October 2015
Online Print Version

International Journal of latest Research in Engineering and Technology (IJLRET)

Print Version, Volume 01, Issue 05
October 2015 Edition
ISSN 2454-5031

IJLRET
www.ijlret.com
International Journal of latest Research in Engineering and Technology (IJLRET)

GENERAL INFORMAION:

IJLRET, International Journal of Latest Research in Engineering and Technology publish monthly journal under ISSN 2454-5031.

All the respective authors are the sole owner and responsible of published research and research papers are published after full consent of respective author or co-author(s). For any discussion on research subject or research matter, the reader should directly contact to undersigned authors.

COPYRIGHT

DISCLAIMER

Contact Information:

Editor: editor@ijlret.com
Website: http://www.ijlret.com

International Journal of latest Research in Engineering and Technology (IJLRET)

Volume 01, Issue 05
October 2015 Edition
ISSN 2454-5031

Contents

International Journal of Latest Research in Engineering and Technology (IJLRET)
ISSN: 2454-5031(Online)
www.ijlret.com‖ Volume 1 Issue 5‖October 2015 ‖ PP 01-04

DYNAMIC AND FATIGUE EVALUATION OF AN ARCH DAM SUBJECTED TO EARTHQUAKE.

LINUMOL JOSE[1], DEEPTHY RAJAGOPAL[2]
[1]*(M.Tech student, Civil engineering, Indira Gandhi Institute of Engineering & Technology for Women-Nellikuzhy,India)*
[2] *(Asst.Proffessor, Civil engineering, Indira Gandhi Institute of Engineering & Technology for Women-Nellikuzhy,India)*

ABSTRACT: **F**or a wide variety of problems in structures, linear and nonlinear solid mechanics, dynamics, and submarines structural stability problems FEM can be used, based on the development in computer technology. As to arch dam stress analysis, the conventional trial-load method has got lots of successful experience, but also has some disadvantages. For instance, it is hard to deal with complicated foundation, topography feature etc. At these aspects, the Finite Element Methods (FEM) has large advantage. Finite element method is an essential and powerful tool for solving structural problems not only in the marine field but also in the design of most industrial products and even in non-structural fields. This paper mainly focuses on the location of major stress concentration points and deflections of the dam under dynamic loads and also to evaluate the fatigue life of the structure by using ANSYS.
KEYWORDS: Double curved arch dams, Dynamic analysis, fatigue analysis, Maximum Credible Earthquake (MCE) ,

1. Introduction

Due to water scarcity, necessary to design structures which were able to regulate the water level. building dams became an appropriate method and for this reason a continuous development of dam structures began. the growing demand for electric energy gained from hydropower also promoted this progress. dams were built of soil and rock for thousands of years, until concrete became one of the most common building material. now a days, structural and topographic as well as economic reasons specify the geometry and building material a dam structure is built with. in the special topographical case of a narrow valley with stable valley sides, engineers started to design dams with a special shape and therefore special structural behavior. simultaneously they reduce the volume of the building significantly by using this new design, which is known as arch dam. Idukki dam consists of three major dams. It has been constructed across the Periyar river in a narrow gorge between two granite hills. It is 169.164m in height and 19.81m thick at the base.

Double curved arch dams feature horizontal as well as vertical arches. Therefore, the geometry of double curved arch dams is the most suitable concerning load transfer and volume reduction and can be fitted to nearly every asymmetric valley shape. The vertical sections are commonly realized with circles, whereas the horizontal arches are often realized with conic sections, basket-handle arches or logarithmic helixes. Despite the fact that double curvature dams have the most convenient structural behavior, the effort, regarding design and construction is enormous. without using modern calculation methods, such as the Finite Element Method, it would be hardly possible to investigate strains, displacements and stresses for such a complex geometry.

Stress results are used to evaluate the dam performance in the response to each loading combination. The evaluation starts with comparison of the computed stresses with strength of the concrete reduced by a factor of safety, but will also involve determination of location, magnitude, extent, and direction of high stresses should some crack-inducing stresses be expected. If all factors of safety are met the dam is considered to perform satisfactorily, even though some minor contraction joint opening may occur.

2. Load Combinations

Arch dams should be evaluated for all appropriate load combinations. Depending on their probabilities of occurrence, three basic loading combinations, Usual, Unusual, and Extreme should be considered. The usual loading combination considers the effects of all loads that may exist during the normal operation of the dam. The unusual loading combination refers to the loads acting on the dam during the flood stage. The extreme loading combination includes any of the usual loading combinations plus the effects of the Maximum Credible Earthquake (MCE). Rare loading conditions which have a remote probability of occurrence at any given time, have a negligible probability of simultaneous occurrence and should not be combined.

Dam Load Classification (Source: KSEB)

B	**Max Operation Level: 732.43m**	UNUSUAL LOAD	In combination with: Concrete Weight, Full Silt Level, Minimum Concrete Temperature, OBE of 0.06g,
C	**Normal Operation Level: 716.28m**	EXTREME LOAD	In combination with: Concrete Weight, Full Silt Level, Minimum Concrete Temperature, DBE of 0.12g,

Concrete Elastic Constants (Source: KSEB)

Load Conditions	Sustained Modulus of Elasticity	Poisson Ratio
Usual Load Conditions, D, E, F	21,000 Mpa	0.17
Unusual Load Conditions, B	28,000 Mpa	0.22
Extreme Load Conditions, A, C	31,000 Mpa	0.27

Coefficient of Thermal Expansion for concrete (Source: KSEB)

Load Conditions	Sustained Structural Conditions
Usual Load Conditions, D, E, F	1.2 x 10-5/°C
Unusual Load Conditions, B	1.2 x 10-5/°C
Extreme Load Conditions, A, C	1.2 x 10-5/°C

