Jan Zizka
Dhinaharan Nagamalai (Eds)

Computer Science & Information Technology

Third International Conference on Database and Data Mining
(DBDM 2015)
Dubai, UAE, April 24 ~ 25 - 2015

AIRCC

Volume Editors

Jan Zizka,
Mendel University in Brno, Czech Republic
E-mail: zizka.jan@gmail.com

Dhinaharan Nagamalai,
Wireilla Net Solutions PTY LTD,
Sydney, Australia
E-mail: dhinthia@yahoo.com

ISSN: 2231 - 5403
ISBN: 978-1-921987-35-9
DOI : 10.5121/csit.2015.50701 - 10.5121/csit.2015.50706

Typesetting: Camera-ready by author, data conversion by NnN Net Solutions Private Ltd., Chennai, India

Preface

Third International Conference on Database and Data Mining (DBDM 2015) was held in Dubai, UAE, during April 24~25, 2015. Third International Conference on Instrumentation and Control Systems (CICS 2015), Second International Conference on Computer Networks & Communications (CCNET 2015), and Second International Conference on Signal Processing (CSIP 2015) were collocated with the DBDM-2015. The conferences attracted many local and international delegates, presenting a balanced mixture of intellect from the East and from the West.

The goal of this conference series is to bring together researchers and practitioners from academia and industry to focus on understanding computer science and information technology and to establish new collaborations in these areas. Authors are invited to contribute to the conference by submitting articles that illustrate research results, projects, survey work and industrial experiences describing significant advances in all areas of computer science and information technology.

The DBDM-2015, CICS-2015, CCNET-2015, CSIP-2015 Committees rigorously invited submissions for many months from researchers, scientists, engineers, students and practitioners related to the relevant themes and tracks of the workshop. This effort guaranteed submissions from an unparalleled number of internationally recognized top-level researchers. All the submissions underwent a strenuous peer review process which comprised expert reviewers. These reviewers were selected from a talented pool of Technical Committee members and external reviewers on the basis of their expertise. The papers were then reviewed based on their contributions, technical content, originality and clarity. The entire process, which includes the submission, review and acceptance processes, was done electronically. All these efforts undertaken by the Organizing and Technical Committees led to an exciting, rich and a high quality technical conference program, which featured high-impact presentations for all attendees to enjoy, appreciate and expand their expertise in the latest developments in computer network and communications research.

In closing, DBDM-2015, CICS-2015, CCNET-2015, CSIP-2015 brought together researchers, scientists, engineers, students and practitioners to exchange and share their experiences, new ideas and research results in all aspects of the main workshop themes and tracks, and to discuss the practical challenges encountered and the solutions adopted. The book is organized as a collection of papers from the DBDM-2015, CICS-2015, CCNET-2015, CSIP-2015

We would like to thank the General and Program Chairs, organization staff, the members of the Technical Program Committees and external reviewers for their excellent and tireless work. We sincerely wish that all attendees benefited scientifically from the conference and wish them every success in their research. It is the humble wish of the conference organizers that the professional dialogue among the researchers, scientists, engineers, students and educators continues beyond the event and that the friendships and collaborations forged will linger and prosper for many years to come.

Jan Zizka
Dhinaharan Nagamalai

Organization

General Chair

Natarajan Meghanathan	Jackson State University, USA
Dhinaharan Nagamalai	Wireilla Net Solutions PTY LTD, Australia

Program Committee Members

Abdolreza Hatamlou	Islamic Azad University, Iran
Abdoulaye Kindy Diallo	Atlassian/ServiceRocket, Malaysia
Aboul Ella Hassanien	Cairo University, Egypt
Ahmad Al-Khasawneh	Hashemite University, Jordan
Ahmad Lotfi	Nottingham Trent University, United Kingdom
Ahmed samir	Ain-Shams University, Egypt
Ali Abid D. Al-Zuky	Mustansiriyah University, Iraq
Ali Riza YILDIZ	Bursa Technical University, Turkey
Amal Al-Dweik	Palestine Polytechnic University, Palestine
Ankit Chaudhary	Truman State University, USA
Arifa Ferdousi	Varendra University, Bangladesh
Arockiasamy S	University of Nizwa, Sultanate of Oman
Ashraf Shahin	Cairo University, Egypt
Athirai	Nanyang Technological Universiy, Singapore
Ayad Ghany Ismaeel	Erbil Polytechnic University, Iraq
Che, Hao	University of Texas at Arlington, USA
Dac-Nhuong Le	Haiphong University, Vietnam
Dhari Ali	University of Technology, Iraq
Ederval Pablo Ferreira da Cruz	Federal Institute of Espirito Santo, Brazil
Emilio Jiménez Macías	University of La Rioja, Spain
Fernando Bobillo	University of Zaragoza, Spain
Gh.A.Montazer	Tarbiat Modares University, Iran
Hacene Belhadef	University of Constantine, Algeria
Hamadouche M	Hamadouche Maamar, Algeria
Hamid Mcheick	Universite du Quebec a Chicoutimi, Canada
Hamid Taghavifar	Urmia University, Iran
Haytham Samy Elwahsh	Kafrelshikh University, Egypt
I-Cheng Chang	National Dong Hwa University, Taiwan
I-Ching Hsu	National Formosa University, Taiwan
Isa Maleki	Islamic Azad University, Iran
Islam Atef	Alexandria University, Egypt
Justinian Anatory	University of Dodoma, Tanzania
Kaamran Raahemifar	Ryerson University, Canada
Koczy T.Laszlo	Budapest University of Technology, Hungary
Malek	Jadara University, Jordan
Manfred Meyer	University of Applied Sciences, Germany
Marco Querini	University of Rome Tor Vergata, Italy

Technically Sponsored by

Networks & Communications Community (NCC)

Computer Science & Information Technology Community (CSITC)

Digital Signal & Image Processing Community (DSIPC)

Organized By

Academy & Industry Research Collaboration Center (AIRCC)

TABLE OF CONTENTS

Third International Conference on Database and Data Mining (DBDM 2015)

Third International Conference on Instrumentation and Control Systems (CICS 2015)

Second International Conference on Computer Networks & Communications (CCNET 2015)

Second International Conference on Signal Processing (CSIP 2015)

EXTENDED FAST SEARCH CLUSTERING ALGORITHM: WIDELY DENSITY CLUSTERS, NO DENSITY PEAKS

Zhang WenKai[1] and Li Jing[2]

[1,2]School of Computer Science and Technology, University of Science and Technology of China, Hefei, 230026, China
zwk@mail.ustc.edu.cn
lj@ustc.edu.cn

ABSTRACT

CFSFDP (clustering by fast search and find of density peaks) is recently developed density-based clustering algorithm. Compared to DBSCAN, it needs less parameters and is computationally cheap for its non-iteration. Alex. at al have demonstrated its power by many applications. However, CFSFDP performs not well when there are more than one density peak for one cluster, what we name as "no density peaks". In this paper, inspired by the idea of a hierarchical clustering algorithm CHAMELEON, we propose an extension of CFSFDP, E_CFSFDP, to adapt more applications. In particular, we take use of original CFSFDP to generating initial clusters first, then merge the sub clusters in the second phase. We have conducted the algorithm to several data sets, of which, there are "no density peaks". Experiment results show that our approach outperforms the original one due to it breaks through the strict claim of data sets.

KEYWORDS

Clustering, Density, Density peaks, K-nearest neighbour graph, Closeness &Inter-connectivity

1. INTRODUCTION

Clustering is known as the unsupervised classification in pattern recognition, or nonparametric density estimation in statistics[1]. The aim is to partition given data set of points or objects into natural grouping(s) according to their similarity to improve understanding on the condition of no priori-knowledge, or be as a method to compress data. Cluster analysis has been widely used in a lot of fields, like computer version ([2], [3], [4]), bioinformatics ([5], [6], [7]), image progressing ([8], [9], [10], [11]), Knowledge Discovery in Databases(KDD), and many other areas ([12]). Thousands of clustering algorithms have been proposed, challenges still remain: differing shapes, high dimensions, how to determine the clusters number, how to define a right clustering, hard to evaluate.

Density-based clustering algorithms which classify points by identifying regions heavily populated with data, such as DBSCAN [13] and GDBSCAN [14], OPTICS [15], and DBCLASD [16], have performed well while handling problems of arbitrary shapes of subclasses. DBSCAN [13] is a representative of density-based methods, by the definition of core points, density connection, in which clusters defined as high density regions separated by low density regions in

Jan Zizka et al. (Eds) : DBDM, CICS, CCNET, CSIP - 2015
pp. 01–17, 2015. © CS & IT-CSCP 2015 DOI : 10.5121/csit.2015.50701

the feature space can be detected without the need for clusters number. However the appropriate threshold MinPts(minimum number of points) for distinguishing core points from border points is hard to select, with a high MinPts, thin clusters(relative low density) would be ignored.

Similar to DBSCAN [13], recently, CFSFDP (clustering by fast search and find of density peaks) [17] was proposed by Alex and Anlessandro to detect non-spherical groups, which does not need to pre-specify the number of clusters of variant shapes either. In addition, CFSFDP needs less parameters. CFSFDP finds the clusters of points by a two phase progression. During the first phase, CFSFDP uses a well-designed decision group to find out the cluster centres, so-called density peaks. During the second phase, each remaining point is assigned to the same cluster as its nearest neighbour of higher density. In a DBSCANs perspective, CFSFDP assumes that every object is density-connected with its nearest neighbour of higher density. Compared to mean-shift methods such as[18], [19], CFSFDP [17] is computationally cheaper for the procedure of maximizing the density field for each data point in the mean-shift approach. By the experiments of identifying the number of subjects in the Olivetti Face Database [20], the team have shown CFSFDP's capacity to solve high dimensional data [17].

However, in our opinion, there are some drawbacks of the beautiful CFSFDP [17], which will limit the application of CFSFDP. Firstly, just as DBSCAN [13], thin clusters would not be captured by the decision graph. Besides, a rigid hidden requirement for getting right clusters is that, each cluster in the data sets must have a density peak and only one peak is promised, otherwise CFSFDP will split natural groups. In this paper, inspired by CHAMELEON [21], we present a novel hierarchical clustering algorithm based on CFSFDP. Thus our approach can find thin clusters. In addition, it eliminates he strict claim of density peaks. To display our efforts, we benchmark our algorithm on the data sets draw from [21],[22], [23], of which there is no unique density peak for each cluster. Our technique gets partitions of these data sets as well as that generated by the methods proposed in the papers where the data set was designed.

We discuss details of CFSFDP [17], CHAMELEON [21] in Section 2. In Section 3, we present drawbacks of CFSFDP and our efforts to overcome these limitations. Section 4 describes our algorithm in detail. In Section 5, we benchmark the performance of our approach on some data sets from other literatures. Finally, a conclusion and direction of future worksare shown in Section 6.

2. BASIC CONCEPTS

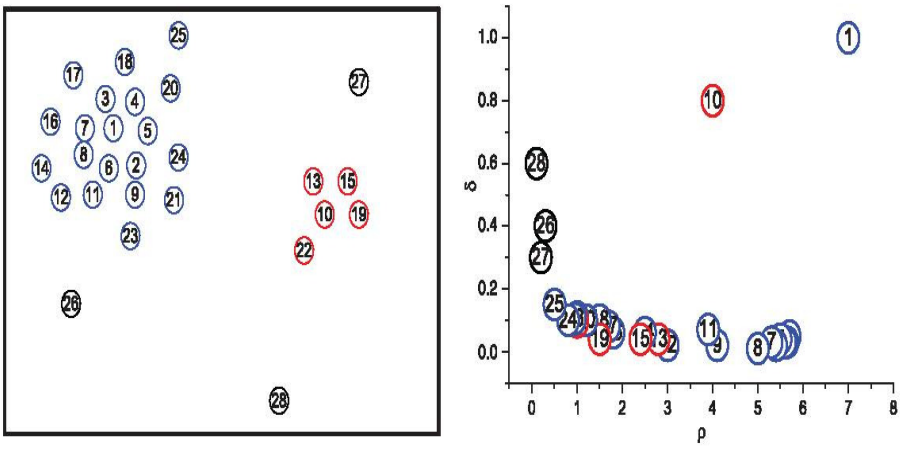
This Section presents two methods and some concepts involved in our technique, what are necessary to understand our approach. If one has been familiar to CFSFDP and CHAMELEON, he could skip to Section 3, where we discuss some disadvantages of CFSFDP in details.

2.1. CFSFDP

CFSFDP [17] generates clusters by assigning data points to the same cluster as its nearest neighbour with higher density after cluster centres are selected by users. The cluster centres are defined as local density maxima, Alex and Alessandro designed a heuristic method for customers to select the genuine cluster centres, what is named as decision graph.

Two important quantities are considered in the decision group: local density ρ_i of each point$i$,its distance δ_i frompoints of higher density. Definition of ρ_i , δ_i follows:

Definition 1: The density of a point i, denoted by ρ_i, is defined as



(1) (2)

$$\rho_i = \sum_j \chi(d_{ij} - d_c).(1)$$

Where $\chi(x) = 1$ if $x < 0$ and $\chi(x) = 0$ otherwise, and d_c is a cutoff distance, which is the only parameter need to be determined by customers. In principle, ρ_i equals to the number of points which are closer than d_c to point i.

Definition 1: The minimum distance of point i from anyother point of higher density, denoted by δ_i, is computed by

$$\delta_i = \min_{j:\rho_j > \rho_i} d_{ij} \ . \qquad (2)$$

Figure1. The CFSFDP in two dimensions. (1) Point distribution. Data points are ranked in order of decreasing density. (2) Decision graph for the data in(1). Different colors correspond to different clusters. [17]

Notice that, for the point of largest density, δ_i is redefinedas $\delta_i = \max_{j \in dataset} d_{ij}$. The simple observation that points ofhigh local density and high density distance are local density maxima, namely density peaks or cluster centres, is the core of this procedure to select cluster centres.

To identify density peaks defined as below, a method named as ”decision graph” is introduced to help users to make a decision. Basically, decision graph is a figure plotting $\delta_i$as afunction of ρ_i , is illustrated by the two-dimensional simpleexample in Fig.1. Points 1 and 10 are the only two points ofhigh δ and high ρ, as a result, they are the cluster centres. Points 26, 27, 28 can be considered as outliers for a relatively high δ and low ρ (which indicates that they are isolated points).

After cluster centres have been found, CFSFDP allocates the rest points to the same cluster as its nearest neighbourhood with higher density.

2.2. CHAMELEON

CHAMELEON discovers clusters of a given data sets gradually by finding groups consisted of most similar points. There are 3 main steps in CHAMELEON: create the k-nearest Neighbour Graph according to data points, partition the graph into sub-classes, then merge the subsets. It's common to model data items as a graph in agglomerative hierarchical clustering techniques[1], CHAMELEON models data based on the widely used k-nearest neighbour graph technique. After the graph was built, a efficient graph partitioning algorithm mMETIS[24]is used to find the initial sub-clusters. As criteria to aggregate the sub clusters, CHAMELEON computes the relative interconnectivity$\mathrm{RI}(C_i, C_j)$and relative closeness $\mathrm{RC}(C_i, C_j)$ between each pair of clusters C_i and C_j. In the merge phase,each round the cluster pair of highest

$$\mathrm{RI}(C_i, C_j) \times RC(C_i, C_j)^{\beta} \qquad (3)$$

will be merged, where β is a user defined parameter to give different importance to the two criterions. The merge progress will be stop when cluster number is equal to the number predefined or there is no cluster pair of which the value of (3) is bigger than user determined threshold.

In (3), the relative inter-connectivity $\mathrm{RI}(C_i, C_j)$betweencluster C_i and clusterC_j is defined as the sum of the weight of edges span the two clusters normalized with respect to the internal inter-connectivity of cluster C_i and cluster C_j. Theinternal inter-connectivity of cluster is computed by addingthe weight of edges that partition the cluster into two roughly equal parts. Meanwhile CHAMELEON defines the relative closeness $\mathrm{RC}(C_i, C_j)$ of cluster C_i and cluster C_j as theaverage weight of edges span the two clusters, which alsois normalized with respect to the internal closeness of each cluster.

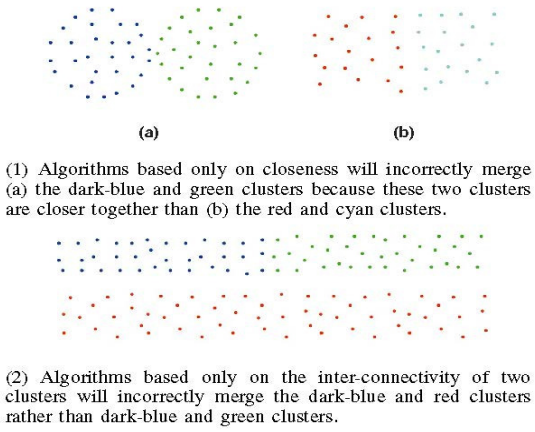


Figure 2. Clustering considering either closeness or inter-connectivity. [21]

The key difference of CHAMELEON and other hierarchical clustering algorithms like CURE [25], ROCK [26] is that it accounts for both inter-connectivity and closeness while identifying the most similar pair of clusters in the third step[21]. The disadvantages of only considering either closeness or inter-connectivity is these schemes could merge wrong pair of clusters. A sample of that was given by George, as shown inFig.2. In Fig.2.(1), algorithms only based on closeness (CURE and related schemes) merge incorrectly because through clusters of (a) — the dark-blue and green clusters — are closer to each other than the case of (b) — the red and cyan clusters, clusters of (b) connect better than those of (a). In Fig.2.(2), the inter-connectivity of the dark-blue

cluster and the red cluster is bigger than that of the dark-blue cluster and the green cluster, however the dark-blue cluster is closer to the red cluster than the green one. Thus an algorithm based only on the interconnectivity(ROCK [26]) merge the dark-blue cluster with the red cluster incorrectly.

3. LIMITATIONS OF CFSFDP AND OUR CONTRIBUTION

In this Section, firstly we discuss our understanding towards limitations of CFSFDP in theory. To help illustrate the limitation intuitively, an instance is presented.

CFSFDP has got high performance classifying several datasets, the success of these examples verifies the availability of CFSFDP in some degree. As shown in Section 2, it generates groups by identifying clusters with density maxima. However, it is impossible to get natural clusters by CFSFDP when the local densities of data points in some or all natural clusters of the data sets is random distributed, such that instead of one density peak two or more density peaks appear in one cluster. In such a case, it is hard for CFSFDP to pick up all the reasonable cluster centres. What's more, even if a reasonable cluster centre set is found by CFSFDP, the natural cluster would be split. Reasonable cluster centre here means that a point as density maximal, with which, the result cluster of CFSFDP clustering procedure (assign each point to the same cluster of its nearest neighbour of higher density than itself) is not consisted of parts of different natural clusters. Reasonable cluster centre set means a set of reasonable cluster centres. The cluster centre set presented in Fig.4.B, is an unreasonable cluster centre set for the red cluster is composed of part of the "arc" cluster and part of left cluster.

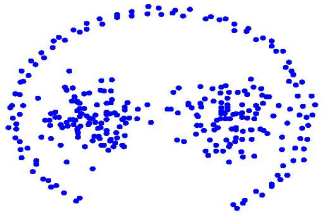


Figure 3. The data set drawn from [22]. There are 3 native clusters of diversedensity, complex shapes, two of the three are surrounded by the third one.

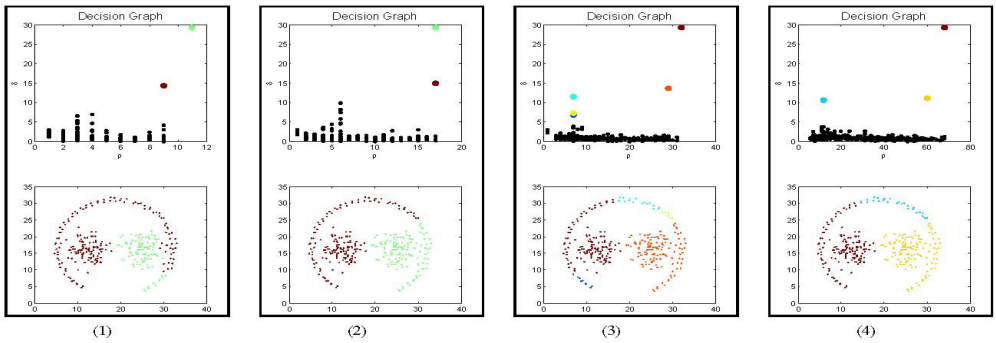


Figure 4. CFSFDP groups the data set in Fig.3, Different colors correspond to different clusters, the figure above in each sub figure is the decision graph, the figure below is the correspond dividing result. (1)d_c = 1:101 such that the largest density equals to about 1% of the number of points size, 2 clusters.(2)d_c = 1:540 such that the largest density is equal to about 2% of the size of data set, 2 clusters. (3)d_c = 2:250, such that the largest density equals toabout 4% of the number of total points, 5 clusters. (4)d_c = 3:814, such that the largest density equals to about 10% of the number of total objects, 3 clusters.

Let's consider the test case in Fig.3. The data set is drawn from [22], there are 3 natural clusters in the data set: the left cluster surrounded by the arc, the right cluster surrounded by the arc, and the arc cluster. We test CFSFDP on this data set by variant dc, the clustering result of some values(1.101, 1.540, 2.250, 3.814) is shown in Fig.4. Except the experiments presented in Fig.4, many other values of dc have been tested, unfortunately CFSFDP can't divide the points into natural groups with any value of d_c. In particular, CFSFDP makes mistake on the arc cluster, parts of which are resigned into wrong clusters.

Firstly, poor ability of decision graph to select cluster centres is one reason of such a bad outcome, because performance of CFSFDP is highly sensitive to the procedure determining cluster centres [17] says, cluster centres of relative high local density ρ & relative high δ are of remarkable location in the decision graph, and indeed its true for some cases, as shown in Fig.1 where cluster centres are of high ρ and high δ in global. Nevertheless, when clusters of more complex data sets, either thin cluster centres of relative low ρ and high δ, or other cluster centres of relative high ρ and low δ, would be overlooked easily. As an example, in Fig.4, the relative thin cluster centres belong to the "arc" cluster are easily ignored for we only select the points of remarkable location as cluster centres. One may argue that with a more suitable value ofd_c,density peaks would be prominent in decision graph and be easy to point out by users. In our opinion, it may be or may not be, it depends on the data sets. A good method to select cluster centres should reduce its dependency upon data sets as much as it can, such that it would detect diverse density or diverse δ cluster centres.

Another cause of wrong clusters detected by CFSFDP is that CFSFDP divides points based on cluster centres, such that CFSFDP might divide the natural cluster if there are more than one cluster centres was determined in a natural cluster. For example, Fig.4 shows the clustering consequence of the[22] data set after choosing reasonable cluster c centre set by our new decision graph. In Fig.4, points of the arc cluster which is split into several clusters. Inspired by the clustering progress of hierarchical clustering algorithm, a novel clustering algorithm consisted of two phases brings out to us: 1) generate subclasses by CFSFDP, 2) merge the sub classes by the similarity between clusters.

3.1. Modelling the Similarity between Clusters

To break through the barriers of agglomerative hierarchical approaches discussed in Section 2, our algorithm looks at their Relative Inter-connectivity $\mathrm{RI}(C_i, C_j)$ and their Relativecloseness $\mathrm{RC}(C_i, C_j)$ while merging cluster C_i and cluster C_j , which is similar with the aggregation phase of CHAMELEON. By considering both of these criteria, ourscheme selects clusters that are well connected as well asclose together to merge. However, there is a shortage inCHAMELEON that is CHAMELEON, which models sub clusters based on the widely used k-nearest neighbour graph, would fail to merge the correct cluster pair in some cases. To solve that, we using a variant of k-nearest neighbour graph to model the sub clusters. Other than the difference of the model to represent sub clusters, we compute the relative inter-connectivity and the closeness almost the same as CHAMELEON, we also use the value defined by (3) to model the similarity between clusters.

In this Section, firstly we show the detail of the relative inter-connectivity and the relative closeness. In the remainder, the detail of drawback in CHAMELEON's merging phase is presented, then our solution is followed.

The relative inter-connectivity $\mathrm{RI}(C_i, C_j)$ between cluster C_i and cluster C_j is given by

$$\mathrm{RI}(C_i, C_j) = \frac{EC_{\{c_I, c_J\}}}{\frac{|c_i|}{|c_i|+|c_j|} EC_{c_I} + \frac{|c_j|}{|c_i|+|c_j|} EC_{C_j}}. \quad (4)$$

Where $EC_{\{C_i, C_j\}} = \sum E(u, v)$ is sum of weight of edgesconnecting the two clusters. $\mathrm{E(u, v)}$ is defined as (4), this isthe main difference between CHAMELEONs merging phaseand our merging phase.

The Relative closeness $\mathrm{RC}(C_i, C_j)$is computed by

$$\mathrm{RC}(C_i, C_j) = \frac{\bar{S}_{EC_{\{C_i, C_j\}}}}{\frac{|c_i|}{|c_i|+|c_j|} \bar{S}_{EC_{C_i}} + \frac{|c_j|}{|c_i|+|c_j|} \bar{S}_{EC_{C_j}}}. \quad (5)$$

Where $\bar{S}_{EC_{\{C_i, C_j\}}} = \frac{EC_{\{C_i, C_j\}}}{|E(u,v)|}$ is the average weightof edges span the two clusters. To get the value of inner criteria, as EC_{C_i}or $\bar{S}_{EC_{C_i}}$, CHAMELEON takes graph partitioningtechnology to bipartite the target cluster, whilewe split the target cluster by CFSFDP. In particular, at first, we utilize the CFSFDP clustering algorithm with the novel decision graph(described in previous part of this Section) to clustering the points of current cluster into two sub classes; then we compute F based on the clustering result. There are two advantages of taking use of CFSFDP rather than graph partitioning methods to split the target cluster. Firstly, CFSFDP is very fast with specific cluster centres, the time complexity is o (n), *k* is the number of points of the target cluster. The selection of cluster centres can be done in constant time based on the computation in the initial clustering phase. Secondly, we use CFSFDP in all phases of our clustering method to keep consistency in theory, in our view, that meets the famous Ockham's Razor principle.

CHAMELEON models the similarity based on the widely used *k*-nearest neighbour graph. The benefit of representing data as a *k*-nearest neighbour graph covers that simplifying topology of data sets by disconnecting the points are far apart, capturing the concept of neighbourhood dynamically, and there are many methods offered to handle graph data [21].However, in the merge phase, we find that CHAMELEON based on the standard k-nearest graph would fail to merge the correct cluster pair in some cases. The example in Figure.5.(1)illustrates, an algorithm that considering the inter-connectivity as CHAMELEON (presented in Section 2) will prefer to merge in correctly the red cluster with the blue cluster, rather than with the green cluster. Notice that, because the blue cluster is much sparser than the red cluster, as a result, the measure of connectivity and closeness depends on point *P*, and almost all the *k*-nearest neighbours of point *P* belong to the red cluster, leading to that both the criteria (inter-connectivity and closeness) of the red cluster and the blue cluster are higher than those of the red cluster and the green cluster.

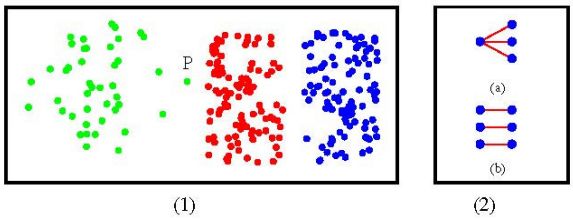


Figure 5. (1) Chameleon prefers to merge incorrectly the red cluster with theblue cluster rather with the green cluster. (2) A simplified instance to explain why Chameleon fails in (1).

Further, we find that CHAMELEON fails because it is almost impossible for CHAMELEON based on the standard k-nearest graph to distinguish case (a) from case (b) inFigure.5.(2).Where the weight of edges is nearly the same. In figure.5, we ignoring the normalization because the inner criteria are the same in this example.

To address that disadvantage, we bring out a variant of *k*-nearest neighbour graph to model the sub clusters. In this paper, the edge connecting point of clusterC_i with point of clusterC_j is given by

$$\mathrm{E}(\mathrm{u},\mathrm{v}) = \max\left(\left\{E(u,p)\middle| u \in C_i \cap p \in C_j \cap\ u \in KN(p) \cap\ p \in KN(u)\right\}\right). \quad (6)$$

Where$\mathrm{E}(\mathrm{u},\mathrm{v})$ represents the edge connecting point *u* with point *v*. $\mathrm{KN}(\cdot)$ is an operator to get the k-nearest neighbor.The $\max(\cdot)$ function is to get the edge of maximum weight,the weight here means the similarity. By only considering the edge of maximum weight, our approach can distinguish the two type k-nearest neighbour relationship. In figure.5. (2). (b), we remove the two slash edges, such that, we can get a smaller inter-connectivity of case (a) than that of case (b). Even if there is still a similar closeness, it is possible to distinguish case (a) from case (b) by giving a higher importance to the inter-connectivity, by specifying a value smaller than *1.o* to β in (3).

