" *International Science Press, IJCTA*, 9(3), 2016, pp. 1705-1711.
- [10]. Dr. Chander Kant, Raksha "Spoof Attack Detection in Fingerprint Authentication using Hybrid fusion" *IJCSCIJ* Volume 4, 1 March 2013 pp. 59-64.
- [11]. Devakumar et al., *International Journal of Advanced Research in Computer Science and Software Engineering* 7(3), March- 2017, pp. 70-76
- [12]. Heeseung Choi, Raechoong Kang, Kyungtaek Choi, and Jaihie Kim "Aliveness Detection of Fingerprints using Multiple Static Features" *International Science Index, Computer and Information Engineering* Vol:1, No:4, 2007 waset.org/Publication/3945
- [13]. Lekshmy. S. Mohan Joby James "Fingerprint spoofing detection using local binary pattern and Hog" *ijastems-issn: 2454-356x* Volume.3, Special Issue.1, April.2017.
- [14]. Shankar Bhausahab Nikam and Suneeta Agarwal "Texture and Wavelet-Based Spoof Fingerprint Detection for Fingerprint Biometric Systems" *First International Conference on Emerging Trends in Engineering and Technology*, 2008.
- [15]. Shankar Bhausahab Nikam, Suneeta Agarwal "Wavelet-based multiresolution analysis of ridges for fingerprint liveness detection" *International Journal of Information and Computer Security* Volume 3 Issue 1, June 2009.
- [16]. Aditya Abhyankar and Stephanie Schuckers "Fingerprint Liveness Detection Using Local Ridge Frequencies and Multiresolution Texture Analysis Techniques" *IEEE International Conference on Image Processing*, 2006, DOI: 10.1109/ICIP.2006.313158.
- [17]. P. Venkata Reddy et. al. "A New Method for Fingerprint Antispoofing using Pulse Oximetry" *First IEEE International Conference on Theory, Applications, and Systems*, 2007, 10.1109/BTAS.2007.4401916.
- [18]. Mojtaba Sepasian, Cristinel Mares, Wamadeva Balachandran "Vitality Detection in Fingerprint Identification" *Wseas Transactions on Information Science and Applications*, Issue 4, Volume 7, April 2010.
- [19]. Reiko Iwai, Hiroyuki Yoshimura "A New Method for Improving Robustness of Registered Fingerprint Data Using the Fractional Fourier Transform" *Int. J. Communications, Network and System*

Sciences, 2010, 3, 722-729
doi:10.4236/ijcns.2010.39096

- [20]. R.Sowmiya, C.Dhivya,B.Nandhini, and T.Anand "Image quality assessment using Biometric Liveness Detection for fake Fingerprint" International Research Journal of Engineering and Technology (IRJET) Volume: 02 Issue: 08 Nov-2015.
- [21]. L. Ghiani et al., "LivDet 2013 fingerprint liveness detection competition 2013," in Proc. Int. Conf. Biometrics (ICB), Jun. 2013, pp. 1–6.

A Novel Approach for Detecting and Matching Iris Crypts For Human Recognition System	S.Yoga Assistant Professor	Computer Science	International Journal of Scientific Research in Computer Science, Engineering and Information Technology	2018	2456 - 3307	UGC Journal No : 64718 https://ijsrcseit.com/paper/CSEIT1833196.pdf
--	----------------------------------	---------------------	---	------	-------------------	--



International Journal of Scientific Research in Computer Science, Engineering and Information Technology

© 2018 IJSRCSEIT | Volume 3 | Issue 3 | ISSN : 2456-3307

A Novel Approach for Detecting and Matching Iris Crypts For Human Recognition System

L. Ponnarasi¹, S. Yoga²

¹Research Scholar, Department of Computer Science, Sakthi college of Arts and Science For Women, oddanchatram, Tamil Nadu, India

²Assistant Professor, Department of Computer Science, Sakthi college of Arts and Science For Women, oddanchatram, Tamil Nadu, India

ABSTRACT

In a variety of applications, the iris is a secure biometric feature that has been extensively employed for human recognition. Though, exploitation of iris recognition in forensic applications has not been informed. A most important cause is being deficient in of human friendly approaches for comparing with iris. Additionally to endorse the utilization of iris recognition in forensics, the resemblance between irises be supposed to made visualizable and understandable. In recent times, a system was proposed, known as “a human-in-the-loop iris recognition system” which was based on detecting and matching iris crypts. Structuring on this system, a new approach for detecting and matching iris crypts automatically is proposed in this work. This detection method is capable to capture iris crypts of different sizes. This matching method is considered to handle possible topological modifications in the detection of the similar crypt in diverse images. This approach does better the well-known visible-feature-based iris recognition method on three dissimilar data sets. Subsequent to iris Crypts detection, Iris images were in use prior to and later than the treatment of eye disease and the outcome illustrates the mathematical divergence accomplished from treatment. Gabor filter is employed to extract the features. This iris recognition was efficiently endured with the majority of ophthalmic disease e.g. corneal oedema, iridotomies and conjunctivitis etc. This developed iris recognition be supposed to employed for resolving the potential issues that might reasonable in key biometric technology and medical diagnosis.

Keywords: Iris Recognition, Forensics, Human-In-The-Loop, Eye Pathology, Ophthalmic Disease, Iridotomies, Conjunctivitis, Visible Feature, Corneal Oedema.

I. INTRODUCTION

Based on biometrics the demand for automated personal identification system has increased with a growing prominence in security. Because the conventional (cards or passwords based) can be broken by stealing cards and forgetting passwords. Thus, there is a requirement for identification systems identify humans which is independent on what person possesses or what person remembers. Biometrics can be separated into two main divisions: physiological and behavioral. The physiological class

is associated to the shape of the body which contains fingerprint, face recognition, palm print, hand geometry, and iris recognition. The behavioral class is associated to the behavior of a person and contains typing rhythm and voice.

In recent times, iris recognition is fetching one of the most vital biometrics employed in recognition when imaging can be performed at distances below two meters. This significance is because of its high reliability for individual identification. Human iris has enormous mathematical advantage that its

pattern inconsistency among different persons is tremendous, since iris patterns acquire a high degree of randomness. Additionally, iris is extremely stable over time. Because the idea of automated iris recognition was developed in 1987, several researchers worked meanwhile that time and they developed different dominant methods. Those methods were based on the texture variations of the iris and can be separated into different techniques e.g. phase-based methods, texture analysis, and intensity variations etc.

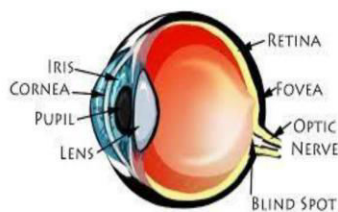


Figure 1. Structure of Iris

Nowadays most of the systems in used and they required unambiguous user collaboration, demanding that the user is placed properly to attain a quality image. These systems give acoustic response to the user to make certain that they are properly situated for image acquisition. In the United Kingdom, the Iris Recognition Immigration System (IRIS) is an intended system that appropriates travelers to authorize through border control stations at various airports rapidly, confirming their identification employing automated roadblocks. CANPASS in Canada is a related program to grant regular travelers to speedily proceed through security verification at airports.

Alternatively, as above study IRIS recognition is most reliable techniques in biometrics for human identification. Thus, the Daugman algorithm is able to acquire a false match rate below 1 in 200 billions. Iris recognition methods have been employed extensively by governments, for example the Aadhaar card project in India. Conversely, biometric feature in law of social control diligences the iris is under evaluation yet. One cause that obstructs the

forensic exploitation of iris is that iris recognition outcomes are not simply understandable to examiners. Thus “Iris Examiner Workstation” may be build equivalently to the “Tenprint Examiner Workstation”, which has been employed in forensics. In fingerprint recognition, a human auditor basis a choice on the number of matched details on two fingerprints. On compare, frequent iris recognition methods, e.g. Daugman’s framework, execute matching on an iris code, which is the outcome of employing a band-pass filter and quantizer to grayscale images. In these circumstances, the entire process becomes visible as a black-box to an examiner without knowing about image processing. Conduct tests have demonstrated that human examiners can act upon better in identity verification with iris images. The result was prepared based on human perception of the on the whole texture. Equivalent to fingerprints, one method to additional endorse the improvement of iris recognition in law enforcement applications is to build the resemblance between irises understandable with the intention that the entire procedure can be supervised and verified by human experts. Explicitly, the decision supposed to be ready based on quantitative matching of visible features in iris images.

II. STATEMENT OF THE PROBLEM

In this research, the eye image to be detect iris crypts by utilizing segmentation process. Iris crypts can emerge in a variety of sizes and shapes in images. In actual fact, it is doubtful from time to time whether numerous proximal crypts are associated. Moreover, slight dissimilarities in the obtained images of the similar iris may change the topology of the detection of the same crypts from image to image. In this modified method mainly there are two tasks: crypt detection and crypt matching. This detection algorithm is developed to handle multi-scale crypts. It employs a key morphological function in a hierarchical manner. Human interpreted training data is utilized to find out the most important parameters, in order that the detected crypts are alike

to those acquired by human examination. In the matching algorithm, a matching model is adopted which is based on the Earth Mover's Distance (EMD). This matching model is somewhat common in use.

- ✓ Using this method the image quality is high and the level process is very fewer.
- ✓ In this process Time consumption is very low
- ✓ Easiest way to find crypts pattern also matching perform based on the distance.

III. SCOPE OF THE RESEARCH

The scope of the research for this system is from the observation that the human iris provides an overall attractive structure on which it supports a technology for noninvasive biometric measurement. Especially it is identified in the biomedical area that the irises are diverse. Iris is an overt body; its outward show is acquiescent to remote examination with the help of a machine-vision system. Based on the observation the iris recognition issues originates from the deficiency of human interpretability in modern biometric techniques, which can be a obstruction to the official employment of iris recognition in forensics. The Gabor filter is employed in an iris recognition framework to extract the features. Generally, two iris images are given and the extracted visible features using Gabor filter on each image, then match the extracted features and evaluate the resemblance sequentially to verify whether the two images are from the same eye.

IV. METHODOLOGY

Input Image: For reading an input image first choose the pathname and then filename set format by using the MATLAB syntax (imread).

Localization: The inner boundary and out boundary of distinctive iris can be chosen as circle. But both two circles are typically not co-centric. Inner circle also will be detecting pupil in iris. Due to the extreme low boundary level contrast the outer boundary of circle is more difficult to detect.

Normalization: The size of the pupil due to the change of variation and illumination elastic deformation. The outcome of pattern matching can be interposed by Iris texture. Then the inner boundaries as well as outer boundaries simple to detect map in iris ring.

Gaussian filter: The filter whose impulse response is a Gaussian function or approximate to it is known as a Gaussian filter. Gaussian filters have the properties that it is not exceed to a step function input while away minimizing the rise and fall time. This activity is directly associated to the conception that the Gaussian filter has the minimum probable group delay. It is conceived the ideal time domain filter, because of the ideal frequency domain filter. These properties are more significant in the fields e.g. oscilloscopes and digital telecommunication systems.

Morphological operation: Morphological image processing is a set of non-linear operations associated with the shape or morphology of traits in an image. Allowing morphological operations trust only on the proportional sequence of pixel values, even not on their numerical values, and consequently are particularly appropriate to the processing of binary images. Morphological operations can also be employed to grayscale images like their light transfer functions are indefinite and thus their complete pixel values are of no or small attention. Morphological techniques investigate an image with a small shape or pattern predicted as a structuring element. The structuring element is located at all probable locations in the image and comparison is done with the consequent neighborhood of pixels. Few operations test whether the element "fits" contained by the neighborhood, whereas others test whether it "hits" or intersects the neighborhood.

Binary Image: A digital image is a binary image that holds just two probable values for each pixel. Though any two colors can be used for binary Image, usually two colors black and white are used for a binary

image. For the object(s) in the image the color used is the foreground color whereas the rest of the image is the background color. This is frequently referred to as "bi-tonal" in the document-scanning industry. Binary images are termed as bi-level or two-level. This intends that each pixel is stored as a single bit such as 0 or 1. The names frequently used for this concept are black-and-white, B&W, monochrome or monochromatic, but possibly will choose any images that contain just one sample per pixel, e.g. grayscale images. The operations are segmentation, thresholding, and dithering. Few input/output devices can only handle bi-level images as laser printers, fax machines, and bi-level computer displays etc. As a bitmap a binary image can be stored in memory. A 640×480 image needs 37.5 KiB of storage. Fax machine and document management solutions normally use this format due to the small size of the image files. With simple run-length compression approaches most of the binary images also compress well.

Segmentation: In computer sight, the process of partitioning a digital image into multiple segments image segmentation is called the Image segmentation. The aim of segmentation is to make simpler and modify the illustration of an image into incredible that is additionally consequential and simpler to examine. Image segmentation is normally utilized to place objects and boundaries (lines, curves, etc.) in images. More accurately, image segmentation is the process of putting a label to each pixel in an image as pixels with the same label allocate convinced distinctiveness. The outcome of image segmentation is a collection of segments that cooperatively deal with the whole image, or a collection of contours extracted from the image. Each of the pixels in a region is similar concerning few qualities or computed property e.g. color, intensity, or texture etc. Adjacent regions are extensively dissimilar regarding the same characteristic(s) as employed to a stack of images, usually in medical imaging, the consequential contours after image segmentation can be utilized to construct 3D

renovation with the aid of interpolation algorithms such as marching cubes.

V. EXPERIMENTAL RESULTS

5.1 Importing and Exporting Images

Image Processing Toolbox chains images produced by a broad range of devices, containing digital cameras, satellite and airborne sensors, medical imaging devices, microscopes, telescopes, and other scientific instruments. It may visualize, analyze, and process these images in various data types, containing single- and double-precision floating-point and signed and unsigned 8-, 16-, and 32-bit integers. There are numerous modes to import and export images into and out of the MATLAB background for processing. Image Acquisition Toolbox can be used to obtain live images from Web cameras, frame grabbers, DCAM-compatible cameras, and other devices. Using Database Toolbox, images can be accessed which are stored in ODBC/JDBC-compliant databases.

5.2 Displaying and Exploring Images

Image Processing Toolbox expands MATLAB graphics to offer image display capabilities which are extremely customizable. It can construct displays with multiple images in a single window, interpret displays with text and graphics, and create specialized displays e.g. histograms, profiles, and contour plots.

Additionally, to display functions, the toolbox provides a suite of interactive tools for exploring images and building GUIs.

5.3 Preprocessing and Post Processing Images

Image Processing Toolbox supports reference-standard algorithms for preprocessing and post-processing responsibilities that resolve frequent system problems, e.g. interfering noise, low dynamic range, out-of-focus optics, and the dissimilarity in color demonstration between input and output devices.