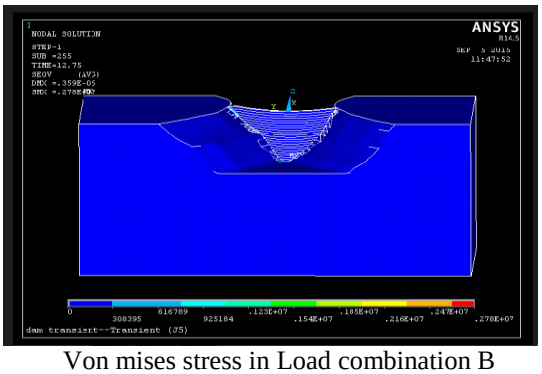
2.1. Evaluation of stress results

Evaluation of the computed stresses should begin with validation of numerical results by careful examination of the deflected shapes and stress distributions due to individual loads. Such data should be inspected for unusual deflected shapes, exceptional high or low displacement and stress magnitudes, and unexpected stress distributions that differ significantly from the results of other arch dams and cannot be explained by intuition. Force equilibrium should also be verified by comparing the sum of reaction forces to the sum of applied loads. Problems usually arise from the input data and modeling errors and should be corrected. Compressive stresses usually meet the criteria but tensile stresses caused by temperature loads, or other unfavourable situations may be significant. When significant tensile stresses are indicated, sections of the arches and cantilevers subjected to excessive tension are assumed to be cracked. This cracking will result in the re-distribution of stresses and loads. For example, localized loss of cantilever action caused by cracking at the base of the dam can be compensated by increased arch action. If cracking appears to be significant, non-linear analysis or linear analysis of the "as cracked" model may be required.

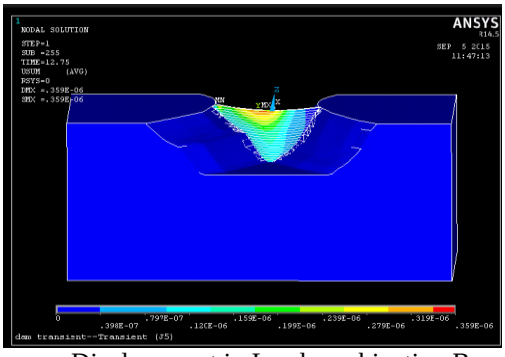
Factors of Safety for Existing Arch Dams(Source: Engineering Guidelines by FERC)

Loading Combination	Compressive Stresses
Usual	2.0
Unusual	1.5
Extreme	1.1

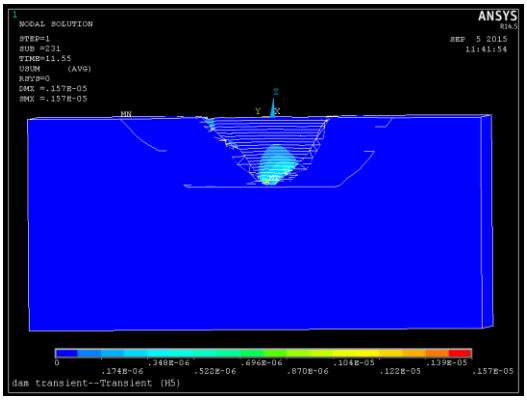
3. Results



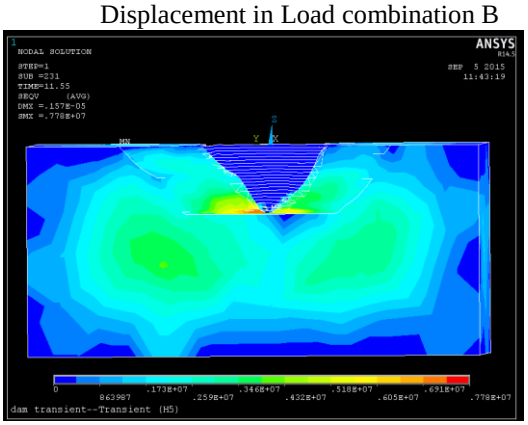
Von mises stress in Load combination B



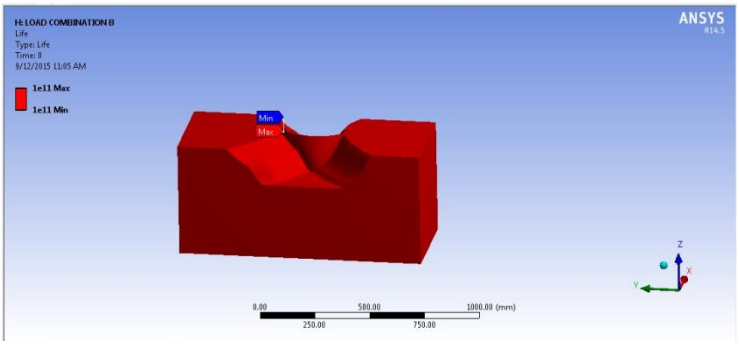
Displacement in Load combination B



Von mises stress in Load combination C



Displacement Load combination C



Fatigue analysis of an arch dam

4. Conclusion

The maximum displacement obtained in the dynamic condition is 0.77 cm. The limiting value of displacement as per KSEB norms is 5 cm. It can be observed that all the displacement values are within the prescribed limit.

The maximum stress obtained in the dynamic condition is 15.76MPa.Limiting stress as per "Engineering guidelines" by FERC is the concrete compressive strength reduced by the prescribed factor of safety. Hence, for usual load combination, limiting stress was found to be 13.25 MPa, for unusual load combination 17.67 MPa and for extreme load combination, the limiting stress was 24.09 MPa. It can be observed that the maximum compressive stress values for the proposed model have not exceeded the limiting values as per FERC.

Acknowledgements

I would like to express my gratitude to God Almighty who has given me the strength to complete this work.