4. THE EXTENDED CFSFDP (E_CFSFDP)

Algorithm 1 E_CFSFDP (X, $\mathbf{d_c}$, $\boldsymbol{N_{neighbor}}$, β)

Inputs:

X= $\{\mathbf{X_1}, \mathbf{X_2}, \boldsymbol{X_3}, \ldots, \boldsymbol{X_n}\}$ Set of data objects

$\mathbf{d_c}$: radius to computeρ.

$\boldsymbol{N_{neighbor}}$: value of neighbour number while modelling the cluster similarity.

β : parameter to control the importance of the two criterion (3).

Output:

C = $\{\boldsymbol{C_1}, \boldsymbol{C_2}, \boldsymbol{C_3}, \ldots, \boldsymbol{C_n}\}$ Set of clusters

{**Phase I**} find initial sub-clusters by CFSFDP

1: compute the similarity matrix.

2: compute the local density ρ and the distance δ of each object.

3: select cluster centres by decision graph.

4: allocate each point to the same cluster as its nearest neighbour of higher ρ.

{**Phase II**} merge the sub-clusters

5: construct the model of sub-clusters.

6: compute the$\mathbf{RI}(\boldsymbol{C_i}, \boldsymbol{C_j}) \times \boldsymbol{RC}(\boldsymbol{C_i}, \boldsymbol{C_j})^{\boldsymbol{\beta}}$ matrix for eachsub-clusters pair $\{\boldsymbol{C_i}, \boldsymbol{C_j}\}$ based on the model built in step4.

7: merge the cluster pair of highest $\mathbf{RI}(\boldsymbol{C_i}, \boldsymbol{C_j}) \times \boldsymbol{RC}(\boldsymbol{C_i}, \boldsymbol{C_j})^{\boldsymbol{\beta}}$.

8: repeat step 6 and step 7 until the termination conditions are meet.

4.1.The Description of E_CFSFDP

While CFSFDP algorithm needs one parameter d_c, our algorithmE_CFSFDP requires three parameters for the extensions presented in Section 3. As in CFSFDP, $d_c$is a value to computethe local density ρ in (1). [17] gives an empirical hint to specifyd_c, choosing the value so that the average number of neighbours is around *1%* to *2%* of the total number of points in the data set. $N_{neighbor}$ is the number of neighbours for modelling the similarity of sub-clusters. β is the factor to control the relative importance of inter-connectivity and closeness in the merge step.
We describe the steps of our algorithm E_CFSFDP in Algorithm 1.

Our algorithm starts at getting the similarity matrix of the data set. Many methods have been proposed by the machine learning community to measure the similarity between different objects, such as the common used Euclidean Distance, Mahalanobis Distance, Cosine Distance, SNN Similarity. One can choose appropriate method based on the application, i.e., Euclidean Distance for 2D data set, SNN Similarity for high dimensional data set. After that, we initialize the clustering group by CFSFDP. What's different from the standard CFSFDP, our scheme chooses as more cluster centres as possible in step 3 to overcome the limitation of CFSFDP described in Section 3.

In Phase II, the algorithm computes the value of necessary variables for merging in Step 5 and Step 6. In Step 5,the variant k-nearest neighbour graph is constructed, where algorithms based on k-d trees can be used [27]. In Step 7,E_CFSFDP combines two clusters into one in each iteration. To do so, the RI$(C_i, C_j) \times RC(C_i, C_j)^{\beta}$matrix needs toupdate in each loop, and the renewal of value refer to thesub cluster has been merged is necessary.

As a hierarchical algorithm, E_CFSFDP will not terminate until there is only one cluster remained. However, in real application, that seems meaningless. So in real application, E_CFSFDP usually terminates at the moment when the number of clusters is equal to the user specified parameter k. And CHAMELEON [21] proposes an alternative scheme to terminate the combination, that is to set two user input threshold T_{RI},T_{RC}, thus the algorithm stops when there is no cluster pair, of which the inter-connectivity RI and the closeness RC are bigger than the respond threshold T_{RI} and T_{RC}. While in the experiments of this paper, we stop thealgorithm by a value *k* of genuine clusters' number.

4.2. Performance Analysis

The runtime complexity of E CFSFDP mainly depends on the two phases in the scheme. In the initial phase, E_CFSFDP is almost the same as CFSFDP.

Let´s consider the runtime of CFSFDP. CFSFDP needs ○ $(n^2)$to compute local density and the distance where *n* is the number of objects in the input data set. For the progress to determine the cluster centres, we take no account of the time for the user to select cluster centres due to that is hard to quantize correctly. CFSFDP needs ○ (n)to construct thedecision graph. After the cluster centers are chose, CFSFD Passigns each point to the same cluster as by the order of descend local density, thus CFSFDP needs ○ $(n \log n)$ to sortthe points with quick sort [28], and CFSFDP's assignment procedure costs ○ (n), such that, the total time complexity of CFSFDP is ○ (n^2) if the similarity matrix has been finished. Inthe first phase, E_CFSFDP costs the same time as the overall time of original CFSFDP, indicates the time complexity of E_CFSFDP's first phase is ○ (n^2). Moreover, our analysis will focus on the computation of inter-connectivity and relative closeness which have been described in Section 3, for the main cost of E_CFSFDP's second phase comes from the computation and the update of these criteria.

The computation of the two criteria is based on the variant-nearest neighbour graph, which can be constructed based on the basic *k*-nearest neighbour graph. For low dimensional dataset, the amount of time to construct the basic *k*-nearest neighbour graph is $o(n \log n)$ if algorithms based on k-d trees [27] areused. For high dimensional data set, schemes based on *k-d* trees are not applicable ([29], [30]), leading to an overall time complexity of $o(n^2)$. To get the variant model presented in Section 3, our algorithm need to check all *k* nearest neighbours for each item, leading to the runtime complexity is $o(kn)$, *k* is the number of neighbours.

To simplify the analysis, we assume that each cluster is of the same size, thus if the number of sub clusters is *m*, the size of each cluster ism/n. E_CFSFDP needs to traverse part of the variant *k*-nearest neighbour graph to compute the interconnectivity and relative closeness for each cluster pair, of which data points belong to the cluster pair, resulting in time complexity of $o(2n/m)$. The amount of time of normalizing the inter-connectivity or the closeness is also $o(2n/m)$ for thebisection costs the same time as that of CFSFDP's alignmentprogress. In Step 5, there are $\binom{n}{2}$ cluster pairs to be solved,causing the amount of time $o(0.5n^2 \times 2n/m) = o(mn)$. Whilemerging iteratively, the runtime depends on the time of findingthe most similar cluster pair and updating the criteria. By usinga heap-based priority queue, the amount of time required tofind the most similar cluster pair is $o(m^2 \log m)$. The timeto update the matrix is $o(n)$ for there are most $m-1$ steps,each step requires the time of $o(2n/m)$.

Thus, the overall complexity of the extended CFSFDP(E_CFSFDP) is $o(n^2 + n \log n + mn)$. Comparedto the runtime of original CFSFDP $o(n^2)$, the largest term $n^2$of the two approaches both comes from the computation oflocal density, the amount of time of our algorithm doesn't increase much, especially for a large value of *n*.

5. EXPERIMENTS

In order to demonstrate the performance of our algorithm E_CFSFDP, we benchmark it on several 2D data sets consist of clusters of diverse density, shapes, which have not no unique density peak for each cluster. The reason why we adopt 2D data sets to benchmark our scheme is that it is easy to visualize the clustering results of 2D data sets, making the comparison of different algorithms much easier. Our experimental study focuses on the comparison of the basic CFSFDP and its extension E_CFSFDP proposed by us, to testify that E_CFSFDP can handle the data set, in which there is no unique density peak for each cluster.

This paper measures the similarity between data points with the famous Euclidean Distance, which is used widely to measure the similarity of spatial data, 2D or 3D. And though we determine parameters (d_c, $N_{neighbor}$, β) based onthe application, we will show that our algorithm E_CFSFDP is robust with respect to parameters by presenting the clustering results of E_CFSFDP with different combinations of value of d_c, $N_{neighbor}$, and β. Besides, for convenience, we use percentage to denote the value of parameter d_c, which is used in (1) to compute the local density, i.e., we say d_c = 2%,indicates that with the d_c, s.t.,max $\rho = 2\% \times size(dataset)$.

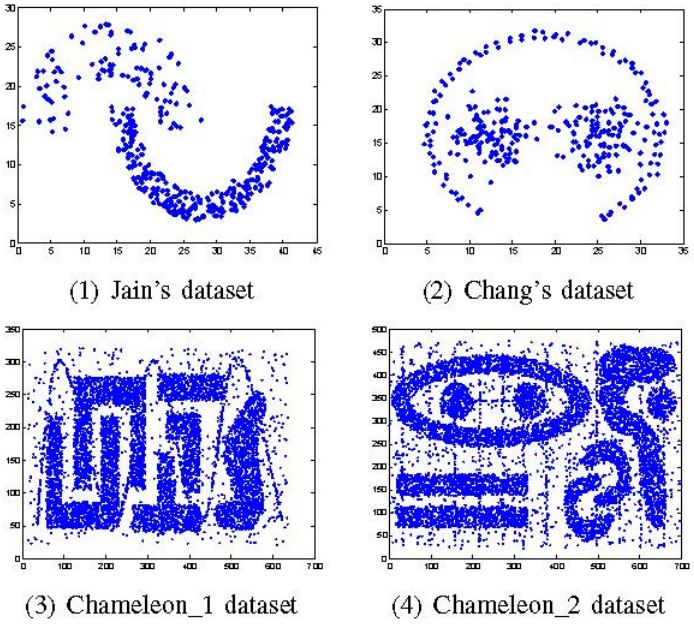


Figure 6. 2D data sets to benchmark E_CFSFDP

5.1. Data Sets

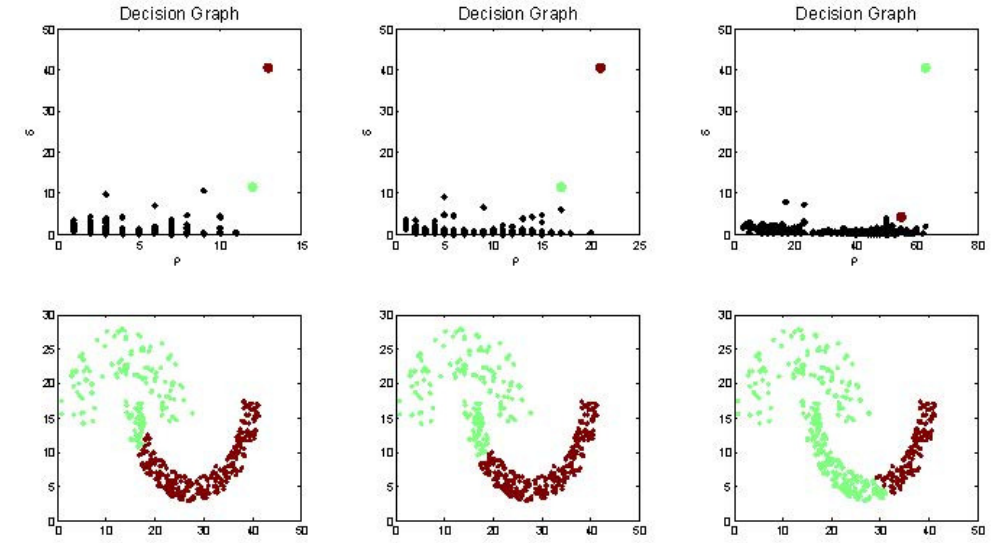
Four data sets proposed by other papers are used to evaluate our algorithm in this paper. In particular, Jain's dataset is taken from [23], Chang's dataset is taken from [22], Chameleon_1and Chameleon_2 are drawn from [21]. As presented in Fig.6,clusters of the 4 data sets are of diverse density, complex shapes. The most important feature of these data sets shared is no obvious single density peak for most clusters, even in Chang's data set, there is one cluster — the arc one — which don't have a unique density peak. We will show you that our algorithm performs better to deal with this case below.

5.2. The Evaluation of Clustering Results

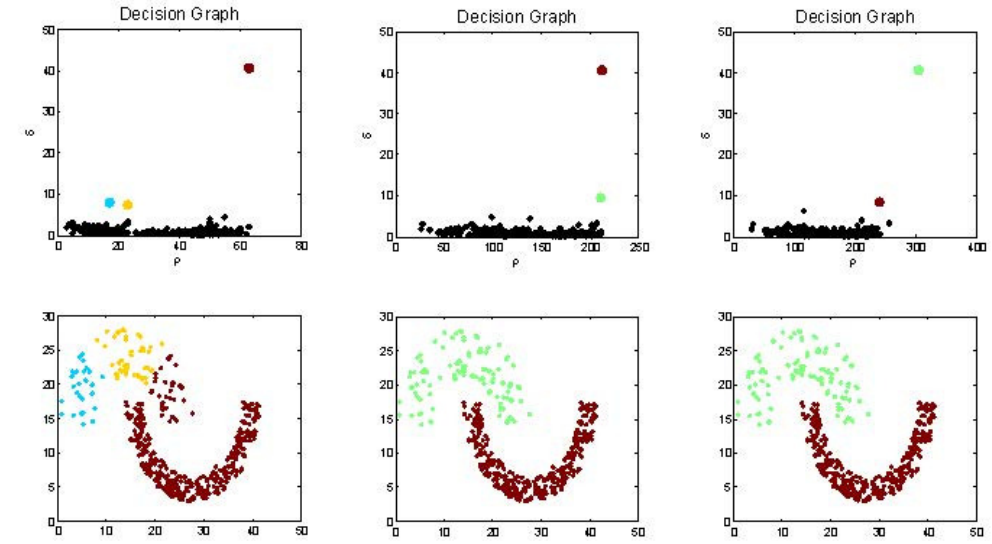
2.6.1. Jain's data set

As shown in Fig. 6, there are 2clusters, 373 data points in the data set, one of the clusters is denser than another. In this case, our approach E_CFSFDP expenses more time to be more robust respect to parameterd_c. Fig.7 presents a lot of clusters found by CFSFDP with different d_c and selections of cluster centers. For CFSFDP's performance is depended on the selection of density peaks, we not only present the clustering result, but also show the decision graph and the selections of cluster centers. As shown, CFSFDP gets perfect result when d_c = {40%, 45%}, however it fails with other values, as d_c = {1%, 2%, 10%}. By more experiments, we find CFSFDP also divides Jain's data set correctly with d_c = {44%, 46%, 47%}.

As an extension of CFSFDP, our scheme not only can find genuine clusters with the same d_c whereas CFSFDP succeeds (E_CFSFDP can just specify the number of sub-clusters to be 2),but also works fine whereas CFSFDP fails to perform well. E_CFSFDP finds the 2 natural clusters in above data set with parameters of any combination ofd_c = {4%, 5%, 6%},$N_{neighbor}$ = {5, 10, 15}, β = {1, 2, 3, 4, 5}. To present the clustering progress better, we present the initial clusters found by E_CFSFDP in Fig. 8. (2), (3), (4).



(1) $d_c = 1\%$, 2 clusters. (2) $d_c = 2\%$, 2 clusters. (3) $d_c = 10\%$, 2 clusters.



(4) $d_c = 10\%$, 3 clusters. (5) $d_c = 40\%$, 2 clusters. (6) $d_c = 45\%$, 2 clusters.

Figure 7. Clusters found in Jain's data set by CFSFDP with different dc. Different colours correspond to different clusters, there is no any relationship of colours in different figures.

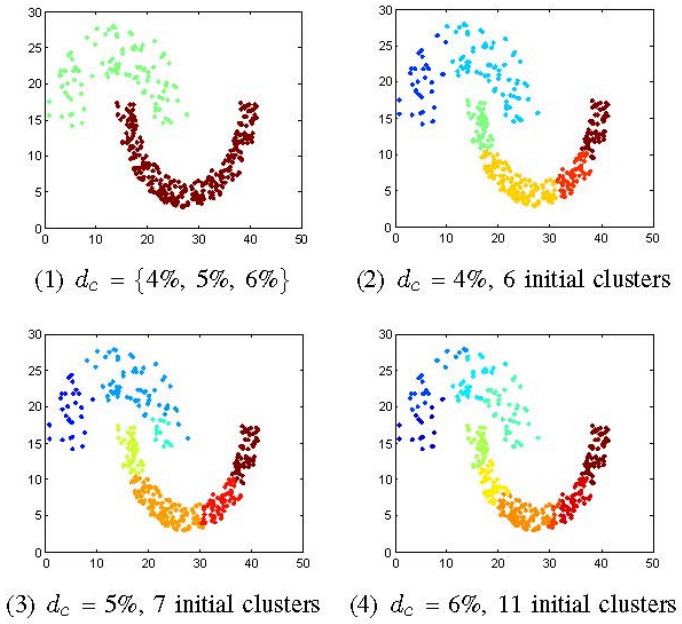


Figure 8. Clusters found by E CFSFDP in Jain's data set. (1) shows the final results of E CFSFDP. (2), (3), (4) show the initial clusters.

2.6.2. Chang's data set

There are 3 natural clusters in the data set, some clustering solutions proposed by CFSFDP have been given in Fig. 4. We also conduct E_CFSFDP on this data set, the result is presented in Fig. 9. In the experiment, we have test CFSFDP and E_CFSFDP on this data set with many parameters, the best result of each approach is given. In this case, E_CFSFDP is the obvious winner.

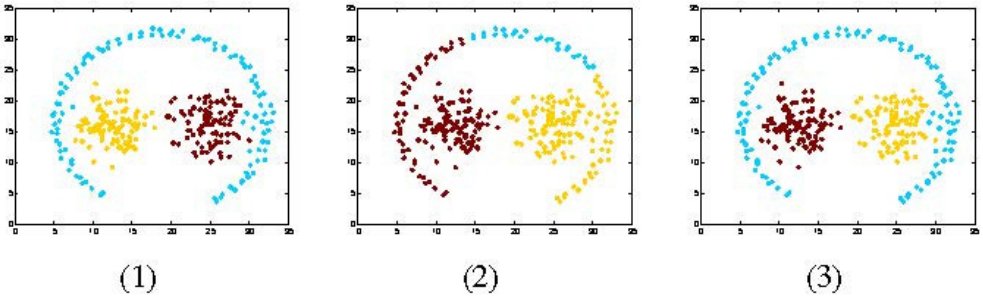


Figure 9. Clustering results for Chang's data set. (1) The original data set; (2)CFSFDP clustering result; (3) E CFSFDP clustering result.

Notice that, on this data set, the clustering result of E_CFSFDP is 100% correct, as shown in Fig. 9. (3). That doesn't mean E_CFSFDP is perfect, because in fact, with other values of d_c, some points between different clusters will be assigned to wrong cluster.

2.6.3. Chameleon 1 and Chameleon 2

There are 8000points, 6 clusters in Chameleon_1 data set and 10000 points,9 clusters in Chameleon_2 data set. By experiments, we find that CFSFDP fails to get natural clusters with any value of dc both on the data sets. Fig.10.(1), (2) show the best clustering results we got by CFSFDP on these data sets. Fig.10.(3), (4)show the best clustering results we got by E_CFSFDP on the two data sets. The two data sets are much denser than the Jain's data set and also Chang's data set, as a result, we use larger values of $N_{neighbor}$, {30, 35, 40}, smaller values ofd_c, {1%,2%, 3%} on these data sets than those of Jain's and Chang's data sets. On these cases, it is clear that E_CFSFDP performs much better than what CFSFDP does.

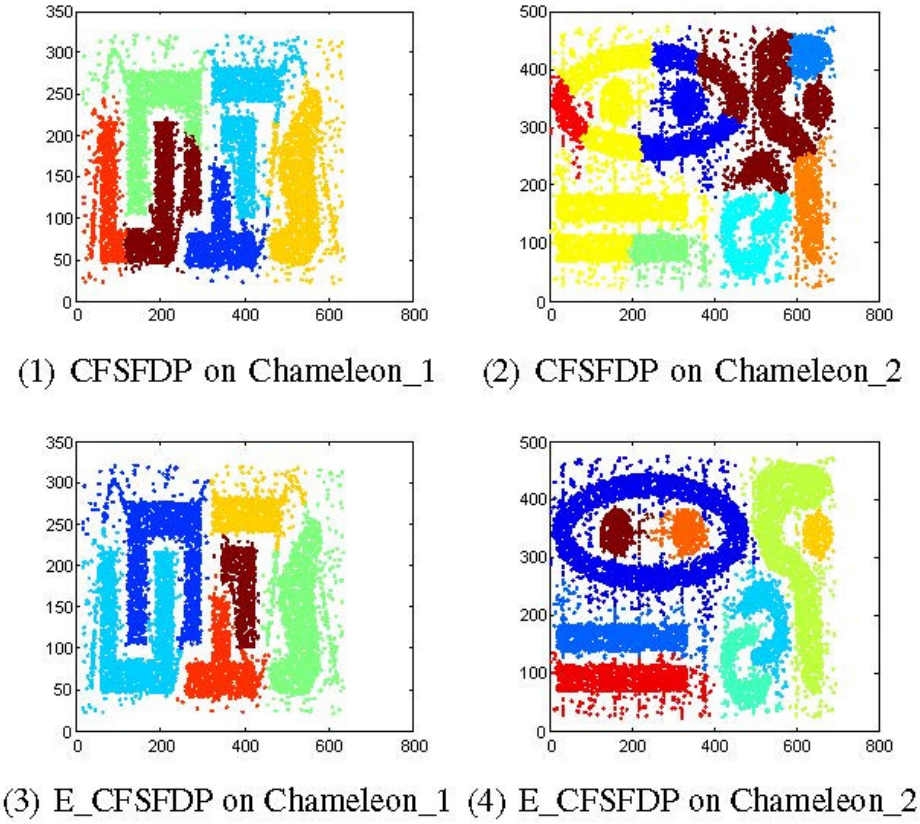


Figure 10. Clustering results of different schemes for the Chameleon data sets.

6. CONCLUSION AND FUTURE WORK

In this paper, we extend a new density-based clustering algorithm CFSFDP [17] to break through the barrier that CFSFDP performs well only when there is unique density peak of each cluster in the data set, what we named as "no density peak". Our solution is taking CFSFDP to get initial clusters, then merge the sub clusters to get final clustering result. The merge progress of our scheme is inspired by the merge phase of CHAMELEON [21]. Instead of the k-nearest neighbour graph, we model the similarity of sub clusters by a variant of k-nearest neighbour graph. Except for that, another difference between our approach and CHAMELEON is that E_CFSFDP is based on density, while CHAMELEON is based on graph partition. In order to demonstrate the

applicability of our algorithm to solve the case ”no density peak”, we conduct CFSFDP and E_CFSFDP on several 2D data sets, there is no density peak for clusters of which. Although our method doesn't increase much more run time complexity than the original algorithm, it's true our algorithm spends more time than the original one.

In future studies, we will focus on reducing the run time of E_CFSFDP. What's more, we will apply it to high dimensional data sets. Another interesting direction is running it in parallel.

ACKNOWLEDGEMENTS

This work was supported by Key Technologies Research and Development Program of China (No.2012BAH17B03),and the Chinese Academy of Science (No. 2014HSSA09).

REFERENCES

[1] P. Berkhin, “A survey of clustering data mining techniques,” in GroupingMultidimensional Data, J. Kogan, C. Nicholas, and M. Teboulle, Eds.Springer Berlin Heidelberg, 2006, pp. 25–71.

[2] H. Frigui and R. Krishnapuram, “A robust competitive clustering algorithmwith applications in computer vision,” Pattern Analysis andMachine Intelligence, IEEE Transactions on, vol. 21, no. 5, pp. 450–465, May 1999.

[3] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Su?sstrunk,“Slic superpixels compared to state-of-the-art superpixel methods,” PatternAnalysis and Machine Intelligence, IEEE Transactions on, vol. 34,no. 11, pp. 2274–2282, Nov 2012.

[4] E. Elhamifar and R. Vidal, “Sparse subspace clustering,” in ComputerVision and Pattern Recognition, 2009. CVPR 2009. IEEE Conferenceon, June 2009, pp. 2790–2797.

[5] W. Li and A. Godzik, “Cd-hit: a fast program for clustering andcomparing large sets of protein or nucleotide sequences,” Bioinformatics,vol. 22, no. 13, pp. 1658–1659, 2006.

[6] A. D. King, N. Pr?ulj, and I. Jurisica, “Proteincomplex prediction via cost-based clustering,” Bioinformatics,vol. 20, no. 17, pp. 3013–3020, 2004.

[7] D. W. Huang, B. T. Sherman, and R. A. Lempicki,“Systematic and integrative analysis of large gene listsusing david bioinformatics resources,” Nat. Protocols,vol. 4, no. 1, pp. 44–57, 2008.

[8] F. Moosmann, E. Nowak, and F. Jurie, “Randomized clustering forestsfor image classification,” Pattern Analysis and Machine Intelligence,IEEE Transactions on, vol. 30, no. 9, pp. 1632–1646, Sept 2008.

[9] A. Ducournau, A. Bretto, S. Rital, and B. Laget, “A reductive approachto hypergraph clustering: An application to image segmentation,” PatternRecognition, vol. 45, no. 7, pp. 2788 – 2803, 2012.

[10] T. Chaira, “A novel intuitionistic fuzzy c means clustering algorithm and its application to medical images,” Applied Soft Computing,vol. 11, no. 2, pp. 1711 – 1717, 2011, the Impact of Soft Computingfor the Progress of Artificial Intelligence.

[11] M. Zheng, J. Bu, C. Chen, C. Wang, L. Zhang, G. Qiu, and D. Cai,“Graph regularized sparse coding for image representation,” ImageProcessing, IEEE Transactions on, vol. 20, no. 5, pp. 1327–1336, May2011.

[12] S.-J. Horng, M.-Y. Su, Y.-H. Chen, T.-W. Kao, R.-J. Chen, J.-L. Lai, andC. D. Perkasa, “A novel intrusion detection system based on hierarchicalclustering and support vector machines,” Expert Systems withApplications, vol. 38, no. 1, pp. 306 – 313, 2011.

[13] M. Ester, H. peter Kriegel, J. S, and X. Xu, "A density-based algorithmfor discovering clusters in large spatial databases with noise," inKDD'96. AAAI Press, 1996, pp. 226–231.

[14] J. Sander, M. Ester, H.-P. Kriegel, and X. Xu, "Density-based clusteringin spatial databases: The algorithm gdbscan and its applications," Data Mining and Knowledge Discovery, vol. 2, no. 2, pp. 169–194, 1998.

[15] M. Ankerst, M. M. Breunig, H.-P. Kriegel, and J. Sander, "Optics:Ordering points to identify the clustering structure," in Proceedings of the 1999 ACM SIGMOD International Conference on Managementof Data, ser. SIGMOD '99. New York, NY, USA: ACM, 1999, pp.49–60.

[16] X. Xu, M. Ester, H.-P. Kriegel, and J. Sander, "A distribution-basedclustering algorithm for mining in large spatial databases," in DataEngineering, 1998. Proceedings., 14th International Conference on, Feb1998, pp. 324–331.

[17] A. Rodriguez and A. Laio, "Clustering by fast search and find of densitypeaks," Science, vol. 344, no. 6191, pp. 1492–1496, 2014.

[18] K. Fukunaga and L. Hostetler, "The estimation of the gradient of adensity function, with applications in pattern recognition," InformationTheory, IEEE Transactions on, vol. 21, no. 1, pp. 32–40, Jan 1975.

[19] Y. Cheng, "Mean shift, mode seeking, and clustering," Pattern Analysisand Machine Intelligence, IEEE Transactions on, vol. 17, no. 8, pp.790–799, Aug 1995.

[20] F. Samaria and A. Harter, "Parameterisation of a stochastic model forhuman face identification," in Applications of Computer Vision, 1994.,Proceedings of the Second IEEE Workshop on, Dec 1994, pp. 138–142.

[21] G. Karypis, E.-H. S. Han, and V. Kumar, "Chameleon: Hierarchicalclustering using dynamic modeling," Computer, vol. 32, no. 8, pp. 68–75, aug 1999.

[22] H. Chang and D.-Y. Yeung, "Robust path-based spectral clustering," PatternRecognition, vol. 41, no. 1, pp. 191 – 203, 2008.

[23] A. Jain and M. Law, "Data clustering: A users dilemma," in PatternRecognition and Machine Intelligence, ser. Lecture Notes in ComputerScience, S. Pal, S. Bandyopadhyay, and S. Biswas, Eds. SpringerBerlin Heidelberg, 2005, vol. 3776, pp. 1–10.

[24] G. Karypis and V. Kumar, "Multilevel k-way hypergraph partitioning,"in Proceedings of the 36th Annual ACM/IEEE Design AutomationConference, ser. DAC '99. New York, NY, USA: ACM, 1999, pp. 343–348.

[25] S. Guha, R. Rastogi, and K. Shim, "Cure: An efficient clusteringalgorithm for large databases," in Proceedings of the 1998 ACMSIGMOD International Conference on Management of Data, ser.SIGMOD '98. New York, NY, USA: ACM, 1998, pp. 73–84.

[26] ——, "Rock: a robust clustering algorithm for categorical attributes," inData Engineering, 1999. Proceedings., 15th International Conferenceon, Mar 1999, pp. 512–521.

[27] "Book reviews," IEEE Computer Graphics and Applications, vol. 10,no. 5, pp. 86–89, 1990.