Image enhancement techniques in Image Processing Toolbox assist to improve the signal-to-noise ratio and accentuate image features by altering the colors or intensities of an image. It can:

- ✓ Perform histogram equalization
- ✓ Perform decorrelation stretching
- ✓ Remap the dynamic range
- ✓ Adjust the gamma value
- ✓ Perform linear, median, or adaptive filtering

5.4 Analyzing Images

Image Processing Toolbox gives a widespread collection of reference-standard algorithms and graphical tools for image analysis tasks e.g. statistical analysis, feature extraction, and property measurement.

Statistical functions analyze the common characteristics of an image by:

- ✓ Computing the mean or standard deviation
- ✓ Determining the intensity values along a line segment
- ✓ Displaying an image histogram
- ✓ Plotting a profile of intensity value

Edge-detection algorithms identify object boundaries in an image. These algorithms contain the Sobel, Prewitt, Roberts, Canny, and Laplacian of Gaussian methods. The dominant Canny method can detect true weak edges without being "fooled" by noise.

5.4 Working with Large Images

Few images are outsized that they are complicated to process and display with standard methods. Image Processing Toolbox offers exact workflows for working with larger images than or else possible. Devoid of loading a large image completely into memory, can create a reduced-resolution data set (R-Set) that partitioned an image into spatial tiles and resample the image at dissimilar resolution levels. This workflow develops performance in image display and navigation. A block processing workflow can be used to apply a function to each distinct block of a large image that considerably reduces use of memory. An additional alternative for functioning

with large images is to make use of the Parallel Computing Toolbox.

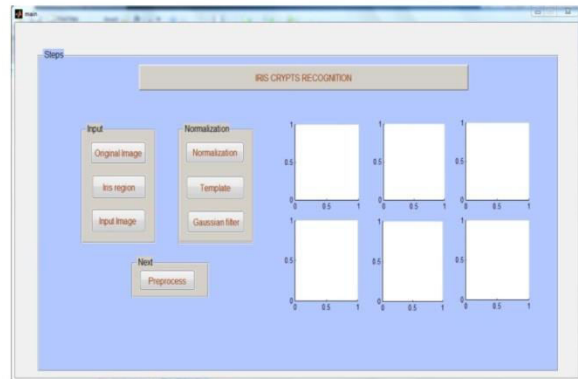


Figure 2. Start page of the Iris Crypts system

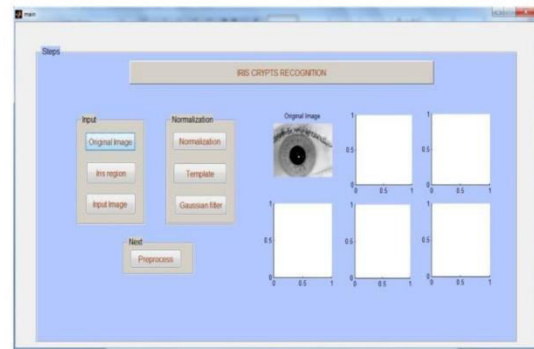


Figure 3. Getting the input image of the system

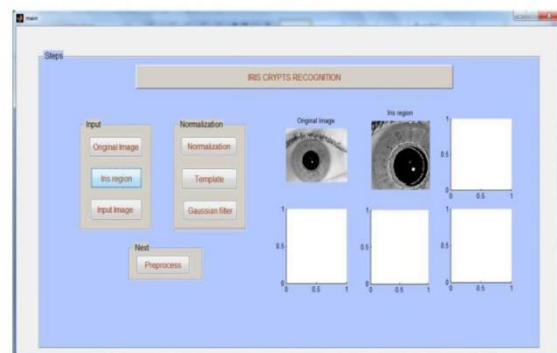


Figure 4. Obtaining the iris region of the original iris image

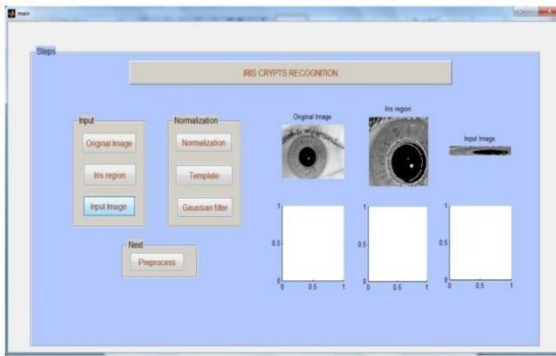


Figure 5. The extracted region of iris is taken as input image

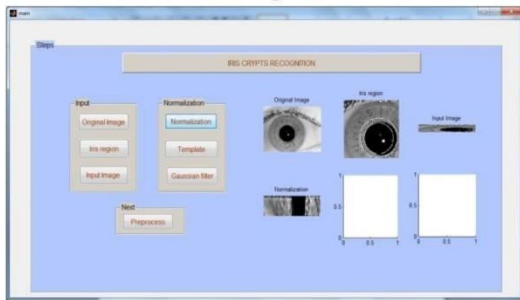


Figure 6. The input image is normalized to remove uncertainties in an image

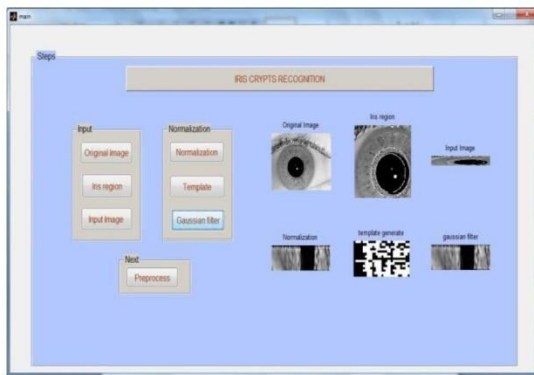


Figure 7. Generating the templates on the extracted iris region and then applying gaussian noise filter to remove low-level features

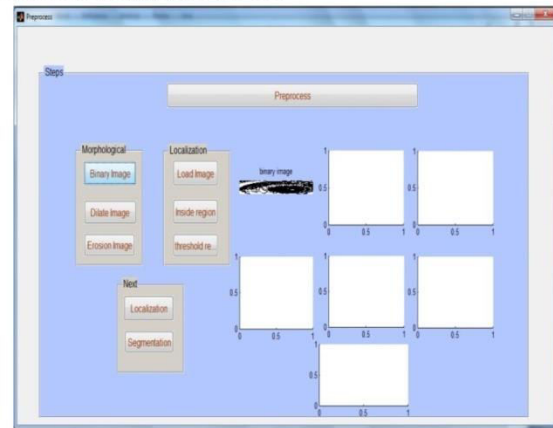


Figure-8: Converting the filtered iris image into binary image

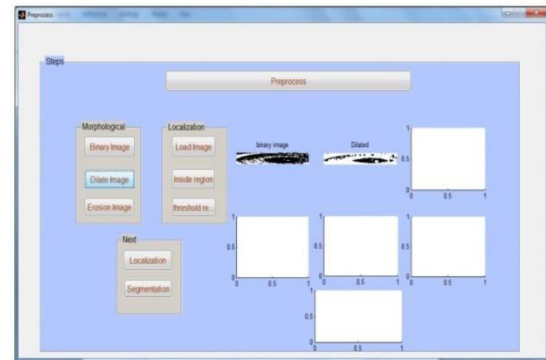


Figure 9. Binary image is further dilated for performing morphological operations

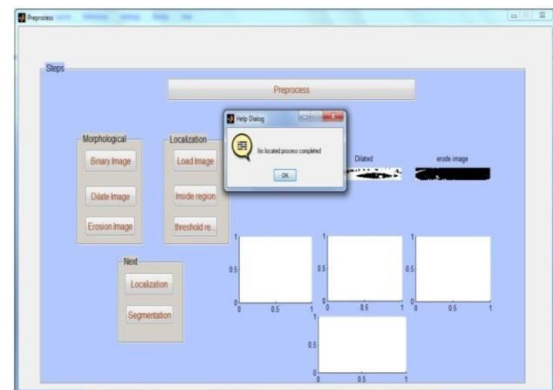


Figure 10. Iris region is successfully completed

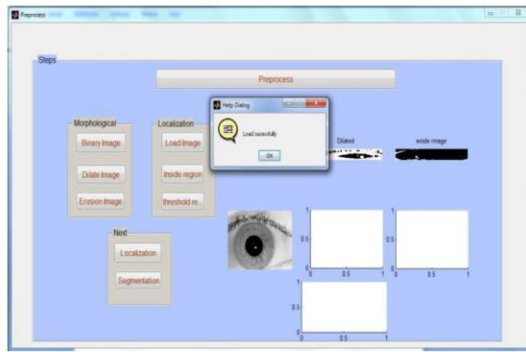


Figure 11. Final Preprocessed Iris Image

VI. CONCLUSION

A novel method for detecting and matching iris crypts for the human-in-the-loop iris biometric system is introduced. The presented method develops predicting outcomes on the three tested datasets, in-house dataset, ICE2005, and CASIA-Iris-Interval. On Comparison with the existing method, this proposed method enhances the iris recognition performance by minimum 22% on the position one hit rate in the circumstance of human identification and by minimum 51% on the equal error rate in provisions of subject verification.

It is noticed that the three datasets under estimation were gathered using dissimilar facilities among diverse population groups. The constraints applied in this method were skilled on a different small set of homemade data. The generalization and usefulness of this method on varied image data can be presented. Additionally, to the extent that, this work is so distant the just estimation of a human-interpretable iris features matching method by using the public datasets (ICE2005 and CASIA-Iris-Interval), that provides a lead contrast with existing methods for example Daugman's framework. Experimental analysis has shown the effectiveness of the proposed system.

VII. REFERENCES

[1]. Pradeep Manikrao Patil "Iris Recognition in Less Constrained Environment" *International Journal of Emerging Technology and*

Advanced Engineering, Volume 3, Issue 7, July 2013.

[2]. Deepa, V.Priyanka and J.Pradeepa "Iris Recognition Based on Iris Crypts" *International Journal of Engineering Sciences & Research Technology*, February, 2017, DOI: 10.5281/zenodo.291855.

[3]. Mrigana walia and Shaily Jain "Iris Recognition System Using Circular Hough Transform" *International Journal of Advance Research in Computer Science and Management Studies*, Volume 3, July 2015.

[4]. Proença H, Filipe S, Santos R, Oliveira J and Alexandre LA "The UBIRIS.v2: a database of visible wavelength iris images captured on-the-move and at-a-distance" *IEEE Transaction* Aug 2010. doi: 10.1109/TPAMI.2009.66

[5]. Mustafa M. Alrifae, Mohammad M. Abdallah and Basem G. Al Okush "A Short Survey of Iris Images Databases" *The International Journal of Multimedia & Its Applications (IJMA)* Vol.9, No.2, April 2017

[6]. Shaik Touseef Ahmad and Sandesh Kumar B. V "Ordinal Feature Selection for IRIS and Palm print Recognition" *International Journal of Ethics in Engineering & Management Education*, Volume 2, Issue 6, June 2015.

[7]. Vineetha John Tharakan and Shaikh Fairouz "Ordinal Features Based Palmprint and Iris Recognition for Military Security" *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, Vol. 6, Issue 4, April 2017.

[8]. Priya. J and A. Alad Manoj Peter "Fusion of Palm Print and Iris for Multimodal Biometric Recognition" *International Journal of Scientific and Research Publications*, Volume 6, Issue 5, May 2016.

[9]. Govindharaj, R. Vinitha, N. V. Dhanapriya, P. Diviyabarady R. Abiramy "Retina Based Personal Identification System using SIFT algorithm: A survey" *Asian Journal of*

- Applied Research (AJAR) Volume 03- Issue 02- pp-18-23, 2017
- [10]. Manisha Sam Sunder and Arun Ross "Iris Image Retrieval Based on Macro-features" International Conference on Pattern Recognition (ICPR), (Istanbul, Turkey), August 2010.
- [11]. Dr.S. Prasath and A.Selvakumar "A Novel Iris Image Retrieval with Boundary Based Feature Using Manhattan Distance Classifier" International Journal of Innovative Technology and Creative Engineering, vol.5 no.7 july 2015.
- [12]. Shweta Malvi and P.M.Agarkar "To Study Iris Recognition Systems, Databases, Complex Patterns and Segmentation Techniques" International Journal of Engineering, Science and Mathematics, Volume 2, Issue 2, June 2013.
- [13]. Nozomi Hayashi and Akira Taguchi "A Novel Feature Extraction for Iris Identification using Morphological Skeleton" 18th European Signal Processing Conference August 2010.
- [14]. Samanpreet Kaur and Er. Mandeep Singh "A Review on various iris recognition techniques in biometrics" international journal of engineering sciences & research technology, Vol. 4.(11), November, 2015.
- [15]. Manisha M. Khaladkar, Sanjay R. Ganorkar "A Novel Approach for Iris Recognition" International Journal of Advanced Research in Computer Engineering & Technology Volume 1, Issue 4, June 2012.
- [16]. Suganthi, M. and P. Ramamoorthy "Principal Component Analysis Based Feature Extraction, Morphological Edge Detection and Localization for Fast Iris Recognition" Journal of Computer Science 8 (9): 1428-1433, 2012.
- [17]. Geetanjali Sharma and Neerav Mehan "Iris based Human Identification using Median and Gaussian Filter" International Journal of Latest Trends in Engineering and Technology Vol.(7), DOI: <http://dx.doi.org/10.21172/1.73.560>
- [18]. Himanshi Budhiraja, Himani Tomar, Harshi Goel, Amar Singh "Fusion of Iris and Fingerprint Biometric for Recognition" International Journal Of Advance Research In Science And Engineering IJARSE, Vol. No.2, Issue No.4, April, 2013.
- [19]. Joaquim de Mira Jr. and Joceli Mayer "Image Feature Extraction for application of Biometric Identification of Iris - A Morphological Approach" IEEE Symposium on Computer Graphics and Image Processing, 2003.
- [20]. Unique Identification Authority of India. Online]. Available: <http://uidai.gov.in>, accessed Nov. 1, 2015.