My project was possible only because of the encouragement I received from all quarters. I take this opportunity to thank our Principal Dr.N.MALMURUGAN for providing me with adequate facilities for carrying out this project. I wish to express my sincere gratitude to Ms. MINU ANNA JOHNY, (Assistant Professor and Head of the Department of Civil Engineering) for her valuable encouragement to carry out this work. I express sincere thanks to M.Tech coordinator Mrs. AHSANA FATHIMA, (Assistant Professor, Department of Civil Engineering) for her valuable guidance and motivation for preparation and presentation work. Also, I have been fortunate in having Ms. DEEPTHY RAJAGOPAL (Assistant Professor, Department of Civil Engineering),

as my guide, whose timely guidance, advices and inspiration helped me in the preparation and presentation of this work in time. I express my sincere gratitude for her invaluable guidance in spite of her almost overpowering schedule. Her views on many matters have inspired me a lot.I would like to extent my special thanks to my parents and friends.

REFERENCES

[1] YANG ZHUAN-YUN, STATIC AND DYNAMIC ANALYSIS OF HIGH ARCH DAM BY THREE DIMENSIONAL FINITE ELEMENT METHOD ,EJEG 2014 VOL. 19 [2014],

[2] Binol Varghese, Arya Saju, Simi John, Finite element analysis of arch dam, International Journal of Research in Engineering and Technology Volume: 03 Issue: 07 | Jul-2014.

[3] In Ho Cho M.EERI, Virtual Earthquake Engineering Laboratory Capturing Nonlinear Shear, Localized Damage and Progressive Buckling of Bar, Earthquake Spectra, Volume 29, No. 1, pages 103–126, February 2013

[4] Ahmed Mohammed Youssef Mohammed, Mohammad Reza Okhovat, and Koichi Maekawa, Damage Evolution of Underground Structural Reinforced Concrete Small-Scale Static-Loading Experiments, World Academy of Science, Engineering and Technology Vol:6 2012-01-28

[5] B. Sevim ,The effect of material properties on the seismic performance of Arch Dams, Natural Hazards and Earth System Sciences August 2011

[6] X.Z. Lu, Y. Li, Z.W. Miao & S.C. Chen, Simulation models of concrete structures f or multi disaster mitigation, Recent Advances in Fracture Mechanics of Concrete - B. H. Oh, et al, 2010 Korea Concrete Institute,

[7] Ionuț-Ovidiu Toma And Gabriela M. Atanasiu , Modern trends in experimental earthquake engineering research, Bul. Inst. Polit. Iaşi, t. LVI (LX), f. 4, 2010

[8] S.J. Hwang, M. Saiid Saiidi, Sara Wadia- Fascetti, JoAnn Browning, Jerry P. Lynch, Kamal Tawfiq, K.C. Tsai, G. Song, and Y.L. Mo, Experiments and simulation of reinforced concrete buildings subjected to reversed cyclic loading and shake table excitation, 4th International Conference on Earthquake Engineering Taipei, Taiwan October 12-13, 2006 Paper No. 175

[9] Engineering guidelines for the evaluation of hydropower projects by Federal

International Journal of Latest Research in Engineering and Technology (IJLRET)
ISSN: 2454-5031(Online)
www.ijlret.com‖ Volume 1 Issue 5‖October 2015 ‖ PP 05-08

Wireless Medium Access Control Protocols

Pooja Sheoran[1], Priyanka[2], Vaibhav Jain[3]
[1]*Department of Electronics and Communication Engineering, Vaish College of Engineering, Rohtak, India*
[2]*Department of Electronics and Communication Engineering, Vaish College of Engineering, Rohtak, India*
[3]*Department of Electronics and Communication Engineering, Vaish College of Engineering, Rohtak, India*

Abstract- As the technology advances and telephone, cable and data networks converges into a single network, MAC (medium access control) protocols play momentous role in WSNs. MAC protocols define rules for accessing the shared medium efficiently. MAC protocols are used for energy conservation and to avoid collisions. In this paper we describe different MAC protocols for WSN. This paper presents the comparative study of different MAC protocols. It shows which protocol will be suitable for specific application and environment. The aim of this paper is to analyze the most energy efficient protocol so that MAC protocols can be categorized and their performance can be compared. There are some problems in existing MAC protocols that is in cellular systems QOS (quality of service) is not good, bandwidth efficiency issues and network synchronisation is difficult. Bluetooth is also having short range issue.

Keywords- Wireless sensor networks, energy efficient MAC protocols, CDT(continual data scheme)

I. INTRODUCTION

As the need of low power microprocessor and the development of new wireless communication techniques increases so WSNs has become the most promising technologies. Designing of WSN's main aim is to minimize energy consumption and maximizing the lifetime of the networks and this can be done by using MAC protocols. The critical issue in wireless sensor networks is energy efficiency and sources of energy loss are idle listening, collision, overhearing etc [11]. In this paper we have compared the different MAC protocols which are suitable for different applications.

II. RELATED WORK

In this section different energy efficient protocols are presented which are energy efficient. These protocols are categorized in to contention based and TDMA based protocols.

Contention based MAC protocols are based on CSMA or CSMA/CA (carrier sense multiple access/collision avoidance). These protocols assigned time slot to each node dynamically. The main concept is to transmit when channel is idle. A ready to send(RTS) and clear to send(CTS) packets are used so that transmission can be collision free.

TDMA based protocols allocate a time slot to each node for data transmission. These protocols are generally more energy efficient than contention based protocols. This MAC can achieve good performance because during the idle times radio transmitter can be turned off in order to conserve energy. As TDMA assign different time slot to each sensor so latency increases with the increase in number of sensors sharing the channel.

S-MAC: S-MAC protocol in [1] combination with energy efficient protocols that are used in WSNs. In this nodes have periodic listen-sleep schedule. And nodes turned off their transmitter during sleep period so that energy can be saved. Focus is mainly on improving the power consumption of S-MAC protocol, by using OLSR routing protocol. Improvement in S-MAC protocol results in increase of sensor nodes lifetime and in further, network connectivity and survivability. In this paper use of S-MAC protocol in combination with OLSR, AODV, DSR and DSDV protocols is evaluated with the aim to reduce energy consumption and thus increase connectivity and survivability of ad-hoc WSNs and by the experimental results it is verified that sensor node lifetimes, using the routing protocol OLSR is longer than remaining schemes using the routing protocols AODV, DSR, DSDV.