[28] C. A. R. Hoare, "Quicksort," The Computer Journal,vol. 5, no. 1, pp. 10–16, 1962.

[29] S. Berchtold, C. B¨ohm, D. A. Keim, and H.-P. Kriegel, "A costmodel for nearest neighbor search in high-dimensional data space,"in Proceedings of the Sixteenth ACM SIGACT-SIGMOD-SIGARTSymposium on Principles of Database Systems, ser. PODS '97.New York, NY, USA: ACM, 1997, pp. 78–86.

[30] S. Berchtold, C. B¨ohm, and H.-P. Kriegal, “The pyramid-technique: Towards breaking the curse of dimensionality,” in Proceedings of the1998 ACM SIGMOD International Conference on Management of Data, ser. SIGMOD ’98. New York, NY, USA: ACM, 1998, pp. 142–153.

AUTHORS

Wen-Kai Zhang now studies for a M.S degree of School of Computer Science and Technology at University of Science and Technology of China. He received the B.E degree at School of Life Science of USTC. His research mainly concentrated on data mining , anomaly detection.

Jing Li is a Professor of School of Computer Science and Technology at University of Science and Technology of China. He received the Ph.D degree at USTC in 1993. His research interests include Distributed Systems, Cloud Computing, Big Data progressing and Mobil Computing.

INTENTIONAL BLANK

AL-HARM EXPANSION MOVIE BASED ON VIRTUAL REALITY

Alanoud Salem[1], Sara Musallam[2], El-Shaimaa Nada[3] and Ahmed Ahmed[4]

[1,2]Department of Information System, Taibah University, Almadinah, KSA
nody_the_mooody@hotmail.com
[3]Department of Computer Science, Taibah University, Almadinah, KSA
Elshaimaanada2012@hotmailcom
[4]Department of Computer Enginnering, Taibah University, Almadinah, KSA
dr.aaahmed@hotmail.com

ABSTRACT

Animated movies are excellent virtual environments for creating models in high quality. Animated movies can include 3D models, sounds and lights effects, and detailed maps. In this paper, a virtual reality movie is applied to Al-Haram Expansion stages including the future stage of expansion. 3DMAX program is used to rich the maximum benefits of using 3D modeling. Maps with details are built by using ARCGIS program in order to understand the real difference between the three different expansions stages clearly and effectively. A novel technique is presented in order to insert 2D maps and other details in the 3D model built by 3DMAX.

The animated movie includes introduction, three basic expansion stages, maps, and conclusion. It is in Arabic with translation to English and sign language is also included. Accurate and documented information is applied in AL-Haram expansion movie that helps people to know about one of the most radical expansions that is occurred in the world.

KEYWORDS

VRML, VR, 3DMAX , ARCGIS, modelling, 3D models

1. INTRODUCTION

Virtual Reality (VR) is a popular system for simulating reality using computer technology that makes the user indulged in the environment. It makes the world to be real enough, using enormous computing power.

Our work presents a 3D modeling show on the stages of the expansion of Al-Haram Mosque in Mecca. The movie contains the different stages of the past, the present and the future and the changes that have been taken place on each stage. Mapping of the expansion stages is also included. 3DMAX and ARCGIS software are chosen in order to create a high quality 3D model with accurate map of each stage.

There is a continuous increase over time to expand the Al-Haram Mosque in Mecca in Saudia Arabia. For the best of our knowledge, there is no such accurate and documented information about the stages of the expansions of Al-Haram Mosque that has been considered in any previous work.

Jan Zizka et al. (Eds) : DBDM, CICS, CCNET, CSIP - 2015
pp. 19–28, 2015. © CS & IT-CSCP 2015 DOI : 10.5121/csit.2015.50702

In this work, Displaying of expansions will include the details of three stages in 7 minutes of 3d model. To create a high quality model, 3DMAX modeling software is used with sound, light features. Static maps are also included to clarify the expansion of each stage using state of art ARCGIS software. The movie will be in Arabic language and will also be translated to English. Our goal is to reduce the time of modeling by using 3DMAX program, that provide visual environment rather than wasting time in coding , to achieve the required interaction between the elements that's in virtual environment by the correct use of tools (sound, light for example) that come with the modeling software , to give the right perception for the events that occur in the Al-Haram by present the scene as virtual environment that simulate the reality in mutable site especially official events like al-hajj , and to clarify the need of the expansion, and to improve the quality of the models made by VRML by merging between 3d modeling of the environment of study and static maps that show the changing details of the this environment.

A novel approach of inserting 2D maps designed by ARCGIS into 3DMAX program, this is done to convert these maps into 3D and to add more features to them such as sound effects, color and reduce errors of small details of the maps.

In next section, a literature review is introduced. In section 3, virtual reality of Al-Haram New Expansion is presented. In section 4, an analysis of online questionnaire is applied. The creating models and proposed technique is presented in section 5. The conclusion and the future work is in section 6.

2. LITERATURE REVIEW

The application of VR causes the classical architecture used for protection and a demonstration to be more scientific, effective and appropriative. It also responds to a request from the property and the current information age. Its goal is to use enormous computing power and chart an exaggeration computer capacity, to make the assumption that the world seems real enough. In this section a description of some related work using VR is presented.

In [1], it presents Crowd Simulation in Emergency Aircraft Evacuation using Virtual Reality CGAMES. VR can be used to produce a realistic environment involving multiple sensory channels VR crowd simulation can be used by airlines to simulate emergency scenarios without the use of live actors. But in this work, there are some considerations such as: The whole system is complex to implement. It didn't simulate each case of emergency that might happen. It didn't consider the external environment enough that can effect in the plan. The programs used in modelling this environment are 3DMAX, Maya program and Virtual Reality Modelling Language (VRML) Virtools.

In [2], it shows the Virtual Reconstruction of FouGuang Temple Based on Virtual Reality. In this work, the history of the humanities and browsing the heritance physical geography roaming is presented. Protection, introduction of humanities history, geography and physical side can play very vital role in the international virtual reality technology. The limitations of this work are: Control buttons on the screen are not provided by the VRML browser and some building components have complex geometric models this leads to some difficulties in modelling all components. It is hard for the user to observe all the details carefully because details of modelling are complex to be modelled. The program used for presenting this 3D environment is 3DMAX.

In [3], it presents Relief Mapping on Facade of Sino Portuguese Architecture in Virtual Reality it presents a mapping technique to reduce the number of polygonal built in model. Virtual reality is applied to a door, wall and house. Graphic card is used to apply great shadows and lighting in

image. Implementation of VR for a model is done to present a Walk through the environment to give information in real time. The program used for building the 3Dmodel is 3DMAX.

The work in [4] aims to investigate the role of Virtual Reality and three-dimensional (3D) computer modelling on learning and teaching in a school of the environment. In order to achieve this aim a number of academic experiences were analyzed to explore the applicability and viability of 3D computer modelling and VR into the modelled environment areas. Although 2D representations have been greatly accepted for obtaining professions and education environment model, 3D computer representations and VR is still required. The general characteristics of this work are: Offering an approach on how three-dimensional computer modelling and virtual reality may be integrated into built environment teaching. A strategic, systematic approach was adopted to raise awareness of VR technologies in the school. A total of 11 semi-structured interviews were conducted with key participants involved in the integration of 3D technologies into the built environment curriculum. Each interview lasted one hour and was audio-taped to facilitate data analysis. After careful consideration of location and needs of users, the facility was centrally situated in the heart of the school, to allow easy access and to promote this technology to students, staff and visitors. Allowing staff and students to view designs in stereoscopic format, from multiple viewpoints, and navigate through space in real time. The aim from the outset was to foster VR applications across all disciplines within the School and to encourage collaboration with local practices and other researchers. The programs used for building this school environment are 3DMAX and VR4MAX.

In [5], it created the 3D model of Multipath by loading models in other software or programming by using Arc Object, because it cannot be directly created by ARCGIS. The methods used to create virtual campus are studied based on 3D modelling, database and image processing, such as using 3DMAX to create 3D model and putting on the texture image, and by ARCGIS special analysis and C# redeveloping specific function. It provides more credibility guarantee for the campus planning and sustainable development. The model of virtual campus compiles with image, text, sound, video to create a vivid 3D campus, which is basis and platform of the digital campus to allow users to access virtual environment. In this work, Data digitization is done as in the following steps:

- Importing the base maps: firstly import into ArcMap the campus planning maps.

- Raster and rectify: because of lacking their geographic coordinates the existing planning map data need to rectify.

- Creating the elements layers: it is used create appropriate layers file in ArcCatalog, base on the ruler of graphics and objects, combined with the practical needs of the virtual campus.

- Vectorization on screen: Adding layer file that created in the ArcCatalog to ArcMap, and activating the ArcScan expansion toolbar.

The programs used in this work are 3DMAX, Arc GIS and Photoshop and the programming languages are C# and SQL.

In [6], it deals with Suzhou University of Science and Technology as an example, designed a real-time virtual campus roaming system. It discussed the three-dimensional modelling of buildings, gates, roads, bridges, surrounding lawns, trees and other objects on campus, and implemented the data change between 3DMAX and VRML Pad. At the same time, it used a model with kinds of texture maps, multi-resolution hierarchy, and achieved the scene using DEF/USE ways of VRML

language. This work focuses on some details such as a variety of scene building, scene model and mapping methods, code conversion, optimization techniques, interactive programming based on the triggers, cruise, controlling viewpoints and communication between JAVA and scene.

3. VIRTUAL REALITY OF AL-HARAM NEW EXPANSION

Although the New expansion of Al-Haram is a great event especially for Muslims, we found no scientific research on the details of the new expansion. There are many videos which show the steps of the expansion of Al-Haram but most of them have a lack of documentation and clarification and even the information cannot be completely trusted.

It is worth to mention here that there are some videos on Al-Haram expansion which has appeared recently by Al-Haram ministry of foreign affairs.

This documentary show in [7] presents the project of expansion space of Almtaf during nine minutes nearly and by Arabic language.

The advantages of this show are:

- The time period to present the project is good.
- The film is supported by date.
- The problem is well defined.

The disadvantages of this show are:

- The film doesn't have more than one language.
- The film contains duplicate scenes.
- Like most other films is talking about some part of expansion.
- The film doesn't contain any illustrative maps.

In our work, 3DMAX software is used to make movie about the stages of the expansions of Al-Haram mosque in Mecca with providing accurate and documented information in clear and simple way. The movie will take 7-minutes and includes introduction, three stages and conclusion. The three stages are King Fahd Expansion that is happened in the past and King Abdullah Expansion in the current period and king Salman Expansion in the future. We selected these stages because these stages made radical changes in Al-Haram.

The movie will be in 3D modelling for all objects in order to achieve a very understandable movie, virtual reality is applied because of its advantages in the speed, ease and information accessibility for the recipient with the best representation.

In addition to the movie, maps will be added at the end of the movie by using ARCGIS software. The maps will be for the three stages which describes the expansion in the form of illustrations and detailed maps. We add effects such as sound, translation to achieve the required interaction between audience and the movie. The main advantages of the presented work

- Conveying the information in a simple, clear in the shortest possible time.
- Presenting accurate and trusted information as the information comes from an accredited Side (Al-Haram development staff).
- Adding a translation of the video to benefit the most people, especially non-native speakers of Arabic.

There are some limitations of our work such as using only two languages (Arabic, English) Multi-languages and sign languages movie will be our next step.

4. ANALYSIS OF APPLIED QUESTIONNAIRE

At the beginning, an online questionnaire has been done to determine the user requirement that will be used in making the movie. A survey contains three groups: Group number one includes general information (Name, Age, Gender, Education).We asked about the age to know which age group have enough information, the Gender will tell us which part of the society have more interest in having information on the expansions, and the benefit of education question is to clarify how the people are effected with their level of Education and how this will be impact in their interest.

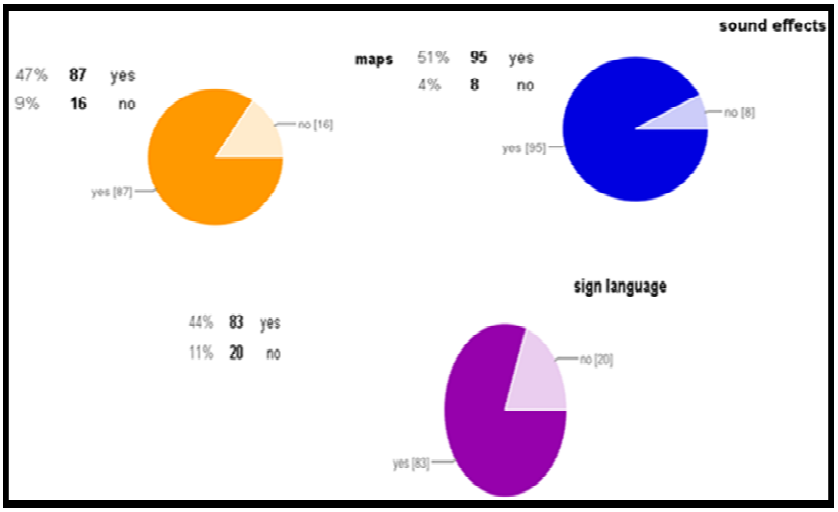


Figure1. Results from group number one

The description of Results from group number one are presented in Figure 1, the results show that 30% of the people who did the survey were female and 25% of the people were male , and most of them their age were above 30 and our survey clarifies that they are interesting in Al-Haram expansion information.

Group number two is about Al-Harm Expansion, it discusses the information that people have ,and their interests about getting information about Al-Haram expansion , how they prefer to get information about Al-Haram Expansion, how they want us to model the movie (3D,map,photos) , what is the appropriate time for them to present the movie , and which language is more suitable for them.

The description of Results from group number two are presented in Figure 2 presents that 52% of people want to get information about Al-Haram expansion in form of movie and 43% of them want to that include maps, photos, and they want it in 3D form, and 26% of people prefer to present data based on King who has been during his reign and the time of its occurrence.
It also shows that 36% of people want the information of Al-Haram expansion to be visible and heard and 43% of them prefer the time to be around 10 minutes and 32% want the language of the movie to be in Arabic.

In Group number three we asked them about the future work they want it to be available in our movie (sign language, sound effects, detailed map). The description of Results from group

number three are presented in Figure 3, it shows that 44% of people prefer to add sign language and 51% of them want to place sound effects and 47% want to see maps in the movie.

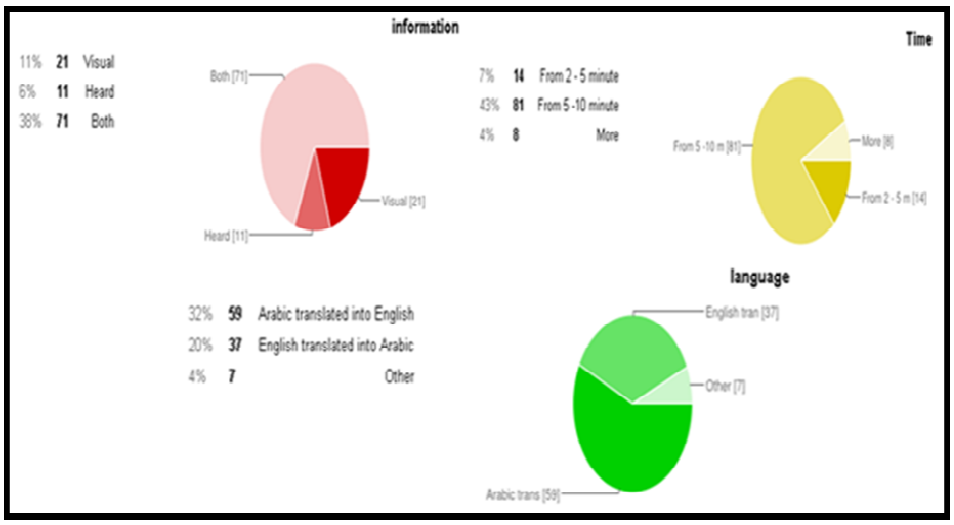


Figure 2. Results from group number two

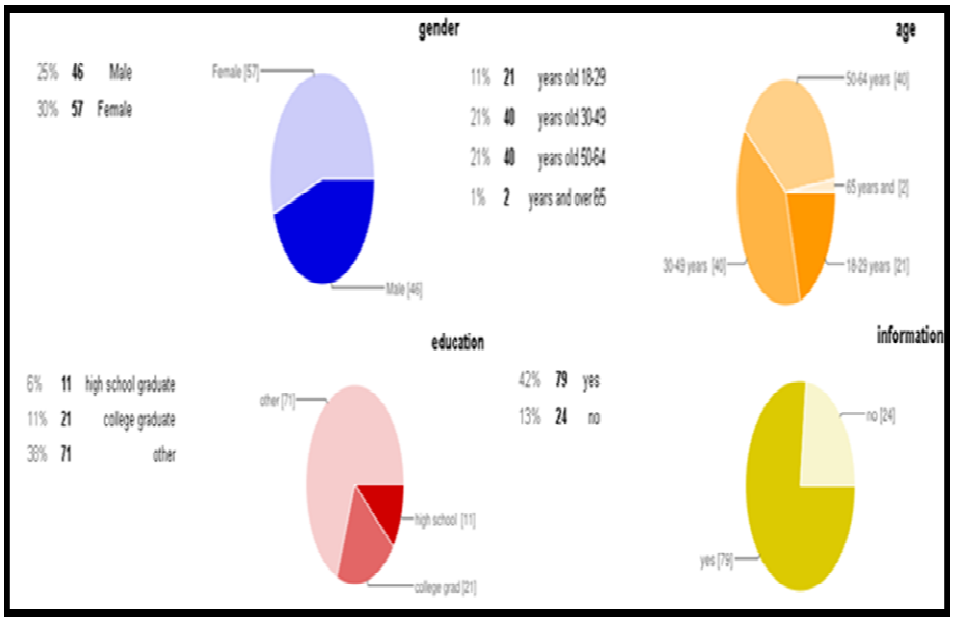


Figure 3. Results from group number three

5. CREATING MODELS FOR THE PROPOSED TECHNIQUE

The description of systems architecture of Al-Haram movie is presented in Figure 4, it contains:

- Introduction: the movie of al Haram new expansion starts with projection to Makkah the voice will tell general information about Makkah and special place of Makkah for muslims.

- Stage 1: It starts with 3D image of king Fahd (the expansion occur in his time) and will display information about the first stage (in the past) of expansion.
- Stage 2: It starts with 3D image of king Abdullah (the expansion occur in his time) and will display information about the second stage (in the present) of expansion.
- Stage 3: It starts with 3D image of king Abdullah and king Salman (the expansion occurs in his time) and will display information about the third (future) stage of expansion.
- Maps: They will display descriptive maps of the three stages.
- Conclusion: the movie will end with the benefits of this expansion in Muslims life and display how it's suited with the increasing number of Muslims.

The proposed technique in this paper is based on inserting Alharam Expansions maps designed by ARCGIS into 3DMAX program to make a complete view of the movie.

The created maps should be converted into3D before importing it, and then modification of the 3D maps is required. Because the small error is difficult to be found and the errors in the original images are hard to be modified, we smooth the maps that have been imported from the ARCGIS. This will increase the accuracy.

In 3DMAX, the models of the buildings are designed and the effects of the buildings are rendered. An example of Al- Kahba is shown in figure 5.
In this work the virtual buildings construction is an important part. Two methods are used in order to model and construct the virtual environment

- First: construction of virtual environment using 3DMAX.
- Second: convert the 3D model to VRML code.

The advantage of this method is that it is easy for constructing complex building but the code will become long and to solve this problem we do not choose a Normal's option, Where Normal's option is used for exporting the geometry of 3DMAX and makes VRML files very larger.

Using 3DMAX to create models of visual environment it taken shorten time of coding and debugging but only when we have a good model design that was created before .
Hybrid between the methods which enable back and forth switching from visually oriented tools to line editors and great high quality of model.

An example of one of created maps using ARCGIS program is presented in figure 6.
The figure shows how to create the map form the projection step to the final step to present a complete map. Figure 7 shows the final map of first stage.
There are many considerations while building Models such as:

- Data pretreatment: information and required data about the expansion is obtained from Al Haram development staff.
- Modeling of building: Complicated features, such as building details and Minarets are very difficult to create. We use AutoCAD with 3DMAX software.
- Optimizing the virtual system: The virtual system must keep proper frame frequency. Some optimizing approaches have been applied in the scene.
- 3D manipulations: 3D manipulations including zoom-out, zoom-in, group movement, shadows, and multi-angle shots have been realized.
- Adapting Virtual Reality for Design: VR is applied to provide several phases of improving the design of the area.
- Creating the environment: Creating the environment with the small details such as shadows and clouds is important to make the scene look more real, and to get the user attention.

- Import audio: to make the movie more attractive, importing audio is require, the audio will be played from the begging of the movie to explain the scenes and to give information to the audience.
- Layout design: The VR models permit 3D visualization and fly through animation to represent environmental quality, based on quantifiable image parameters, and assist the understanding of the complex. Without a VR system, the details could be evaluated only after completion of construction. In the design stage, it gives the ability to visualize potential modifications to the space.

The VR of the final first stage is shown in figure 8.

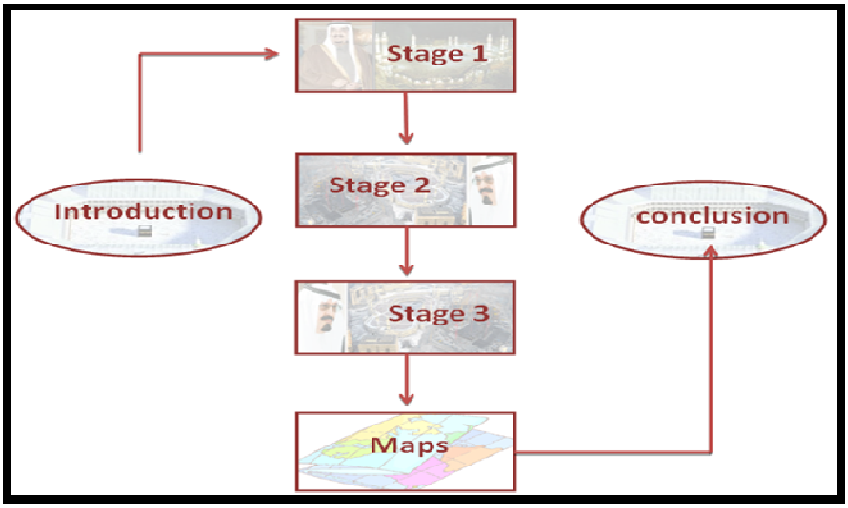
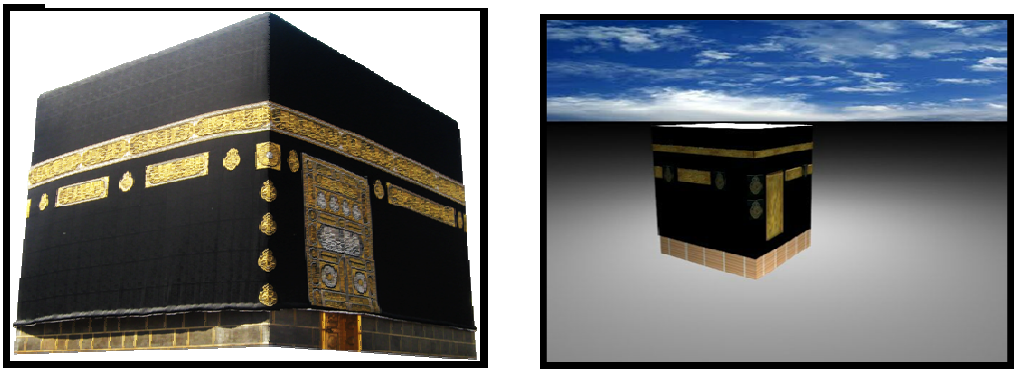


Figure 4. System architecture of Al-Haram movie



The real image implementation in 3DMAX

Figure 5. Creating environment that looks like the real image

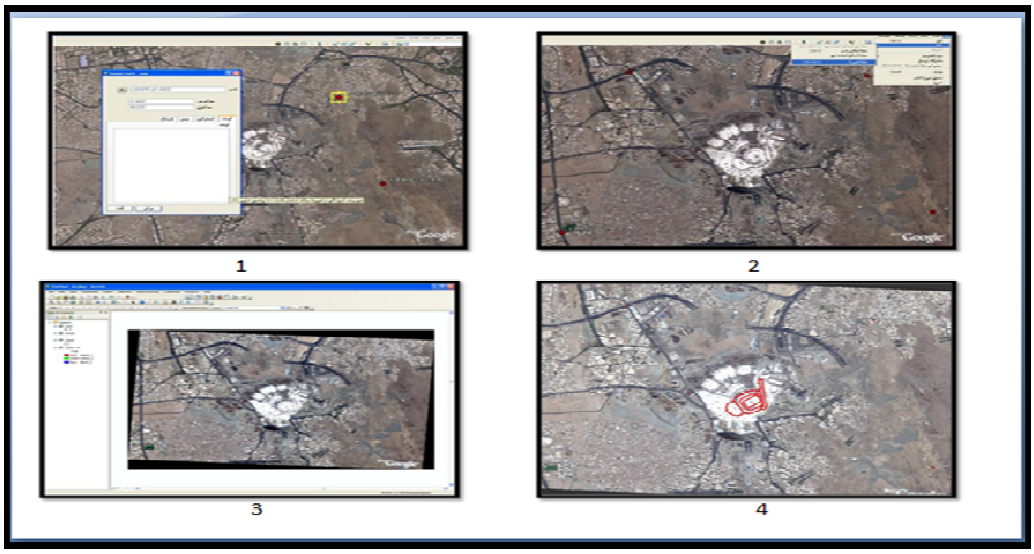


Figure 6. Designed maps using ARCGIS

6. CONCLUSION

In this paper, VR of Al-Haram Expansion Movie is presented. The last recent three stages of Al-Haram expansion are shown with 3D maps of every stage. A novel technique of converting the 2D maps designed by ARCGIS program into 3D maps designed by 3DMAX program is presented. Many features are added in the movie such as sound, audio, translation. The movie is presented in Arabic and is translated in English. Sign language is added.

In the future, based on the use requirements we decide to add a real time view, languages other than Arabic and English, and we will expand the stages to include the first early stage of the expansion stages.

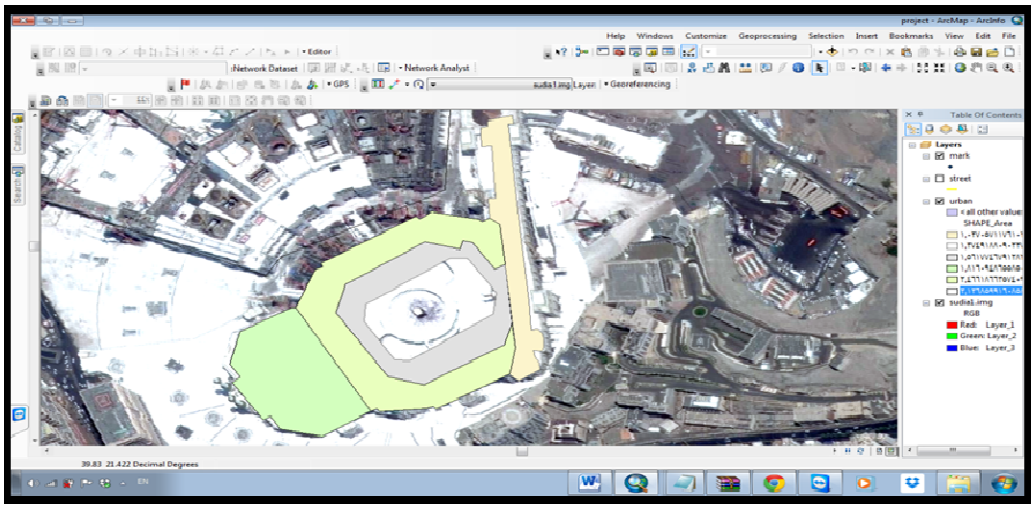


Figure 7. The final map of first stage

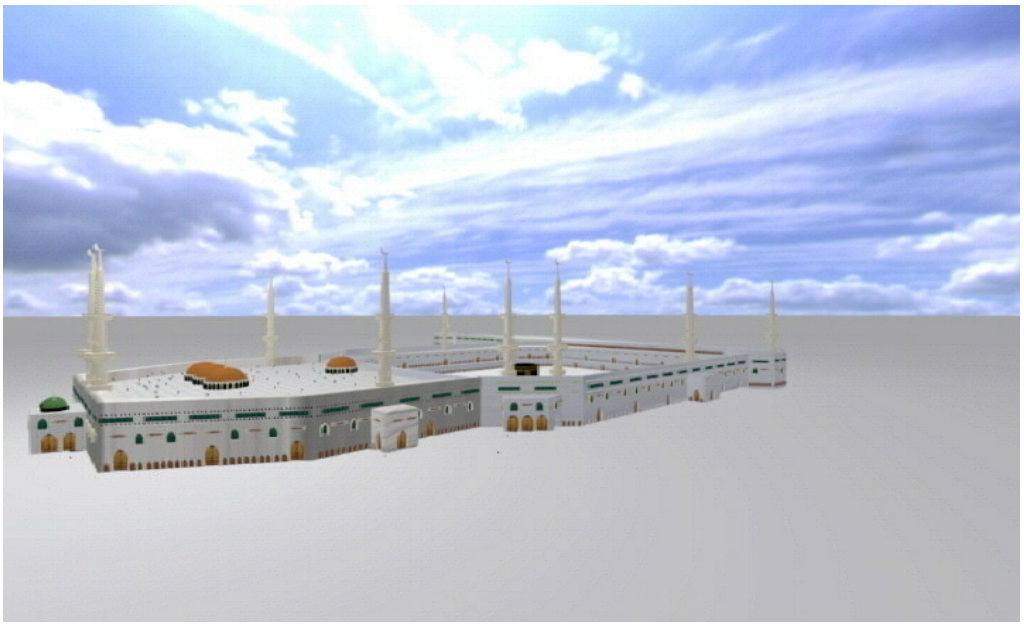


Figure 8. VR of the final first stage

REFERENCES

[1] Sharad Sharma, Stephen Otunba, and Jingxin Han .(2011) "Crowd Simulation in Emergency Aircraft Evacuation using Virtual Reality" . USA: Department of Computer Science.