A Literature Review on Security Enhanced Multi-Factor Biometric Authentication System Using FFF and KSVM	N. Nanthini Assistant Professor	Computer Science	International Journal of Scientific Research in Computer Science, Engineering and Information Technology	2018	2456 - 3307	UGC Journal No : 64718 https://ijsrcseit.com/paper/CSEIT1833231.pdf
--	---	------------------	--	------	-------------	--



International Journal of Scientific Research in Computer Science, Engineering and Information Technology

© 2018 IJSRCSEIT | Volume 3 | Issue 3 | ISSN : 2456-3307

A Literature Review on Security Enhanced Multi-Factor Biometric Authentication System Using FFF and KSVM

¹P. Pandimeena, ²N. Nanthini

¹ Research Scholar, Department of Computer Science, Sakthi College Of Arts and Science for Women, Oddanchatram, India

² Assistant Professor, Department of Computer Science, Sakthi College Of Arts and Science for Women, Oddanchatram, India

Department of ECE, Sri Eshwar College of Engineering, Coimbatore, Tamil Nadu, India

ABSTRACT

We focus on multimodal biometric system by combining finger knuckle and finger vein using feature level fusion optimization. Biometric characteristics (Eyes, Finger vein, Finger Knuckle, Face, Ear, and Palm) like. Here used unique and secure password (like Finger Vein, Finger Knuckle). In this paper, the authors propose a multimodal biometric system by combining the finger knuckle and finger vein images at feature-level fusion using fractional firefly (FFF) optimization. Biometric characteristics, like finger knuckle and finger vein are unique and secure. Initially, the features are extracted from the finger knuckle and finger vein images using repeated line tracking method. Then, a newly developed method of feature-level fusion using FFF is used. This method is utilized to find out the optimal weight score to fuse the extracted feature sets of finger knuckle and finger vein images. Thus, the recognition is carried out by the fused feature set using layered k-SVM (k-support vector machine) which is newly developed by combining the layered SVM classifier and k-neural network classifier. The experimental results are evaluated and the performance is analyzed with false acceptance ratio, false rejection ratio and accuracy. The outcome of the proposed FFF optimization system obtains a higher accuracy.

Keywords : Feature Level Fusion, FFF Optimization, Repeated Line Tracking method, Layered K-SVM, K-neural network classifier.

I. INTRODUCTION

Nowadays, many of the multimodal biometric systems are in use and gained a lot of importance due to its uniqueness and effectiveness. The multimodal biometric systems include hand geometry, signature, retinal pattern, iris, voice-print, finger knuckle, fingerprint, finger vein, face and so on. The advantages and disadvantages of the biometric systems are based on the three main factors, such as user acceptance, accuracy and applicability. The accuracy of the iris pattern, retinal pattern and face is

minimal, when compared to the finger knuckle and the finger vein traits. User acceptance is also very high for the finger knuckle and the finger vein compared to the other biometric traits. The performance is also good for the finger knuckle and the finger vein due to the finger geometry. In addition to, security, non-traceability, speed, user friendly, accuracy and so on are the advantages of the finger vein.

The integration of the feature sets is used to enhance the outcome of the recognition of the biometric system by the corresponding multiple modalities. The integration of the feature is done in three ways, such as feature-level fusion, score-level fusion and decision-level fusion. The integration of the feature set is difficult, when (i) the feature sets of multiple modalities are incompatible, (ii) unknown relationship between the feature space of multiple modalities and, (iii) curse of dimensionality problem. Commonly, three level fusion before and after matching criteria are used for fusing the features. In score-level fusion, the integration of feature vector is done with the matching score output of the individual matches, and then, the feature vectors are accepted or rejected by an information

- Combining the fractional theory and firefly algorithm as fractional firefly (FFF) optimisation algorithm for feature-level fusion based on the finger knuckle and finger vein images.
- FFF optimisation algorithm is proposed to find out the optimal weight score level for the feature-level fusion. Thus, this optimisation is used to fuse the feature set of both finger knuckle and vein image by the weight score level.
- A new classifier called, k-SVM (k-support vector machine) is developed for the recognition of person by combining the k-NN (k-neural network) classifier and SVM classifier.

A. Challenges

On the basis of the literature review conducted, multimodal recognition have been actively studied with various machine learning techniques but for the unsurpassed recognition, the feature-level fusion-based recognition is the fine choice considering the matching score-level as well as the decision-level fusion. The developing of multimodal recognition techniques using feature-level fusion have not been studied much in the literature even though it contains more advantages than the score- and decision-level fusion. In feature-level fusion, the

concatenation of the feature vector with reasonable accessibility is an important challenge in the biometric recognition system. Even if the features of the multimodalities are not compatible, the concatenation must be appropriate for the recognition.

Furthermore, fusion of the feature with the ultimate robust recognition is crucial challenge considerable in the multimodal biometric recognition system. While using feature-level fusion, the biometric recognition system must not degrade along with the quality of the feature sets. Proper processing over the feature must be employed for the thriving function of the recognition system. Another important challenge with respect to feature-level fusion is to develop the reliable recognition system. The fusion level must be selected in a way improving the recognition accuracy of the recognition system without degrading the system performance.

B. New FFF Optimization

A novel optimization method is proposed for feature-level fusion using FFF optimisation, which comprises fractional theory and firefly algorithm. In the firefly algorithm, variation of light intensity and the formulation of attractiveness are the two significant issues. It is a Meta heuristic algorithm for global optimisation, which is inspired by flashing behaviour of firefly insects. For simplicity, assume that the attractiveness of a firefly is determined by its brightness or light intensity, which in turn is associated with the encoded objective function. The brighter one will attract the other; so the less bright one is moved towards the brighter one. In the simplest case, for the optimisation problems, the brightness I of a firefly at a particular location x can be chosen as $I(x) \propto f(x)$. In this paper, the fireflies are initialised randomly. For the next iteration, the fireflies are newly generated by finding the movement of firefly with another firefly, which is expressed using the fractional theory. The fractional theory can be rather interesting for filtering and edge

detection and also enhance the quality of images. When differential and integral calculus plays a significant role in mathematics, experts investigated the computation of non-integer order derivatives and integrals. Thus, the integration of firefly optimisation and fractional theory is used here to calculate the appropriate value for α and β .

II. METHODOLOGIES USED FOR MULTIMODAL BIOMETRIC SYSTEM

A. Authentication using finger vein recognition based on Matlab

This thesis aims to developing a system for acquiring images of finger veins and processing them using MATLAB for the purpose of authentication. It includes designing of hardware for image acquisition, coding the matching algorithm for processing the finger vein pattern and training and testing of algorithm module. Typical Finger vein recognition system consists of image acquisition module, image preprocessing, feature extraction, and matching.

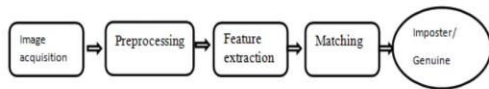


Figure-1: Authentication Processing

B. Image acquisition

Finger Vein patterns can be viewed through an image sensor sensitive to infrared light. Infrared light passing through the tissues of the human body is blocked by hemoglobin. As hemoglobin exists densely in blood vessels, infrared light passing through veins appears as dark shadow lines in the captured image.

C. Pre-processing

The first step of the proposed multimodal biometric recognition is pre-processing which makes the input training images better suitable for the subsequent steps. The important processes such as, normalization, filtering and resizing are carried out under pre-processing steps. Once the input images are read out,

it undergoes the normalization steps to convert the range of pixels within the particular range. The, median filtering is applied to smooth the input images which makes the input images much visible. Also, this process is helpful for the feature extraction to easily identify the vein parts. Then, resizing is performed to convert all the input images into fixed size through interpolation scheme.

Preprocessing step includes image segmentation in which captured image is divided into multiple parts. Each of the pixels in a segment will be similar with respect to some properties, such as color, texture or intensity. The aim of segmentation is to change the representation of an image into something that is easier to analyze. Image segmentation is used to locate objects and boundaries in an image. Segmentation is the process by which we are assigning a label to every pixel in an image. Pixels sharing the same label will have certain similar visual characteristics.

D. Vein and knuckle print extraction using repeated line tracking

In this method, the extraction of finger knuckle and vein print using a repeated line method is discussed. The line tracking operation starts at any pixel in the source image. We defined the current pixel position in an image as the current tracking point and this point is moved from pixel to pixel along the dark line direction in the finger knuckle and finger vein images. Thus, the method of feature extraction from the image is described as follows. $F_{i,j}$ is the intensity of the pixel i, j in the finger knuckle image. Similarly, $F_{m,n}$ is the intensity of the pixel m, n in the finger vein image. Z_{fk} and Z_{fv} are the set of pixels in the finger knuckle and finger vein images, respectively. S_1 is considered as the locus space. Thus, the knuckle and vein print are extracted by the following four steps:

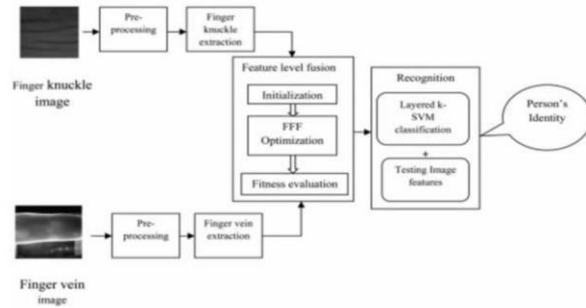


Figure-2 Block diagram of the proposed methodology

The finger vein image features are extracted using wavelet transform and line detection. Wavelet transform is a mathematical function which divides a function into its different frequency components. Wavelet transform analyzes each individual component with a resolution that matches its scale. HAAR wavelet transform multiplies a function against the HAAR wavelet with various shifts and stretches. HAAR transform is easy to implement and is able to analyze the local features. These characteristics make HAAR wavelets applicable for Finger vein recognition algorithm. At last, matching with database is a final decision making step to get a result from the finger vein recognition algorithm. In the matching stage two types of errors are considered FAR (False Acceptance Rate), FRR (False Rejection Rate). FRR is the rate of occurrence of a scenario in which two fingerprints from same finger fails to match (the matching score is below the threshold) while FAR is the rate of occurrence of a scenario in which two fingerprints from different fingers will match (matching score is greater than the threshold). EER is the error rate at which the FAR equals the FRR and is therefore, suitable for measuring the overall performance of biometric recognition system.

Sample image and its feature extracted image are shown below.

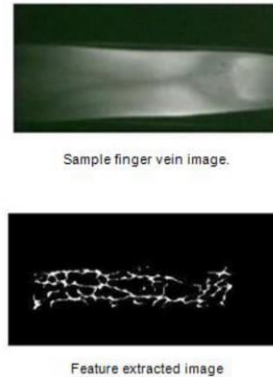


Figure-3: Feature extracted image

In the verification stage, newly captured finger vein image is applied to preprocessing stages, and at last vein image is replaced with the feature extracted image. Finally that extracted image is sent to an authentication stage. This stage will match the newly feature extracted image with the database image, after matching it will create a match score of each finger vein images in the database. Depending on the match score authentication is carried out. This project implements a highly secured authentication system based on using finger vein recognition.

E. Feature-level fusion by FFF optimisation

Fusion at the feature level is least explored even though they are expected to provide better recognition results and much easier to compute. The matching score-level and decision-level supplies less information to be exploited for personnel authentication than the feature-extraction level. Also, the feature-level fusion carries much richer information about the raw biometric data than the matching score or decision level. This is the driving force for the proposed scheme.

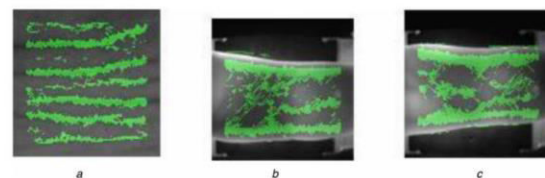


Figure-4: Vein extracted image

F. Recognition using layered k-SVM classifier

The extracted features of finger knuckle and finger vein are fused by the FFF optimisation. Then, the classification is performed using layered k-SVM classifier. Here, SVM classifier and k-NN classifier are combined to perform binary classification and then, $N - 1$ k-SVM classifiers are connected serially to perform multi-level classification. Here, SVM classifier is a binary classifier which is classified by either 0 or 1. Similarly, k-NN classifier is popular technique for data classification based on the neighbours of the input test data. The reason of selecting the k-NN classifier is that it can perform better for multi-classification because the classification is purely based on the distance between

the training data and test sample. Also, SVM is preferably chosen here because of the good performance for the high dimensional data. In proposed work, we used an N number of persons for biometric recognition. Thus, recognition is done by the layered k-SVM classifier which consists of $N - 1$ number of classifiers.

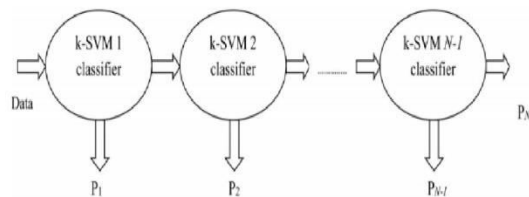


Figure-5:Architecture of layered k-SVM classifier

Table-1: Comparison table on the literature survey

S.No	Title	Description	Merit	Demerit
1	Spoof Attacks on Multimodal Biometric Systems	In addition, latest results have questioned that, contrary to a common claim, multimodal systems can be cracked by spoofing only one trait. Those results were obtained using simulated.	Investigate this significant security issue, focusing on behavior of fixed and trained score fusion rules, using real spoof attack samples under different spoof attack scenarios.	In particular, most widely used fixed rules can be less robust, even if the quality of fake biometric trait is low.
2	Presentation Attack Detection Algorithm for Face and Iris Biometrics	A novel solution to detect a presentation attack based on exploring both statistical and Cepstral features.	Binarized Statistical Image Features (BSIF) and Cepstral features that can reflect the micro changes in frequency using 2D Cepstrum analysis.	Generating face and iris artifacts is not only easy but also cost effective.
3	Fake Biometric Detection to Iris, Fingerprint using Image Quality Assessment	To make sure the actual presence of a true legitimate attribute in distinction to a faux self-manufactured artificial or reconstructed sample may be a significant drawback in biometric	The target of the planned system is to boost the safety of biometric recognition frameworks, by adding animate ness assessment during a quick, easy, and non-	Low complexness options are most well-liked over those that need a high machine load. Unknown, because the detection system solely has access to