Timeout MAC [7]: As the traffic load fluctuates S-MAC protocol does not work properly so the T-MAC protocol is used to finish the time period of active node. When all the network traffic has finished, T-MAC[6] allows the node to go into sleep mode. Timeout (TA) duration is given to each node. If no transmission occur for TA duration then node switch off its radio so that energy can be saved and move into sleep state. The timeout should be more enough so that it can overcome the early sleeping problems. T-MAC also introduces flow control in data flow by using FRTS (future request to send).

D-MAC: It is also an energy efficient protocol and it uses a staggered wake up pattern to transmit data to the base station. It considered the nodes at different level of tree. The nodes which are at same level of tree would wake-up simultaneously to receive data. The nodes which are at next level would wake-up just after the lower level's receiving

period. And this data transfer forms staggered pattern and the data packet reaches from root to leaves in one cycle. It also adopt prediction method [5] when multiple child nodes want to send data to one parent in one cycle only.

I-MAC: Intelligent MAC is having intelligent sleep and wake-up schedule. It is based on CSMA/CA (carrier sense multiple access with collision avoidance) and consist of BEB (binary exponential back-off) algorithm to reduce the collisions. Each slot is assigned a time duration of seconds. When a station transmit data, it initialize its back-off counter[10]. The counter is decreamented, froze or resumed if medium is idle for seconds, busy or idle for seconds after the last transmission finishes respectively. The station transmits when the back-off counter reaches to zero. The receiver will wait for /2 seconds when there is successful reception of packets and then an ACK will be send to the transmitter. If it will not get any ACK then packet will be retransmitted. It is necessary that as soon as the transmission finished its next slot should synchronize with the slot boundary of all idle nodes.

TEEM[3]: Traffic aware energy efficient MAC protocol is contention based protocol. This protocol is similar to S-MAC. In TEEM, the durations of sleep/listen modes are adaptive by utilizing traffic information of each node. When the nodes have no data traffic to, they are put into sleep state. TEEM is more energy efficient than all other because it is having shorter and adaptive listen period and it saves more energy than S-MAC.

R-MAC[9]: Receiver driven TDMA based MAC not only eliminates the need to wake-up a receiver node by sensor node but also reduces collision among sensor node. In this the receiver node assigns its timeslot to the neighbour sender nodes and thus forms clusters of sender and receiver nodes in the network. Receiver nodes only wake-up to listen the transmission during their assigned time slot otherwise asleep for all other time slots. If any sender will not use its assigned time slot, the receiver node can go back to sleep after its timeout period of no-channel activity so that idle listening overhead may be reduced.
G-MAC [9]: Energy adaptive WSN MAC protocol i.e. Gateway MAC (G-MAC) which gives centralized function of cluster management to distribute cluster energy resources and extend network lifetime. G-MAC achieves energy saving by performing all required traffic scheduling operations while most of the nodes are sleeping in both heavy and light density traffic environments. G-MAC dynamically proposed TDMA slots according to the demand of network traffic without imposing any cluster wide message overhearing and increase the network lifetime by 250% for unicast network GMAC cluster centering functions offers significant energy savings by considering the advantages of both contention and contention-free protocol.

AEEMAC[2]: Adaptive energy efficient MAC protocol, the available energy efficient medium access control protocols for WSNs focusing on their energy conserving methods and present AEEMAC as a simple but effective energy[14] efficient MAC protocol as an optimization over SMAC. AEEMAC incorporates three additional optimizations to improve energy efficient at MAC layer are
1. Adaptive sleeping and reusing of channel
2. Use of combined 'SYNC-RTS' control packet
3. Use of 'ACK-RTS' control packet in bidirectional and multihop data transmission.

Adaptive sleeping and reusing of channel scheme and two 'combined transmission' schemes are proposed, where control messages can be piggybacked in the messages with reservation slots. Collision probability of RTS can be reduced by this scheme as it save slots resource. The simulation studies show that AEEMAC achieves better energy performance than SMAC. AEEMAC reduces energy consumption while providing good end-to-end delay, packet delivery ratio and throughput in comparison to SMAC.

ALLEE-MAC[12]: An adaptive low latency and energy efficient MAC protocol. Energy efficient is a critical issue in WSNs. To reduce energy consumption periodic listen/sleep mechanism is used. But if fixed listen/sleep scheme is used it introduces additional latency in packet delivery and energy wastage for unnecessary idle listening. This paper proposed a low latency and energy efficient MAC protocol with an adaptive listen/sleep mechanism, named ALLEE-MAC. Cross layered design approach is adopted in ALLEE-MAC for achieving two novel schemes. These are continual data scheme and early- sleep scheme. To evaluate the performance of proposed new protocol simulations have been done, through which we get ALLEE-MAC can really reduce packet end- to- end delay in case of heavy traffic and save more energy compared with S-MAC/AL.

PW-MAC[8]: Predictive wake up MAC. It is based on asynchronous duty cycle. It enables senders to predict receive wakeup time that could minimize energy consumption, although facing the challenges of OS delay, clock drift and unpredictable network. Even in the presence of wireless collisions, it achieves high energy efficiency by an efficient prediction based retransmission mechanism.
To evaluate the performance of PW-MAC with Wise MAC, R-MAC and X-MAC, experiments are conducted on a testbed of MICAZ note.