[2] Lu Shi Zhu. (2008) "Virtual Reconstruction of FouGuang Temple Based on Virtual Reality" . Nanchang: IEEE.

[3] Kalarat, K. (2014) "Relief Mapping on Facade of Sino Portuguese Architecture in Virtual Reality". Fourth Bangkok : IEEE.

[4] Margaret, Emine , (2007) International Conference on Construction Applications of Virtual Reality. UK: Northumbria.

[5] Shiliang Zhanga, Xiaobinhuangb , (2012) "Development of virtual Campus System Based on ARCGIS. Department of Information and Engineering Ningde Normal University Ningde, Fujian", Province, China: by Elsevier Ltd. Selection and/or peer-review under responsibility of ICMPBE International Committee.

[6] Zhao Mei-Hong.(2012) "Design and Implementation of Modeling Virtual Campus", Vol. 1, No.4. North University of China: IACSIT.

[7] Information available from their website (http://www.mofa.gov.sa/Pages/Default.aspx , 2005).

NETWORK FAULT DIAGNOSIS USING DATA MINING CLASSIFIERS

Eleni Rozaki

Cardiff School of Computer Science & Informatics,
Cardiff University, Cardiff, UK
E.Rozaki@cs.cardiff.ac.uk

ABSTRACT

Mobile networks are under more pressure than ever before because of the increasing number of smartphone users and the number of people relying on mobile data networks. With larger numbers of users, the issue of service quality has become more important for network operators. Identifying faults in mobile networks that reduce the quality of service must be found within minutes so that problems can be addressed and networks returned to optimised performance. In this paper, a method of automated fault diagnosis is presented using decision trees, rules and Bayesian classifiers for visualization of network faults. Using data mining techniques the model classifies optimisation criteria based on the key performance indicators metrics to identify network faults supporting the most efficient optimisation decisions. The goal is to help wireless providers to localize the key performance indicator alarms and determine which Quality of Service factors should be addressed first and at which locations.

KEYWORDS

Fault diagnosis, Weka classifiers, Rules, Decision trees, Bayesian networks.

1. INTRODUCTION

In this paper, we propose and develop an automated scheme for monitoring and fault detection based on data mining techniques to provide converged services with functionalities that facilitate its fault-handling and operational management. With the use of Weka classifiers (rules, trees, Bayes), one of the main challenges is to carry out effective and accurate monitoring of services with the detection and classification of possible faults that may arise at runtime so that appropriate measures rules and decision trees can be made to ensure correct fault-handling and compliance with the Quality of Service (QoS) constraints established between service providers and end-users [1].

The FCAPS framework was defined and implemented in 1997 as a management framework for network provides. The FCAPS framework stands for fault, configuration, accounting, performance, and security. Within this framework, performance management is the process of quantifying, measuring, analysing, and controlling the performance of the components of a network. Fault management is the process of acquiring, revealing, and then counteracting faults that exist in the network. Configuration management is the process of obtaining and controlling the configuration parameters of a network. Accounting management refers to usage statistics, cost allocation, and pricing. Security management is the process of monitoring access to a network based on defined policies [2].

Jan Zizka et al. (Eds) : DBDM, CICS, CCNET, CSIP - 2015
pp. 29–40, 2015. © CS & IT-CSCP 2015 DOI : 10.5121/csit.2015.50703

We are working in the diagnosis stage and investigating how to automate the process of setting up a fault classification algorithm based on KPI alarms to identify network faults. At the same time, we investigate classification rules and decision tree algorithms between different alarms of KPIs based on causes and symptoms of network malfunctions and demonstrate both system performance and localisation of the KPI alarms. [3].

Within the academic literature, a great deal of time and effort has been devoted to methods of fault management and fault diagnosis within wireless networks. A variety of methods, such as hierarchical support vector machine, the creation of assessment indexes, and neural algorithms based on faculty connectors [3][4][5].

There has been a limited amount of research regarding the degree of fault alarm under data mining techniques. This is problematic given that defect management on telecommunication networks has been studied extensively utilizing the correlation techniques. In this way, the current paper helps to move forward the existing knowledge regarding fault diagnosis.

2. DATA MINING PROCESS

The process of mining data so that it can be gathered and used to make decisions does not involve a single activity. Instead, data mining involves a series of steps that must be undertaken in order to allow large amounts of data to be reduced and put into a format that can be used for the purpose of decision-making [7].

The first step in the data mining process is known as data cleaning. In this step, data are examined to determine which data are relevant for the given task or purpose. Irrelevant data and data that may be considered to be invalid are removed. Next data integration occurs in which data from multiple sources are combined into a single source or dataset. It is often the case that the data that are combined are heterogeneous [8].

The next step in the data mining process is data selection in which further examination of the data occurs in order to determine which data should be taken from the dataset to be used for the task at hand. Once the data selection has been completed, then data transformation can occur. This stage in the data mining process involves transforming the selected data into appropriate formats for the mining procedure. [8].

The next step in the process is the actual data mining. In this step, various techniques for analysing data are used. Then, pattern evaluation is undertaken in which interesting or unique patterns in the data are identified. The identification of patterns occurs in relation to the larger issue or problem that is attempting to be addressed [8].

With the use of Weka, the final step in the data mining process is knowledge representation. This process of incorporating Weka tools into the data process requires analysing and pre-processing the data, as well as examining the validity of the data. Then, class attributes must be defined followed by the actual extraction of the features that will be used for classification. Next, a subset of features is selected to be used as part of the knowledge construction [8].

The next part of the knowledge representation using Weka is the investigation of any imbalances in the selected data and a determination of how those imbalances may be counteracted. Then, a subset of instances is chosen. The subset of instances might be the records upon which the learning will be based. Following the selection of the subset of instances, the classifier algorithm is applied for the learning process. Finally, a testing method is chosen in order to estimate the performance of the algorithm [8].

3. DATA MINING TECHNIQUES

There are a large number of tools that can be used in data mining to handle a variety of tasks, such as data pre-processing, regression, clustering, feature selection, and even visualization. However, it is important to understand that the various tasks that can be performed with data mining tools require different algorithms. At the same time, a programming environment is needed to run the algorithms [9]. In this study, the Waikato Environment for Knowledge analysis (Weka) has been chosen because of its ability to run a variety of different algorithms and its general robustness [10].

An important part of this study is understanding the classifiers that are used for the network faults detection. The classifiers that are of importance to this study are briefly explained.

3.1. J48 tree

The J48 tree functions through the creation of decision trees in which data are examined with the goal of reducing the data into small subsets under the smallest subset can be found at which point an optimised outcome has been determined. A leaf node is created within the decision tree to indicate that a particular class should be chosen [9][11].

3.2. LAD Tree

The Logical Analysis of Data (LAD) tree works by creating a classifier for a binary target. The classifier is created through a process in which a logical expression is learned that has the ability to determine positive and negative samples in a dataset. The underlying theory of the LAD tree is that any binary point that contain some positive patterns but not negative patterns could be classified as positive. At the same time, any binary point that has negative patterns but no positive patterns can be classified as negative. In a similar manner to the J48 tree, a large set of patterns occurs and then subsets are determined until a subset is determined that satisfies the patterns in the data [11].

3.3. JRip

The JRip classifier, which was introduced in 1995 by W. W. Cohen, is an optimised version of an older classifier that Cohen created [12]. JRip was implemented with the prepositional rule learner known as Repeated Incremental Pruning to Produce Error Reduction (RIPPER), meaning that it has the ability to replace or revise rules. With JRip, it is possible to isolate some of the data being examined for training. At the same time, a decision can be made regarding the set of rules generated for selected attributes.

3.4. PART

The PART algorithm [24] is an algorithm that does not function to generate accurate rules through global optimisation. Instead, PART functions to build a rule and then remove the instances that it covers in order to create a recursive rule until there are no instances that remain. Another way of thinking about the PART algorithm is that creates decision lists. With the decision lists, new data are compared to the rules in the decision list. Upon the first match of a rule, data are assigned to that rule. In this way, PART makes the best leaf into a rule rather than the optimised leaf [9].

3.5. Naïve Bayes

The Naïve Bayes classifier [17] operates in accordance with the Bayes rule of conditional probability. The classifier uses all of the attributes of a dataset by analysing them individually. The idea is that all of the attributes are equally important. No single attribute is more important than another attribute. At the same time, all of the attributes of the data are considered to be independent of each other [8].

3.6. Bayes Networks

Bayesian networks, or Bayesnet, are also known as belief probabilistic networks. It is the structure of a Bayesian network that determines the relationships and dependencies between the variables in a dataset. Because of the fact that the structure of Bayesian networks determine dependencies and relationships between variables, the use of such networks have been proposed for use in the diagnosis of faults in some cellular networks [13].

4. DATA PREPARATION

The work of data preparation begins with analysing and pre-processing of database features, as well as an assessment of the correctness of the data that are to be analysed. The work of data preparation involves determining the specific data that can be used to identify network malfunctions. The selection of the data to be analysed involves several considerations, such as which data are representative of the KPIs (symptoms), data availability (inputs), the attributes of the data, and establishing limits of the network faults [9][12]. The next important issue is the generation of the alarms, as well as the relationship between the KPIs and alarms. Before discussing alarms, it is appropriate to briefly explain the difference between an alarm and a fault. An alarm is an indication of a problem or issue in a network. An alarm can be thought of as the outward symptom of a problem. In contrast, a fault is the actual cause of a problem in a network.

The fault is the actual issue that needs to be corrected in order to restore a network to optimisation [2]. The alarms are generated when the nodes of the KPIs cross a specific limit or threshold that is pre-determined. It must be recognized that the thresholds are not set randomly. Instead, the thresholds are set based on their ability to provide an indication of a problem. At the same time, the threshold levels must be such that the probability of false alarms is minimised [2].

After each KPI is associated with an alarm each alarm is associated with a specific security level. An example might be a binary alarm that has the possible conditions or states or On and Off. If the KPI should be On, but it is found to be Off, then an alarm occurs. A KPI can have as many states or conditions as necessary. Some KPIs might have three states, while others might have four or even five states [2].

The KPIs are derived with the help of counters using different formulations. A single counter helps to provide a very limited indication of the larger network. However, with the use of several counters, it is possible to gain a much broader view of the network. In this study, the evaluation is presented on the basis of four major KPIs used as an input to create a classification algorithm to perform the data for the pre-processing process [13]. Table 1 shows the four KPIs that are used are Call set up success rate (CSSR), Call Drop Rate (CDR), Handover Failures (HOF) and Call Traffic Rate (TR). In addition, each KPI can take one of three states: Normal, Critical, or Warning.

Table 1. KPI alarm metrics

Symptom	State alarm indicator		
KPI	**Normal (NORM)**	**Critical (CR)**	**Warning (WARN)**
DCR	DCR<2	4> DCR >=2	DCR >=4
CSSR	CSSR >=98%	90%> CSSR =>98%	CSSR <=90%
TR	TR<60%	60%> TR> =70%	TR> =70%
HOF	TR<10%	10> HOF >15	HOF >=15

Table 1 also shows the specific metrics that have been established as thresholds for each alarm.

The KPI metrics of Handover Failure (HOF) is the ratio of unsuccessful intersystem handovers from 3G to 2G over the total number of such attempts. The call dropped rate (CDR) is the ratio of dropped calls over the total successfully established calls [15]. The call set-up success rate (CSSR) is the rate of call attempts until successful assignment. The call drop rate (CDR) is the rate of calls not completed successfully [16]. The traffic rate indicator is the actual rate or percentage of traffic on the network [14].

4.1. Fault classification algorithm

We selected the most popular Key Performance Indicators to create a Fault Classification Algorithm. The algorithm parameters must be set up for the input data. The rules generated for the fault classification algorithm are given bellow:

```
compute #Result=0.

do if DCR>2 AND CSSR>= 90.
-   do if DCR<=4 AND CSSR<98.
-     #KPIClass=CR.
-   else if TR> 60 AND HOF>10.
-      do if TR<70 AND HOF<25
-            #KPIClass=CR.
-      else if DCR>4 OR CSSR<90
-            do if  HOF>=25 OR TR>70.
                #KPIClass=WARN.
-           else if KPIClass=NORM.
-           end if..

FREQUENCIES
  VARIABLES=#KPI
  /FORMAT=DFREQ
  /ORDER=  ANALYSIS .
execute.
```

Figure 1. Fault classification algorithm

The parameters will include equality or inequality for the state alarm indicators and greater than or less than for the KPI alarms metrics. The relational nodes can be placed into decisions trees or into rules. However, plans that use relations are more often placed in rules rather than in decision trees.

4.2. Explanation of Output Variables

The next step in the process is the selection of a subset of instances in Weka for the parameters of KPIs. It is important to note that the selection of the subset of instances was not random. Instead, the selection of subsets of instances for this study was based on previous studies and investigations of fault diagnosis in networks using data mining classifiers [14][16][17].

The instance of BSC represents the base station controllers, while Id represents the cell Ids. The Standalone Dedicated Control Channel Congestion Rate (SDCCHCR) indicates the probability of accessing a stand alone dedicated control channel available during call set up [14] [17]. TCH Seizure Attempts is the number of traffic channels that are allocated for traffic [17]. The (SDCCH) Access Rate shows the percentage of call access success rate received in the base station location [17]. TCH Availability is the rate of Traffic Channels availability in the network, while the TCH Drop is the total drop rate in the network and also the TCHCR is the rate of TCH congestion in a network area [14]. Finally, the TCHTR is the rate of incoming and outgoing traffic of a channel [16].

The Handover Success parameter shows the percentage of success handovers on the network of all handover attempts [15]. The SDCCH Drops Excessive TA presents the number of drops due to excessive timing advance [17]. Based on those relational nodes, the system would identify the optimised status of the TCH (traffic channel), RAN (radio accessibility network performance audit) and show their optimised states [17]. The Traffic Channel Congestion Rate of the SDCCH drops measures the total number of RF losses as a percentage of the total number of call attempts for the SDCCH channels on the network. The SDCCD Availability Rate is the percentage of SCDDH channels that are available on a network at a given time [14].

The SCDDHD Sudd Lost Con is the rate of connections that are suddenly lost on a network [17]. Additionally the TCH Drop Rate is the rate of traffic assignment failures, while TCHCR is the congestion rate related to call setup traffic [17]. The HOFR statistic is intended to give an indication of the rate as a percentage of handover failures in relation to total handovers [14]17]. The TCHSS shows the percentage of TCHs that are successfully seized [17]. The SDCCH Drops is the number of drops due to low signal strength or network congestion [14] [17]. The Handover Success Rate is the percentage of rate of handover successes. Finally, the KPI Alarms are the symptoms that occur due to the establishing limits that are set by the classifiers of the system.

5. IMPLEMENTATION OF THE SYSTEM

The implementation of alarm detection and fault localizations may generate a large number of alarms. The problem with the generation of a large number of alarms is the potential for a high rate of false alarms. The ability for a human to visualize all of the alarms that may be generated is difficult or even impossible. In this way, the alarms generally need to be aggregated so that visualization can occur more easily for the person that is responsible for examining them [2].

Aside from the aggregation of alarms for the purpose of easier visualization, it is also possible to implement an automatic fault identification module that can correlate the observed alarms in order to indicate the root cause of the alarms. With the automatic fault identification module, a human may not even be needed as an automatic recovery system can be used to use the alarms to quickly determine the fault and correct it [2].

For this study, 2,100 instances are inserted that represent the Cell Ids and Base Station locations. In addition, 25 attributes are input that show the data and key performance indicators including

the data base. The dataset representation in ARFF (Attribute-Relation File Format) is shown below:

@Period,@BSC,@Id,@SDCCHCR,@TCHSeizureAttempts,@TCHAccessRate@TCHAvailability,@TCHDrop,@TCHCR,@TCHTR,@HandoverSuccess,@SDCCHDropsExcessiveTA,@RAN,@SDCCHDR,@SDCCHAvailabilityRate,@HF,@SDCCHDropSuddLostCon,@TCHAssignmentFailureRate,@TCHDropRate,@TCHCongestionRate,@HOFR,@TCHSS,@SDCCHDrops,@HOSR@KPIAlarms

Once classification is settled down, representation of data can use data visualization techniques of decision trees and rules classifiers. Figure 2 shows the JRP optimisation rules.

```
JRIP rules:
===========

(TCHDropRate <= 1.96) and (TCHSeizureAttempts <= 59.78) and (SDCCHCR <= 5.57) => KPIAlarms=NORM (319.0/0.0)
(TCHSeizureAttempts <= 12.93) and (TCHDropRate <= 5) => KPIAlarms=NORM (9.0/0.0)
(TCHDropRate <= 4.99) and (TCHTR <= 51.09) and (TCHAvalability <= 99.22) => KPIAlarms=CR (639.0/1.0)
(TCHSeizureAttempts <= 59.14) and (TCHTR >= 49.52) and (TCHSS >= 90.01) => KPIAlarms=CR (126.0/1.0)
(TCHTR <= 15.98) and (HF >= 714) and (SDCCHCR <= 7.97) => KPIAlarms=CR (42.0/0.0)
 => KPIAlarms=WARN (965.0/1.0)

Number of Rules : 6
```

Figure 2. JRP Optimisation rules

The JRP optimisation rules show that six rules were extracted from Weka regarding the JPR classifier. The rules that were generated show the conditions of critical optimization faults and the KPI metrics to show when a cell can be considered to be optimized. Optimization is considered to occur when the Normal (NORM) condition is achieved. The last rule that was generated tells the system that if the cell is not classified as either normal (NORM) or critical (CR), then it needs to be checked for a warning (WARN) status.

```
Classifier output
TCHDropRate <= 1.96 AND
TCHTR <= 38.91 AND
HOFR <= 6.76: NORM (314.0)

TCHDropRate > 9.98: WARN (433.0)

TCHTR > 51.11: WARN (230.0)

SDCCHDR > 193 AND
TCHDropRate > 3.01: WARN (103.0)

TCHCR <= 4.16 AND
TCHAvalability > 90.66 AND
TCHDropRate <= 5 AND
TCHDropRate > 2.35 AND
TCHDrop > 6.2: CR (608.0)

HOFR > 7.99 AND
TCHAvalability > 90.66 AND
TCHSS > 89.98 AND
TCHDropRate > 1.6: CR (179.0/1.0)

TCHDropRate > 5: WARN (190.0)

TCHCR <= 3.46 AND
TCHAvalability <= 98.99: CR (20.0)

TCHTR <= 30.32: NORM (14.0)

: WARN (9.0)
<
```

Figure 3. Part optimisation decision list

Figure 3 shows the optimisation decision list. A total of six rules were extracted from Weka regarding the Part classifier. As before, the rules that were generated illustrate the conditions of critical optimization faults and the KPI metrics to show when a cell can be considered to be optimized. Optimization is considered to occur when the Normal (NORM) condition is achieved. The last rule that was generated tells the system that if the cell is not classified as either normal (NORM) or critical (CR), then it needs to be checked for a warning (WARN) status.

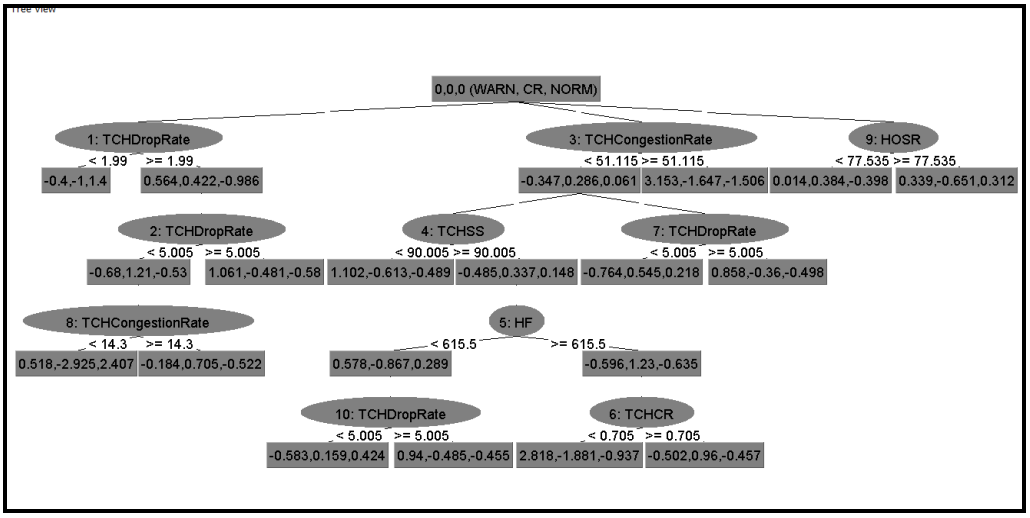


Figure 4. Optimisation decision LADtree

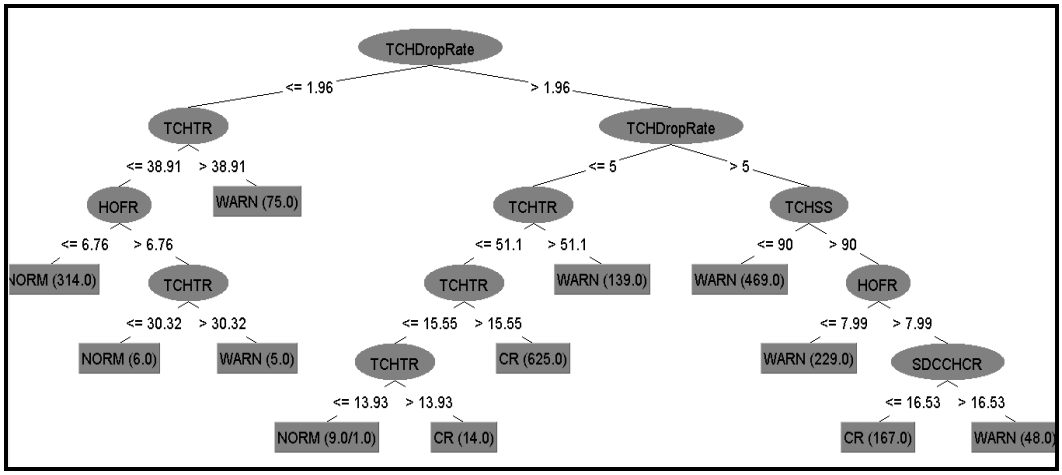


Figure 5. Optimisation decision tree J48

Figures 4 and 5 show the optimization decision trees for J48 and LAD Tree. The number of Leaves is 12 for J48 class and the Size of the tree is 23. The attribute selection from the J48 and LADTree classifiers is based on the variables related to the utilization of the network such as Traffic Channels and Handover success and Number. The LAD Tree provides a Logical Analysis of Data as the classification method. The LAD Tree shows a prediction for faults detection that is based on 21 predictor nodes at the tree side and 14 predictor nodes for the optimisation results at the leaves of LAD Tree [11].

Perhaps just as importantly, the decision trees demonstrate the ability to provide a visual representation of the faults that can be understood by a human operator whose job it might be to analyse the data and determine the cause of the lack of network optimisation. By examining the decision trees, it is possible to quickly determine which cells are indicated as being normal (NORM) and which cells are indicated as being critical (CR). By looking at the specific classifiers, the human operator can truly make a quick decision about the fault in the network to correct that fault and return the network to maximum optimisation.

6. FAULT DIAGNOSIS RESULTS DISCUSSION

An average of 860 cells of the data base were classified in a state alarm of "Warning" , 785 found to be "Critical" and only 320 cells were considered as "Normal". The most popular KPI metrics extracted by the system are based on TCH Availability, Congestion and Success Rate, Handover Success and Failure Rate, SDCCH and TCH Drop Rate and final TCH attempts.

6.1. Comparison using different classifiers

Table II also shows the accuracy alarm prediction by the proposed technique. The table shows that the performance measure results of J48 and PART classifiers have the same performance in terms of accuracy in classification. Both methods achieved an accuracy rate of 99.9%. With the rules and Decision trees classifiers, the accuracy rate is actually more than 99.9%. However, the Bayesian classifiers showed a decline in their accuracy levels as compared to the other methods [12].

This scale by which to measure the performance of algorithms is important in relation to the method that was tested in this study because distributed fault detection and isolation with the use of a Bayes classifier has demonstrated that even if the accuracy is not too high, it is the total output of the use of the classifier that provides important information that can be used by wireless providers to isolate and localize the network malfunctions.

Table 2. Performance measure results

Weka classifiers	Correctly classified Instances	Incorrectly Classified Instances	Kappa Statistics	Mean absolute error	Root mean squared error	Accuracy
Decision Trees						
J48	2099	1	0.99	0.0006	0.01	99.9%
LADTree	2097	3	0.99	0.0243	0.05	99.8%
Rules						
JRIP	2097	3	0.99	0.0019	0.03	99.8%
Part	2099	1	0.99	0.0006	0.01	99.9%
Bayes						
BayesNet	1839	261	0.80	0.08	0.28	87.5%
NaiveBayes	1701	399	0.69	0.13	0.34	81%

Table III shows the evaluation of final statistics. The scale is based on the level of accuracy of the algorithm with regards to correctly identifying faults as compared to false results [12]. The final statistics show the accuracy of the model and visualize precision and recall for ROC curve analysis (true positive rate vs false positive rate). [10]. TP rate also demonstrates the sensitivity of the model such as a scale of actual positive values and the FP rate shows the specificity such as the negative tuples that are incorrectly labelled. [11]. In this regard, the slightly less accurate information of BayesNet and NaiveBayes classifiers across the entire system still results in a high

level of ability to determine the fault or faults that need to be corrected in order to return a network to optimisation [12].

7. CONCLUSIONS

This paper proposes a systematic approach for fault diagnosis that is based on a KPI data analysis model using the Weka environment for those who are responsible for the monitoring and optimisation of GSM networks. The importance of the method of fault diagnosis that has been proposed in this paper is that it is one that has not been widely studied. In this regard, a potentially new method, or at least one that has not been greatly considered, is available for network operators who want an automated means by which to quickly identify faults and return networks to optimisation in order to provide a high quality of service to users.

Table 3. Comparison of final statistics weighted avg

Classifiers	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
J48	1.00	0.00	1.00	1.00	1.00	0.99	1.00	1.00
LADTree	0.99	0.00	0.99	0.99	0.99	0.99	0.99	0.99
JRIP	0.99	0.00	0.99	0.99	0.99	0.99	0.99	0.99
Ridor	1.00	0.00	1.00	1.00	1.00	0.99	1.00	1.00
BayesNet	0.87	0.034	0.913	0.876	0.884	0.833	0.993	0.98
NaïveBayes	0.810	0.109	0.846	0.810	0.809	0.709	0.935	0.90

The state alarm indicator values shown in the method used in this study provide a strong means by which to perform network diagnosis efficiently and quickly. The rules and decision trees set up that are extracted by the data mining process simplify the optimisation process. The decision trees allow network operates to visually examine multiple network parameters, such as traffic channels and handovers statistics. The end result is that the task performance management network optimisation through fault location finding, and root cause analysis is much easier [6].
It is important to iterate the method of fault diagnosis shown in this paper is not only highly accurate, but it is also efficient. A method of fault diagnosis can be high, but a lack of efficiency in using the method can be ineffective for network operators who must return a network to optimum performance levels in a matter of seconds or minutes, rather than hours [18]. Wireless networks have become more complex with increasing numbers of users [19]. This only adds to the burden of finding automatic methods of fault diagnosis that are efficient, both in terms of accuracy of output and actual usage.

The method of fault diagnosis described in this paper is indeed computationally efficient regarding the rules, decision trees and Bayes classifiers. It also represents optimisation solutions defined by rules. Furthermore, this method of fault diagnosis retains the ability to identify individual KPI metrics and alarm factors severely affecting reliability of fault location calculation. In the end, the results presented in this paper allow for the conclusion that the proposed method has produced very promising results in the classification of multi-class optimisation faults.

8. FUTURE WORK

The results of this study are not only important in terms of what can be concluded from them about the method of network fault diagnosis that was presented, but also the broader context of network fault management. The analysis of network fault diagnosis in this study, and most other studies, requires some knowledge of statistical analysis. The question that arises, however, is whether the people who are responsible for fault diagnosis in practice have the statistical analysis that is needed to efficiently perform the same procedures and analysis conducted in this study and others. While the focus of so much research in this area has been on methods of fault detection, it might be time to study those who have to implement these methods and use them to make real-world decisions.