		identification, which needs the event of recent and efficient protection measures.	intrusive manner, through the utilization of image quality assessment.	the input sample.
4	Biometrics In Abc: Counter-Spoofing Research	Automated Border Control (ABC) is concerned with fast and secure processing for intelligence-led identification.	It indicates that the new developing trend is fusion of multiple biometrics against spoofing attacks.	Ideal ABC may have a nature of non-intrusive, efficiency, and effectiveness.
5	Fake Biometric Detection for Iris, Fingerprint and Face Recognition	To ensure the actual presence of a real legitimate trait in contrast to a fake self-manufactured synthetic or reconstructed sample is a significant problem in biometric authentication, which requires the development of new and efficient protection measures.	Novel software-based fake detection method that can be used in multiple biometric systems to detect different types of fraudulent access attempts.	The problem of fake biometric detection can be seen as a two-class classification problem where an input biometric sample has to be assigned to one of two classes: real or fake.
6	A Survey Based on Fingerprint, Face and Iris Biometric Recognition System, Image Quality Assessment and Fake Biometric	This paper introduce three biometric techniques which are face recognition, fingerprint recognition, (Multi Biometric System) and also introduce the attacks on that system and by using Image Quality Assessment for Liveness Detection how to protect the system from fake biometrics.	Hardware –based schemes generally present a higher fake detection rate, at the same time software-based techniques are in general less expensive (like no extra device is needed), and less intrusive since their implementation is clear to the user.	Fingerprints have been used from long time for identifying individuals.
7	Fingerprint Liveness Detection in the Presence of Capable Intruders	Fingerprint Liveness detection methods have been developed as an attempt to overcome the vulnerability of fingerprint biometric	The design by modeling the distribution of the live samples and predicting as fake the samples very unlikely	Traditional approaches have been quite optimistic about the behavior of the intruder assuming the use of a previously

		systems to spoofing attacks.	according to that model.	known material.
8	Face Spoof Detection with Image Distortion Analysis	Automatic face recognition is now widely used in applications ranging from de-duplication of identity to authentication of mobile payment.	Popularity of face recognition has raised concerns about face spoof attacks (also known as biometric sensor presentation attacks), where a photo or video of an authorized person's face could be used to gain access to facilities or services.	Different classifiers needed for different spoof attacks.
9	Image Quality Assessment for Fake Biometric Detection: Application to Face and Fingerprint Recognition	Develop a system to enhance the security of biometric recognition framework, by providing a two stage security using finger print and face detection applications.	High level security. More efficient and accurate. Protection against the fake biometric traits. Varied usability, right from small scale companies to the government organizations.	Poor generalization ability (vulnerable to the variations in acquisition conditions).
10	Biometric Anti spoofing Methods: A Survey in Face Recognition	Spoofing, referred to by the term presentation attack in current standards, is a purely biometric vulnerability that is not shared with other IT security solutions.	It refers to the ability to fool a biometric system into recognizing an illegitimate user as a genuine one by means of presenting a synthetic forged version of the original biometric trait to the sensor.	The spoofing artefact is a 3D mask of the genuine client's face, increasing the difficulty to find accurate counter measures against them.

III. CONCLUSION

In this paper, a multimodal biometric recognition system based on the finger knuckle and finger vein was proposed. An important aspect of the proposed system was the development of FFF optimisation for feature-level fusion. After input images were pre-

processed, the FKP was extracted from the knuckle image and vein was extracted from finger vein images using the repeated line tracking method. Then, the features were extracted from the finger knuckle and vein by applying the grid operation to the image. Subsequently, the proposed system was fused the obtained feature set with the help of weight score level, which was obtained by feature-

level fusion using FFF optimisation method. Then, recognition was performed by the fused feature set using layered k-SVM classifier. The proposed system was evaluated with the existing systems and the performance was analysed by the metrics, FAR, FRR, EER and accuracy. From the outcome, we found that the accuracy was obtained for the proposed method. In future, the proposed method can be extended to develop the different objective functions to find the optimal weight score.

IV. REFERENCES

- [1]. Jain, A.K., Hong, L., Kulkarni, Y.: 'A multimodal biometric system using fingerprint, face and speech'. Proc. of Int. Conf. on Audio- and Video-based Biometric Person Authentication, 1999, pp. 182–187
- [2]. Saini, R., Rana, N.: 'Comparison of various biometric methods', *Adv. Sci. Technol.*, 2014, 2, (1), pp. 24–30
- [3]. Perumal, E., Ramachandran, S.: 'A multimodal biometric system based on palmprint and finger knuckle print recognition methods', *Inf. Technol.*, 2015, 12, (2), pp. 118–127
- [4]. Neware, S., Mehta, K., Zadgaonkar, A.S.: 'Finger knuckle surface biometrics', *Eng. Technol. Adv. Eng.*, 2012, 2, (12), pp. 452–455
- [5]. Lu, L., Peng, J.: 'Finger multi-biometric cryptosystem using feature-level fusion', *J. Signal Process., Image Process. Pattern Recogn.*, 2014, 7, (3), pp. 223–236
- [6]. Kale, K.V., Rode, Y.S., Kazi, M.M., et al.: 'Multimodal biometric system using fingernail and finger knuckle'. Proc. of Int. Symp. On Computational and Business Intelligence, 2013, pp. 279–283
- [7]. Jacob, A.J., Bhuvan, N.T., Thampi, S.M.: 'Feature level fusion using multiple fingerprints', *Comput. Sci.-New Dimens. Perspect.*, 2011, 4(1), pp. 13–18
- [8]. Kang, B.J., Park, K.R.: 'Multimodal biometric method based on vein and geometry of a single finger', *IET Comput. Vis.*, 2010, 4, (3), pp. 209–217
- [9]. Michael, G.K.O., Connie, T., Teoh, A.B.J.: 'A contactless biometric system using multiple hand features', *Visual Commun. Image Represent.*, 2012, 23, pp. 1068–1084
- [10]. Ross, A., Govindarajan, R.: 'Feature level fusion in biometric systems'. Proc. of Biometric Consortium Conf. (BCC), 2004
- [11]. Yang, W., Huang, X., Zhou, F., et al.: 'Comparative competitive coding for personal identification by using finger vein and finger dorsal texture fusion', *Inf. Sci.*, 2014, 268, pp. 20–32
- [12]. Park, G., Kim, S.: 'Hand biometric recognition based on fused hand geometry and vascular patterns', *Sensors*, 2013, 13, pp. 2895–2910.
- [13]. Rattani, A., Kisku, D.R., Bicego, M., et al. 'Feature level fusion of face and fingerprint biometrics'. Proc. of Int. Conf. on BTAS, 2007, pp. 1–6.
- [14]. Srivastava, D.K., Bhambhu, L.: 'Data classification using support vector machine', *J. Theor. Appl. Inf. Technol.*, 2009, 12, (1), pp. 1–7
- [15]. Dass, S.C., Nandakumar, K., Jain, A.K.: 'A principled approach to score level fusion in multimodal biometric systems'. Proc. of Audio- and Video-Based Biometric Person Authentication, 2005, pp. 1049–1058
- [16]. Feifei, C.U.I., Gong ping, Y.A.N.G.: 'Score level fusion of fingerprint and finger vein recognition', *Comput. Inf. Syst.*, 2011, 7, (16), pp. 5723–5731
- [17]. Jain, A.K., Ross, A., Prabhakar, S.: 'An introduction to biometric recognition', *Circuits Syst. Video Technol.*, 2004, 14, (1), pp. 4–20
- [18]. Yang, J., Zhang, X.: 'Feature-level fusion of fingerprint and finger-vein for personal identification', *Pattern Recogn. Lett.*, 2012, 33, pp. 623–628
- [19]. Park, Y.H., Tien, D.N., Lee, E.C., et al.: 'A multimodal biometric recognition of touched

- fingerprint and finger-vein'. Proc. of Int. Conf. on Multimedia and Signal Processing, 2011, vol. 1, pp. 247–250
- [20]. Miura, N., Nagasaka, A., Miyatake, T.: 'Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification', Mach. Vis. Appl., 2004, 15, pp. 194–203
- [21]. Kumar, A., Ravikanth, C.: 'Personal authentication using finger knuckle surface', IEEE Trans. Inf. Forensics Sec., 2009, 4, (1), pp. 98–110
- [22]. Kumar, A., Zhou, Y.: 'Human identification using finger images', IEEE Trans. Image Process., 2012, 21, (4), pp. 2228–2244
- [23]. Miura, N., Nagasaka, A., Miyatake, T.: 'Extraction of finger vein patterns using maximum curvature points in image profiles', IEICE Trans. Inf. Syst., 2007, 8, pp. 1185–1194
- [24]. Yang, W., Yu, X., Liao, Q.: 'Personal authentication using finger vein pattern and finger-dorsa texture fusion'. Proc. of the 17th ACM Int. Conf. on Multimedia, 2009, pp. 905–908
- [25]. Prabhakar, S., Pankanti, S., Jain, A.K.: 'Biometric recognition: security and privacy concerns', IEEE Secur. Priv., 2003, 1, (2), pp. 33–42
- [26]. Deepak, A., Shirsat, S.: 'Multimodal biometric recognition system'. Proc. Of Int. Conf. on recent Innovations in Engineering and Management, 2016, pp.237–244
- [27]. Yang, X.-S.: 'Firefly algorithm, stochastic test functions and design optimisation', Int. J. Bio-Inspired Comput., 2010, 2, (2), pp. 78–84

Distributed Intrusion Detection System for Cognitive Radio Networks Based on Weighted Fair Queuing Algorithm	S. Kavitha Assistant Professor	Computer Science	International Journal of Scientific Research in Computer Science, Engineering and Information Technology	2018	2456-3307	UGC Journal No : 64718 https://ijsrcseit.com/paper/CSEIT1833232.pdf
---	-----------------------------------	------------------	--	------	-----------	---



International Journal of Scientific Research in Computer Science, Engineering and Information Technology

© 2018 IJSRCSEIT | Volume 3 | Issue 4 | ISSN : 2456-3307

Distributed Intrusion Detection System for Cognitive Radio Networks Based on Weighted Fair Queuing Algorithm

¹M. Indhumathi, ²S. Kavitha

¹Research Scholar, Department of Computer Science, Sakthi College of Arts and Science for Women, Oddanchatram, India

²Head & Associate Professor, Department of Computer Science, Sakthi College of Arts and Science For Women, Oddanchatram, India

ABSTRACT

Reliable detection of intrusion is the basis of safety in cognitive radio networks (CRNs). So far, few scholars applied intrusion detection systems (IDS) to combat intrusion against CRNs. In order to improve the performance of intrusion detection in CRNs, a distributed intrusion detection scheme has been proposed. In this paper, a method base on Dempster-Shafer's (DS) evidence theory to detect intrusion in CRNs is put forward, in which the detection data and credibility of different local IDS Agent is combined by D-S in the cooperative detection center, so that different local detection decisions are taken into consideration in the final decision. The effectiveness of the proposed scheme is verified by simulation, and the results reflect a noticeable performance improvement between the proposed scheme and the traditional method.

Keywords : Safety, cognitive radio networks, intrusion detection, IDS Agent, cooperative detection center, Dempster-Shafer's evidence theory

I. INTRODUCTION

1.1 INTRODUCTION ABOUT TO IDS

1.1.1 IDS Defined

Intrusion detection is the process of identifying computing or network activity that is malicious or unauthorized. Most all Intrusion Detection Systems (IDS) have a similar structure and component set. This consists of a sensor (or agent) that monitors one or more data sources, applies some type of detection algorithm, and then initiates zero or more responses. Usually there is a management system that provides for monitoring, configuration and analysis of intrusion data.

1.1.2 Evolution of IDS

The first IDS were host-based, and looked at system operating logs performing simple pattern matches

against a small set of signatures. This approach quickly expanded to systems that looked at network traffic, initially also for simple patterns. As IDS gained a level of protocol-awareness, they were able to look for certain single packet traffic types known to be malicious, examining the source and destination IP addresses, along with source and destination ports. Further sophistication brought an awareness of network sessions and the ability to examine dialogs between systems for multi-packet activity. More recent IDS can examine and respond to entire conversations between hosts, using knowledge of protocols and network sessions to analyze traffic for malicious activity based on how that traffic would appear at the destination-a task often requiring specialized network drivers to operate at full wire-speed (For a good discussion of the evolution and genealogy of IDS, see article by Inella). The emerging class of IDS take this one step

further by combining log analysis, along with information from other IDS and anti-virus software to correlate events in an effort to identify and respond to intrusions in real time.

1.1.3 Relation of IDS to dIDS

From the above, it is clear that as IDS grow in function and evolve in power, they also evolve in complexity. Agents of each new generation of IDS use agents of the previous generation as data sources, applying ever more sophisticated detection algorithms to determine ever more targeted responses. Often, one or more IDS and management system(s) may be deployed by an organization within its own network, with little regard to their neighbors or the global Internet. Just as all individual networks and intranets connect to form "The Internet", so can information from stand-alone

A secure computer system provides guarantees regarding the confidentiality, integrity, and availability of its objects (such as data, purpose, or services). However, systems generally contain design and implementation flaws that result in security vulnerabilities. An intrusion can take place when an attacker or a group of attackers exploits the vulnerabilities and thus damages the confidentiality, integrity or availability guarantees of a system. Intrusion Detection Systems (IDSs) detect some set of intrusions and execute some predetermined actions when an intrusion is detected. Over the last one and half decade, research in the field of intrusion detection has been heading towards a distributed framework of systems that do local detection and provide information to perform global detection of intrusions.

These distributed frameworks of intrusion detection have some advantages over single monolithic frameworks. Most of these distributed systems are hierarchical in nature. The local intrusion detection components look for local intrusions and pass their analysis results to the upper levels of the hierarchy.

The components at the upper levels analyze the refined data from multiple lower level components and attempt to establish a global view of the system state. However, such IDSs are not fully distributed systems because of the centralized data analysis performed at the higher levels of the hierarchy. An agent-based architecture is proposed for performing intrusion detection in a distributed environment. By employing a suitable communication mechanism, the resource overhead is minimized in the distributed intrusion detection process.

1.2 INTRODUCTION TO DIDS

Some of the existing distributed IDS frameworks are discussed briefly. DIDS is a distributed intrusion detection system consisting of host managers and LAN managers doing distributed data monitoring, and sending notable events to the DIDS director. These managers also do some local detection, passing the summaries to the director. The director analyzes the events to determine the security state. AAFID is distributed IDS developed in CERIAS at Purdue University. It employs agents at the lowest level of the hierarchy for data collection and analysis and transceivers and monitors at the higher levels for controlling agents and obtaining a global view of activities. It provides a subscription-based service to the agents.

A prototype called the Hummingbird System is developed at University of Idaho. It is a distributed system that employs a set of Hummer agents, each assigned to a single host or a set of hosts. Each Hummer interacts with other hummers in the system through a manager, a subordinate, and the peer relationships. It enables a system administrator to monitor security threats on multiple computers. Architecture of an intrusion detection system using a collection of autonomous agents has been proposed in. In cooperation and communication model proposed by the authors, agents request and receive information solely on the basis of their interests. They can specify new interests as a result of a new

event or alert. This avoids unnecessary data flow among the agents.