III. Comparison of Different MAC Protocols

MAC protocol/ type	PROTOCOL QUALITY	DRAWBACKS
T-MAC	It introduce the timeout window to finish the active period of a node so that energy can be saved.	The major problem is that nodes sleep in accordance to their activation time and for long messages data may get lost.
D-MAC	It adds dynamic duty cycle to decrease the latency for delay sensitive applications[6].	It does not utilize collision avoidance methods. Due to which, when nodes with same schedule transmit simultaneously, collision occurs.
I-MAC	The major quality is its intelligent sleep and wake-up procedure[15]	It is limited in the real time test bed.
TEEM	It is more energy efficient due to shorter and adaptive listen period.	Collision may occur if a node transmit RTS/CTS when neighbour node is also switching its transmitter to listen in sending mode.
R-MAC	It uses scheduling to transmit control and data packets to avoid collisions.	It is not robust in error prone transmission channel.
S-MAC	It gives good scalability and collision avoidance by combined scheduling and contention scheme.	It does not work well when traffic load fluctuates.
G-MAC	Ability to schedule heavy and light density traffic to achieve significant energy saving and eliminates network wide idle listening.	It requires delicate tradeoff in energy, latency and throughput.
AEEMAC	Improves energy efficiency by adaptive sleeping and reusing of channel, combined SYNC-RTS and ACK-RTS control packets.	The three optimization schemes used makes it more complex.
ALLEEMAC	To save more energy and short end to end delay it uses two novel schemes i.e CDT and early –sleep.	End-to-end delay of ALLEE-MAC is always a bit larger than that of SMAC when the interval of message is larger than two seconds.
PW-MAC	The main concept is for sender to wake up just before receiver is ready[11]. For this every node uses pseudo random wake up schedule.	In some applications accurate prediction is not possible.

IV. CONCLUSION

This paper gives the different MAC protocols used in WSN. TMAC introduce the timeout window to finish the active period of a node so that energy can be saved. But it is not suitable for long data length. DMAC adds dynamic duty cycle to decrease the latency for delay sensitive applications. It does not consider collision avoidance methods. Intelligent wake up and sleep schedule are used in I-MAC. In TEEM, the durations of sleep/listen modes are adaptive by utilizing traffic information of each node. RMAC uses scheduling to transmit control and data packets to avoid collisions. In SMAC nodes have periodic listen-sleep schedule. And nodes turned off their transmitter during sleep period so that energy can be saved. But it does not work when traffic fluctuates. In GMAC ability to schedule heavy and light density traffic to achieve significant energy saving and eliminates network wide idle listening. AEEMAC improves energy efficiency by adaptive sleeping and reusing of channel, combined SYNC-RTS and ACK-RTS control packets. ALLEEMAC save more energy and short end to end delay it uses two novel schemes i.e CDT and early –sleep. The main concept in PW-MAC is for sender to wake up just before receiver is ready.

This paper gives a detailed view of various protocols. In future further detailed study on protocols could be done revealing other important properties.

REFERENCES

[1] A Nikolaos A. Pantazis, Apostolos Pantazis, Stefanos A. Nikolidakis, Dimitrios D. Vergados, " A Performnce Evalution of SMAC Protocol in Combination with energy efficient Protocol for WSN", 18th International conference on Telecommunication, IEEE,2011.

[2] Alak Ray, Nityananda Sarma, " AEEMAC: Adaptive Energy Efficient MAC Protocol for WSN".

[3] Changsu Suh And Young- Bae Ko, " A traffic Aware, Energy Efficient MAC Protocol For WSN", IEEE, 2005.

[4] G M Shafiullah, Adam Thompson, Peter J Wolfs, Shawkat Ali, " Energy Efficient TDMA MAC Protocol for WSN applications", IDCS' 08,IEEE, 2008.

[5] Himanshu Singh and Bhaskar Biswas, " Comparison Of CSMA Based MAC Protocols Of WSN", IJANS Vol. 2, April 2012.

[6] Ilker Demirkol, Cem Ersoy, and Fatih Alagöz, "MAC Protocols for Wireless Sensor Network".

[7] Lamia CHAARI and Lotfi KAMOUN, " Wireless Sensor Networks MAC Protcols Analysis", Journal Of Telecommunications, Vol. 2, ISSUE 1, April 2010.

[8] Lei Tang,Yanjun Sun, Omer gurewitz, David B. Johnson, " PWMAC: An Energy Efficient Predictive-Wakeup Mac Protocol For WSN", IEEE INFOCOM, 2011.

[9] Michael I. Brownfield, Kaveh Mehrjoo, Almohonad S. Fayez, Nathaniel J. Davis IV , " Wireless Sensor Network Energy adaptive MAC Protocol", IEEE CCNC, 2006.

[10] Mohammad Masummuzzaman Bhuiyan, Iqbal Gondal and Joarder Kamruzzaman, "IMAC: Energy efficient Intelligent MAC Protocol for WSN", APCC,October 2011.

[11] Rohan Parmar, Dr. R. C. Poonia, "A Literature Survey On MAC Protocols for WSN", IJCSIT, VOL. 5 (3),2014.

[12] Wee Lum Tan, Wing Cheong Lau and Onching Yue, "An energy efficient and receiver driven MAC protocol for wireless sensor networks", IEEE, 2007.

[13] Zhenzhou Tang, Qian Hu, " ALLEE- MAC :An Adaptive low Latency and Energy efficient MAC Protocol For WSN", 6th advanced International conference On Telecommunications,2010.

[14] Zhenzhou Tang, Qian Hu,, "An Adaptive Low Latency Cross-layer MAC Protocol for Wireless Sensor Networks", the 8th international Conference on Pervasive Intelligence & Computing, Chengdu, China, Dec 2009.

International Journal of Latest Research in Engineering and Technology (IJLRET)
ISSN: 2454-5031(Online)
www.ijlret.com‖ Volume 1 Issue 5‖October 2015 ‖ PP 09-14

REVIEW ON PREDICTION SYSTEM FOR HEART DIAGNOSIS USING DATA MINING TECHNIQUES

Divya Kundra
Shaheed Udham Singh College of Engineering & Technology, Tangori, Distt Mohali.

Er. Navpreet Kaur
Shaheed Udham Singh College of Engineering & Technology, Tangori, Distt Mohali.