Another area for future research might be to examine the business decisions that underlie network optimisation efforts on the part of network providers. While the focus of so much of this area of research is on the techniques and methods, the reality is that there are important business decisions that dictate which techniques and methods are used, as well as how they are used [20]. It would be useful to understand the decision-making process and the business realities that determine the specific methods that companies that operate wireless networks use to perform fault diagnosis.

REFERENCES

[1] D.Adrada, E. Salazar, J. Rojas, and J. C. Corrales, (2014) "Automatic code instrumentation for converged service monitoring and fault detection", 28th International Conference on Advanced Information Networking and Applications Workshops, pp708-713.

[2] L. Monacelli and R. Francescangeli, (2011) "Fault management for volp applications over wireless and wired ngn networks: an operational prospective", IEEE 36th Conference Local Computer Networks, pp711-718.

[3] A. Rosich, H. Voos, and L. Pan, (2014) "Network design for distributed model-based fault detection and isolation", IEEE Conference on Control Applications, pp1226-1231.

[4] L. Zhang, X. Meng and H. Zhou, (2009) "Network fault diagnosis using hierarchical svms based on kernel method", 2nd International Workshop on Knowledge Discovery and Data Mining, pp753-756.

[5] R. Ji, J. Gao, G. Xie, G.T. Flowers, and C. Chen, (2014) "A fault diagnosis method of communication connectors in wireless receiver front-end circuits", 60th Holm Conference on Electrical Contacts, pp1-6.

[6] Y-Y. Wang, K-W. Wu, C-M. Huang, and C-C. Chan, (2014) "Quality management and network fault diagnosis for iptv service", 16th Asia-Pacific Network Operations and Management Symposium, pp1-4.

[7] B. M. Patil, D. Toshniwal, and R. C. Joshi, (2009) "Predicting burn patient survivability using decision tree in weka environment", IEEE International Advance Computing Conference, pp1353-1356.

[8] P. K. Srimani, and K. Balaji, (2014) "A comparative study of different classifiers on search engine based educational data", International Journal of Conceptions on computing and Information Technology, Vol. 2, pp 6-11.

[9] W. Nor Haizan, W. Mohamed, M. N. M. Salleh, and A. H. Omar, (2012) "A comparative study of reduced error pruning method in decision tree algorithms", IEEE International Conference on Control System, Computing and Engineering, pp392-397.

[10] I. A. Witten and E. Frank, (2005) Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations, San Francisco, CA: Morgan-Kaufmann.

[11] T. C. Sharma and M. Jain, (2013) "WEKA approach for comparative study of classification algorithm", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 2, pp1925-1931.

[12] W. Shahzad, S. Asad, and M. A. Khan, (2013) "Feature subset selection using association rule mining and JRip classifier", International Journal of Physical Sciences, Vol. 18, pp885-896.

[13] R. Barco, V. Wille, L. Diez, and P. Lazaro, (2006) "Comparison of probabilistic models used for diagnosis in cellular networks", IEEE 63rd Vehicular Technology Conference, Vol. 2, pp981-985.

[14] M. Panda and S. P. Padhy, (2009) "Traffic analysis and optimization of gsm network", IJSCI International Journal of Computer Science Issues, Vol. 1, pp28-31.

[15] C. Skianis, (2013) "Introducing automated procedures in 3g network planning and optimization", The Journal of Systems and Software, Vol. 86, pp1596-1602.

[16] B. Haider, M. Zafrullah, and M. K. Islam, (2009) "Radio frequency optimization & QoS evaluation in operational gsm network", Proceedings of the World Congress on Engineering and Computer Science, Vol. 1, pp1-6.

[17] M Ali, A. Shehzad, and M. A. Akram, (2010) "Radio access network audit & optimization in gsm (radio access network quality improvement techniques)," International Journal of Engineering & Technology, Vol. 10, pp55-58.

[18] Y. Temprado, C. Garcia, F. J. Molinero, and J. Gomez, (2008) "Knowledge discovery from trouble ticketing reports in a large telecommunications company," International Conference on Computational Intelligence for Modelling Control & Automation, pp37-42.

[19] R. Barco, L. Diez, V. Wille, and P. Lazaro, (2009) "Automatic diagnosis of mobile communications networks under imprecise parameters", Expert Systems with Applications, Vol, pp489-500.

[20] V. P. Bresfelean, (2007) "Analysis and predictions on students' behavior using decision trees in weka environment", Proceedings of the ITI 2007 29th International Conference on Information Technology Interfaces, pp51-56.

AUTHOR

Eleni Rozaki (S-22, M-13) received the M.Sc. degree in Quantitative methods and Informatics from University of Bari Italy in 2008. Currently, she is working toward the PhD degree, in the area of network optimisation at Cardiff University, United Kingdom. Her current research interests include network management, decision support systems and data mining techniques.

TORQUE RIPPLES MINIMIZATION ON DTC CONTROLLED INDUCTION MOTOR WITH ADAPTIVE BANDWIDTH APPROACH

Fatih Korkmaz[1],Yılmaz Korkmaz[2],İsmail Topaloğlu[1] and Hayati Mamur[1]

[1]Department of Electric and Electronic Engineering,
Çankırı Karatekin University, Çankırı, Turkey
fkorkmaz@karatekin.edu.tr
[2]Faculty of Technology,
Department of Electric and Electronic Engineering,
Gazi University, Ankara, Turkey

ABSTRACT

Field oriented control (FOC) and direct torque control (DTC), also called vector control, are most famous control methods in high-performance motor applications. If we want to specify the basic handicaps of both methods: the FOC has parameter dependence while the DTC has high torque ripples. This paper proposes a new adaptive bandwidth approach to reduce torque ripples in DTC controlled induction motor drives. With the proposed method, instead of fixed bandwidth, adaptive bandwidth approach is investigated in hysteresis controllers on the DTC method. Both the conventional DTC(C-DTC) method and adaptive bandwidth DTC (AB-DTC) for induction motor are simulated in MATLAB/SIMULINK and the results are presented and discussed to verify the proposed control. The comparisons have shown that, torque ripples have been reduced remarkably with the proposed AB-DTC method.

KEYWORDS

Direct torque control, Adaptive hystresis controller, Induction motor control, Vector control

1. INTRODUCTION

Three-phase induction motors (IMs) are the most common motors used in industrial control systems and commercial applications. Simple and rugged design, low-cost, low maintenance and direct connection to an AC power source are the main advantages of IMs[1].

In the past, IMs were preferred only constant speed applications because of speed adjustment on IMs were not only hard to realize but also need high costs. So, DC motors were the optimum option for the variable speed applications. But over the years, this situation has changed depending on developments in power electronics and semiconductor technology. As a result, many kinds of IM variable speed driver have been produced and now the IMs are very good alternative for variable speed applications.

Today, vector controlled drivers are the most popular variable speed driver for the IMs and vector control methods can basically be grouped under two headings: FOC and DTC. The FOC was first introduced by Blaschke [2] in the 1970's. It was unrivalled in industrial induction motor drivers until DTC was introduced by Takahashi [3] in the middle of the 1980's. It was a good alternative

Jan Zizka et al. (Eds) : DBDM, CICS, CCNET, CSIP - 2015
pp. 41–48, 2015. © CS & IT-CSCP 2015 DOI : 10.5121/csit.2015.50704

to FOC due to some well-known advantages, such as simple control structure, no need much motor parameters so independency of parameter changes, fast dynamic response. Besides these advantages, DTC scheme still had some disadvantages like high torque and current ripples, variable switching frequency behaviour and implementation difficulties owing to necessity of low sampling time [4].

Despite, the DTC was originally developed for induction motor drives, it has also been applied other motor types like PMSM and BLDC [5-6]. When we look at the recent studies about the DTC, we can see that many kinds of approaches have been investigated to overcome the high torque and current ripple problems on DTC method.

In these studies, researchers proposed different ways. Some studies suggest using different switching techniques and inverter topologies [7-8], in another group of researchers, different observer models have been suggested [9-10]. On the other hand, intelligent control methods like fuzzy logic have been explored by several researchers for its potential to improve the speed regulation of the drive system. [11-12]

In this paper, a new hysteresis controller algorithm is presented to improve the dynamic torque performance of the DTC controlled IM. To illustrate the effect of the proposed system, conventional and proposed systems are simulated in Matlab/Simulink environment and results have been analyzed. Simulation studies have proved that this method reduces the torque ripple of the DTC method.

2. BASICS OF DTC

Direct torque control (DTC) directly controls the flux linkage and electromagnetic torque, considering the motor, voltage source inverter, and the control strategy at the system level. A relationship is established between the torque, the flux and the optimal inverter switching so as to achieve a fast torque response. It exhibits better dynamic performance than conventional control methods, such as vector control, is less sensitive to parameter variations, and is simpler to implement [13].

The DTC bases on the selection of the optimum voltage vector which makes the flux vector rotate and produce the demanded torque. In this rotation, the amplitude of the flux and the torque errors are kept within acceptable limits by hysteresis controllers [14]. The rotation of the stator flux vector and an example of the effects of the applied inverter switching vectors are given in the Fig.

1. The DTC allows for very fast torque responses, and flexible control of the induction motor. The flux and torque errors are kept within acceptable limits by hysteresis controllers [15].

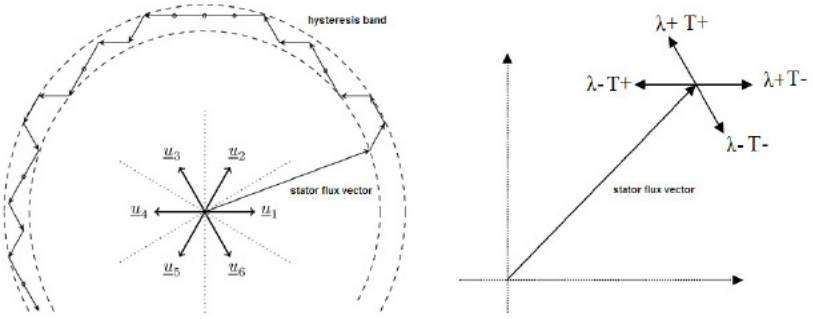


Figure 1. The rotation of the stator flux vector and an example for the effects of the applied inverter switching vectors [16].

The block diagram of the conventional DTC controlled motor is given in Fig. 2.

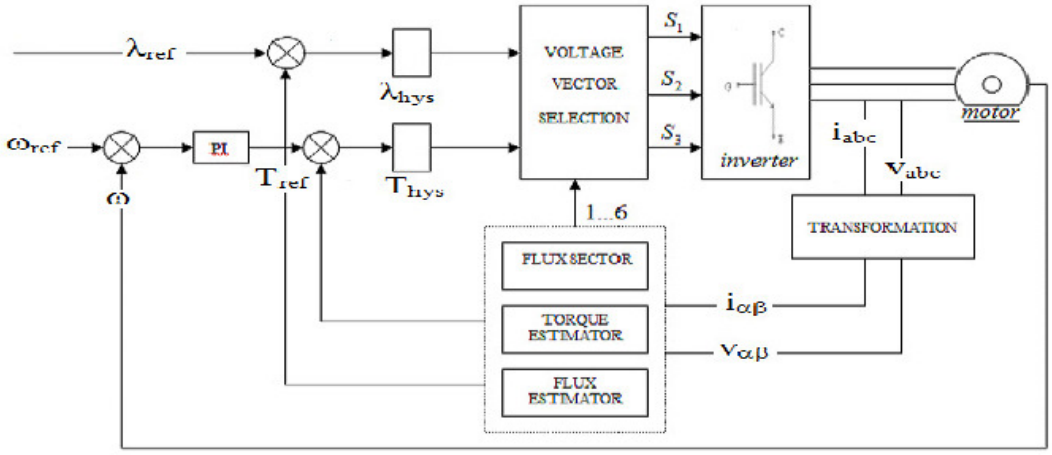


Figure 2. The block diagram of the conventional DTC

The DTC algorithm controls the stator flux and the torque by using measured currents and voltages.

The instantaneous values of the flux and torque can be obtained by using the transformation of the measured currents and the voltages of the motor. The stator flux is calculated as given in Eq.1-3 in a stationary reference frame.

$$\lambda_\alpha = \int (V_\alpha - R_s i_\alpha) dt \tag{1}$$

$$\lambda_\beta = \int (V_\beta - R_s i_\beta) dt \tag{2}$$

$$\lambda = \sqrt{\lambda_\alpha^2 + \lambda_\beta^2} \tag{3}$$

Where, λ_α-λ_β are stator fluxes, i_α-i_β are stator currents, V_α-V_β are stator voltages, $\alpha-\beta$ components and R_s is the stator resistance. Motor torque can be calculated as given in Eq.4.

$$T_e = \frac{3}{2} p(\lambda_\alpha i_\beta - \lambda_\beta i_\alpha) \tag{4}$$

Where, p is the motor pole pairs. The stator flux vector region is an important parameter for the DTC, and it can be calculated as given in Eq.5:

$$\theta_\lambda = \tan^{-1}\left(\frac{\lambda_\beta}{\lambda_\alpha}\right) \tag{5}$$

The torque and flux errors, which are obtained by comparing the reference and observed values, are converted to control signals by hysteresis comparators. The switching table is used to determine the optimum switching inverter states, and it determines the states by using the hysteresis comparator outputs and the flux region data [17].

3. ADAPTIVE BANDWIDTH APPROACH

Hysteresis control is one of PWM methods used for generating pulses to order the power switches of the inverter. Among the various current control techniques, it is widely used due to the fast response, simple implementation, negligible tracking error, inherent robustness to load parameter variations and proper stability [18].

In DTC method, two different hysteresis controllers are used to determine the changes in stator flux and electromagnetic torque. In constant bandwidth approach, small bandwidth values results in a higher switching frequency. So, it results in low harmonic copper losses in the motor while switching losses in the inverter are high. Conversely, in a large bandwidth values case, switching losses decrease in the inverter while the harmonic copper losses increase in the motor [19]. So selection of optimum amplitude of flux and torque hysteresis band is important for the drive but there are no certainties to determine optimum amplitude of hysteresis bandwidth.

The main idea of adaptive hysteresis bandwidth approach as given in Fig. 3.

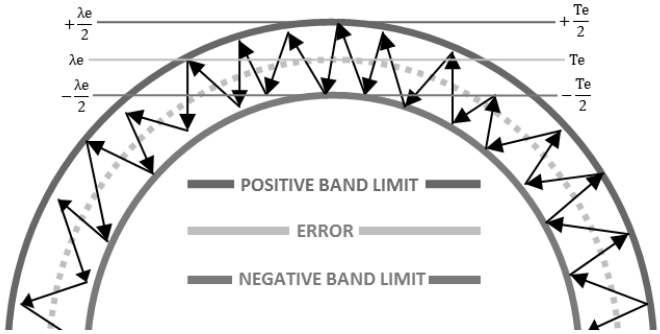


Figure 3. Adaptive hysteresis bandwidth

In this approach hysteresis bandwidth is determined by error values in the previous step for stator flux and electromagnetic torque. So, for the next step of the control process, hysteresis bandwidth is adapted with change in error. It means, if error on flux / torque is high, hysteresis bandwidth will be extended, on the contrary, if error on flux / torque is low, hysteresis bandwidth will be reduced. In this way, hysteresis bandwidth is designed to be flexible and the control algorithm tries to minimize previous step errors at every step. In other words, flux and torque errors are associated with each other with this approach.

4. SIMULATIONS AND RESULTS

The C-DTC and the proposed AB-DTC drive systems have been developed using Matlab/Simulink and the AB-DTC system with the Simulink blocks is shown in Fig. 4.

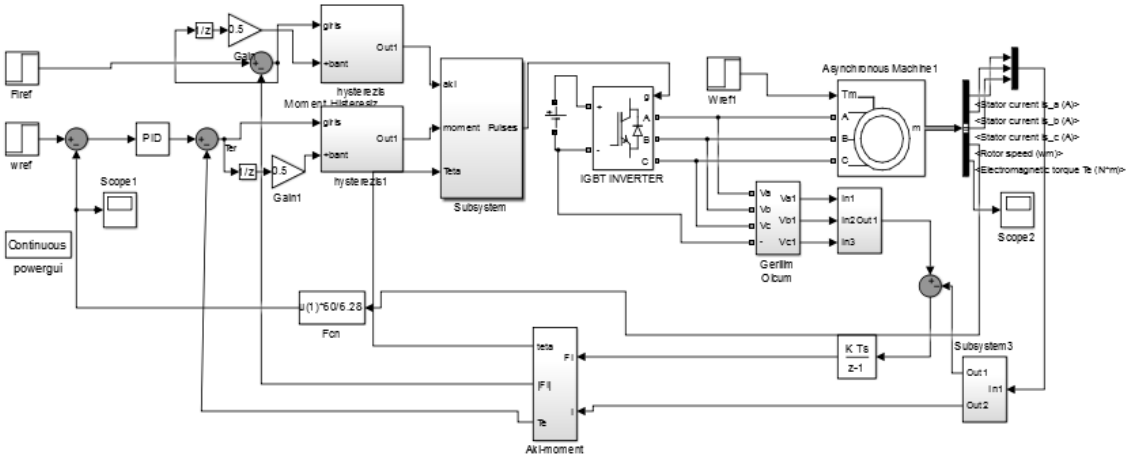


Figure 4. The proposed AB-DTC system with the Simulink blocks

The both control systems have been simulated in order to demonstrate the validity of the proposed AB-DTC system. The sampling interval is 50 µs. The magnitudes of the torque and flux hysteresis bands in the C-DTC are 0.1 Nm and 0.02 Wb, respectively. The nominal motor parameters are mentioned in Table.

Table. Specifications and parameters of the IM

Symbol	Quantity	Value
P	Power (kW)	1.1
V_n	Line to Line voltage (V)	230
p	Number of poles	2
f	Frequency(Hz)	50
R_s	Stator resistance(Ω)	8
Lm	Mutual inductance (H)	0.55
L	Inductance (H)	0.018

To investigate and compare the performances of the C-DTC and the AB-DTC algorithms, unloaded and rated loaded (3 Nm) working conditions are employed. The speed reference is set at 2800 rpm for both conditions. The electromagnetic torque responses of motor for unloaded and loaded conditions are shown in Fig. 5 and Fig. 6., respectively.

As it can be seen if Fig. 5, for unloaded working conditions, the C-DTC torque ripple changes in ±1.5 Nm band while the AB-DTC changes in ±1 Nm. It means the motor torque ripple has been reduced about 65% with the proposed DTC method for unloaded working conditions.

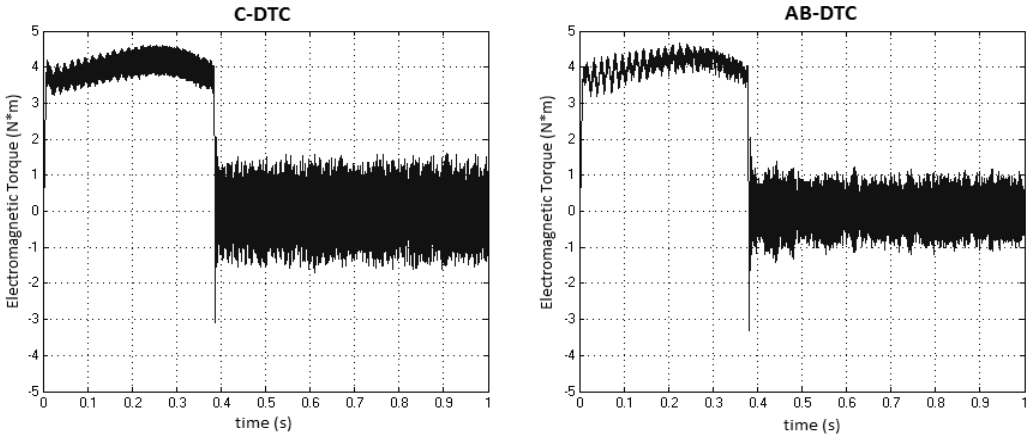


Figure 5. Electromagnetic torque responses at unloaded working conditions

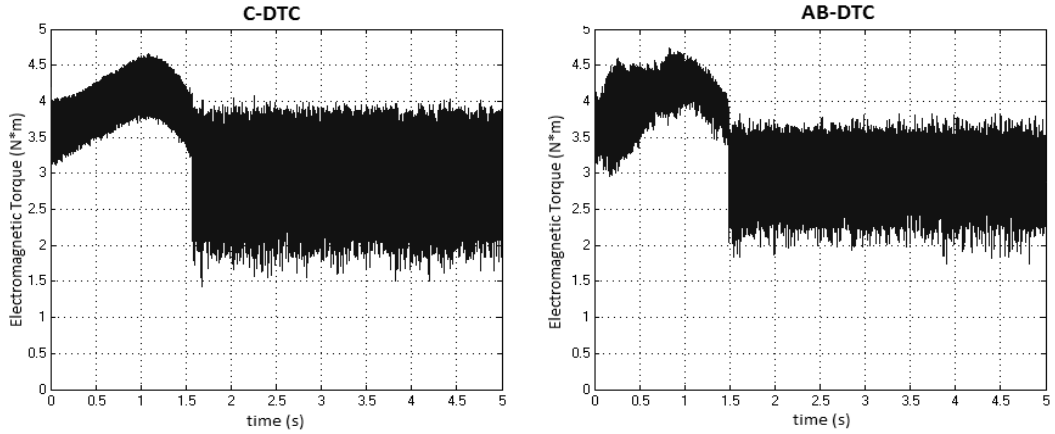


Figure 6. Electromagnetic torque responses at loaded working conditions

If we investigate loaded working conditions, as it can be seen in Fig. 6, the C-DTC torque ripple band changes between about 4 Nm /1.75 Nm (bandwidth 2.25 Nm) and the AB-DTC torque ripple band changes between about 3.6 Nm/2.25 Nm (bandwidth 1.35 Nm). It means the motor torque ripple has been reduced about 60% with the proposed DTC method for loaded working conditions. So, in general, it must be pointed out that the torque ripple of the motor has been reduced about 60% with the proposed DTC approach.

5. CONCLUSIONS

The conventional DTC offers several inherent advantages such as faster dynamic performance and robust controller structure compared to the Field Oriented Control for IMs. However, most faced problem in the C-DTC is high torque ripple. This paper presents an adaptive control algorithm on hysteresis controllers which are directly effect on system performance. In this approach hysteresis controller band limits are not constant and band limits are determined by error values in the previous step for stator flux and electromagnetic torque. The proposed approach is verified by the Matlab simulation and the obtained results shows that the proposed DTC approach can significantly reduce the torque ripple and improve the drive performance.

REFERENCES

[1] Rakesh Parekh, (2003) "AC Induction Motor Fundamentals", Microchip Technology Inc, AN887, DS00887A, pp 1-24.

[2] Blaschke, F.,(1972) "The Principle of Field Orientation Applied to The New Transvector Closed-Loop Control System for Rotating Field Machines", Siemens-Rev., Vol. 39, 217–220.

[3] Takahashi, I. & Noguchi. T. (1986) "A new quick-response and high efficiency control strategy of an induction motor," IEEE Transactions on Industrial Applications, vol.I A-22 ,No.5. , pp. 820–827.

[4] Fatih Korkmaz. & M. Faruk Çakır. & Yılmaz Korkmaz. & Ismail Topaloglu, (2012) "Fuzzy Based Stator Flux Optimizer Design For Direct Torque Control" International Journal of Instrumentation and Control Systems (IJICS) Vol.2, No.4, pp 41-49.

[5] Fatih Korkmaz. & M.Faruk Cakır. & İsmail Topaloğlu. & Rıza Gurbuz, (2013) "Artificial Neural Network Based DTC Driver for PMSM", International Journal of Instrumentation and Control Systems (IJICS) Vol.3, No.1, pp 1-7.

[6] Masmoudi, M. & El Badsi, B. & Masmoudi, A.,(2014) "Direct Torque Control of Brushless DC Motor Drives With Improved Reliability," Industry Applications, IEEE Transactions on , vol.50, no.6, pp.3744,3753. doi: 10.1109/TIA.2014.2313700

[7] D. Casadei. & G. Serra. & A. Tani,(2001) "The use of matrix converters in direct torque control of induction machines", IEEE Trans. on Industrial Electronics, vol.48 , no.6 , pp. 1057–1064.

[8] D. Casadei. & G. Serra. & A. Tani,(2000) "Implentation of a direct torque control algorithm for induction motors based on discrete space vector modulation", IEEE Trans. on Power Electronics, vol.15 , no. 4 , pp. 769–777.

[9] Z. Tan. & Y. Li . & Y. Zeng,(2002) "A three-level speed sensorless DTC drive of induction motor based on a full-order flux observer", Power System Technology, Proceedings. PowerCon International Conference, vol. 2, pp. 1054- 1058.

[10] G. Yav. & L. Weiguo, (2011) "A new method research of fuzzy DTC based on full-order state observer forstator flux linkage", Computer Science and Automation Engineering (CSAE), 2011 IEEE International Conference, vol. 2 , pp.104-108.

[11] S. Benaicha. & F. Zidani. & R.-N. Said. & M.-S.-N. Said,(2009) "Direct torque with fuzzy logic torque ripple reduction based stator flux vector control", Computer and Electrical Engineering, (ICCEE '09), vol.2 , pp. 128–133.

[12] N. Sadati. & S. Kaboli. & H. Adeli. & E. Hajipour . & M. Ferdowsi,(2009) "Online optimal neuro-fuzzy flux controller for dtc based induction motor drives", Applied Power Electronics Conference and Exposition (APEC 2009), pp.210–215.

[13] Liu, Y. & Zhu, Z.Q. & Howe, D. (2005) "Direct torque control of brushles DC drives with reduced torque ripple", IEEE Transactions on Industry Applications, 41(2). pp 599-608.

[14] P. Vas,(2003) "Sensorless vector and direct torque control", Oxford University Press.

[15] Kaboli, S. & Zolghadri, M.R. & Haghbin, S. & Emadi, A.,(2003) "Torque ripple minimization in DTC of induction motor based on optimized flux value determination", IEEE Ind. Electron. Conf., pp. 431–435.

[16] Fatih Korkmaz. & Ismail Topaloglu. & Hayati Mamur (2013) “Fuzzy Logic Based Direct Torque Control Of Induction Motor with Space Vector Modulation”, International Journal on Soft Computing, Artificial Intelligence and Applications (IJSCAI), Vol.2, No. 5/6, pp 31-40.

[17] Yılmaz Korkmaz. & Fatih Korkmaz. & Ismail Topaloglu . & Hayati Mamur, (2014) “Comparing of Switching Frequency on Vector Controlled Asynchronous Motor”, International Journal on Soft Computing, Artificial Intelligence and Applications (IJSCAI), Vol.3, No. 3/4, pp 19-27.

[18] Seyed Mehdi Abedi. & Hani Vahedi, (2013) “Simplified Calculation of Adaptive Hysteresis Current Control to be Used in Active Power Filter”, Trends in Applied Sciences Research, 8(1), pp 46-54.

[19] H. Ibrahim OKUMUS. &Mustafa AKTAS,(2010) “Adaptive hysteresis band control for constant switching frequency in DTC induction machine drives”, Turk J Elec Eng & Comp Sci, Vol.18, No.1, pp 59-69.

QOS-BASED PERFORMANCE EVALUATION OF CHANNEL-AWARE/QOS-AWARE SCHEDULING ALGORITHMS FOR VIDEO-APPLICATIONS OVER LTE/LTE-A

Najem N. Sirhan, Gregory L. Heileman, Christopher C. Lamb, Ricardo Piro-Rael

Electrical and Computer Engineering Department,
University of New Mexico, Albuquerque, New Mexico, USA
najem83@unm.edu heileman@ece.unm.edu cclamb@unm.edu
fdisk122r@gmail.com

ABSTRACT

Long Term Evolution (LTE) is defined by the Third Generation Partnership Project (3GPP) standards as Release 8/9. The LTE supports at max 20 MHz channel bandwidth for a carrier. The number of LTE users and their applications are increasing, which increases the demand on the system BW. A new feature of the LTE-Advanced (LTE-A) which is defined in the 3GPP standards as Release 10/11 is called Carrier Aggregation (CA), this feature allows the network to aggregate more carriers in-order to provide a higher bandwidth. Carrier Aggregation has three main cases: Intra-band contiguous, Intra-band non-contiguous, Inter-band contiguous. The main contribution of this paper was in implementing the Intra-band contiguous case by modifying the LTE-Sim-5, then evaluating the Quality of Service (QoS) performance of the Modified Largest Weighted Delay First (MLWDF), the Exponential Rule (Exp-Rule), and the Logarithmic Rule (Log-Rule) scheduling algorithms over LTE/LTE-A in the Down-Link direction. The QoS performance evaluation is based on the system's average throughput, Packet Loss Rate (PLR), average packet delay, and fairness among users. Simulation results show that the use of CA improved the system's average throughput, and almost doubled the system's maximum throughput. It reduced the PLR values almost by a half. It also reduced the average packet delay by 20-40\% that varied according to the video bit-rate and the number of users. The fairness indicator was improved with the use of CA by a factor of 10-20%.

KEYWORDS

Long Term Evolution (LTE), LTE-Advanced (LTE-A), Carrier Aggregation (CA), Quality of Service (QoS), downlink scheduling.