However, most of these intrusion detection systems have the following drawbacks: (i) **Analysis hierarchy**: as there is a hierarchy in data analysis these systems are very difficult to modify. Changes may have to be made at many levels if any new distributed attack is developed. (ii) **Data refinement**: when a module from a lower level sends results of analysis to a higher level, some data refinement is done. However, the knowledge of what events are important in a system-wide level is difficult to anticipate at the lower levels of the hierarchy, and thus data refinement may result in loss of important information.

1.3 Wireless Sensor Networks Applications



Figure-1: Wireless Sensor Networks Applications

- (i) These networks are used in environmental tracking, such as forest detection, animal tracking, flood detection, forecasting and weather prediction, and also in commercial applications like seismic activities prediction and monitoring.
- (ii) Military applications, such as tracking and environment monitoring surveillance applications use these networks. The sensor nodes from sensor networks are dropped to the field of interest and are remotely controlled by a user. Enemy tracking, security detections are also performed by using these networks.
- (iii) Health applications, such as Tracking and monitoring of patients and doctors use these networks.

(iv) The most frequently used wireless sensor networks applications in the field of Transport systems such as monitoring of traffic, dynamic routing management and monitoring of parking lots, etc., use these networks.

(v) Rapid emergency response, industrial process monitoring, automated building climate control, ecosystem and habitat monitoring, civil structural health monitoring, etc., use these networks.

Wireless Sensor network (WSN) is a network system comprised of spatially distributed devices using wireless sensor nodes to monitor physical or environmental conditions, such as sound, temperature, and motion. Cluster-based data transmission in WSNs has been investigated by researchers to achieve the network scalability and management, which maximizes node lifetime and reduce bandwidth consumption by using local collaboration among sensor nodes. A CH aggregates the data collected by the leaf nodes (non-CH sensor nodes) in its cluster, and sends the aggregation to the base station (BS). To prevent quick energy consumption of the set of CHs, LEACH randomly rotates CHs among all sensor nodes in the network, in rounds. LEACH achieves improvements in terms of network lifetime. In this paper, for convenience, we call this sort of cluster-based protocols as LEACH-like proto-cols. Researchers have been widely studying CWSNs in the last decade in the literature.

The multi-constrained QoS routing is NP-hard and heuristic algorithms are proposed to find solution for the problem. But these algorithms are too complex and cannot obtain best global solution. QoS may be more accurately determined by using fuzzy logic instead of static values. Fuzzy Inference System (FIS) accepts more number of uncertain and imprecise data as inputs and thereby achieves flexibility, robustness, and low cost solution. But, FIS uses human-determined membership functions (MFs) that are fixed. Therefore, they are rarely optimal in

terms of reproducing the desired outputs. Tuning membership functions of parameters is a time consuming task. Neural networks overcome most of the complex problems to adapt dynamically to the system operating conditions, and to make correct decisions, if the signals are uncertain. But the integration of neural network into the fuzzy logic system makes it possible to learn from prior obtained data sets. This paper proposes an approach which integrates both neural and fuzzy techniques to select a server from a number of group members belonging to any cast group by considering QoS constraint route and server with higher stability in MANETs. This section presents some of the related works, software agent concept and our contributions.

The set of destinations is identified by unique any cast address and provide the same services. Searching for services on networks often depends on the broadcast or multicast mechanism to acquire the information, which usually results in large overhead. It will be a serious problem in ad-hoc wireless networks, where the bandwidth is limited and each node moves arbitrarily. Any casting scheme in ad hoc wireless networks can simplify access management in distributed service, improve the robustness and performance of an ad hoc network when mobility and link disconnections are frequent, and reduces the communication overhead.

The source node does not need to know about picking a single server and is determined by routing scheme. The server in any cast routing may be chosen by minimum hops, delay or other metrics. Any casting along the minimum hops path may result in inefficient use of network resources, because it forwards packets along already congested shortest path, and also may not satisfy the Quality of Service (QoS) requirements for multimedia and real time application services. Set of mobile or semi mobile nodes with no available pre-established communications is a MANET Forming a short-term network. Each mobile node in the network acts as a

computer switching program that transfers incoming messages to outgoing links via the most efficient route possible, e.g. over the Internet i.e., a router. This kind of networks are characterize by the relationships between parts linked together in a system such as a computer network topologies, continuation of bandwidth constrain and variable capacity links, energy constrain operations and are highly intensity to security threats. Due to all these characteristics routing is a major issue in ad hoc networks. The routing protocols for ad hoc networks have been classified as: (a) Proactive or table driven for example Destination Sequenced Distance Vector (DSDV) and Optimized Link State Routing (OLSR), (b) Reactive/On-demand, e.g. Dynamic Source Routing Protocol, Ad hoc On-Demand Distance Vector routing protocol, Temporally Ordered Routing Algorithm. In table driven or proactive routing, each node has one or more tables that include the latest information of the routes to any node in the network. Each row has the subsequently hop for reaching to a node or subnet and the cost of this route. Different table-driven protocols vary in the way the information about alter in topology is spread through all nodes in the network. The two kinds of table keep informed in proactive protocols are the periodic update and the triggered update.

1.4 COGNITIVE RADIO NETWORK

Cognitive radio (CR) is a form of wireless communication in which a transceiver can intelligently detect which communication channels are in use and which are not, and instantly move into vacant channels while avoiding occupied ones. This optimizes the use of available radio-frequency (RF) spectrum while minimizing interference to other users.

In its most basic form, CR is a hybrid technology involving software defined radio (SDR) as applied to spread spectrum communications. Possible functions of cognitive radio include the ability of a transceiver to determine its geographic location,

identify and authorize its user, encrypt or decrypt signals, sense neighboring wireless devices in operation, and adjust output power and modulation characteristics.

II. LITERATURE REVIEW

Jaydip Sen, A Survey on Wireless Sensor Network Security [1] Wireless sensor networks (WSNs) have recently attracted a lot of interest in the research community due their wide range of applications. Due to distributed nature of these networks and their deployment in remote areas, these networks are vulnerable to numerous security threats that can adversely affect their proper functioning. This problem is more critical if the network is deployed for some mission-critical applications such as in a tactical battlefield. Random failure of nodes is also very likely in real-life deployment scenarios. Due to resource constraints in the sensor nodes, traditional security mechanisms with large overhead of computation and communication are infeasible in WSNs. Security in sensor networks is, therefore, a particularly challenging task. This paper discusses the current state of the art in security mechanisms for WSNs. various types of attacks are discussed and their countermeasures presented. A brief discussion on the future direction of research in WSN security is also included. In addition to traditional security issues like secure routing and secure data aggregation, security mechanisms Deployed in WSNs also should involve collaborations among the nodes due to the decentralized nature of the networks and absence of any infrastructure. In real-world WSNs, the nodes cannot be assumed to be trustworthy apriori. Researchers have therefore, focused on building a sensor trust model to solve the problems which are beyond the capabilities of traditional cryptographic mechanisms. In this chapter, we present a survey of the security issues in WSNs. First we outline the constraints of WSNs, security requirements in these networks, and various possible attacks and the corresponding countermeasures. Then a holistic view

of the security issues is presented. These issues are classified into six categories: cryptography, key management, secure routing, secure data aggregation, intrusion detection and trust management. The advantages and disadvantages of various security protocols are discussed, compared and evaluated. Some open research issues in each of these areas are also discussed.

Efficient and Secure Routing Protocol for Wireless Sensor Networks through SNR based Dynamic Clustering Mechanisms [2] Advances in Wireless Sensor Network Technology (WSN) have provided the availability of small and low-cost sensor with capability of sensing various types of physical and environmental conditions, data processing and wireless communication. In WSN, the sensor nodes have a limited transmission range, and their processing and storage capabilities as well as their energy resources are limited. Triple Umpiring System (TUS) has already been proved its better performance on Wireless Sensor Networks. Clustering technique provides an effective way to prolong the lifetime of WSN. In this paper, we modified the Ad hoc on demand Distance Vector Routing (AODV) by incorporating Signal to Noise Ratio (SNR) based dynamic clustering. The proposed scheme Efficient and Secure Routing Protocol for Wireless Sensor Networks through SNR based dynamic Clustering mechanisms (ESRPSDC) can partition the nodes into clusters and select the Cluster Head (CH) among the nodes based on the energy and Non Cluster Head (NCH) nodes join with a specific CH based on SNR Values. Error recovery has been implemented during Inter cluster routing itself in order to avoid end-to-end error recovery. Security has been achieved by isolating the malicious nodes using sink based routing pattern analysis. Extensive investigation studies using Global Mobile Simulator (GloMoSim) showed that this Hybrid ESRP significantly improves the Energy efficiency and Packet Reception Rate (PRR) compared to SNR unaware routing algorithms like Low Energy Aware Adaptive Clustering Hierarchy (LEACH) and Power-Efficient Gathering

in Sensor Information Systems (PEGASIS). Sensor Network Wireless is widely considered as one of the most important technologies for the twenty-first century. The sensing electronics measure ambient conditions related to the environment surrounding the sensors and transform them in to an electrical signal. In many WSN applications, the deployment of sensor

Node informing that the misuse IDS system is operational. The messages sent to the Central IDS Node are formatted using the extended signed IDMEF format. In addition, the upper tier process listens for commands from the Central IDS Node. It receives parameters for the rate limiting of alert messages, configuration for the Snort process and new attack signatures.

III. METHODOLOGY

(i) Security Agents

3.1. Misuse Detection Agent

As we previously mentioned, each security agent consists of two tiers. The lower tier comprises of the process that handles the misuse detection within our network. Snort [6] has been chosen as the misuse IDS software for our system. Snort is a libpcap-based [7] software that can be used as a sniffer, packet logger or network intrusion detection system. In our case, we used Snort as a misuse intrusion detection tool. The detection of malicious packets is based on known attack signatures. Snort is able to detect a variety of attacks such as DoS/DDoS attacks, Portscans, HTTP, DNS, SMTP, IMAP, POP3 attacks and Virus/Worm attacks.

Alerts generated from Snort are passed to the upper tier of our agent. The upper tier of the Misuse Detection Agent receives alert messages from the lower tier and stores them for a defined period of time in a buffer. For every different case of attack, that is, source IP address and port, target IP address and port and known attack signature, the upper tier process uses a unique alert identification. Rate limiting is achieved independently for different types of attacks, sending the alert message only once in the specified period of time.

Agent's upper tier process is also responsible for sending the heartbeat messages to the Central IDS

3.2. Anomaly Detection Agent

For the lower tier of the Anomaly Detection Agent we developed a prototype anomaly detection tool [8] that currently focuses on DoS Attacks. The prototype tool consists of two main modules: the collector and the detector. The collector module is responsible for asynchronously receiving flow data from the Netflow-enabled [9] router; information is analyzed, mean values and adaptive thresholds are calculated and stored in a local data structure.

The tool extracts and stores packet and flow counters per destination IP address, as well as total counters and mean values for each pair of input-output interfaces. The detector process is responsible for calculating the metrics for the interface pairs stored by the collector, and comparing the results to detection thresholds. It is periodically activated, implements extensive logging of detection events and generates notifications with security alerts which are sent to the upper tier.

The upper tier process receives the alerts and sends them to the Central IDS Node using the signedIDMEF Format. Moreover, the Central IDS Node adjusts Anomaly Detection Agent's parameters (metrics and thresholds for the DoS attack detection algorithm).

3.3. SNMP Query Agent

As the other two agents, the SNMP Query Agent is comprised of two tiers. The lower tier process is a

custom SNMP client that performs SNMP queries at the routers of the network. Values like CPU and memory usage, active and inactive flows are polled from routers at specific intervals. The upper tier accepts the values from the SNMP queries and forwards them to the Central IDS Node after formatting them using the signed-IDMEF data model. The upper tier process is also responsible for sending heartbeat messages to inform the Central IDS Node that the SNMP client is operational. Instructions from the Central IDS Node are sent to the SNMP Query Agent, giving information about the router and the SNMP objects to be polled.

(ii) Intrusion Detection System

Intrusion detection mechanism can detect malicious behavior on the network and identify malicious users. So Intrusion detection mechanism can protect the reliability of the network, especially it is more important in distributed cognitive radio network which absents center facilities. The traditional intrusion detection system (IDS) was proposed by Denning in 1987. It is composed of main body, object, audit record, activity profile and exception record and activity rules. A more detailed description of IDES is given as follows.

There are six main parts in the IDS model [12].

1) Subject: Active initiator in the system operation, the entity that moves on the target system, such as the process of the computer operating system, the service connection of the network and so on.

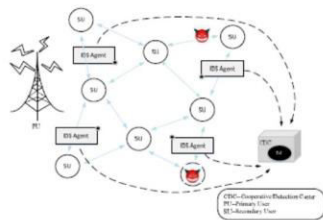


Figure-2:. System of Distribute Intrusion Detection in CRNs

2) Object: Resources that are managed by the system- files, devices, commands, for example.

3) Audit Records: when a subject operates on an object, such as user registration, command execution,

and file access, data will be produced by the target system.

4) Activity Profile: Preserve the information about the normal activity, and the specific implementation depends on the detection method.

5) Anomaly Record: Used to indicate the occurrence of an abnormal event. The format is <Event, Time-Stamp, Profile>

6) Activity rule: An action that should be taken when an audit record, update profile, an exception record, detect relate anomalies to some suspected intrusions or abnormal behavior is produced.

Actually, the Denning model is can be described as a rule based pattern matching system. After generating an audit record, it will match against the profiles. Then type information will determine the rule to report anomalies detection. It's largely system-independent for the rules and profile structures to establish profile templates. Not all of the IDS can be fully consistent with the model.

IV. CONGESTION CONTROL FOR WIRELESS SENSOR NETWORK

4.1 Congestion Control

The flow type is of high importance to guide a real congestion control. Flow types may include a single packet, few packets, a large number of packets, which require light control, medium level control, and tight control, respectively. When a large number of nodes transmit information, their flows will cross at intermediate nodes. This high number of sources increases the congestion but helps improving the reliability. For example, in tree architectures, every intermediate node can suffer from congestion causing packet loss, which in turn decreases network performance and throughput and cause energy waste. It is very difficult to predict the intersection points due to network dynamics (addition or removal of sensors or a change in the report rate), variability in radio channel quality over time. All these can

transform uncongested parts of the network to under-provisioned and congested regions. The area around the intersection will become a hotspot and there is a possibility of congestion (buffer overflow) and contention (links interference). For these reasons, a congestion control algorithm for data packet transmission is necessary.

