

POWER QUALITY IMPROVEMENT USING ACTIVE FILTERS: A REVIEW

Annu Govind¹, Omendra Govind²

¹ Research Scholar, Department of Electrical Engineering, Jamia Millia Islamia, New Delhi, (India)

² Advanced Level Telecom Training Centre (AL TTC), BSNL, Ghaziabad, (India)

ABSTRACT

This paper covers review of the various configurations and control strategies available for active filters. Various configurations of APFs have been proposed to compensate aforementioned power quality problems. Complexity of the control algorithm can also be reduced with the involvement of artificial intelligence like fuzzy logic, neural network etc. with the conventional approach to resolve the poor power quality problems not only in steady-state but also in dynamic/transients situation.

Keywords: Active Power Filter (APF), Harmonics, Power Quality, Nonlinear, Reactive Power Compensation.

I. INTRODUCTION

The widespread use of non-linear loads in industries and extensive proliferation of energy-efficient power electronic based equipment's have led to many power quality problems in electrical power system and becoming a great concern for utilities and customers. These Non-linear loads draw harmonics and reactive power (VARs) components of current from the utility. In three-phase systems, they could also induce imbalance and draw excessive neutral current. All these issues create serious problems for power quality.

The effective compensation of harmonics, reactive power, neutral current and supply current balancing with other power quality improvement are essential for the utilities as well as the end users. This attracted power electronics and power system engineers to develop dynamic and adjustable solutions to the power quality problems by various custom power devices like active power filter (APF), hybrid filter, unified power quality compensator (UPQC) etc. Many control algorithms with new / improved designs are presented by various researchers in context to this. In this paper a review on the various topologies and control strategies of active filters is presented.

This paper is presented in five sections. First section introduces the power quality issues followed by sections on technologies used in APFs to resolve them, classification of APFs based on converter type, topologies and number of phases and various control strategies. At last, conclusion along with future scope is presented.

II. ACTIVE POWER FILTERS

Many passive and active harmonic filters have been investigated to satisfy the power quality problems. Passive filtering has been preferred for harmonic compensation in the electrical system due to low cost, simplicity, reliability, and control less operation. Active power filters (APF) have many advantages over the passive filters.

They can suppress not only the supply current harmonics, but also the reactive currents and without causing harmful resonances with the power distribution systems like passive filters.

In the beginning, the APFs were used for suppression of harmonics generated by thyristor based converters and inverters used in HVDC transmission system. However, the design could not become technologically and economically practicable until the last two decades when fast and cost effective semiconductor devices such as Insulated Gate Bipolar Transistors (IGBTs) and Metal Oxide Semiconductor Field Effect Transistors (MOSFETs), and high performance and cost effective Digital Signal Processors (DSPs) became available. Modern active filters are superior in filtering performance since they inject voltage / current harmonics produced by nonlinear loads of same magnitude but of opposite sign so they cancel each other and sinusoidal waveforms are obtained at the power line. Depending on the APF type, controllable reactive power compensation for power factor correction, voltage regulation, load balancing, voltage-flicker reduction, harmonic damping, harmonic isolation and / or their combinations could be provided.

Now a day's improved control strategies for AF with design of high efficiency and large capacity converters has been developed for three-phase three- or four- wire power circuit. These wide ranges of objectives are achieved either individually or in combination, depending upon the requirements and control strategy and configuration which have to be selected appropriately.

With the revolution in microelectronics, microprocessors, microcontrollers, DSP technology and with the advent of fast self-commutating solid-state devices like IGBTs and sensor technologies have enhanced performance of active filter. This paper presents various types of active filters and their classification based on topologies, control strategies and number of phases and brief explanation of mainly used techniques.

III. CLASSIFICATION OF ACTIVE POWER FILTERS

Active Filters can be classified based on converter type, topologies and number of phases. On converter based CSI or VSI bridge structure. On topology based shunt, series, hybrid and UPQC. On the bases of no. of phases two-wire (single phase) and three- or four-wire three-phase systems.

3.1 Converter Based Classification

In general, there are two types of converters used in active filters - CSI or VSI bridge structure. A current source PWM converter (CSI) is equipped with a dc inductor and a voltage source PWM converter (VSI) is equipped with a dc capacitor. Fig. 1. shows CSI bridge structure.

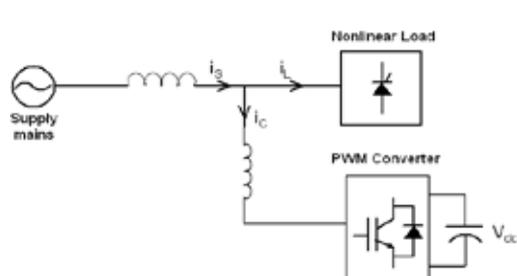


Fig. 1 Voltage Source PWM Converter Based Shunt Active Power Filter

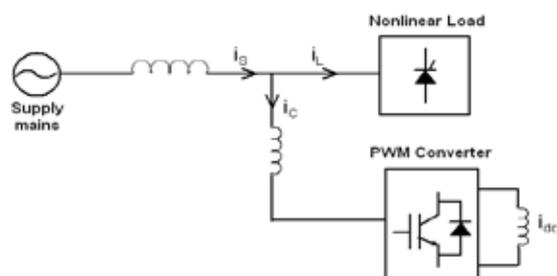


Fig. 2 Current Source PWM Converter Based Shunt Active Power Filter

Fig.2. shows VSI based filter that has a self-supporting dc voltage bus with a DC capacitor. It becomes dominant over CSI because of its high efficiency, lightweight, low cost and expandability to multilevel and multi-step versions.