1. INTRODUCTION

The LTE was introduced as an evolution to the Universal Mobile Telecommunication Systems (UMTS) to provide cellular network users with high data rates in both the uplink and downlink direction, decreased latency, and good spectrum utilization [1]. The spectrum utilization could be achieved by the use of the right scheduling algorithm that meets with the environment's conditions and the users' requirements demands. There are many scheduling algorithms that exist

Jan Zizka et al. (Eds) : DBDM, CICS, CCNET, CSIP - 2015
pp. 49–65, 2015. © CS & IT-CSCP 2015 DOI : 10.5121/csit.2015.50705

in the literature that are used in the LTE scheduling process. These algorithms can be classified in five main groups: channel-unaware, channel-aware/QoS-unaware, channel-aware/QoS, semi-persistent for VoIP support, and energy-aware [2]. When the number of users and their applications increases, such as video-streaming and video-conferencing, this requires higher data rates and decreased latency, which declines the service that the LTE provides to its users. This challenge of providing a reliable service up to the users' requirements demands can not be solved entirely by choosing the right scheduling algorithm, because the performance of these scheduling algorithms is bounded by the existing LTE capabilities, such as the system's bandwidth. The LTE supports at max 20 MHz channel bandwidth. However, the LTE-A can support more channel bandwidth according to the release as specified in the 3GPP's technical specifications. In Release 10 (R10), the maximum aggregated bandwidth is 40MHz. And it is also 40MHz in Release 11 (R11), but with much more CA configurations [3]. This was the motivation to evaluate the QoS performance of three Channel-aware/QoS-aware scheduling algorithms for video-applications over the LTE Release 8/9 and LTE-A Release 10/11. The LTE-Sim-5 can only simulate the LTE network without the use of CA, so the performance of these algorithms over LTE-A could not be evaluated without making these modifications to the LTE-Sim-5, which motivated us to modify it in order to make these evaluations.

The structure of this paper is as follows: in section 2, we explained the LTE network architecture. In section 3, we explained the LTE radio spectrum. In section 4, we explained the carrier aggregation. In section 5, we explained the general model of LTE scheduling. In section 6, we explained the LTE scheduling algorithms which we evaluated in this paper. In section 7, we explained the simulation environment and listed its parameters. In section 8, we used the simulation results to measure the QoS parameters which we displayed in line charts and then analysed. In the last section, we provided a concluding remarks.

2. LTE NETWORK ARCHITECTURE

The LTE network architecture can be divided into two main parts: the Radio Access Network (RAN), and the Evolved Packet Core (EPC) as in Figure 1. The RAN consists of an Evolved NodeB (eNodeB) and User Equipment (UE). The eNodeB is the connection point for the UE with the core network. It hosts the PHYsical (PHY), Medium Access Control (MAC), Radio Link Control (RLC), and Packet Data Control Protocol (PDCP) layers that include the functionality of user-plane header-compression and encryption. It also offers Radio Resource Control (RRC) functionality that corresponds to the control plane. Scheduling, admission control, and radio resource management are also performed in the eNodeB. The EPC part consists of five main components: the Policy Control and Charging Rules Function (PCRF), the Home Subscriber Server (HSS), the PDN-Gateway (P-GW), the Serving Gateway (S-GW), and the Mobility Management Entity (MME). The PCRF is a logical node responsible for policy control decision-making, and controlling the flow-based charging functionalities in the Policy Control Enforcement Function (PCEF) that is being hosted at the P-GW. It also decides how a certain data flow will be treated in the PCEF by providing the QoS authorization, QoS class identification, and determine the bit rates in accordance with the user's subscription profile. The HSS is the database of the LTE network, it contains all the users' subscription QoS profile, information about the Packet Data Networks (PDNs) the user can connect to, dynamic information that relates the identity of the MME to which the user is currently attached or registered to, and it may also integrate the Authentication Center (AuC) that generates the vectors for authentication and security keys. The P-GW is the gateway which is responsible for QoS enforcement for Guaranteed Bit Rate (GBR) bearers, flow-based charging according to rules from the PCRF, and the allocation of IP addresses to users. In addition it filters user's IP packets into different QoS-based bearers based on Traffic Flow Templates (TFTs). It also serves as the mobility anchor for inter-working with non-3GPP networks such as WiMAX and WiFi. The S-GW is the gateway

that serves as the local mobility anchor for the data bearers while users are moving between eNodeBs, in which all their IP packets are transferred through it. It temporarily buffers user's downlink data when it is in the idle state, while the MME initiates paging of the UE to re-establish the bearers. It performs administrative functions in the visited network such as collecting information for charging and legal interception. It also serves as the mobility anchor for inter-working with 3GPP networks such as General Packet Radio Service (GPRS) and Universal Mobile Telecommunication Systems (UMTS). The MME is the main node in the EPC, it manages the Authentication and Security, and the subscription profile and service connectivity of users. It is responsible for all the mobility management tasks such as inter eNodeBs handovers, inter MMEs handovers, and keeping track of the location of all users [4].

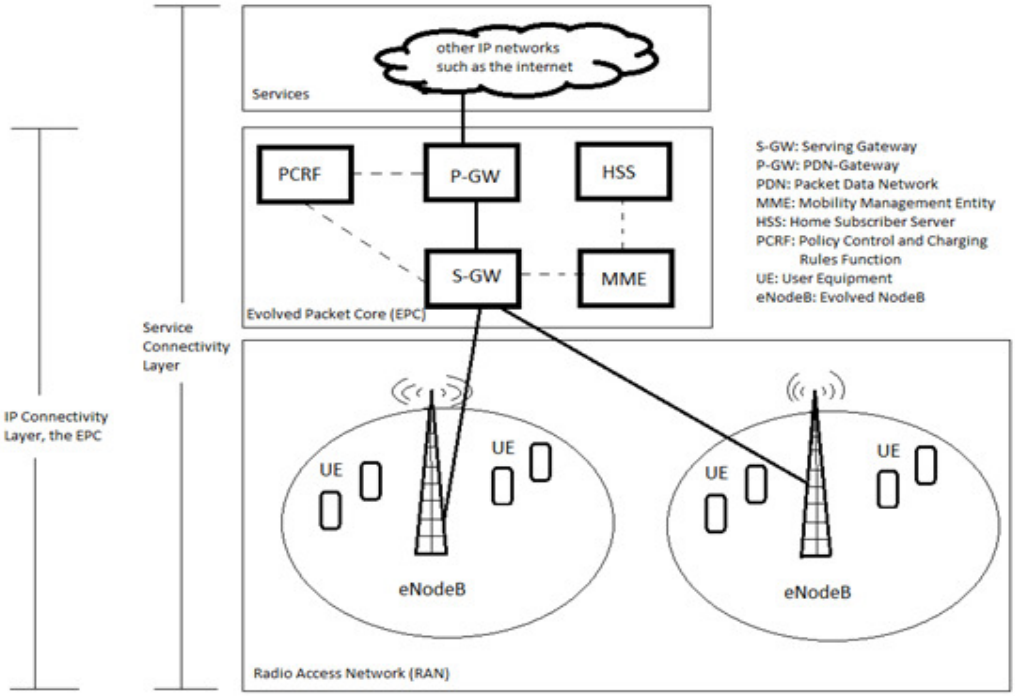


Figure 1. LTE network architecture

3. LTE RADIO SPECTRUM

The LTE radio spectrum can be represented in a two dimensional array: one dimension exists in the time domain, the other one in the frequency domain as shown in Figure 2. In the time domain, LTE transmissions are organized into radio frames of length 10 ms, each frame is equally divided into 10 sub-frames “1 ms each”'. This sub-frame is the minimum scheduling unit in LTE, each sub-frame consists of two equal time slots “0.5 ms each”. There are two types of time slots, one is the normal cyclic prefix which is combined of 7-OFDM symbols, the other is the extended cyclic prefix which is combined of 6-OFDM symbols. In LTE, each downlink sub-frame is divided into two regions: the first one is the control region which consists of two or three OFDM symbols, the remaining part is used for data as shown in figure 3 [5].

In the frequency domain, the bandwidth is divided into sub-carriers, each sub-carrier has a width of 15 KHz, and it is occupied by one OFDM symbol, which is the smallest physical resource in LTE. It is called a resource element, as shown in Figure 3. A group of resource elements form a resource block which is extended for a one slot time, consequently the resource block will have a

width of 180 KHz "12*15KHz", and it is the minimum scheduling unit for LTE users. Each sub-carrier is being modulated using either Quadrature Phase-Shift Keying (QPSK), 16-Quadrature Amplitude Modulation (16-QAM), or 64-QAM. The number of bits each sub-carrier can occupy at a period of one OFDM symbol time depends on the modulation type [6].

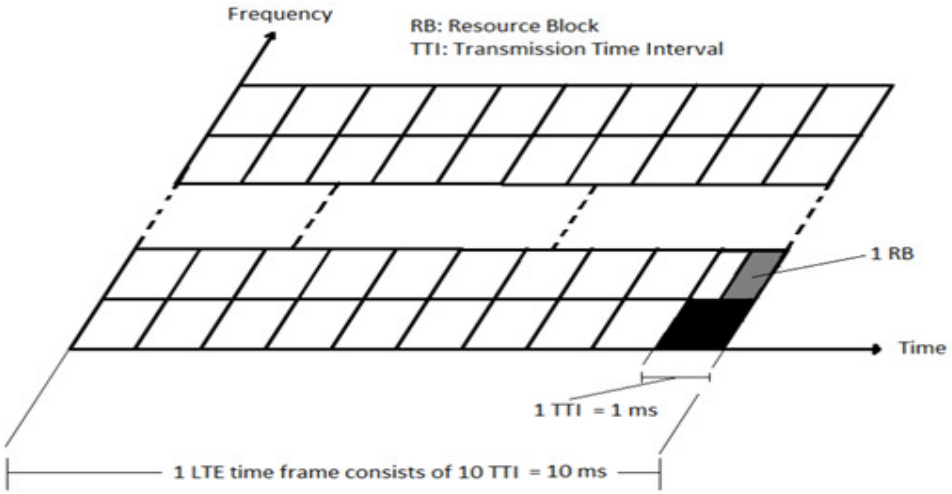


Figure 2. LTE radio spectrum

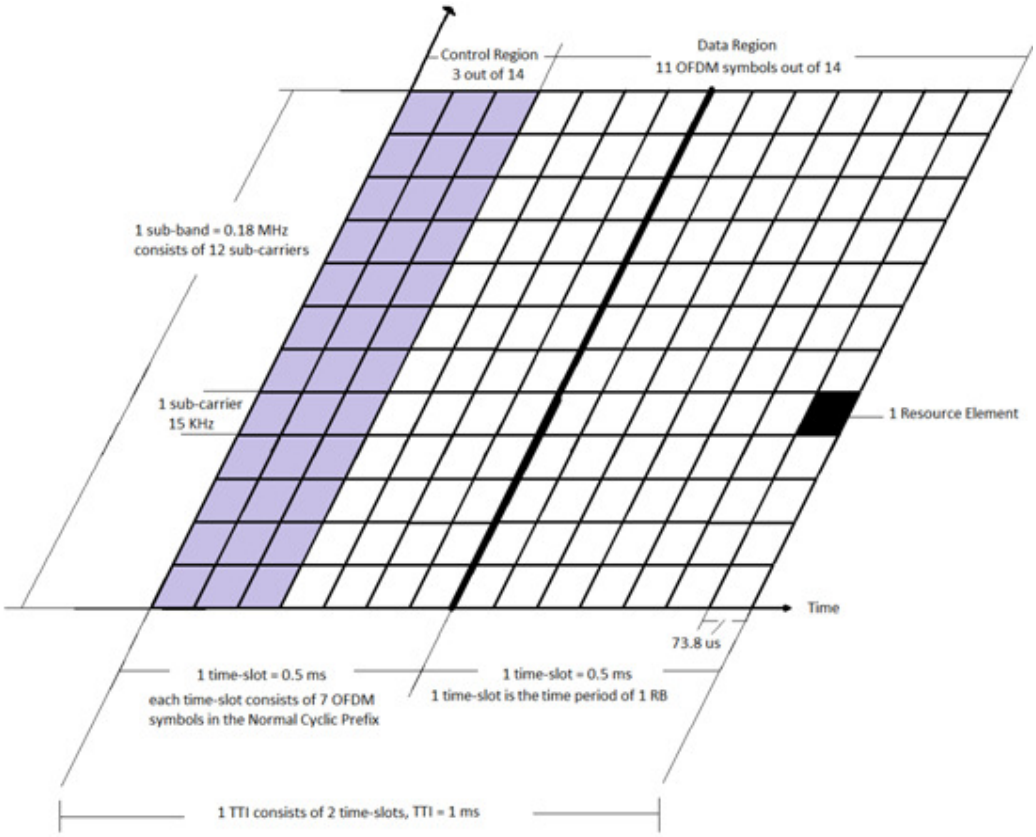


Figure 3. The smallest scheduling LTE unit in the time-frequency domain

4. CARRIER AGGREGATION

The 3GPP Release 8/9 supports at max 20 MHz channel bandwidth for a carrier. The issue of supporting more bandwidth for a carrier seems to be a straight forward solution to support more data rate. Hence the concept of carrier aggregation was introduced where multiple carriers of 20 MHz (or less) would be aggregated for the same UE. Figure 4 shows the principle of carrier aggregation [7].

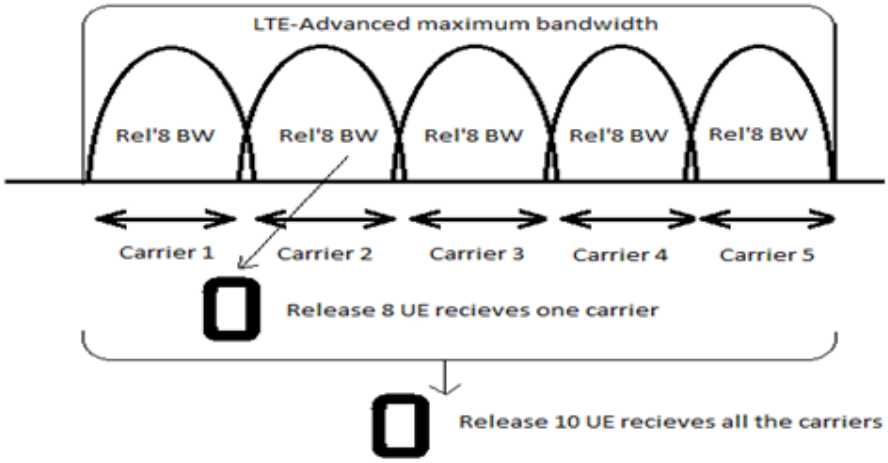


Figure 4. The principle of carrier aggregation

Carrier aggregation in the downlink and the uplink are entirely independent as long as the number of uplink carriers cannot exceed the number of downlink carriers. Each aggregated carrier is called as Component Carrier (CC). 3GPP defined three types of allocation that meets different operator's spectrum scenarios: Intra-band continuous, Intra and Inter-band non-continuous as Figure 5 shows [8].

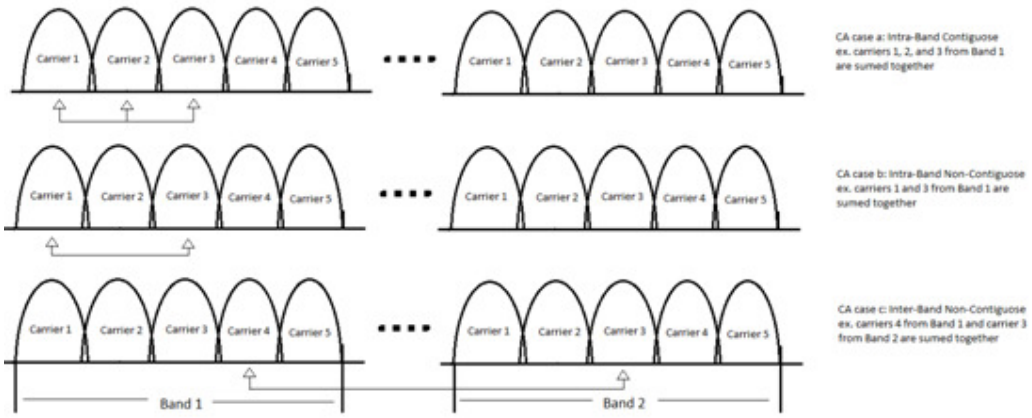


Figure 5. Carrier aggregation cases

5. LTE SCHEDULING

Packet schedulers that allocate radio resources to users' packets are deployed by the MAC layer which is hosted at the eNodeB, while users' applications and connections are scheduled by the

application layer. For allocating the resources to users, a comparison has to be made based on a previously defined metric. This metric could be seen as the priority of each user for a specific resource block. This metric comparison is done every Transmission Time Interval (TTI) in order to perform the allocation decision which is sent to the users over the PDCCH. One of the main characteristics of the Physical Downlink Shared Channel (PDSCH) is to be shared among users, which means that on every TTI which equals 1 ms, portions of the radio spectrum should be distributed among users. The user with the highest metric will be allocated this specific resource block. The Physical Downlink Control Channel (PDCCH) contains the Downlink Control Information (DCI) messages that informs the users about resource blocks allocated for data transmission on the PDSCH in the downlink direction, and about the radio resource that were allocated to their data transmission on the Physical Uplink Control Channel (PUSCH) in the uplink direction [2].

The general model of how a downlink packet scheduler interacts with users is shown in Figure 6. In [2], they have divided the whole scheduling process in a sequence of operations that are repeated every TTI. First, the UE decodes the reference signal and computes the Channel Quality Indicator (CQI) then sends it back to the eNodB. Second, the CQI information is being used by the eNodeB for making the allocation decisions and filling up a RB allocating mask. Third, the Adaptive Modulation and Coding (AMC) module selects the best Modulation and Coding Scheme (MCS) that should be used for the data which will be transmitted by the scheduled users. Fourth, all the above information are sent to the UEs on the PDCCH. Finally, each UE reads the PDCCH and accesses to the proper PDSCH if it has been scheduled. This model is slightly different in the case of uplink because the eNodeB does not require extra information about the uplink channel quality.

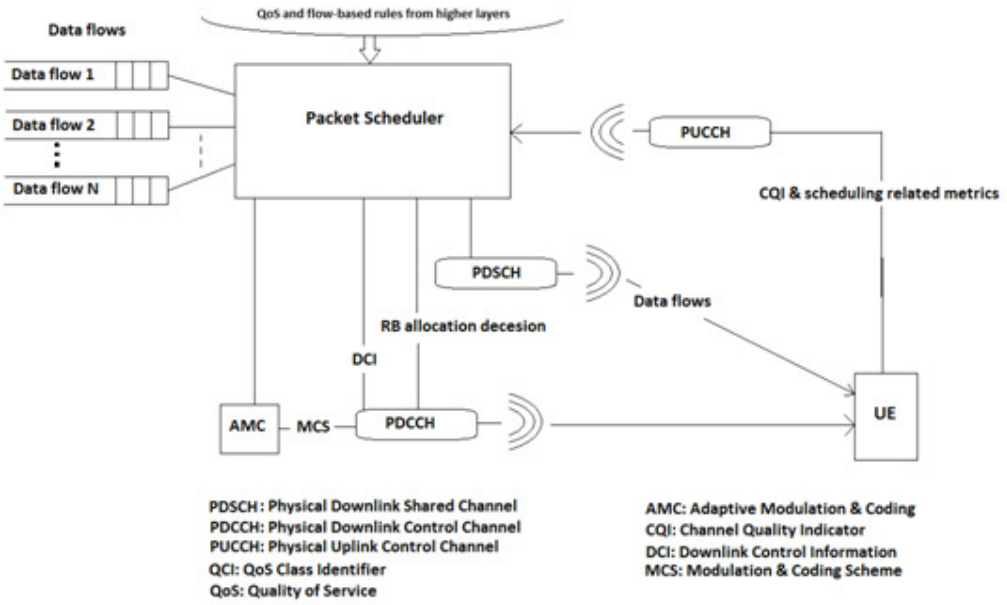


Figure 6. LTE scheduling general model

6. LTE SCHEDULING ALGORITHMS

The LTE scheduling algorithms that were studied in this paper are: the Modified Largest Weighted Delay First (MLWDF), the Logarithmic Rule Algorithm (Log-Rule), and the Exponential Rule Algorithm (Exp-Rule). In all these algorithms the Proportional Fairness (PF) scheduler is used in-order to achieve channel awareness, which makes a trade-off between users'

fairness and spectrum efficiency [2]. It schedule users in a fair way by taking into account both the experienced channel state and the past data rate when assigning radio resources. It aims to

$$k = argmax \frac{r_i(t)}{R_i(t)}$$

obtain satisfying throughput and at the same time, guarantee fairness among flows. The equation that users are selected based on the following formula [9]:

Where $r_i(t)$ is the achievable data rate according to the instantaneous channel quality of user *i* at *t-th* TTI, and $R_i(t)$ is the average data rate of user *i* over a time window, and it is calculated based on the following formula [9]:

$$R_i(t) = (1 - \beta) * R_i(t-1) + \beta * r_i(t-1)$$

Where β is a variable ranging from 0 to 1.

6.1. Modified Largest Weighted Delay First (MLWDF)

The MLWDF scheduling algorithm is designed to support multiple real time data users by taking into account their different QoS requirements. For example, in the case of video services, the instantaneous channel variations and delays are taken into account. It tries to balance the weighted delays of packets and to utilize the knowledge about the channel state efficiently. It chooses user *j* at time *t* based on the following formula [10]:

$$j = max_i \alpha_i \frac{\mu_i(t)}{\bar{\mu}} W_i(t)$$

Where $\mu_i(t)$ is the data rate corresponding to user *i*'s channel state at time *t*, $\mu_i(t)$ is the mean data rate supported by the channel, $Wi(t)$ is the HOL packet delay and $\alpha_i > 0$, i = 1, ... , N are weights that represent the required level of QoS.

The MLWDF's delay is bounded by the Largest Weighted Delay First (LWDF) scheduler. The LWDF metric is based on the system parameter, representing the acceptable probability for the *i-th* user, in which a packet is dropped due to deadline expiration, and this metric is calculated based on the following [2]:

$$m_{i,k}^{LWDF} = \alpha_i . D_{HOL,i}$$

$$\alpha_i = -\frac{log \delta_i}{\tau_i}$$

Where α_i is calculated based on the following formula:

The MLWDF is also expressed in terms of the PF scheduler as in the following formula:

$$m_{i,k}^{MLWDF} = \alpha_i D_{HOL,i} . m_{i,k}^{PF} = \alpha_i D_{HOL,i} . \frac{d_k^i(t)}{\bar{R}^i(t-1)}$$

6.2. Logarithmic Rule Algorithm (LOG-Rule)

The delay of this scheduling algorithm is bounded by the following logarithmic formula [2]:

$$m_{i,k}^{LOGrule} = b_i log(c + \alpha_i D_{HOL,i}).\Gamma_k^i$$

Where α_i, b_i, c are tunable parameters, and the spectral efficiency for the *i-th* user on the *k-th* sub-channel is represented by:

Γ_k^i

6.3. Exponential Rule Algorithm (Exp-Rule)

The delay of this scheduling algorithm is bounded by the following Exponential formula [2]:

$$m_{i,k}^{EXPrule} = b_i exp(\frac{\alpha_i D_{HOL,i}}{c + \sqrt{(1/N_{rt}) \sum_j D_{HOL,j}}}).\Gamma_k^i$$

7. SIMULATION ENVIRONMENT

The simulation environment consists of one Macro cell that is served by one transmitter "eNodeB". The bandwidth was varied from 20MHz "to represent the LTE maximum bandwidth – the case in which CA was not used" to 40MHz "to represent the LTE-A bandwidth – the case in which the CA was used", the number of users was varied from 15, 30, 60, 90, 120, 150, 180, 210 and the video bit-rate was varied from 128Kbps, 242Kbps, and 440Kbps. Detailed parameters of this simulation are listed in Table 1.

Table 1. Simulation Parameters.

Parameter	**Value**
Simulator	LTE-Sim-5
Simulation time	20 sec
Scheduling Algorithms	Exp-Rule, Log-Rule, MLWDF
Network Layout	1 Macro cell "Urban environment"
Transmitter	1 eNodeB
Cell Radius	1 Km
Carrier frequency	2120, 2130 MHz
Bandwidth	2130-2110=20, 2150-2110=40 MHz
Carrier Aggregation case	Inter-band contiguous
Frame structure	FDD
Number of users	15, 30, 60, 90, 120, 150, 180, 210
Users' Distribution	Random
User speed	3 Km / h
Traffic type	Video
Bit rate	128, 242, 440 kbps
Maximum delay	0.1 sec
Buffer type	Infinite buffer
Propagation model	Macro Urban channel realization

8. SIMULATION RESULTS

The LTE-Sim-5 [11] was used in this paper after modifying it to support the first case of the CA. While the LTE-Sim-5 simulator is in the process of simulating a scenario with a pre-defined

conditions, it takes into account both the signalling and data traffic. However, it only displays the data traffic in its traces. These data traffic traces are used to measure the QoS parameters, the system's average throughput, Packet Loss Rate (PLR), average packet delay, and fairness among users. These measurements are displayed in all the following figures by taking the number of users as its X-axis factor and the QoS parameter as the Y-axis factor.

8.1. System's Average Throughput

System's average throughput is defined as the amount of the total received packets for all users per second. The system's average throughput with and without the use of carrier aggregation for the three scheduling algorithms at different video bit-rates are displayed in the following figures “fig. 7, 8, and 9.

According to the obtained results that are displayed in figures “fig. 7, 8, and 9”, increasing the number of users will increase the system's average throughput until it reaches its maximum value. This increase is due to transmitting more data from the eNodeB to the new added users. The maximum value of the system's average throughput differs based on the system's capabilities.

The use of CA will increase the bandwidth which will decrease the values of the spectral efficiency. This is defined in the literature as “the number of successfully transmitted bits normalized by the consumed resources in time and in bandwidth” [12]. This will prompt the eNodeB to transmit more data to users leading to higher system's average throughput which will in turn lead to increasing the spectral efficiency. This is shown in figures “fig. 7, and 8” in which the use of CA has slightly increased the system's average throughput even before reaching the maximum value of the system's average throughput in the case where the CA was not used.

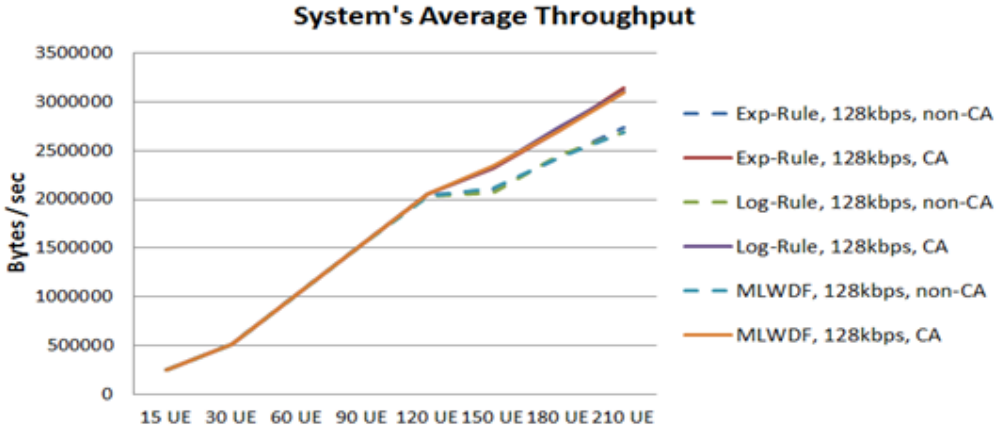


Figure 7. System's average throughput at video bit-rate 128Kbps, with and without the use of CA

According to the obtained results that are displayed in figure 7, it is shown that at this scenario's conditions, the three algorithms showed similar performance in both cases.

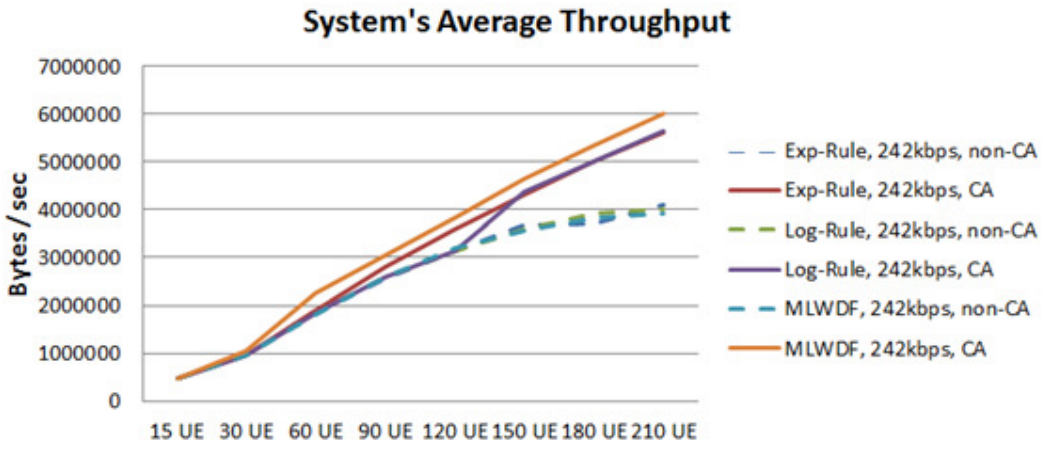


Figure 8. System's average throughput at video bit-rate 242Kbps, with and without the use of CA

According to the obtained results that are displayed in figure 8, in both cases when the CA was and was not used, it is shown that the system's average throughput increased gradually by increasing the number of users. However, in the case in which the CA was not used and the increase of the system's average throughput started to reach its maximum value of between (4 and 4.5) MBps, this increase almost stopped. At this scenario's conditions, when the CA was not used, the three algorithms showed similar fluctuating performance. However, when the CA was used, the use of the MLWDF led to a slightly higher system's average throughput than the Exp-Rule, and the Log-Rule.