Contention-based Congestion: when many nodes within range of one another attempt to transmit simultaneously, losses occur due to interference and packet loss is engendered. If the packet generation rate is sufficiently small, simultaneous transmission becomes independent of the rate. Rather, it depends on the exact time generation of the packet. Explicit local synchronization (or also named phase shifting) among neighbours can reduce this type of loss, but it cannot eliminate the problem as non-neighbouring nodes can still interfere (hidden nodes). The contention may happen between different flows in the same area, and between different packets of the same flow, especially in the case of high density networks. Consequently, the nodes' channel capacity becomes time-variant. **Buffer-based Congestion:** each node uses a buffer for the packets waiting to be sent. The overflow of this buffer causes congestion and packets loss. This is due to high reporting rate that varies in time due to dynamic channel conditions. The many-to-one nature (or converge cast) of WSNs causes congestion, in addition to the other causes shared with general wireless networks.

4.2 Congestion Detection Strategies

Many congestion detection mechanisms are used and tested. The most used are: packet loss, queue length, packet service time, the ratio between packet service time and packet inter-arrival time, delay. In many cases, a single parameter cannot indicate congestion accurately.

Packet loss: It can be measured at the sender if ACKs (Acknowledgements) are used; this suggests reliability to be ensured by the protocol. It can also

be measured at the receiver with sequence numbers use. Further, CTS (Clear To Send) packet loss can be used as congestion indication.. Not overhearing the parent's forwarding on the upstream link, by a child node over the downstream link, can be used as an indication for packet loss as well. The time to repair losses (if reliability ensured) can be used as a congestion indication. Loss ratio is also used in some protocols.

Queue length: As each node has a buffer; its length can serve a simple and good indication of congestion a fixed threshold is used and the congestion is signalled as soon as the buffer length exceeds this threshold , the remaining buffer length from the overall size is used. In the difference between the remaining buffer and the traffic rate is used as congestion indication. The traffic rate represents the excess rate, which is the difference between the output rate and the sum of sourced and forwarded rates. In the buffer length is used in addition to the difference of output and input time, which is quite similar to output and input rate. In buffer length and capacity of the node are used together. The number of non-empty queues can indicate congestion level. When there is a congestion, this number is larger than 0. This number increases with network load. If the link layer applies retransmissions, link contention will be reflected through buffer length.

Queue length and Channel load: In case of increase in packets collision, and after several unsuccessful MAC (Medium Access Control) retransmissions, packets are removed. Consequently, the decrease in buffer occupancy due to these drops may mean the absence of congestion when only buffer state is used for congestion detection. Therefore, for accurate congestion detection, a hybrid approach is required using queue length and channel loading as a congestion indication. Channel busyness ratio or channel load is the ratio of time intervals when the channel is busy (successful transmission or collision) to the total time. In the authors use the busyness channel ratio, similarly to channel load, but apply it

to a subset of nodes, and queue length for another set of nodes. The node activates channel monitoring only when it receives a packet to forward. Therefore, there is no overhead to measure channel loading.

4.3 Channel Busyness Ratio and Throughput Measurement

Throughput is addition to channel busyness to take into account the effects of hidden nodes problem in multi-hop environment. The throughput quantifies the number of successful transmissions.

Packet service time: The inverse of packet service rate, it is the interval between packet arrival at the MAC layer and its successful transmission. It covers packet waiting, collision resolution, and packet transmission time at the MAC layer.

The congestion control cannot be decoupled from the MAC protocol, and adequate protocol should first be used to avoid congestion. In applications where the event cannot be known a priori, random access contention based MAC protocols are necessary (CSMA "Carrier Sense Multiple Access"-based). Continuous periodic applications with high rate a TDMA "Time Division Multiple Access"-like scheme is more appropriate.

V. RESULTS AND DISCUSSION

5.1 EXPERIMENTAL SETUP

Congestion in a network may occur if the load on the network the number of packets sent to the network is greater than the capacity of the network the number of packets a network can handle. Congestion control refers to the mechanisms and techniques to control the congestion. The congestion control having a different type of models but it have some disadvantage .To overcome the drawbacks, we proposed evolutionary algorithm, ant colony algorithm to get the optimal solution for the congestion control. In order to avoid congestion delays, the ant colony optimization paradigm is

explored to find a optimize routes and to proposed routing algorithms are simple yet efficient. The routing optimization is driven by the minimization of total latency during packets transmission between the tasks.

5.2 DYNAMIC SOURCE ROUTING (DSR)

Genetic algorithms are a part of evolutionary computing. It is also an efficient search method that has been used for path selection in networks. These stochastic search algorithms are based on the principle of natural selection and recombination. GA has been an efficient search method based on principles of natural selection and genetics. They are being applied successfully to find acceptable solutions to problems in business, engineering, and science.

We can find good solution for adequate amount of data at hand, but the complexity of data increases as GA takes time to find the solution. GA works well for network model to find the optimal path. In this, the source and the destination nodes are sure to participate in every generation. Other nodes or the genes become a part of the chromosome if they find an optimal path between the source and destination. GA is composed with a set of solutions, which represents the chromosomes. This composed set is referred to population. Population consists of set of chromosome which is assumed to give solutions. From this population, we randomly choose the first generation from which solutions are obtained. These solutions become a part of the next generation. Within the population, the chromosomes are tested to see whether they give a valid solution. This testing operation is nothing but the fitness functions which are applied on the chromosome. Operations like selection, crossover and mutation are applied on the selected chromosome to obtain the progeny. Again fitness function is applied to these progeny to test for its fitness. Most fit progeny chromosome will be the participants in the next generation. The disadvantage of this protocol is that the route maintenance

mechanism does not locally repair a broken link. Stale route cache information could also result in inconsistencies during the route reconstruction phase. The connection setup delay is higher than in table-driven protocols. Even though the protocol performs well in static and low-mobility environments, the performance degrades rapidly with increasing mobility. Also, considerable routing overhead is involved due to the source-routing mechanism employed in GAs. This routing overhead is directly proportional to the path length.

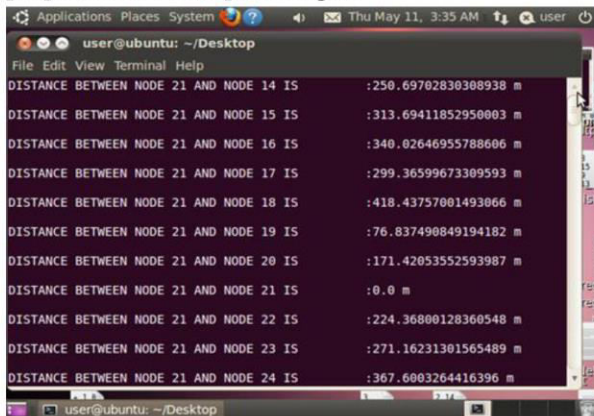


Figure-3: Calculate the Distance between all the Nodes

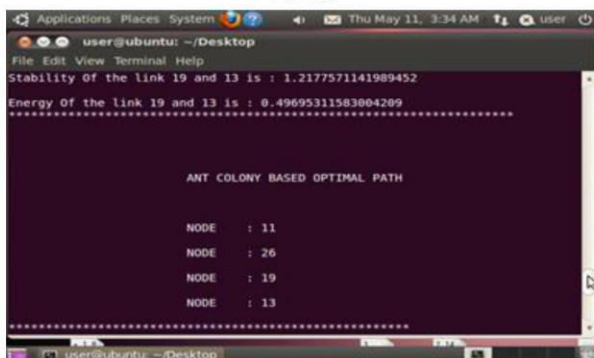


Figure-4: After the Selection of Source and Destination Calculate the Distance between their Neighbouring Nodes

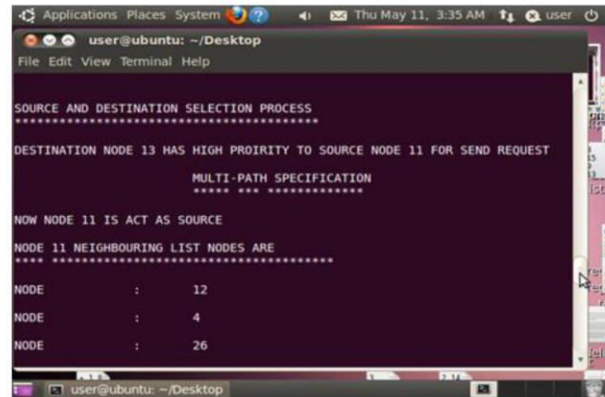


Figure-5: List out the Neighbouring Node

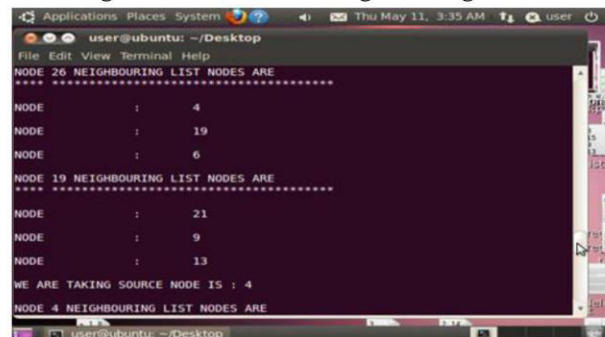


Figure-6: Finally Display the Optimal Path between Source and Destination

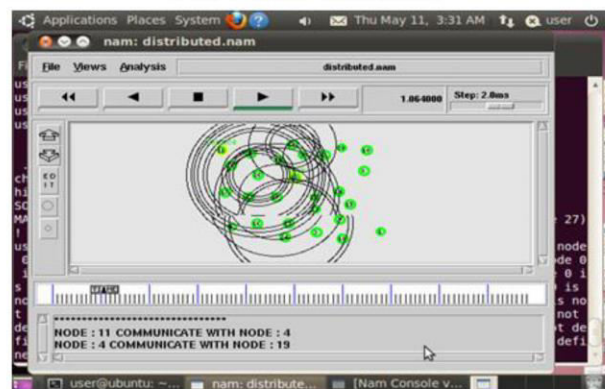


Figure-7: Neighbouring Path between the Source and Destination

VI. CONCLUSION

I propose distributed IDS for CRNs based on evidence theory. Aim to get more accurate final detection result making at CDC, we apply D-S theory of evidence to combine different detection data and credibility from every IDS Agents. Simulations presented show that the proposed system performs more excellent than the traditional Weighted Fair Queuing (WFQ) Combination algorithm.

VII. FUTURE ENHANCEMENT

In my future work, I would like to work on any cast routing protocols to check the efficiency under high throughput applications, e.g. multimedia applications by employing negotiation parameters in route request packet in finding nearest server through non congestion paths.

VIII. REFERENCES

- [1]. T. Hara, V.I. Zadorozhny, and E. Buchmann, *Wireless Sensor Network Technologies for the Information Explosion Era, Studies in Computational Intelligence*, vol. 278. Springer-Verlag, 2010.
- [2]. Y. Wang, G. Attebury, and B. Ramamurthy, "A Survey of Security Issues in Wireless Sensor Networks," *IEEE Comm. Surveys & Tutorials*, vol. 8, no. 2, pp. 2-23, Second Quarter 2006.
- [3]. A.A. Abbasi and M. Younis, "A Survey on Clustering Algorithms for Wireless Sensor Networks," *Computer Comm.*, vol. 30, nos. 14/15, pp. 2826-2841, 2007.
- [4]. W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "An Application-Specific Protocol Architecture for Wireless Micro sensor Networks," *IEEE Trans. Wireless Comm.*, vol. 1, no. 4, pp. 660-670, Oct. 2002.
- [5]. A. Manjeshwar, Q.-A. Zeng, and D.P. Agrawal, "An Analytical Model for Information Retrieval in Wireless Sensor Networks Using Enhanced APTEEN Protocol," *IEEE Trans. Parallel & Distributed Systems*, vol. 13, no. 12, pp. 1290-1302, Dec. 2002.
- [6]. S. Yi et al., "PEACH: Power-Efficient and Adaptive Clustering Hierarchy Protocol for Wireless Sensor Networks," *Computer Comm.*, vol. 30, nos. 14/15, pp. 2842-2852, 2007.
- [7]. K. Pradeepa, W.R. Anne, and S. Duraisamy, "Design and Implementation Issues of Clustering in Wireless Sensor Networks," *Int'l J. Computer Applications*, vol. 47, no. 11, pp. 23-28, 2012.
- [8]. L.B. Oliveira et al., "Sec LEACH-On the Security of Clustered Sensor Networks," *Signal Processing*, vol. 87, pp. 2882-2895, 2007.
- [9]. K. Zhang, C. Wang, and C. Wang, "A Secure Routing Protocol for Cluster-Based Wireless Sensor Networks Using Group Key Management," *Proc. Fourth Int'l Conf. Wireless Comm., Networking and Mobile Computing (WiCOM)*, pp. 1-5, 2008.
- [10]. P. Banerjee, D. Jacobson, and S. Lahiri, "Security and Performance Analysis of a Secure Clustering Protocol for Sensor Networks," *Proc. IEEE Sixth Int'l Symp. Network Computing and Applications (NCA)*, pp. 145-152, 2007.
- [11]. K. Zhang, C. Wang, and C. Wang, "A Secure Routing Protocol for Cluster-Based Wireless Sensor Networks Using Group Key Management," *Proc. Fourth Int'l Conf. Wireless Comm., Networking and Mobile Computing (WiCOM)*, pp. 1-5, 2008.

Fuzzy Based Genetic Operators for Cyber Bullying Detection Using Social Network Data	M. Chitra Devi Assistant Professor	Computer Science	International Journal of Scientific Research in Computer Science, Engineering and Information Technology	2018	2456-3307	UGC Journal No : 64718 https://ijsrcseit.com/paper/CSEIT1833243.pdf
--	---------------------------------------	------------------	--	------	-----------	---



International Journal of Scientific Research in Computer Science, Engineering and Information Technology

© 2018 IJSRCSEIT | Volume 3 | Issue 4 | ISSN : 2456-3307

Fuzzy Based Genetic Operators for Cyber Bullying Detection Using Social Network Data

¹M.Devi ²M. Chitra Devi

¹Research Scholar, Department of Computer Science, Sakthi Arts and Science College For Women, Oddanchatram, Tamil Nadu, India

²PG Head & Associate Professor, Department of Computer Science, Sakthi Arts and Science College For Women, Oddanchatram, Tamil Nadu, India

ABSTRACT

Social media getting more and more popular in our day today life. By the popularity of the social media affects the people who involving into it. This makes the technology to work or to feel smarter and makes us lazier. On resulting to this robust and discriminative numerical representation learning of text messages is a critical issue. Hence here we propose a learning method to tackle this issue which is named as Semantic Enhanced Marginalized Denoising Auto Encoder (smsda). Semantic extension of the popular deep learning model stacked denoising auto encoder plays a major role in this method whereas semantic extension consists of semantic dropout noise and sparsity constraints. The semantic dropout noise is designed based on domain knowledge and the word embedding technique. Our proposed method is able to exploit the hidden feature structure of bullying information and learn a robust and discriminative representation of text. Comprehensive experiments on two public cyber bullying corpora (Twitter and myspace) are conducted, and the results show that our proposed approaches outperform other baseline text representation learning methods.