ABSTRACT: Data mining is the process of analyzing large sets of data and then extracting the meaning of the data. It helps in predicting future trends and patterns, allowing business in decision making. Data mining applications can answer business questions that take much time to resolve traditionally. Large amount of data which is generated for the prediction of heart disease is analyzed traditionally and is too complicated and voluminous to be processed. Data mining provides the techniques and methods for the transformation of data into useful information for decision making. These techniques make the process fast and it takes less time for the prediction system to predict the heart disease with more accuracy. In this paper we survey different papers in which one or more algorithms of data mining used for the prediction of heart disease. Result from using neural networks is 99.62% in one paper [6] . By Applying data mining techniques to heart disease data which needs to be processed, we can get effective results and achieve reliable performance which will help in decision making in healthcare industry. It will help the medical practitioners to diagnose the disease in less time and predict the probable complications well in advance. Identifying the major risk factors of Heart Disease categorizing the risk factors in an order which causes damages to the heart such as high blood cholesterol, diabetes, smoking, poor diet, obesity, hyper tension, stress, etc. Data mining functions and techniques are used to identify the level of risk factors to help the patients in taking precautions in advance to save their life.

Keywords: heart disease ,data mining, data mining techniques, data mining tools, data mining applications, methodology

INTRODUCTION

Data mining is the analytical process to explore specific data from large volume of data. It is a process that finds previously unknown patterns and trends in databases. This information is further used to build predictive models. The main objective of our paper is to learn the different data mining techniques which are used in the prediction of heart diseases using any data mining tool. Heart is the most vital part of the human body as life is dependent on efficient working of heart. A Heart disease is caused due to narrowing or blockage of coronary arteries. This is caused by the deposition of fat on the inner walls of the arteries and also due to build up cholesterol. Heart diseases can be caused due to number of factors:

High blood pressure: when the heart pumps blood, the force of the blood pushes against the walls of the arteries causing pressure. If the pressure rises and stays high over the time it is called high blood pressure or hypertension which can harm the body in many waysi.e. Increasing the risk of heart stroke or developing heart failure, kidney failure etc.

High cholesterol: cholesterol is a waxy substance found in the fatty deposits in the blood vessels. Increase in the fatty deposits (high cholesterol) does not allow sufficient blood to flow in through the arteries causing heart attacks.

Unhealthy diet: eating too much fast food increases blood pressure and cholesterol level causing the risk of heart attacks.

Smoking: it damages the lining of arteries and builds up a fatty material called atheroma which narrows the arteries causing heart attacks.

Lack of physical activity: lack of exercise increases cholesterol level in blood vessels which further increases the risk of heart attacks.

Obesity: obese people are more likely to have high blood pressure, high cholesterol level and diabetes (increase in blood sugar level) which increases the risk of heart strokes in human body.Nowadays, data mining is gaining popularity in health care industry as this industry generates large amount of complex data about hospital resources, medicines, medical devices, patients, disease diagnosis etc. This complex data needs to be processed and analysed for knowledge extraction which will further help in decision making and is also cost effective.

World health organisation has estimated 17.5 million people died from cardio vascular diseases in 2012, representing 31 percent of all global deaths. Out of these, an estimated 7.4 million were due to coronary heart disease and 6.7 million were due to stroke. WHO estimated by 2030, almost 23.6 million people will die due to heart disease as written in [1].

Thus, a beneficial way to predict heart diseases in health care industry is an effective and efficient heart disease prediction system. This system will find human interpretable patterns and will determine trends in patient records to improve health care.

LITERATURE SURVEY

Over the years, numerous works have been done related to heart disease prediction system using different data mining algorithms by different authors. They tried to achieve efficient methods and accuracy in finding out diseases related to heart by their work including datasets and different algorithms along with the experimental results and future work that can be done on the system to achieve more efficient results. This paper aims at analyzing different data mining techniques that has been introduced in recent years for heart disease prediction system by different authors.

M.A.Nishara Banu and B.Gomathy[2] used C4.5 algorithm, MAFIA and K means clustering in the year 2014 using 13 attributes in the dataset achieving 89 percent accuracy.

Aqueel Ahmed et al. [3] show the classification techniques in data mining and show the performance of classification among them. In this classification accuracy among these data mining has discussed. In this decision tree and SVM perform classification more accurately than the other methods and was able to achieve 91% accuracy

Ms.Ishtake et al. [4] developed a prediction system for heart diagnosis using decision tree, Neural Network and Naive Bayes techniques using 15 attributes in the year 2013.

Chitra R.et al. [5] developed the computer aided heart disease prediction system that helps the physician as a tool for heart disease diagnosis. From the analysis it is concluded that neural network with offline training is good for disease prediction in early stage an d good performance can be obtained by pre-processed and normalized dataset.

Nidhi Bhatlaet al. [6] projected the study of different data mining techniques that can be employed in automated heart disease prediction systems. The analysis shows that Neural network with 15 attributes has shown the highest accuracy. On the other hand, Decision tree has also performed well with 99.62% accuracy by using 15 attributes.

Shadab et al. [7] used Naive Bayes technique in the year 2012 using 15 attributes in the dataset for the heart diagnosis in heart prediction system.

COMPARATIVE STUDY OF DATA MINING TRENDS FROM PAST TO FUTURE

PAST:- in the previous years, statistical and Machine learning techniques were used on numerical data stored in traditional databases and the computing resources were 4G PL and various related techniques[8].

PRESENT: - these days, along with the statistical and Machine learning techniques, artificial intelligence and pattern reorganisation techniques are also used [8].

FUTURE:-in future, for complex data objects which includes high dimensional, high speed data streams, sequence, noise in the time series and for multi instance objects, soft computing techniques like fuzzy logic, neural networks and genetic programming is used. Computing resources used would be multi-agent technologies and cloud computing [8].