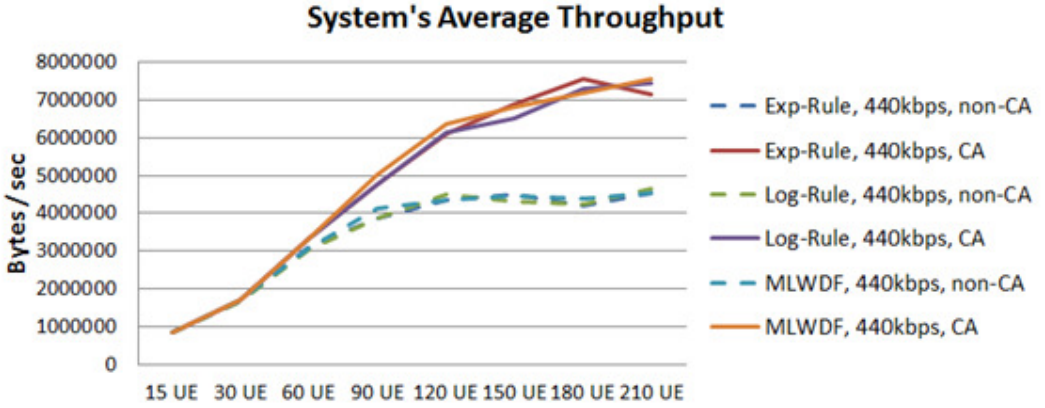


Figure 9. System's average throughput at video bit-rate 440Kbps, with and without the use of CA

According to the obtained results that are displayed in figure 9, it is shown that the maximum value of the system's average throughput was almost doubled by the use of CA. At this scenario's conditions, the three algorithms showed a similar fluctuating performance in both cases.

8.2. Packet Loss Rate (PLR)

Packet Loss Rate (PLR) is measured by dividing the difference between the total transmitted and received packets for all users over the total transmitted packets. The greater the system's load the higher the value of the PLR. The system's load could be increased by increasing the number of users in the same cell, or by increasing the video bit-rates. The increase in the PLR values with the increase of the system's load is due to the increasing number of packets in the waiting queues that are competing for the same resource blocks. This will lead to a higher rate of dropped packets from these queues. Since both factors, the number of users and the video bit-rates, are proportionally related to increasing the PLR values, increasing one will limit the increase of the

other in a limited resources system. For example, the significant increase of the PLR values started to take place when the number of users exceeded 120 users at a video bit-rate of 128kbps, and when it exceeded 60 users at a video bit-rate of 242kbps, and when it exceeded 30 users at a video bit-rate of 440kbps. The increases of the PLR values with the increase of the number of users and the video bit-rates are shown in figures "fig. 10, 11, and 12". The results in these three figures, shows a decrease of the PLR values with the use of CA. This is because using the CA will increase the available resources and decrease the values of the spectral efficiency. This will prompt the eNodeB to allocate more resource blocks to the same user, which will reduce the amount of packets in the waiting queues, allowing more packets to be served and fewer packets to be dropped. This causes a decrease in the values of the PLR as shown in figures "fig. 10, 11, and 12".

The PLR with and without the use of carrier aggregation for the three scheduling algorithms at different video bit-rates are displayed in the following figures "fig. 10, 11, and 12".

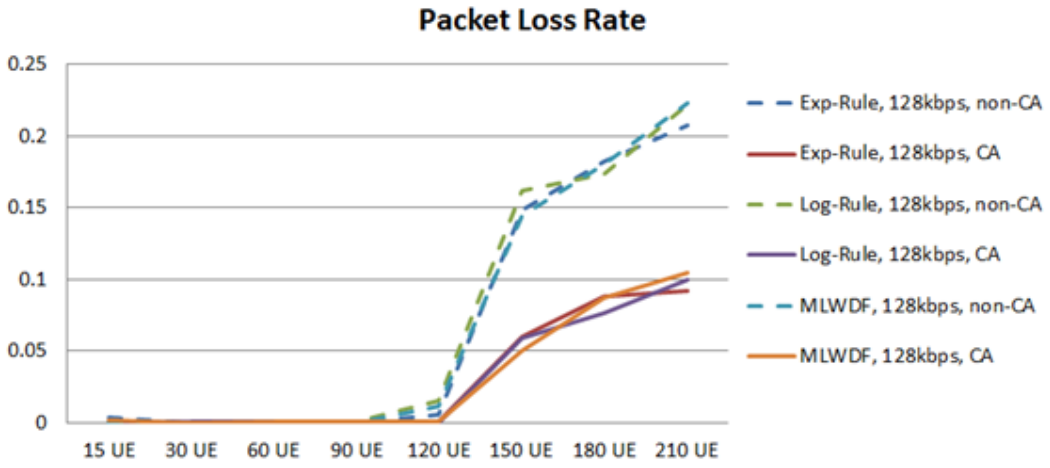


Figure 10. Packet Loss Rate (PLR) at video bit-rate 128kbps, with and without the use of CA

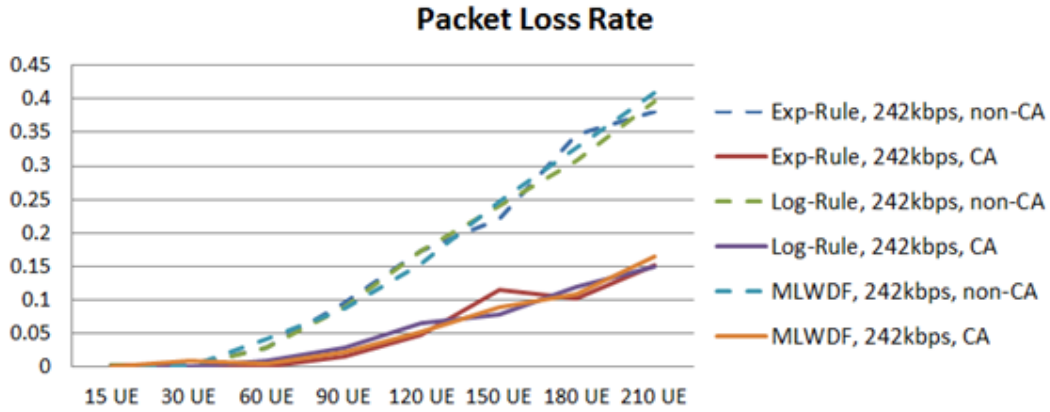


Figure 11. Packet Loss Rate (PLR) at video bit-rate 242kbps, with and without the use of CA

According to the obtained results that are displayed in figure 10, it is shown that the number of lost packet started to increase significantly when the number of users exceeded 120 UEs. However, when the CA was not used, this increase was twice what it was when the CA was used. At this scenario's conditions, the three algorithms showed similar fluctuating performance in both cases.

According to the obtained results that are displayed in figure 11, the number of lost packets started to increase significantly after the number of users exceeded 60 UEs. However, when the CA was not used, the PLR was about two to three times what it was when the CA was used. At this scenario's conditions, the three algorithms showed similar fluctuating performance in both cases.

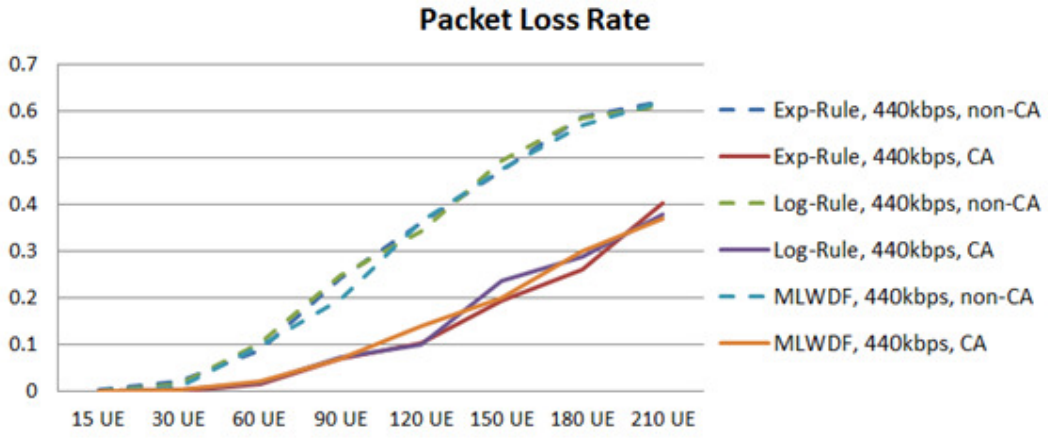


Figure 12. Packet Loss Rate (PLR) at video bit-rate 440kbps, with and without the use of CA

According to the obtained results that are displayed in figure 12, the number of lost packets started to increase critically after the number of users exceeded 30 UEs. However, when the CA was not used, the differences between the values of the PLR in both cases were on an average of twice what it was when the CA was used. At this scenario's conditions, the three algorithms showed similar fluctuating performance in both cases.

8.3. Average Packet Delay

The packet delay is the time that it takes a packet to travel from the source to its destination. It includes the propagation and waiting time of the packet. The Average Packet Delay is measured by dividing the sum of the total packet delays that were successfully received over the number of total packets. The use of the CA causes a significant beneficial reduction of the average packet delay. This is because it reduces the propagation time which is found by dividing the packet length by the link bandwidth. Also, it reduces the waiting time for the packets in the waiting queues at the eNodeB.

The average packet delay with and without the use of carrier aggregation for the three scheduling algorithms at different video bit-rates are displayed in the following figures “fig. 13, 14, and 15”.

According to the obtained results that are displayed in figure 13, when the CA was not used, the average packet delay kept significantly increasing with the increase of the number of users into the the cell. While this was occurring, the PLR values remained almost negligible at this scenario's conditions, as seen in fig. 10. This led to a high queueing delay which peaked at 120 UE. However, when the number of users exceeded 120 UE, the waiting time for some of the packets in the waiting queues started to exceed the threshold defined as 0.1 sec, resulting in a higher rate of dropped packets. Consequently, the PLR values started to increase significantly. And since the average packet delay does not include in its calculations the dropped and lost packets, the values of it reduced significantly. This problem did not occur when the CA was used, because the available resources were almost doubled and the packets at the waiting queues were able to be served more quickly. At this scenario's conditions, the three algorithms showed similar fluctuating performance in both cases.

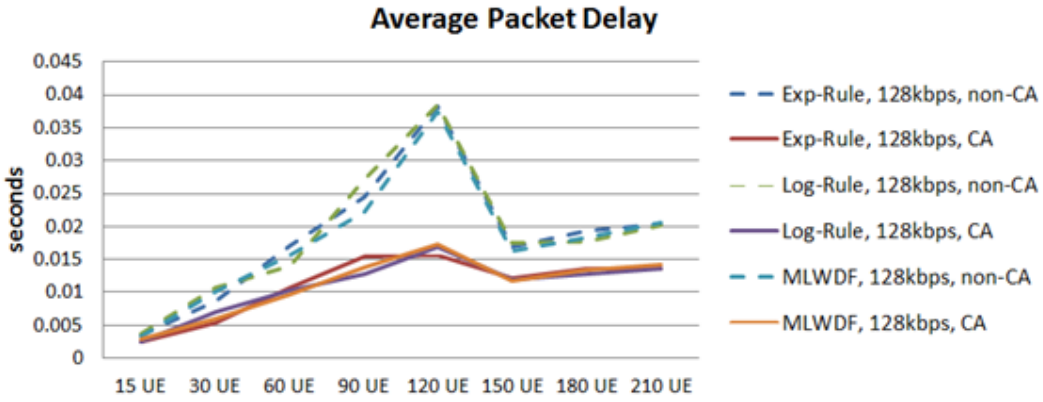


Figure 13. Average packet delay at video bit-rate 128kbps, with and without the use of CA

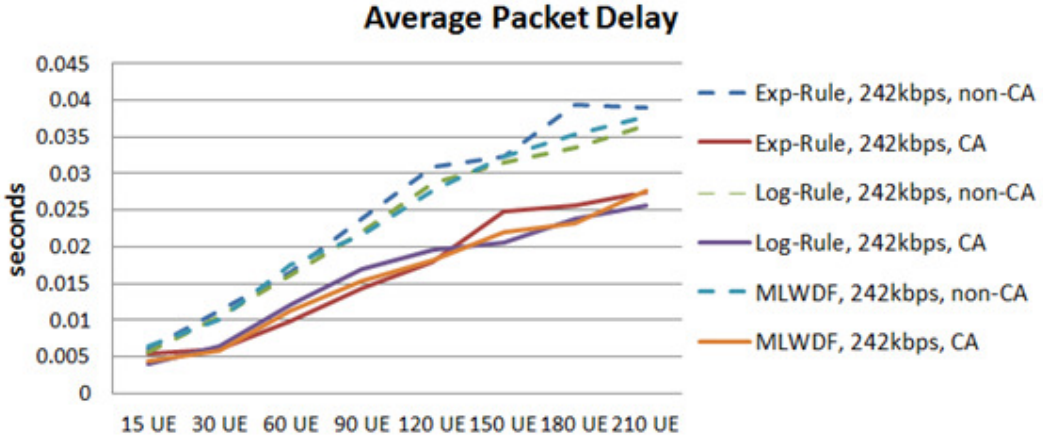


Figure 14. Average packet delay at video bit-rate 242kbps, with and without the use of CA

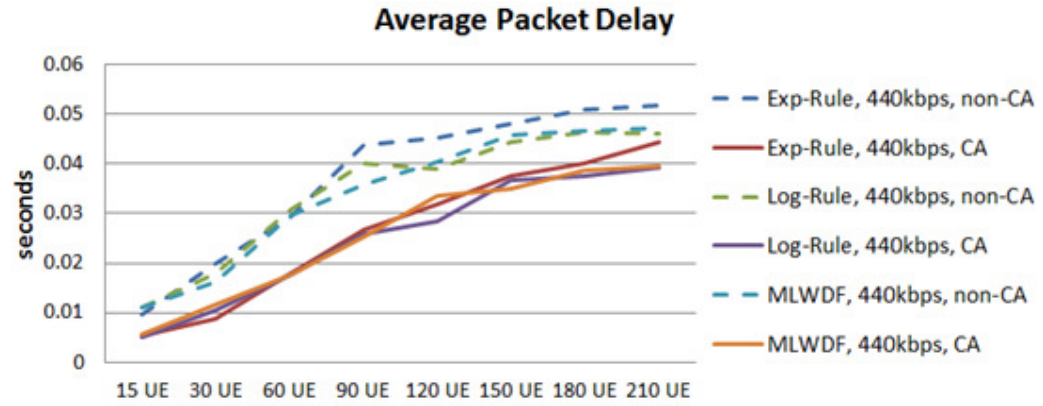


Figure 15. Average packet delay at video bit-rate 440kbps, with and without the use of CA

According to the obtained results that are displayed in figure 14, it is shown that the use of CA reduced the average packet delay to 60% of what it was when the CA was not used. The three algorithms showed similar fluctuating performance in both cases. However, the MLWDF results indicate that this algorithm has a more reliable performance in terms of increasing the average packet delay with increasing the number of users.

According to the obtained results that are displayed in figure 15, it is shown that the use of CA reduced the average packet delay to 50-60% of what it was when the CA was not used. This

effect of the reduction continued until the number of UEs exceeded 90. After adding more users to the Macro-cell, this reduction started to decrease until the value of the average packet delay was almost 80% of what it was when the CA was not used. At this scenario's conditions, the three algorithms showed similar fluctuating performance in both cases with more performance stability to the MLWDF. The Exp-Rule had a slightly higher average packet delay than the other two algorithms.

8.4. Fairness

Jain's fairness index is used in this paper to determine if the scheduling algorithms are distributing fair portions of the spectrum to the users. It is measured by the following formula [13]:

$$F = \frac{(\sum_{k=1}^{K} r_k)^2}{K \sum_{k=1}^{K} r_k}$$

Where r_k denotes the throughput of user k.

The Jain's fairness index for the three scheduling algorithms at different video bit-rates, with and without the use of carrier aggregation, are displayed in the following figures “fig. 16, 17, and 18”.

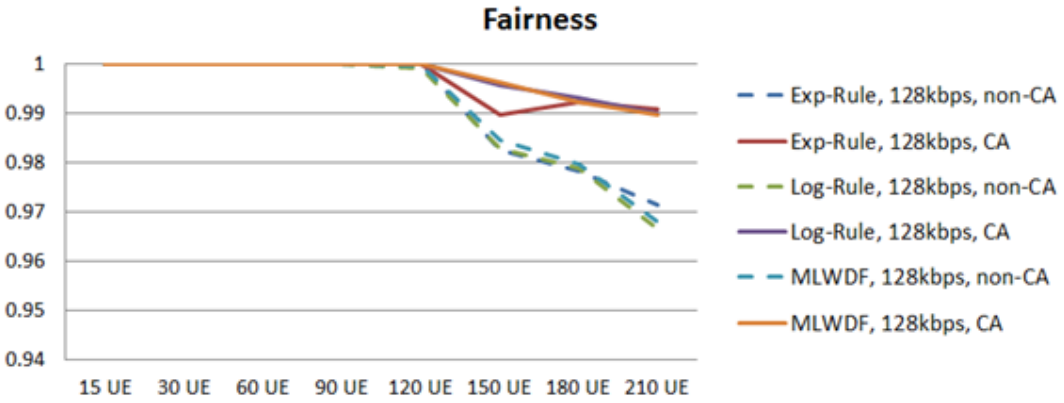


Figure 16. Fairness index for the three algorithms at video bit-rate 128kbps, with and without the use of CA

According to the obtained results that are displayed in figure 16, at this scenario's conditions, the three algorithms showed similar fluctuating fairness in both cases. An excellent fairness took place among users for the three algorithms when the CA was used. However, when the CA was not used, the fairness indicator dropped slightly after the number of users exceeded 120 UEs.

According to the obtained results that are displayed in figure 17, at this scenario's conditions, the three algorithms showed similar fluctuating fairness in both cases. An excellent fairness was forced among users for the three algorithms when the CA was used. However, when the CA was not used, the fairness indicator dropped 5-20% from what it was when the CA was used, this drop started to take place after the number of users exceeded 120 UEs.

According to the obtained results that are displayed in figure 18, at this scenario's conditions, the three algorithms showed similar fluctuating fairness in both cases. When the CA was used, and when the number of users started to exceed 90 UE, the fairness indicator started to drop until it reached a value of 0.8 when the number of users was 210 UE. When the CA was not used the drop of the fairness indicator was more significant, it reached to a value of 0.6 when the number of users was 210 UE.

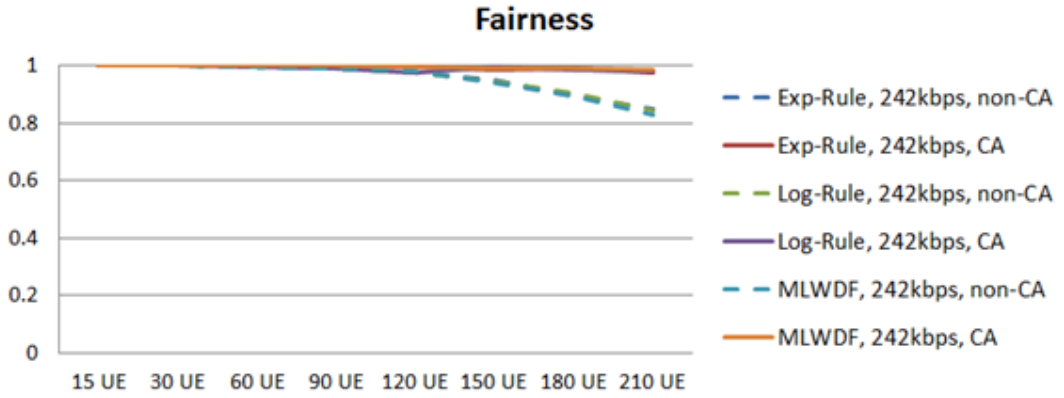


Figure 17. Fairness index for the three algorithms at video bit-rate 242kbps, with and without the use of CA

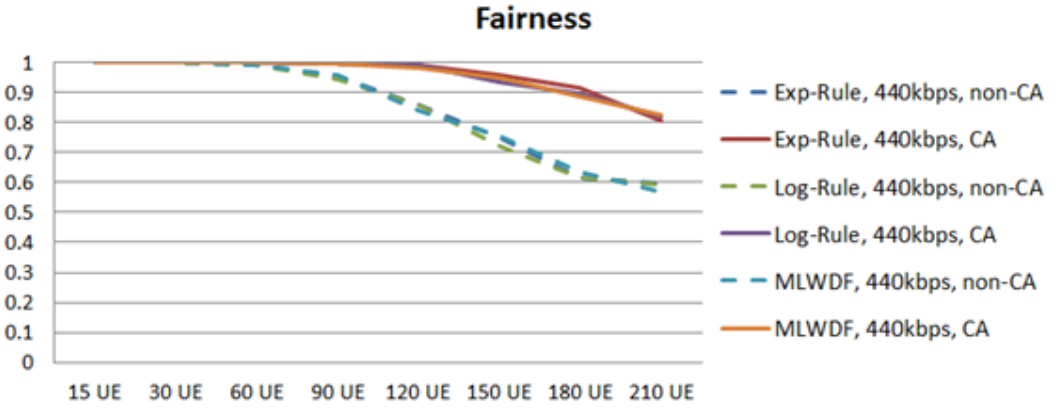


Figure 18. Fairness index for the three algorithms at video bit-rate 440kbps, with and without the use of CA

9. CONCLUSION

This paper has provided a comparative study on three Channel-aware/QoS-aware scheduling algorithms over LTE/LTE-A for video-applications. The comparison aimed to study the behaviour of the selected algorithms when the network load reaches the capacity limitations of the LTE macro-cell and to compare it when the network is supported by the CA feature to operate as LTE-A. In addition, there was a comparison among the scheduling algorithms in both cases. The evaluation process was based on simulating different scenarios by varying the number of users, the video bit-rates, and the system's bandwidth. The LTE-Sim-5 was used in the simulation process both with and without modifications. The QoS performance evaluation was in terms of the QoS parameters, the system's average throughput, Packet Loss Rate (PLR), average packet delay, and fairness among users. Simulation results show that the system's average throughput was significantly improved by the use of CA. The capacity limitation boundaries were also doubled. The use of the MLWDF slightly improved the system's average throughput while the CA was used. The PLR was significantly reduced almost 50% by the use of CA at a video bit-rate of 128kbps, by 50-70% at a video bit-rate of 242kbps, and by 40-60% at a video bit-rate of 440kbps. The average packet delay was reduced by the use of CA at a video bit-rate of 128kbps, and at a video bit-rate of 242Kbps by 30-40%, and at a video bit-rate of 440kbps by 20-30%. The MLWDF showed more performance stability in terms of increasing the average packet delay with increasing the number of users. The Exp-Rule had a slightly higher delay than the LOG-rule and the MLWDF. The fairness indicator was improved with the use of CA by a factor of 10-20%.

These results show that the use of CA is worth being investigated by researchers, implemented by the manufacturers, and deployed by the service providers.

APPENDIX A: LIST OF ACRONYMS

3GPP Third Generation Partnership Project
AMC Adaptive Modulation and Coding
AuC Authentication Center
CA Carrier Aggregation
CC Component Carrier
CQI Channel Quality Indicator
DCI Downlink Control Information
eNodeB Evolved NodeB
EPC Evolved Packet Core
Exp-Rule Exponential Rule
FDD Frequency Division Duplex
GBR Guaranteed Bit Rate
GPRS General Packet Radio Service
HOL Head of Line
HSS Exponential Rule
LTE Long Term Evolution
LTE-A LTE-Advanced
Log-Rule Logarithmic Rule
LWDF Largest Weighted Delay First
MAC Medium Access Control
MCS Modulation and Coding Scheme
MLWDF Modified Largest Weighted Delay First
MME Mobility Management Entity
OFDM Orthogonal Frequency Division Multiplexing
PCEF Policy Control Enforcement Function
PCRF Policy Control and Charging Rules Function
PDCCH Physical Downlink Control Channel
PDSCH Physical Downlink Shared Channel
PDCP Packet Data Control Protocol
PDNs Packet Data Networks
PF Proportional Fairness
PLR Packet Loss Rate
PUSCH Physical Uplink Control Channel
P-GW PDN-Gateway
QAM Quadrature Amplitude Modulation
QoS Quality of Service
QPSK Quadrature Phase-Shift Keying
R8/9 Release 8/9
R10/11 Release 10/11
RAN Radio Access Network
RB Resource Block
RLC Radio Link Control
RRC Radio Resource Control
S-GW Serving Gateway
TFTs Traffic Flow Templates
TTI Transmission Time Interval
UE User Equipment
UMTS Universal Mobile Telecommunication Systems

VoIP Voice over IP
Wi-Fi wireless fidelity
WiMAX Worldwide Interoperability for Microwave Access

REFERENCES

[1] "3gpp, "lte." [online]. available:." http://www.3gpp.org/technologieskeywords-acronyms/98-lte

[2] F. Capozzi, G. Piro, L. A. Grieco, G. Boggia, and P. Camarda, "Downlink packet scheduling in lte cellular networks: Key design issues and a survey," Communications Surveys & Tutorials, IEEE, vol. 15, no. 2, pp. 678–700, 2013.

[3] "3gpp, "carrier aggregation explained." [online]. Available:." http://www.3gpp.org/technologies/keywords-acronyms/101-carrier-aggregation-explained

[4] H. Holma and A. Toskala, LTE for UMTS-OFDMA and SC-FDMA based radio access. John Wiley & Sons, 2009.

[5] S. P. E. Dahlman and J. Skold, 4G LTE/LTE-Advanced for Mobile Broadband. Elseiver, 2011.

[6] "Anritsu, "lte resource guide." [online]. available:." http://www.3gpp.org/technologies/keywords-acronyms/98-lte

[7] A. T. H. Holma, LTE Advanced: 3GPP Solution for IMT-Advanced. Wiley, 2012.

[8] "Anritsu, understanding lte-advanced carrier aggregation. [online]. Available:." http://downloadfile.anritsu.com/RefFiles/en-GB/Promotions/Understanding-Carrier-Aggregation-web.pdf

[9] B. Liu, H. Tian, and L. Xu, "An efficient downlink packet scheduling algorithm for real time traffics in lte systems," in Consumer communications and networking conference (CCNC), 2013 IEEE, pp. 364–369, IEEE, 2013.

[10] M. Iturralde, T. Ali Yahiya, A. Wei, and A.-L. Beylot, "Performance study of multimedia services using virtual token mechanism for resource allocation in lte networks," in Vehicular Technology Conference (VTC Fall), 2011 IEEE, pp. 1–5, IEEE, 2011.

[11] G. Piro, L. A. Grieco, G. Boggia, F. Capozzi, and P. Camarda, "Simulating lte cellular systems: an open-source framework," Vehicular Technology, IEEE Transactions on, vol. 60, no. 2, pp. 498–513, 2011.

[12] Y. Chen and B. Walke, "Analysis of cell spectral efficiency in 3gpp lte systems," in Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium on, pp. 1799–1804, IEEE, 2013.

[13] C. Bae and D.-H. Cho, "Fairness-aware adaptive resource allocation scheme in multihop ofdma systems," Communications Letters, IEEE, vol. 11, no. 2, pp. 134–136, 2007.

INTENTIONAL BLANK

A MODIFIED METHOD FOR PREDICTIVITY OF HEART RATE VARIABILITY

Mazhar B. Tayel[1] and Eslam I AlSaba[2]

[1,2] Department of Electrical Engineering,
Alexandria University, Alexandria, Egypt
profbasyouni@gmail.com
eslamibrahim@myway.com

ABSTRACT

Heart Rate Variability (HRV) plays an important role for reporting several cardiological and non-cardiological diseases. Also, the HRV has a prognostic value and is therefore quite important in modelling the cardiac risk. The nature of the HRV is chaotic, stochastic and it remains highly controversial. Because the HRV has utmost importance, it needs a sensitive tool to analyze the variability. In previous work, Rosenstein and Wolf had used the Lyapunov exponent as a quantitative measure for HRV detection sensitivity. However, the two methods diverge in determining the HRV sensitivity. This paper introduces a modification to both the Rosenstein and Wolf methods to overcome their drawbacks. The introduced Mazhar-Eslam algorithm increases the sensitivity to HRV detection with better accuracy.

KEYWORDS

Heart Rate Variability, Chaotic system, Lyapunov exponent, Transform domain, and Largest Lyapunov Exponent.