Keywords: Semantic Enhanced Marginalized Denoising Auto-Encoder, cyberbullying.

I. INTRODUCTION

Internet has become very popular and used around the world in our day to day life. By the growing of internet the cyber security is becoming the most important factor. Currently web 2.0 allows us to access the online related services and some users have been affected by the cybercrimes like cyber bullying experiences internationally. By these kinds of issues the growth of social media gets the negative impacts from the various users. We propose an effective predator and victim identification with semantic enhanced marginalized denoising auto-encoder approach to detect cyber-bullying message from social media through the weighing scheme of feature of selection. We present Model to extract the

cyber bullying network, which is used to identify the most active cyber bullying predators and victims to ranking algorithms the existing filters generally work with the simple key word search and are unable to understand the Semantic meaning of the text. So we propose Semantic Enhanced Marginalized Denoising Auto-Encoder.

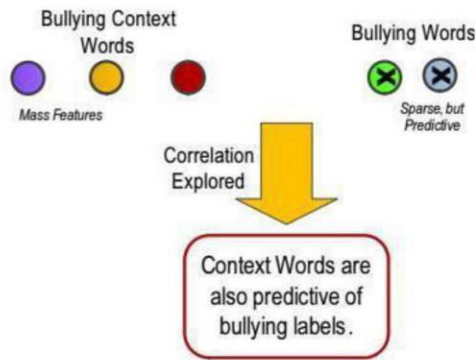


Figure-1: Cyberbullying

Cyberbullying is an increasingly important and serious social problem, which can negatively affect individuals. It is defined as the phenomena of using the internet, cell phones and other electronic devices to willfully hurt or harass others. Due to the recent popularity and growth of social media platforms such as Facebook and Twitter, cyberbullying is becoming more and more prevalent. It has been identified as a serious national health concern by the American Psychological Association 1 and the White House 2 . In addition to that, according to the recent report by National Crime Prevention Council, more than 40% of the teens in the US have been bullied on various social media platforms.

II. RELATED WORK

A Sexual Predator Identification competition took place for the first time at PAN-2012. Given a set of chat logs the participants had to identify the predators among all users in the different conversations or the part (the lines) of the conversations which are the most distinctive of the predator behavior. In conclusion, it is impossible to identify predators using a unique method but it is necessary the use of different approaches. Moreover the most effective method for identifying distinctive lines of the predator behavior in a chat log appeared to be those based on filtering on a dictionary or LM basis [4]. Yin et al., was the sole submission in the misbehavior detection task of CAW 2.0. Using three from the five datasets which were provided by the

organizers of the content analysis workshop, they proposed a supervised learning approach for detecting harassment with a focus on detecting intentional annoyance. By employing a SVM classifier with the linear kernel and combining TF-IDF measure as local features, sentiment features, and contextual features of documents proved that identification of online harassment provide significantly improved performance when TF-IDF is supplemented with sentiment and contextual feature attributes. The results show improvements over the baselines. In a recent study on cyberbullying detection, Kontostathis et al., [6] taking a collection of posts from the website Formspring.me, which allows users to post questions anonymously (a question-answer website where users openly invite others to ask and answer questions) proposed a "bag-of-words" language model, which based on the text in online posts, in order to detect instances of cyberbullying. Moreover, they exploited a supervised machine learning called Essential Dimensions of LSI (EDLSI) approach in order to identify additional terms of cyberbullying in Formspring.me data. The data was labeled using a web service, Amazon's Mechanical Turk.

The Mechanical Turk (MTurk) is a crowdsourcing Internet marketplace that enables individuals or businesses (known as Requesters) to co-ordinate the use of human intelligence to perform tasks that computers are currently unable to do. It is one of the sites of Amazon Web Services. The goal was to identify the most commonly used cyberbullying term.

III. STATEMENT OF THE PROBLEM

This research is discussed as follows:

- (i) Three kinds of information such as text, demography and social features are used for detecting the cyberbullying messages. Hence, text based cyberbullying detection framework is required.

- (ii) Each autoencoder layer is intended to learn an increasingly abstract representation of the input.
- (iii) Fuzzy rules are used for labeling the cyberbullying messages.
- (iv) In addition to genetic algorithm is used for optimizing the parameters for labeling systems.
- (v) The correlation information discovered by fuzzy rule generation helps to reconstruct bullying features from normal words, and this in turn facilitates detection of bullying messages without containing bullying words.

IV. CYBERBULLY ACTIVITIES

In the proposed framework for detecting cyberbully activities, following steps have been included:

- Data Pre-processing
- Feature Extraction
- FuzGen learning algorithm
- Naïve classifier technique

4.1. Data Pre-Processing

The data pre-processing is an important phase in representing data in feature space to the classifiers. Social network data are noisy, thus pre-processing has been applied to improve the quality of the research data and subsequent analytical steps, and this includes removing stop words, unwanted characters, etc.

4.2. Feature Extraction

This module is used for extracting the data required from the processed data. The part of speech for every word in the conversation is obtained using natural language processing technique and then features like Noun, Adjective and Pronoun are extracted from the tagged output and statistics on occurrence of word in the text are also extracted.

4.3. FuzGen Learning Technique

The learning module incorporates the adaptive component of the system by means of a GA with

fuzzy set genes. GAs are adaptive search and optimization algorithms that work by mimicking the principles of natural genetics (Deb, 1996). In the proposed system, the function to be optimized is a hypothetical representation of cyberbully terms in the Social Network.

In the following, the elements of the GA model, namely: the fuzzy gene types and the GA operators are presented.

4.3.1. The Fuzzy Set Genes

A gene G is, $G = (t, g, \text{ and } c)$, where
 t is frequency of the term,
 g identifies the gene type and
 c is a non-negative real number

When $G(t=c)$, gene type represents the occurrences of a cyberbully term.

When $G(t < c)$. This gene type is completely satisfied by dataset that have no occurrences of the cyberbully term t .

When $G(t \geq c)$. Genes of this type are satisfied completely by dataset with at least c occurrences of the cyberbully term t .

4.3.2. The GA Operators

Selection, crossover, and mutation are the genetic operators of evolutionary process. Choice of chromosomes from population to reproduce is done by selection. Using crossover an offspring chromosome is produced by taking sequences of genes from each of two parent chromosomes selected and combining them. The mutation is the random alteration of a gene in the chromosome selected.

4. 4 Advantages

The main advantages of this research are:

- (i) These robust features are learned by reconstructing original input from corrupted (i.e., missing) ones. The new feature space can improve the performance of cyberbullying detection even with a small labeled training corpus.

- (ii) These specialized modifications make the new feature space more discriminative and this in turn facilitates bullying detection.
- (iii) Comprehensive experiments on real-data sets have verified the performance of our proposed model.

V. CYBERBULLYING ALGORITHM

Input: Conversation dataset from Social Network.

Step-1: Current population is assigned to the initial population.

Step-2: Evaluate the current population with the fuzzy rule set given as knowledge base.

Step-3: The fitness value of the current population is calculated using the function EvalPop ().

Step-4: The current population is considered as best population since it is the initial population.

Step-5: The fitness value of the current population is assigned as the best fitness value.

Step-6: The size of the term set retrieved from input is assigned as null.

// For Parent selection

Step-7: The size of the current term set is compared with the size of evolved term set, Ne, if the size of N is less than Ne, then the following steps takes place.

Step-8: The offspring population is initialized as null

Step-9: If the size of offspring population is less than current population then following steps will be executed

Step-10: Parents are selected by using the tournament selection mechanism and children are created by using mutation and cross over mechanism, where Tournament selection is a method of selecting an individual from a population of individuals in a genetic algorithm.

Step-11: Once the offspring population is created, it is joined to current population.

Step-12: End of while loop.

Step-13: Evaluate the Fuzzy rule set for the offspring population.

Step-14: Once the offspring population is created, it is joined to current population.

Step-15: Token competition is carried out to obtain the best individuals from the joint population.

Step-16: The Joint population is assigned to the current population.

Step-17: The Fitness value of the joint population is calculated using the function EvalCurPop(). // Updating the Best fitness value and Best population for obtaining classified output.

Step-18: Fitness values of the current population are checked with the best fitness value. If the current fitness value is greater than following steps occur

Step-19: The best fitness value is updated with the current fitness value.

Step-20: The best population is updated with the current population.

Step-21: End of if loop.

Step-22: The size of the current term set is incremented.

Step-23: End of while loop.

Output: Identified Cyberbully terms and their type from the input dataset

VI. RESULT AND DISCUSSION

We used two social media datasets, namely Twitter and MySpace for the problem we study. Both datasets contain labeled social media post. Twitter is a micro blogging website which allows users to post 140 characters messages called "Tweets". The retweets are removed from the dataset. The posts in this dataset have been manually labeled as bully or normal. MySpace is a social networking website which allows a registered users to view pictures, read chat and check other users profile information. The MySpace dataset used in the experiments is crawled from MySpace's groups feature. Each post in the dataset is manually labeled as normal or bully.

	Twitter	Myspace
No. of posts	7321	3245
No. of Features	3709	4236
No. of Positive Posts	2102	950
No. of Negative Posts	5219	2295
No. of users	7043	1053
Average posts per user	1.04	2.98

Table-1: Verifying Sentimental score Distribution

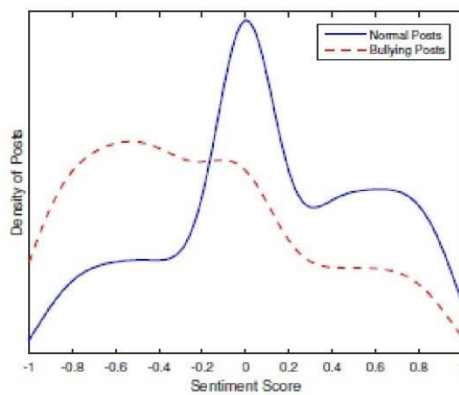


Figure-1: Sentiment Score Distribution of Normal Posts and Bullying Posts in the Twitter Dataset.

Figure-1 shows the sentiment score distribution of the normal and the bullying posts. In Figure-1, the X-axis shows the sentiment polarity score and Y-axis shows the density of users. From the Figure-1 we can observe that two distributions are centered around different mean values. This suggests that there is a clear difference between the sentiment of the normal posts and the bullying posts, and bullying posts tend to have more negative sentiment than normal posts. The sentiment distribution pattern is similar in MySpace dataset.

As the interest and utilization of OSNs are expanding on a regular routine, there emerges the need to fundamentally break down the networks in an efficient manner. The current issues are analyzed:

- Influence Propagation
- Community Detection
- Expert Finding
- Link Prediction
- Recommender systems
- Predicting trust and distrust among individuals
- Opinion mining
- Influence Propagation

(i) Influence Propagation

Domingos and Richardson gave the first algorithmic way to manage influence propagation. At that point, Kempe et al contemplated influence propagation so as to engender on two basic spread models, named Independent Cascade (IC) Model and Linear Threshold (LT) Model, which prompted the advancement of the Greedy Algorithm for influence propagation. They managed the influence propagation issue from an alternate point of view in particular various scalability issues. Chen et al. proposed another proliferation model like the greedy algorithm yet with a superior proficient result. Saito et al were the first to

concentrate how to take in the probabilities for the IC model from an arrangement of past propagations. Goyal et al additionally had made an investigation of the issue of learning impact probabilities utilizing an example of the General Threshold Model (GTC).

(ii) Community Detection

A relative examination on different group location calculations can be found in. Starting study on group or gathering identification was engaged predominantly on the connection structure of OSNs while disregarding the substance of social collaborations, which is likewise pivotal for exact and significant group extraction. It is just as of late that couple of analysts has tended to the issue of finding topically important groups from an OSN.

Pathak et al have proposed a Community-Author-Recipient Topic (Truck) model which utilizes both connection and content data for group location. Liu et al moreover have based a model taking into account Topic-Link Latent Dirichlet Distribution (LDA) however which works just with report systems. Zhao et al have tended to theme situated group discovery through social articles and join investigation in informal organizations. Sachan et al have proposed Topic User Community Model (TUCM) as Topic User Recipient Community Model (TURCM) which offers high time consumption.

(iii) Expert Finding

Analysis on expert ranking estimation is generally taking into account either domain based learning driven systems or space learning free systems or both. The expert ranking issue is likewise looked into on email communication relations. Zhang et al have proposed proliferation based methodology in view of Probabilistic Latent Semantic Analysis (PLSA) for expert finding in social organizations. Authors have utilized the RarestFirst and Enhanced Steiner calculations for expert finding while authors have changed the RarestFirst estimation and discovered the Simplified RareFirst (SRareFirst) estimation. Smirnova et al have proposed a client model for expert finding in light of objective client behavior. Jin et al discovered the ExpertRank calculation which depends on dissecting closeness and power for ranking experts in interpersonal organizations.

(iv) Link Prediction

Liben-Nowell and Kleinberg have managed link forecasting in interpersonal organizations however which works with just a static depiction of a system. Hasan et al have proposed a few characterization models for connection expectation which gives an examination of a few elements by demonstrating their rank of significance as acquired by distinctive estimation. Fouss et al have introduced a connection expectation system in light of a Markov-chain model of random walk however which does not scale well for huge databases. Zheleva et al have utilized a

parallel algorithm in which family was utilized for connection expectation.

(v) Recommender systems

Recommender frameworks (RF) have created in parallel with the web. A decent overview on different RS can be found. They were at first in light of demographic, content based and collaborative sifting. Collaborative sifting is the most widely recognized system utilized for RS. Linden et al introduced their work on thing to item shared sifting for amazon.com suggestions. On the other hand, the development of RS has demonstrated the significance of half and half systems of RS, which blend diverse systems with a specific end goal to get the points of interest of each of them.

(v) Predicting Trust and Distrust among Individuals

Various orders have taken a gander at different issues identified with trust. The first errand was the EigenTrust estimation that expects to lessen the number of inauthentic record downloads in a P2P system. Guha et al proposed systems for engendering of trust and distrust, each of which is suitable in specific circumstances. PowerTrust is a trust proposal framework that totals the positive and negative feelings between the clients into the neighborhood trust scores, comparably to EigenTrust. Other work that studies an informal community with positive and negative feelings is introduced. DuBois et al introduced a paper for foreseeing trust and distrust in light of way likelihood in arbitrary diagrams. Kim et al have additionally proposed a technique for anticipating trust and distrust of clients in online networking sharing groups. Ortega et al proposed a novel framework planned to spread both positive and negative assessments of the clients through a system, in such way that the assessments from every client about others impact their worldwide trust score.