DIFFERENT DATA MINING TOOLS USED IN HEART DISEASE PREDICTION SYSTEM WITH ACCURACY

Abhishek et al in the year 2013 used data mining tool Weka 3.6.4 in heart disease prediction system usingJ48 technique achieved 95.56% accuracy and using Naive Bayes achieved 92.42%. [9]

Rashedur et al in the year 2013 used Neural network technique using Weka data mining tool and achieved 79.19% and to compare various classification techniques, he used another technique fuzzy logic using TANGRA data mining tool and achieved 83.85% accuracy. [10]

Nidhi et al in the year 2012 used data mining tool Weka 3.6.6 in the analysis of heart disease prediction system and achieved 99.52% using Naive Bayes. She also used TANGRA data mining tool but could achieve up to 52.33% only using decision trees. She also tried .NET data mining tool and achieved up to 96.5% using neural networks. [6]

Resul et al in the year 2009 used SAS base software 9.1.3 achieving 97.4% using neural networks. [11]

DATA MINING TECHNIQUES ALSO USED IN DIAGNOSIS OF OTHER DISEASES

Humar et al in the year 2008 used classification, Backpropagation, Fuzzy neural network techniques for diabetes and heart diseases. [12]

Marcel et al in the year2007 used in Bayesian classification for Characinoid heart disease. [13]

Mohammad et al in the year 2012 used C4.5 and C5.0 algorithm for heart disease and breast cancer diagnosis. [14]

M.Akhil et al in the year 2012 used associative classification and genetic algorithm for the diagnosis of breast cancer, Pima Indian Diabetes and heart disease. [15]

DATA MINING AND ITS TECHNIQUES

DATA MINING

It is main concerned with extracting useful information from large amount of databases. Data mining techniques and tools are used to find unknown patterns and trends from the data set. Its main objective is to automatically find the patterns in the dataset with minimal user effort and input. Data mining's main contribution is in decision making and in forecasting future trends of market. Many organisations use data mining as a tool these days for data analysis as it easily evaluates patterns and trends of market and produce effective results.

DATA MINING TECHNIQUES

ASSOCIATION: it is the best known and well researched method for data mining. Association is also called relation technique because patterns which are discovered from the dataset are based on the relationship between the items. For example, when association technique is used in heart disease prediction system, it tells us the relationship between all the attributes and sort out all the patients with all the risk factors which are required for the heart disease predictions.

CLASSIFICATION: it is a data mining technique which is used to classify each item in a data set into one of predefined set of classes or groups. It is a classic data mining technique which is based on machine learning.

CLUSTERING: It is a data mining technique which creates useful cluster of objects that have similar characteristics using automatic technique. There is a slight difference between clustering and classification. Clustering defines classes and put objects in them while classification assigns objects into predefined classes. Clustering helps to make clusters or list of patients having same risk factor.

PREDICTION: it is a data mining technique which discovers relationship between independent variables and relationship between dependent and independent variables.

SEQUENTIAL PATTERNS: it is a data mining technique that discovers similar patterns or regular events in transaction data over a business period.

DECISION TREE: it is the most used data mining techniques and its model is easily understandable. The root of the decision tree is a simple question or condition that has multiple answers. Each answer leads to a set of questions or conditions which helps to determine the data so that we can take a final decision based on it.

DATA MINING APPLICATIONS

Data mining is used in various fields such as retail industry, telecommunication industry, healthcare industry, financial data analysis, intrusion detection, sports and also in analyzing student's performance.

RETAIL INDUSTRY: data mining is a great application in retail industry as it collects large amount of data which includes transportation, sales and consumption of goods and services. This data expand rapidly due to increase in purchase and sales in business. Data mining helps to identify customer's buying patterns and trends that lead to improved quality of customer service and customer's satisfaction.

TELECOMMUNICATION INDUSTRY:

Telecommunication industry is the most growing industry as it provides various services such as fax, pager, cellular phones and e-mails. With the development of computer, telecommunication services have integrated with the communication technologies and works more effectively. Data mining helps to identify telecommunication patterns, fraud activities, make better use of resources and improve quality of service.

HEALTHCARE INDUSTRY: Data mining is very useful in healthcare industry in diagnosis of heart diseases, breast cancer and diabetes. It helps in identifying patterns and trends in patient's records having same risk factor and helps in decision making.

FINANCIAL DATA ANALYSIS: financial data in banking is reliable and of high quality which facilitates systematic data analysis in financial industry. It helps in loan payment prediction and customer credit policy analysis. It also helps in clustering of customers for target marketing.

INTRUSION DETECTION: intrusion is any kind of action that threatens the confidentiality or integrity of network resources from any outside party. With the increased usage of internet and availability of the tools and tricks of intrusion and attacking network, intrusion detection has become an important issue for network administration. Data mining helps in the development of data mining algorithm for intrusion detection and analysis of stream data so that intrusion threats can be avoided.

SPORTS: in sports, vast amount of statistics are gathered for each player, team, game and season. Data mining is used in the prediction of performance of players, selection of players and forecast of future events.

STUDENT'S PERFORMANCE: data mining is used to evaluate student's performance using classification technique for data classification. Attendance, class test, seminar and assignment marks are collected from the student record to predict the performance of the student at the end of the semester.

DATA MINING TOOLS:
There are various data mining tools used for data mining purpose. These are WEKA, TANAGRA, MATLAB and .NET FRAMEWORK.

WEKA: it is a data mining tool which was developed in New Zealand by the University of Waikato that implements data mining algorithms using JAVA language. WEKA is a collection of machine learning algorithms and their application to the data mining problems. These algorithms are directly applied to the dataset. WEKA supports data file in ARFF format. WEKA is open source software and hence, it is not dependent on any platform. It includes algorithms for data processing, classification, regression, clustering, association and also visualization tools.

TANAGRA: it is open source software as researchers can access to the source code and add their own algorithms and compare their performances, if it conforms to the software distribution license. It includes several data mining algorithms from statistical learning, machine learning, data analysis and database area.

MATLAB: It is a data mining tool built in high level language. It provides interactive environment for visualization, numerical computation and programming. The built in math functions, language and tool explore various approaches and helps to reach a solution faster than with the spreadsheet of traditional programming languages like C,C++ and JAVA. It analyse data, develop algorithms, and create models and applications.

.NET FRAMEWORK: it is a software framework developed by Microsoft which runs primarily on Microsoft windows. It provides secure communication and consistent applications. it provides language interoperability(each language can code written in other languages) across several programming languages.