1. INTRODUCTION

Cardiovascular diseases are a growing problem in today's society. The World Health Organization (WHO) reported that these diseases make up about 30% of total global deaths. Those heart diseases have no geographic, gender or socioeconomic boundaries [3]. Therefore, early stage detection of cardiac irregularities and correct treatment are very important. This requires a good physiological understanding of cardiovascular system.

Studying the fluctuations of heart beat intervals over time reveals a lot of information called heart rate variability (HRV) analysis. A reduction of HRV has been reported in several cardiological and non-cardiological diseases. Also, HRV has a prognostic value and is therefore quite significant in modelling the cardiac risk. HRV has already proved his usefulness and is based on several articles that have reviewed the possibilities of HRV [1 - 5].

The fact that HRV is a result of both linear and nonlinear fluctuations opened new perspectives as previous research was mostly restricted to linear techniques. Some situations or interventions can change the linear content of the variability, while leaving the nonlinear fluctuations intact. In addition, the reverse can happen: interventions, which up till now have been believed to leave cardiovascular fluctuations intact based on observations with linear methods, can just as well modify the nonlinear fluctuations. This can be important in the development of new drugs or treatments for patients. This paper introduces a modification algorithm to overcome the

Jan Zizka et al. (Eds) : DBDM, CICS, CCNET, CSIP - 2015
pp. 67–77, 2015. © CS & IT-CSCP 2015 DOI : 10.5121/csit.2015.50706

drawbacks arising in both Rosenstein and Wolf method using the same approach of Lyapunov exponent. That analysis the nonlinear behaviour of the HRV signals.

2. HEART RATE VARIABILITY

Heart Rate Variability (HRV) is a phenomenon that describes temporal variation in intervals between consecutive heartbeats in sinus rhythm. HRV refers to variations in beat-to-beat intervals corresponding to instantaneous HRs. HRV is a reliable reflection of many physiological factors modulating normal rhythm of heart. In fact, they provide a powerful means of observing interplay between sympathetic and parasympathetic nervous systems. It shows that a structure generating a signal is not only simply linear, but also involves nonlinear contributions.

Spontaneous variability of HR has been related to three major physiological originating factors: quasi-oscillatory fluctuations thought to arise in blood-pressure control, variable frequency oscillations due to thermal regulation, and respiration. Frequency selective analysis of cardiac inter-beat interval sequences allows separate contributions to be isolated. Using this method, a laboratory and field study of effects of mental work load on the cardiac interval sequence has been carried out [4].

The diagnosticity of HR is restricted by several factors like environmental stressors and physical demands that may be associated with a task. These tasks may have different physiological consequences and change in HR may depends on these factors more than mental workload. Backs (1998) [6] focused on the fact that observed HR could be caused by different underlying patterns of autonomic nervous system activity. If different information processing demands affect the heart via different modes of autonomic control, it could increase diagnosticity of HR [6]. Backs‟ study addressed the validity of the autonomic component, using data from a large study, in which many central and peripheral psycho-8 physiological measures were collected simultaneously while performing single and dual tasks which had different physical demands. The measures collected were the residual HR, parasympathetic and sympathetic activity, respiratory sinus arrhythmia (RSA), and THM (Traube-Hering-Mayer) wave using principle component analysis (PCA), image factoring, impedance cardiogram (ZKG) and ElectroCardioGram (ECG). The sympathetic and parasympathetic systems were examined for independence. From the study, the PCA factors computed on raw ECG data provided useful information like different autonomic modes of control were found that were not evident in heart period. The objective was to verify if factors extracted using residual HR as a marker variable validly reflected cardiac sympathetic activity and if the solutions obtained from raw and baseline corrected data were in compliance with each other. This information about the underlying autonomic activity may increase the diagnosticity of HR. Figure 1 shows the example of HR variation of normal subject (control case).

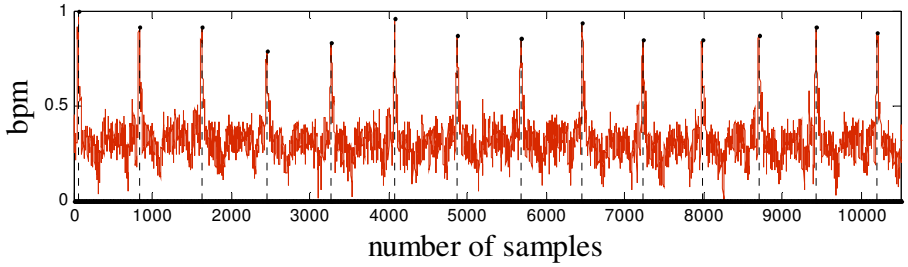


Figure 1. Heart rate variation of normal subject.

3. LYAPUNOV EXPONENT

Lyapunov exponent Λ is a quantitative measure of the sensitive dependence on the initial conditions. It defines the average rate of divergence or convergence of two neighbouring trajectories in the state-space. Consider two points in a space, X_0 and $X_0 + \Delta x_0$, each of which will generate an orbit in that space using system of equations as shown in figure 2.

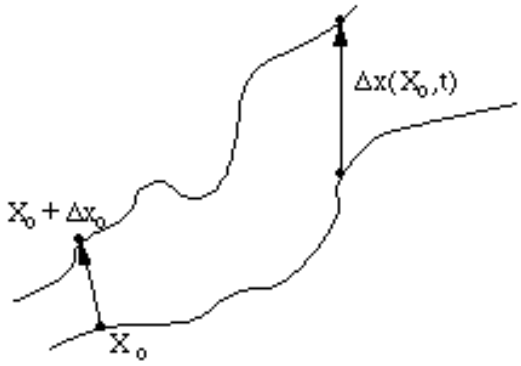


Figure 2. Calculation for two neighbouring trajectories.

These orbits are considered as parametric functions of a variable. If one of the orbits is used as a reference orbit, then the separation between the two orbits will be a function of time. Because sensitive dependence can arise only in some portions of a system, this separation is a location function of the initial value and has the form $\Delta x(X_0, t)$. In a system with attracting fixed points or attracting periodic points, $\Delta x(X_0, t)$ decreases asymptotically with time. For chaotic points, the function $\Delta x(X_0, t)$ will behave unpredictably. Thus, it should study the mean exponential rate of divergence of two initially closed orbits using the formula

$$\Lambda = \lim_{\substack{t\to\infty \\ |\Delta x_0|\to 0}} \frac{1}{t} \ln \frac{\Delta x(X_0,t)}{|\Delta x_0|} \qquad (1)$$

This number, called the Lyapunov exponent "Λ". An exponential divergence of initially nearby trajectories in state-space coupled with folding of trajectories, to ensure that the solutions will remain finite, is the general mechanism for generating deterministic randomness and unpredictability. Therefore, the existence of a positive Λ for almost all initial conditions in a bounded dynamical system is the widely used definition of deterministic chaos. To discriminate between chaotic dynamics and periodic signals, Λ s are often used. The trajectories of chaotic signals in state-space follow typical patterns. Closely spaced trajectories converge and diverge exponentially, relative to each other. *A negative exponent* $(\Lambda < 0)$ the orbit attracts to a stable fixed point or stable periodic orbit. Negative Lyapunov exponents are characteristic of dissipative or non-conservative systems. Such systems exhibit asymptotic stability. The more negative the exponent, the greater the stability. Super stable fixed points and super stable periodic points have a Lyapunov exponent of $\Lambda = -\infty$. This is something similar to a critically damped oscillator in that the system heads towards its equilibrium point as quickly as possible. A *zero exponent* $(\Lambda = 0)$ the orbit is a neutral fixed point (or an eventually fixed point). A Lyapunov exponent of zero indicates that the system is in some sort of steady state mode. A physical system with this exponent is conservative. Such systems exhibit Lyapunov stability. Take the case of two identical simple harmonic oscillators with different amplitudes. Because the frequency is independent of the amplitude, a phase portrait of the two oscillators would be a pair of concentric

circles. The orbits in this situation would maintain a constant separation. Finally, a *positive exponent* implies the orbits are on a *chaotic attractor.* Nearby points, no matter how close, will diverge to any arbitrary separation. These points are unstable [3]. The flowchart of the practical algorithm for calculating largest Lyapunov exponents is shown in figure 3.

In the following, the Largest Lyapunov Exponent algorithms are discussed.

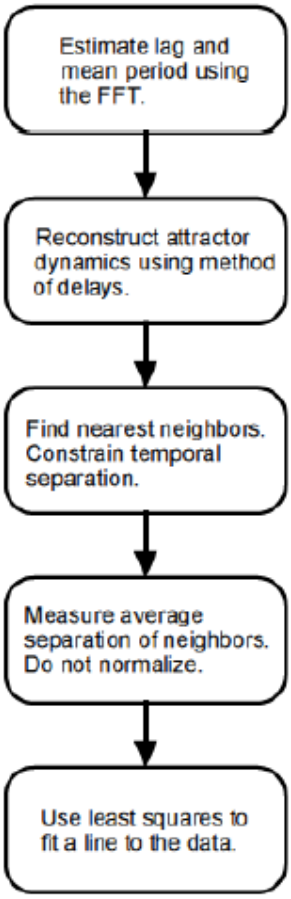


Figure 3. Flowchart of the practical algorithm for calculating largest Lyapunov exponents

3.1. Wolf's Algorithm

Wolf's algorithm [7] is straightforward and uses the formulas defining the system. It calculates two trajectories in the system, each initially separated by a very small interval. The first trajectory is taken as a reference, or 'fiducial' trajectory, while the second is considered 'perturbed'. Both are iterated together until their separation is large enough.

In this study, the two nearby points in a state-space x_i and $x_i + \Delta x$, that are function of time and each of which will generate an orbit of its own in the state, the separation between the two orbits Δx will also be a function of time. This separation is also a function of the location of the initial value and has the form $\Delta x(x_i, t)$, where t is the value of time steps forward in the trajectory. For chaotic data set, the mean exponential rate of divergence of two initially close orbits is characterized by

$$\Lambda = \lim_{K\to\infty} \frac{1}{K} ln \frac{|\Delta x(x_i,t)|}{|\Delta x|} \quad (2)$$

The maximum positive Λ is chosen as Λ_m.

Largest Lyapunov exponent's quantify sensitivity of the system to initial conditions and gives a measure of predictability. This value decreases for slowly varying signals like congenital heart block (CHB) and Ischemic/dilated cardiomyopathy and will be higher for the other cases as the variation of RR is more [8, 9]. The LLE is 0.505 for the HR signal shown in Figure 1.

3.2. Rosenstein's Algorithm

Rosenstein's algorithm [10] works on recorded time-series, where the system formulas may not be available. It begins by reconstructing an approximation of the system dynamics by embedding the time-series in a phase space where each point is a vector of the previous m points in time (its 'embedding dimension'), each separated by a lag of j time units. Although Taken's theorem [11] states that an embedding dimension of 2D + 1 is required to guarantee to capture all the dynamics of a system of order D, it is often sufficient in practice to use m = D. Similarly, although an effective time lag must be determined experimentally, in most cases j = 1 will suffice.

Given this embedding of the time-series, for each point founding its nearest neighbour (in the Euclidean sense) whose temporal distance is greater than the mean period of the system, corresponding to the next approximate cycle in the system's attractor. This constraint positions the neighbours as a pair of slightly separated initial conditions for different trajectories. The mean period was calculated as the reciprocal of the mean frequency of the power spectrum of the time-series calculated in the usual manner using the Fast Fourier Transform (FFT). As shown in figure 3.

Now it can be performed a process similar to Wolf's algorithm to approximate the Largest Lyapunov Exponent (LLE). The first step of this approach involves reconstructing the attractor dynamics from the RR interval time series. After reconstructing the dynamics, the algorithm locates the nearest neighbour of each point on the trajectory. The nearest neighbour, $x_{j'}^{/}$ is found by searching for the point that minimizes the distance to the particular reference point, x_j. This is expressed as:

$$d_j(0) = \min||x_j - x_{j'}|| \qquad \forall x_{j'} \quad (3)$$

where $d_j(0)$ is the initial distance from the j^{th} point to its nearest neighbour and ‖ ... ‖ denotes the Euclidean norm. An additional constraint is imposed, namely that nearest neighbours have a temporal separation greater than the mean period of the RR interval time series. Therefore, one can consider each pair of neighbours as nearby initial conditions for different trajectories. The largest Lyapunov exponent (LLE) is then estimated as the mean rate of separation of the nearest neighbours. More concrete, it is assumed that the j^{th} pair of nearest neighbours diverge approximately at a rate given by the largest Lyapunov exponent Λ_m:

$$d_j(j) \approx d_j(0) e^{\Lambda_m(i.\Delta t)} \quad (4)$$

By taking the ln of both sides of this equation:

$$\ln d_j(j) \approx \ln d_j(0) + \Lambda_m(i.\Delta t) \quad (5)$$

which represents a set of approximately parallel lines (for j = 1, 2, . . . ,M), each with a slope roughly proportional to Λ_m.

The natural logarithm of the divergence of the nearest neighbour to the j^{th} point in the phase space is presented as a function of time. The largest Lyapunov exponent is then calculated as the slope of the 'average' line, defined via a least squares fit to the 'average' line defined by:

$$y(t) = \frac{1}{\Delta t}\langle \ln d_j(t) \rangle \quad (6)$$

where $\langle \ldots \rangle$ denotes the average over all values of j. This process of averaging is the key to calculating accurate values of Λ_m using smaller and noisy data sets compared to other Lyapunov algorithms. The largest Lyapunov exponent (LLE) is 0.7586 for the HR signal shown in Figure 1.

3.3. The Λ_m calculations

The Λ_m exponents were calculated using the Wolf and Rosenstein algorithms implemented as previously recommended. For both algorithms, the first two steps were similar. An embedded point in the attractor was randomly selected, which was a delay vector with d_E elements

$$[x(t), x(t+\tau), x(t+2\tau), \ldots, x(t+(d_E-1)\tau)] \quad (7)$$

This vector generates the reference trajectory. It's nearest neighbour vector

$$[x(t), x(t_0+\tau), x(t_0+2\tau), \ldots, x(t_0+(d_E-1)\tau)] \quad (8)$$

was then selected on another trajectory by searching for the point that minimizes the distance to the particular reference point. For the Rosenstein algorithm, it is imposed the additional constraint that the nearest neighbour has a temporal separation greater than the mean period of the time series defined as the reciprocal of the mean frequency of the power spectrum.

The two procedures then differed. For the Wolf algorithm, the divergence between the two vectors was computed and as the evolution time was higher than three sample intervals, a new neighbour vector was considered. This replacement restricted the use of trajectories that shrunk through a folding region of the attractor. The new vector was selected to minimize the length and angular separation with the evolved vector on the reference trajectory. This procedure was repeated until the reference trajectory has gone over the entire data sample and k_1 was estimated as:

$$\Lambda_m = \frac{1}{t_M - t_0}\sum_{k=1}^{M} \ln \frac{L'(t_k)}{L(t_{k-1})} \quad (9)$$

where $L'(t_k)$ and $L(t_{k-1})$ are the distance between the vectors at the beginning and end of a replacement step, respectively, and M is the total number of replacement steps. Note this equation uses the natural logarithm function and not the binary logarithm function as presented by Wolf [7]. This change makes Λ_m exponents more comparable between the two algorithms.

For the Rosenstein algorithm, the divergence d (t) between the two vectors was computed at each time step over the data sample. Considering that $N-(d_E-1)\tau$ embedded points (delay vectors) composed the attractor, the above procedure was repeated for all of them and Λ_m were then estimated from the slope of linear fit to the curve defined by:

$$y(t) = \frac{1}{\Delta t}\langle \ln d_j(t) \rangle \quad (10)$$

where $\langle \ln d_j(t) \rangle$ represents the mean logarithmic divergence for all pairs of nearest neighbours over time. . This process of averaging is the key to calculating accurate values of Λ_m using smaller and noisy data sets compared to Wolf algorithms.

3.4.The modified Mazhar-Eslam Lyapunov Exponent

The modified algorithm steps are same Rosenstein's algorithm steps. However, the Rosenstein's algorithm uses the Fast Fourier Transform (FFT) as shown in figure 3; the modified algorithm uses the Discrete Wavelet Transform (DWT) as shown in figure 4. That it is because DWT advantages comparing with FFT. In the next some reasons to choose DWT instead of FFT.

3.4.1. Preference for choosing DWT instead of FFT

Although the FFT has been studied extensively, there are still some desired properties that are not provided by FFT. This section discusses some points are lead to choose DWT instead of FFT. The first point is hardness of FFT algorithm pruning. When the number of input points or output points are small comparing to the length of the DWT, a special technique called pruning is often used [12]. However, it is often required that those non-zero input data are grouped together. FFT pruning algorithms does not work well when the few non-zero inputs are randomly located. In other words, sparse signal does not give rise to faster algorithm.

The other disadvantages of FFT are its speed and accuracy. All parts of FFT structure are one unit and they are in an equal importance. Thus, it is hard to decide which part of the FFT structure to omit when error occurring and the speed is crucial. In other words, the FFT is a single speed and single accuracy algorithm.

The other reason for not selecting FFT is that there is no built-in noise reduction capacity. Therefore, it is not useful to be used. According to the previous ,the DWT is better than FFT especially in Lyapunov exponent calculations when be used in HRV, because each small variant in HRV indicates the important data and information. Thus, all variants in HRV should be calculated.

3.4.2. Modified calculation of Largest Lyapunov Exponent Λ

The modified method depends on Rosenstein algorithm's strategy with replacing the FFT by DWT to estimate lag and mean period. However, the modified method use the same technique of Wolf technique for Λ_m calculating except the first two steps and the final step as they are taken from Rosenstein's method. The Lyapunov exponent (Λ) measures the degree of separation between infinitesimally close trajectories in phase space. As discussed before, the Lyapunov exponent allows determining additional invariants. The modified method of LLE (Λ) is calculated as

$$\Lambda = \sum_{i=1}^{j} \frac{\Lambda_i}{j} \quad (11)$$

Note that the $\Lambda_i s$ contain the maximum Λ and variants Λs that indicate to the helpful and important data. Therefore, the modified Lyapunov is a more sensitive prediction tool. Thus, it is robust predictor for real time, in addition to its sensitivity for all time whatever the period. It is found that the modified largest Lyapunov exponent (Λ_m) is 0.4986 for the HR signal shown in Figure 1. Thus, it is more accurate than Wolf and Rosenstein Lyapunov exponent.

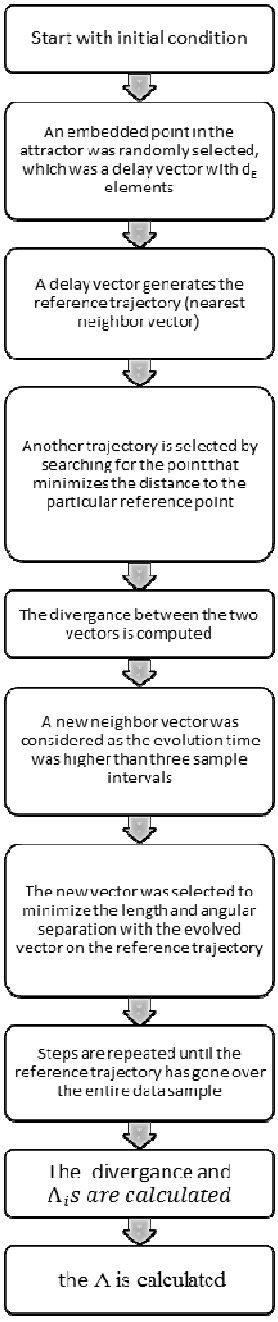


Figure 4. The flowchart of the modified algorithm.

The figure 4 shows the steps for calculating modified Mazhar-Eslam Lyapunov exponent. The modified method is the most useful and sensitive comparing to Wolf and Rosenstein methods. The table (1) discusses the different results in normal case between Modified, Wolf, and

Rosenstein methods. The Rosenstein method is the lowest sensitive method because of its quite high error comparing to the optimum. The Wolf method takes a computational place of sensitive. However, the modified method shows more sensitivity than Wolf method as the modified error is lower than Wolf as shown in figure 5. The error for each case is calculated as

$$error\ (r) = |normal\ (optimum) - case| \tag{12}$$

Thus, the accuracy for Wolf and Modified method should be calculated. The accuracy is calculated as

$$accurcy = (1 - r) \times 100\% \tag{13}$$

Figure 6 shows the accuracy of Wolf and modified method for control or normal case. It is clear that the modified Mazhar-Eslam method is more accurate than Wolf by 0.36%. This result comes because the modified takes all Λs unlike the Wolf method as it takes only the largest. Each interval of HRV needs to be monitored and taken into account because the variant in HRV indicates to another case.

Table 1. Lyapunov result for normal case shown in figure 1

Optimum	Modified method	Wolf method	Rosenstein method
0.5	0.4986	0.505	0.7586
Error	0.0014	0.005	0.2586

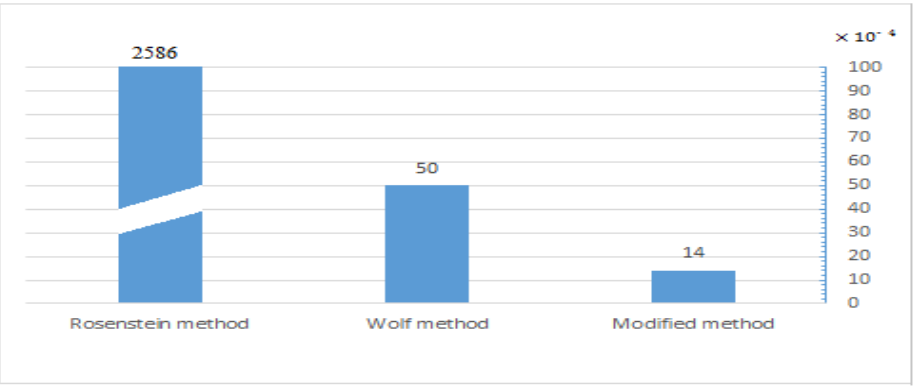


Figure 5: Methods error for the normal case.

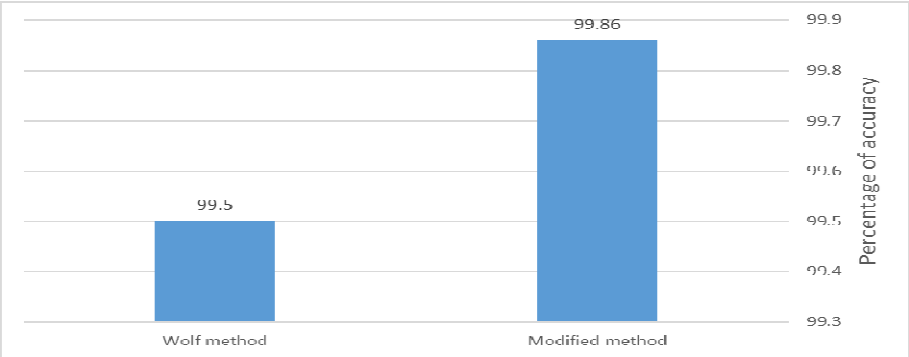


Figure 6: Accuracy percentage of Wolf and Modified method for the normal case.

4. CONCLUSIONS

Heart Rate Variability (HRV) is reported in several cardiological and non-cardiological diseases. Also, it has a prognostic value and is therefore very important in modelling the cardiac risk. HRV is chaotic or stochastic remains highly controversial. In order to have utmost importance, HRV needs a sensitive tool to analyze it like Lyapunov exponent as it is a quantitative measure of sensitivity. While the Λ_m from both Lyapunov exponent algorithms were nearly equal for small heart rate (HR) data sets. The Rosenstein algorithm provided less sensitive Λ_m estimates than the Wolf algorithm to capture differences in local dynamic stability from small gait data sets. The data supported the idea that this latter outcome results from the ability and inability of the Wolf algorithm and Rosenstein algorithm, respectively, to estimate adequately Λ_m of attractors with an important rate of convergence as those in gait. Indeed, it was found that the Wolf algorithm makes an excellent use of the attractor divergences for estimating Λ_m while the Rosenstein algorithm overlooks the attractor expansion. Therefore, the Wolf algorithm appears to be more appropriate than the Rosenstein algorithm to evaluate local dynamic stability from small gait data sets like HRV. Increase in the size of data set has been shown to make the results of the Rosenstein algorithm more suitable, although other means as increasing the sample size might have a similar effect. The modified Mazhar-Esalm method combines Wolf and Rosenstein method. It takes the same strategy of Rosenstein method for initial step to calculate the lag and mean period, but it uses Discrete Wavelet Transform (DWT) instead of Fats Fourier Transform (FFT) unlike Rosenstein. After that, it completes steps of calculating Λs as Wolf method. The modified Mazhar-Eslam method care of all variants especially the small ones like that are in HRV. These variants may contain many important data to diagnose diseases as RR interval has many variants. Thus, the modified Mazhar-Eslam method of Lyapunov exponent Λ takes all of Λs. That leads it to be robust predictor and that appear in different results between modified Mazhar-Eslam, Wolf, and Rosenstein. The Mazhar-Eslam is more accurate than Wolf and Rosenstein Lyapunov exponent. The accuracy of Wolf and modified Mazhar-Eslam method for control or normal case is more accurate than Wolf by 0.36%.

REFERENCES

[1] A.E. Aubert and D. Ramaekers. Neurocardiology: the benefits of irregularity. The basics of methodology, physiology and current clinical applications. ActaCardiologica, 54(3):107–120, 1999.

[2] Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Circulation, 93:1043–1065, 1996.

[3] U. RajendraAchrya, K. Paul Joseph, N. Kannathal, Choo Min Lim, Jasjit S. Suri. "Heart Rate Variability: a review" Med Bio EngComput (2006) 44:1031-1051

[4] P.K. Stein and R.E. Kleiger. Insights from the study of heart rate variability. Annual Review of Medicine, 50(1):249–261, 1999.

[5] Y. Gang and M. Malik. Heart rate variability analysis in general medicine. Indian Pacing and Electrophysiology Journal, 3(1):34–40, 2003.

[6] Backs, R. W. (1998). A comparison of factor analytic methods of obtaining cardiovascular autonomic components for the assessment of mental work load. [heart rate analysis]. Ergonomics. 41, 733-745.

[7] WOLF, A., SWIFT, J., SWINNEY, H., AND VASTANO, J. Determining lyapunov exponents from a time series Physica D: Nonlinear Phenomena 16, 3 (July 1985), 285–317.

[8] Acharya UR, Kannathal N, Krishnan SM (2004) Comprehensive analysis of cardiac health using heart rate signals.

[9] Acharya UR, Kannathal N, Seng OW, Ping LY, Chua T (2004) Heart rate analysis in normal subjects of various age groups. Biomed Online J USA 3(24)

[10] ROSENSTEIN, M. T., COLLINS, J. J., AND DE LUCA, C. J. A practical method for calculating largest lyapunov exponents from small data sets. Phys. D 65, 1-2 (1993), 117–134.

[11] Takens F 1981 Detecting strange attractors in turbulence Springer Lecture Notes in Mathematics vol 898, pp 366–81

[12] H.V. Sorensen and C.S. Burrus.Efficient computation of the DFT with only a subset of input or output points. IEEE Transactions on Signal Processing, 41(3): 1184-1200, March 1993.

AUTHORS

Mazhar B. Tayel was born in Alexandria, Egypt on Nov. 20th, 1939. He was graduated from Alexandria University Faculty of Engineering Electrical and Electronics department class 1963. He published many papers and books in electronics, biomedical, and measurements.
Prof. Dr. Mazhar Bassiouni Tayel had his B.Sc. with honor degree in 1963, and then he had his Ph.D. Electro-physics degree in 1970. He had this Prof. degree of elect. and communication and Biomedical Engineering and systems in 1980. Now he is Emeritus Professor since 1999.
From 1987 to 1991 he worked as a chairman, communication engineering section, EED BAU-Lebanon and from 1991 to 1995 he worked as Chairman, Communication Engineering Section, EED Alexandria. University, Alexandria Egypt, and from 1995 to 1996 he worked as a chairman, EED, Faculty of Engineering, BAU-Lebanon, and from 1996 to 1997 he worked as the dean, Faculty of Engineering, BAU - Lebanon, and from 1999 to 2009 he worked as a senior prof., Faculty of Engineering, Alexandria. University, Alexandria Egypt, finally from 2009 to now he worked as Emeritus Professor, Faculty of Engineering, Alexandria University, Alexandria Egypt. Prof. Dr. Tayel worked as a general consultant in many companies and factories also he is Member in supreme consul of Egypt. E.Prof. Mazhar Basyouni Tayel

Eslam Ibrahim ElShorbagy AlSaba was born in Alexandria, Egypt on July. 18th, 1984. He was graduated from Arab Academy for Science Technology and Maritime Transport Faculty of Engineering and Technology Electronics and Communications Engineering department class 2007. He published many papers in signal processing.
Eslam Ibrahim AlSaba had his B.Sc. with honor degree in 2007, and then he had his MSc. Electronics and Communications Engineering degree in 2010. He is a PhD. Student in Alexandria university Faculty of Engineering Electrical department Electronics and Communications Engineering section from 2011 until now. From 2011 until now, he works as a researcher in Alexandria University. In addition, now he works as a lecturer in Al Baha International college of Science, KSA.

AUTHOR INDEX