(vi) Opinion Mining

The majority of works in this examination concentrated on classifying texts as per their

sentiment polarity, which can be positive, negative or neutral. Authors gave a top to bottom study of supposition mining and sentiment analysis. The issue was concentrated utilizing directed considering so as to learn logical feeling influencers, for example, invalidation (e.g., not and never) and contrary (e.g., yet and in any case). Wilson et al have considered a few distinctive learning estimations, for example, boosting, rule learning, and Support Vector Machines that can consequently recognize subjective and objective (impartial) dialect furthermore among weak, medium and strong subjectivity.

VII. CONCLUSION

The paper addresses the text-based cyber bullying detection problem, where we have developed semantic enhanced marginalized denoising auto encoder as a specialised illustration learning model for cyber bullying detection. In addition, word embeddings have been wont to automatically expand and refine bullying word lists that's initialized by domain information. The performance of our approaches has been experimentally verified through cyber bullying methods. As a next step we area unit coming up with to additional improve the strength of the learned illustration by considering ordination in messages.

VIII. REFERENCES

- [1]. A. M. Kaplan and M. Haenlein, "Users of the world, unite! The challenges and opportunities of social media," *Business horizons*, vol. 53, no. 1, pp. 59–68, 2010.
- [2]. R. M. Kowalski, G. W. Giumetti, A. N. Schroeder, and M. R. Lattanner, "Bullying in the digital age: A critical review and metaanalysis of cyberbullying research among youth." 2014.
- [3]. M. Ybarra, "Trends in technology-based sexual and non-sexual aggression over time And linkagest non technology aggression," *National Summiton Interpersonal Violence and Abuse Across the Lifespan: Forging a Shared Agenda*, 2010.
- [4]. B. K. Biggs, J. M. Nelson, and M. L. Sampilo, "Peer relations in the anxiety– depression link: Test of a mediation model," *Anxiety, Stress, & Coping*, vol. 23, no. 4, pp. 431–447, 2010.
- [5]. S. R. Jimerson, S. M. Swearer, and D. L. Espelage, *Handbook of bullying in schools: An international perspective*. Rout ledge/Taylor & Francis Group, 2010.
- [6]. G. Gini and T. Pozzoli, "Association between bullying and psychosomatic problems: A meta-analysis," *Pediatrics*, vol. 123, no. 3, pp. 1059–1065, 2009.
- [7]. A. Kontostathis, L. Edwards, and A. Leatherman, "Text mining and cybercrime," *Text Mining: Applications and Theory*. John Wiley & Sons, Ltd, Chichester, UK, 2010.
- [8]. J.-M. Xu, K.-S. Jun, X. Zhu, and A. Bellmore, "Learning from bullying traces in social media," in *Proceedings of the 2012 conference of the North American chapter of the association for computational linguistics: Human language technologies*. Association for Computational Linguistics, 2012, pp. 656–666.
- [9]. Q. Huang, V. K. Singh, and P. K. Atrey, "Cyber bullying detection using social and textual analysis," in *Proceedings of the 3rd International Workshop on Socially-Aware Multimedia*. ACM, 2014, pp. 3–6.
- [10]. D. Yin, Z. Xue, L. Hong, B. D. Davison, A. Kontostathis, and L. Edwards, "Detection of harassment on web 2.0," *Proceedings of the Content Analysis in the WEB*, vol. 2, pp. 1–7, 2009.
- [11]. K. Dinakar, R. Reichart, and H. Lieberman, "Modeling the detection of textual cyberbullying." in *The Social Mobile Web*, 2011.
- [12]. V. Nahar, X. Li, and C. Pang, "An effective approach for cyberbullying detection," *Communications in Information Science and Management Engineering*, 2012.

- [13]. M. Dadvar, F. de Jong, R. Ordelman, and R. Trieschnigg, "Improved cyberbullying detection using gender information," in Proceedings of the 12th - Dutch-Belgian Information Retrieval Workshop (DIR2012). Ghent, Belgium: ACM, 2012.
- [14]. P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," *The Journal of Machine Learning Research*, vol. 11, pp. 3371–3408, 2010.
- [15]. P. Baldi, "Autoencoders, unsupervised learning, and deep architectures," *Unsupervised and Transfer Learning Challenges in Machine Learning*, Volume 7, p. 43, 2012.

Fetal Brain Border Detection from MRI using Chain Code Algorithm	Gayathri.S.P	Computer Science	International Journal Of Computer Sciences And Engineering	May 2018	E-ISSN: 2347-2693	https://www.ijcsonline.org/pdf/spl_paper_view.php?paper_id=388&PID072.pdf
--	--------------	------------------	--	----------	-------------------	---

Fetal Brain Border Detection from MRI using Chain Code Algorithm

S.P. Gayathri^{1*}, K.Somasundaram², R.Siva Shankar³

^{1*}Department of Computer Science, Sakthi College of Arts and Science for Women, Oddanchatram, Dindigul, India.

²Department of Computer Science and Applications, The Gandhigram Rural Institute (Deemed to be University), Gandhigram, Dindigul, India.

³Department of Computer Applications, Madanapalle Institute of Technology and Science (UGC-Autonomous), Madanapalle-517325, Andhra Pradesh, India.

*Corresponding Author: gayathrisp12@gmail.com, Tel.: +91-9790952789

Available online at: www.ijcseonline.org

Abstract—Magnetic Resonance Imaging of fetal brain facilitates to evaluate in-utero fetal brain development. Segmentation of fetal brain from MRI is a challenging task due to significant changes in terms of geometry as well as tissue morphology. In order to make ease of segmentation, fetal brain border is detected using chain code algorithm. Detected brain border is used further through in which query images will be extracted from fetal MRI database. This work will be extended with feature extraction and 3D modeling of fetal brain in a little while.

Keywords—fetal brain, chain code, feature extraction.

I. INTRODUCTION

This work makes use of chain codes as features for image boundary detection. The boundary is further used to automatically retrieve the fetal brain shapes from a MRI database. In general, Shape descriptors [1] are classified as: region based methods and boundary based methods. The area within the image region is considered for shape description by region based methods. Hence, more computation time and storage space are required. The contour based methods require only the pixels lying on the boundary region. So, it require less time and space for description. In this paper, we used contour based shape description using chain code [2] for boundary detection in order to achieve images retrieval. Chain code methods have numerous advantages. The chain codes [3],[4] are compact and translation invariant for representation of a binary object. In addition, chain code methods can be applied to calculate any shape feature as since it is a complete representation of any object or curve. The remaining part of the paper is organized as follows.

In section II, we present the boundary detection using chain code methodology, Section III describes results and discussion, Section IV concludes research work with future directions.

II. METHODOLOGY

Boundary Detection and Chain code

Chain code [5] is constructed with preprocess of boundary detection. Chains represent the boundaries of any discrete shape such as binary object. The objective is to find the boundary pixels in the given image. The contour is recognized as a connected pixels lying along the boundary of the object. The neighborhood for image pixel is connected as 4-connected neighbor and 8-connected neighbor, as shown in Figure 1.

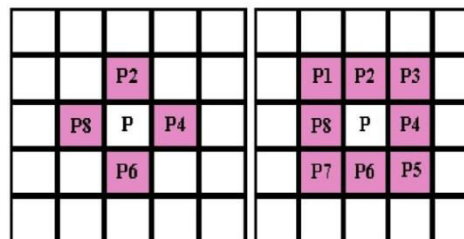
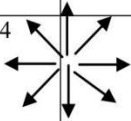


Figure1. Pixel connectedness (a) 4- neighbour pixels
(b) 8-neighbor pixels

The 4-connectedness allows adjacent pixels in vertical and horizontal directions, whereas the 8-connectedness includes pixels in diagonal directions also. 4-connectedness failed to produce the diagonal points where such points are very

helpful for boundary detection in most of the applications. In order to solve this issue, we use 8-connectedness for chain code. The direction of the chain code is given in Table 1. The boundary of a binary image is computed by chain code [6] as array of integer values $X = \{x_0, x_1, \dots, x_{n-1}\}$ where x_i from the set $\{0, 1, \dots, 7\}$ for $i=0, 1, \dots, n-1$. The length of the chain code depends on the number of elements in the array. The starting and ending point of the chain code is represents x_0 and x_{n-1} . The process starts from the initial point to ending point. The tracing through this directions may change based on the requirement and shape of the images, and thus the chain code [7] obtained. The positions of subsequent pixels are given in Table 2.

Table 1. Direction of chain code

3	2	1
4		0
5	6	7

Current pixel coordinate (i,j)		
Code	Next row	Next column
0	i	j+1
1	i-1	j+1
2	i-1	j
3	i-1	j-1
4	i	j-1
5	i+1	j-1
6	i+1	j
7	i+1	j+1

III. RESULTS AND DISCUSSION

Experiments done by detecting fetal brain boundary (Figure 2) using fetal brain mask which comprises difference in curvature, corner sharpness and anomaly along the boundary curves.

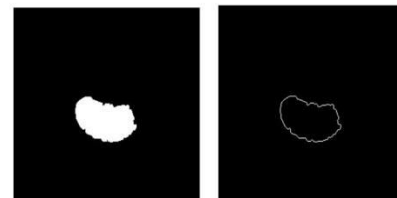


Figure 2 (a) Fetal brain Mask (b) Brain Boundary

Table 2. Position of subsequent pixels

0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	5	0	4	0	5	1	1
0	0	0	5	0	1	1	0	0	0	0
0	0	0	5	1	0	0	0	0	0	0
0	0	5	1	0	0	0	0	0	0	0
0	6	1	0	0	0	0	0	0	0	0
0	6	1	0	0	0	0	0	0	0	0
0	6	1	0	0	0	0	0	0	0	0

Figure.3 Sample code for deriving chain code starting from the initial location X_0

Next chain codes for the boundary pixels are calculated. The number of direction changes for each pair of pixels in the contour is determined. For example, from the figure.3 we note that the chain code method starts from point X_0 . The sample chain code is obtained and thus the values are 5 4 5 5 5 6 6 in 8-directional. Here, from point 1 to point 2 the chain code is 5. The change in chain code at point 2, a corner will be detected. The chain code for point 2 to point 3 is 4, here the chain code again change so in point 2 yet again a corner will be detected. In the same way, the method proceeds for all boundary pixels to obtain the chain code. The sample chain code extracted for the fetal brain boundary is given in figure 4.

Figure 4 Sample chain code extracted for fetal brain boundary image

Columns 1 through 19
5 4 4 5 4 5 4 5 5 6 6 5 6 5 6 6 6 6 6 6
Columns 20 through 38
6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 7 6 6 6 6
Columns 39 through 57
6 7 6 7 7 7 7 0 0 7 0 0 0 0 0 0 0 1 0

VII. CONCLUSION and Future Scope

Our work shows that the chain code approach is used to achieve fetal brain boundary detection. By using this brain boundary, retrieval of fetal brain from the database is simple and fast. Further this work will be extended for feature extraction and finally to produce a 3D model for fetal brain MRI.

REFERENCES

- [1] Rafael C. Gonzalez, Richard E. Woods and Steven L. Eddins, "Digital Image Processing Using MATLAB", 2nd edition, pp 440-488.
- [2] Jahne, B., "Practical handbook on image processing for scientific and technical applications", CRC Press, p.494,2004.
- [3] Neeta Nain, Vijay Laxmi, Ankur Kumar Jain and Rakesh Agarwal, "Morphological Edge Detection and Corner Detection Algorithm Using Chain-Encoding", IPCV'2006, pp 1-5,2008.
- [4] Baji F, Mocanu M, "Chain Code Approach for Shape based Image Retrieval", Indian Journal of Science and Technology, vol.11,2018.
- [5] Boodoo-Jahangeer, N.B. and Baichoo, "Face recognition using chain codes", Journal of Signal and Information Processing, vol.4,p.154,2013.
- [6] Parmar DA., "Content MRI Brain Image Retrieval using Shape Descriptors and Relevance Vector Machine (RVM)", International Journal of Advance Research in Computer Science and Management Studies, vol.4,2016.

- [7] Anjan Bikash Maity, Sandip Mandal, Ranjan Podder, "Edge Detection Using Morphological Method and Corner Detection Using Chain Code Algorithm", IJCSI International Journal of Computer Science Issues, vol.8, 2011.

Authors Profile

Dr.S.P.Gayathri received her Master of Computer Science from Seethalakshmi Ramasamy College of Arts and Science, Trichy, India. From 2004 to 2007, she was a Lecturer in Department of Computer Science, Ramaprabha College of Arts and Science, Dindigul, TN, India. From 2007 to 2011 December, she worked as Assistant professor in Department of Computer Science and Applications in Gandhigram Rural Institute (DU), Dindigul, TN, India. She graduated Ph.D. in Computer Science and Applications from Gandhigram Rural Institute - Deemed University, Dindigul, India. Her research interest is Digital and Medical Image Processing. She is presently working as Associate professor in Sakthi College of Arts and Science for Women, Oddanchatram, TN, India



Dr.K.Somasundaram, received his Master of Science (M.Sc) degree in Physics from the University of Madras, Chennai, India in 1976, the Post Graduate Diploma in Computer Methods from Madurai Kamaraj University, Madurai, India in 1989 and the Ph.D degree in theoretical Physics from Indian Institute of Science, Bangalore, India in 1984. He is presently working as Professor at the Department of Computer Science and Applications, Gandhigram Rural Institute, Dindigul, India. From 1976 to 1989, he was a Professor with the Department of Physics at the same Institute. He was senior Research fellow of Council of Scientific and Industrial Research (CSIR), Govt. of India. He was previously a Researcher at the International Centre for Theoretical Physics, Trieste, Italy and Development Fellow of Commonwealth Universities at Edith Cowan University, Perth, Australia. His research interests are image processing, image compression and Medical imaging. He is Life member of Indian Society for Technical Education, India and Life member in Telemedicine society of India. He is also a member of IEEE USA.



Dr.Siva Shankar Ramasamy did MCA and Ph.D from Gandhigram Rural University, Tamil Nadu, India. He worked in National Institute of Technology-Trichy-620015, India. He is a Life Member of CSI since 2015. His main research work focuses on Medical image segmentation and his recent research area is IOT, Automation in Agriculture. He can be reached through his mail arjunshankar@gmail.com