DATA MINING METHODOLOGY:

data mining is the fundamental part of knowledge discovery database. Knowledge discovery database is the process of discovering hidden knowledge from massive amounts of data that we are technically capable of gathering and storing. It is a process which contains sequence of following steps:

PROBLEM DEFINITION: in this, we define the problem that is to be solved and the noise or the irrelevant data is removed.

DATA INTEGERATION: here multiple data sources are combined to gather and prepare information.

MODEL BUILDING: here the data is retrieved from the database which is relevant to the analysis task and is transformed into the form which is appropriate for mining by performing certain operations.

DATA MINING: in this, different algorithms are applied to extract patterns.

KNOWLEDGE DEPLOYMENT: knowledge representation techniques are used to present mined knowledge to the user.

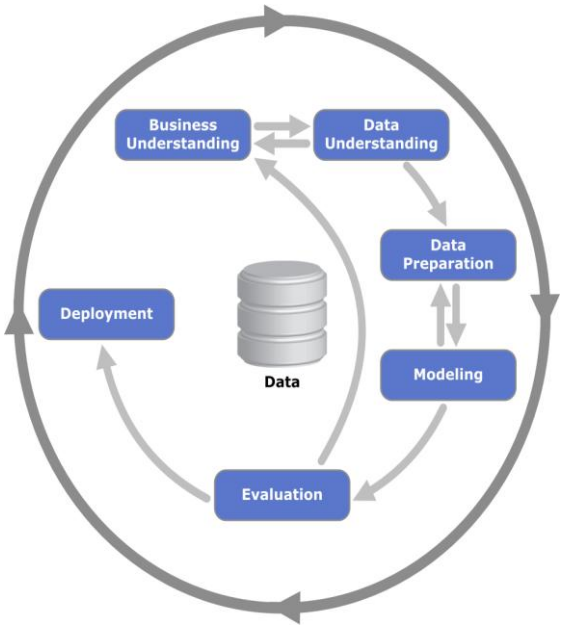


Fig 1: concept of data mining

CONCLUSION

Our goal is to analyse different data mining techniques in a way that they help us to predict and employ the heart disease prediction system which predicts the possibility of heart disease risk of patients for next ten years. To spread awareness amongst patients to take precautionary actions and hence live longer; and to assist doctors to diagnose and predict the probable complications well in advance. As the heart disease patients are increasing world-wide each year and huge amounts of data is available for research, researchers are using data mining techniques in the diagnosis of heart disease. Analysis presented by different researcher's shows that different data mining techniques and classifiers are defined in this work which has emerged in recent years for efficient and effective heart disease diagnosis. The analysis shows that using different techniques and taking different number of attributes we get different accuracies for predicting heart diseases. Taking analysis from some papers published by the researcher's it is shown that techniques such as neural networks give an accuracy of 100% in predicting heart diseases where as some may show that using the technique of Decision Tree we can get an accuracy of 99.62%. So, different techniques used have shown different accuracies depending upon number of attributes taken and tool used for implementation.

REFERENCES

[1]. Ms. Chaitrali S. Dangare, Dr. Mrs. Sulabha S. Apte, "A data mining approach for prediction of heart disease using neural networks, international journal of computer engineering and technology",2012

[2]. M.A.Nishara Banu and B.Gomathy," Disease Forecasting System Using Data Mining Methods",2014

[3]. Aqueel Ahmed, Shaikh Abdul Hannan,"Data Mining Techniques to Find Out Heart Diseases", International Journal of Innovative Technology and Exploring Engineering(IJITEE)ISSN: 2278-3075, Volume-1,Issue-4,September 2012.

[4]. Ms. Ishtake S.H, Prof. Sanap S.A., "Intelligent Heart Disease Prediction System Using Data Mining Techniques", International J. of Healthcare & Biomedical Research, 2013

[5]. Chitra R and Seenivasagam V, "REVIEW OF HEART DISEASE PREDICTION SYSTEM USING DATA MINING AND HYBRID INTELLIGENT TECHNIQUES", ISSN: 2229-6956(ONLINE) ICTACT JOURNAL ON SOFT COMPUTING, JULY 2013, VOLUME: 03, ISSUE: 04, 2013

[6]. Nidhi Bhatla and Kiran Jyoti, "An Analysis of Heart Disease Prediction using Different Data Mining Techniques", International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181,Vol. 1 Issue 8, October – 2012

[7]. Shadab Adam Pattekari and Asma Parveen, prediction system for heart disease using naïve bayes, International Journal of Advanced Computer and Mathematical Sciences, 2012

[8]. Venkatadri.M, Dr. Lokanatha C. Reddy a review on data mining from past to the future. International Journal of Computer Applications, 2011.

[9]. Abhishek taneja, Heart Disease Prediction System Using Data Mining Techniques, Oriental Scientific Publishing Co., India, 2013.

[10]. Rashedur M. Rahman, Farhana Afroz, Comparison of Various Classification Techniques Using Different Data Mining Tools for Diabetes Diagnosis, Journal of Software Engineering and Applications, 2013.

[11]. Nidhi Bhatla Kiran Jyoti, An Analysis of Heart Disease Prediction using Different Data Mining Techniques, International Journal of Engineering Research & Technology (IJERT), 2012.

[12]. Humar Kahramanli, Novruz Allahverdi, Design of a hybrid system for the diabetes and heart diseases, Elsevier, 2008.

[13]. Marcel A.J. van Gerven, Predicting carcinoid heart disease with the noisy-threshold classifier, Elsevier, 2007.

[14]. Mohammad Taha Khan, Dr. Shamimul Qamar and Laurent F. Massin, A Prototype of Cancer/Heart Disease Prediction Model Using Data Mining, International Journal of Applied Engineering Research, 2012

[15]. M.Akhil jabbar, Dr.Priti Chandra, Dr.B.L Deekshatulu, Heart Disease Prediction System using Associative Classification and Genetic Algorithm, International Conference on Emerging Trends in Electrical, Electronics and Communication Technologies, 2012