3.2 Topologies Based Classification

Based on the topology, AF can be classified as series, shunt, hybrid active filters and UPQC. The appropriate topology is used as per the compensation required by the active filter. Parallel or shunt APF is the fundamental system configurations and it has been used in of three-phase three- or four- wire. Fig. 1 and 2 shows shunt APF, which consists of a controllable voltage and current source respectively. It is mainly used to eliminate current harmonics, reactive power compensation and balancing unbalanced input currents. Shunt active filters carries only the compensation current plus a small amount of active fundamental current which is supplied to compensate for system losses. This cancels harmonics and/or reactive components of the nonlinear load current at the point of common coupling (PCC). When it is employed to three-phase four-wire systems, it also compensates the neutral current (zero sequence current) component.

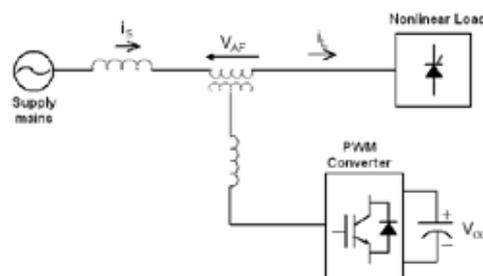


Fig. 3 Principle configuration of a VSI based series active power filter

Series active filter produces a PWM voltage waveform which is subtracted / added, on an instantaneous basis, from / to the supply mains voltage to maintain a pure sinusoidal voltage waveform across the load as shown in Fig. 3. It is similar to shunt APF, except that the interfacing inductor of shunt APF is replaced with the interfacing transformer. Its advantage over shunt active filters is that they are superlative for eliminating voltage-waveform harmonics, and for balancing three-phase voltages

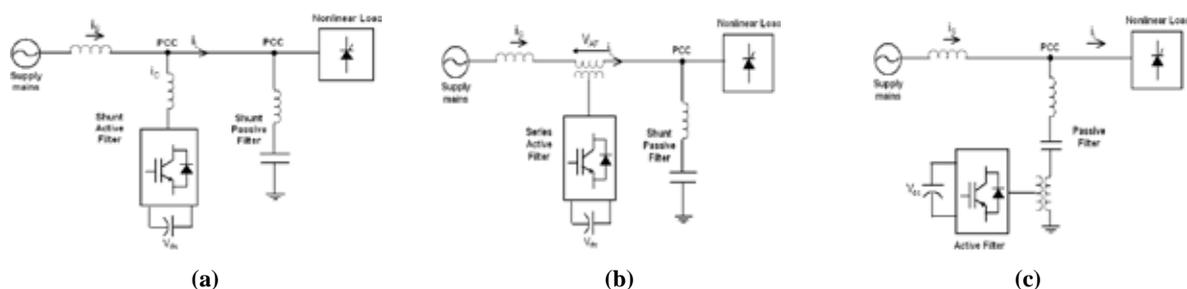


Fig. 4 Hybrid Active Filter In Combination With

- (a) Shunt Active and Shunt Passive Filter, (b) Series Active and Shunt Passive Filter, And (c) Active Filter Connected In Series With Shunt Passive Filter.**

Hybrid active filter configurations are typically the combination of basic active shunt / series and a passive filter. Hybrid APFs inherits the advantages of both active and passive filters. The active filter cancels the lower order harmonics, while the passive filter is responsible for higher order harmonics. Therefore, the main objective of hybrid active filter is to improve the filtering performance of high-order harmonics while providing

a cost-effective low order harmonics mitigation. Most prominent hybrid APFs are shown in Fig. 4 (a), (b) and (c).

UPQC, also known as universal AF, is a combination of active series and active shunt filters. Fig. 5 shows basic schematic of UPQC the dc-link storage element (either inductor or dc-bus capacitor) is shared between two current source or voltage-source bridges operating as active series and active shunt compensators.

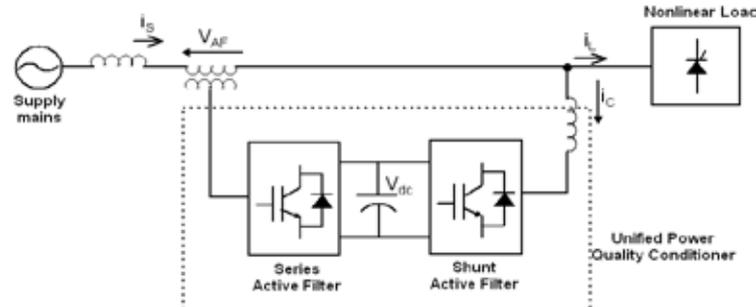


Fig. 5 Schematic diagram of Unified Power Quality Conditioner

Depending upon the nature of supply and/or the load system used, APFs are classified as single phase APFs or three phase APFs (three-phase three-wire and three-phase four-wire).

From various topologies summarized so far, the appropriate type of APFs should be installed at customer premises and utilities points for power quality improvements.

IV. CONTROL STRATEGIES

In APF design and control, calculation of the compensation current and reference signal generation is the main task for achieving the goal of power quality. The accurate estimation of reference signal depends on the control strategies used for the implementation of APF design. Control scheme is implemented in three stages.

- Sensing and conditioning of control signals
- Estimation of compensating signal
- Generation of switching signal for switching devices

The whole control scheme can be realized using analog, digital and advance programmable devices like single-chip microprocessor, digital signal processors (DSPs), field programmable gate array (FPGA), etc.

4.1 Sensing and Conditioning of Control Signals

AC mains voltage, DC link voltage, load currents, supply currents, compensator current, etc. are sensed by using proper voltage and current sensors circuitry. Voltage signals are sensed using either potential transformers (PT) or different Hall-effect voltage sensors or isolation amplifiers. Current signals are sensed using current transformers (CT) and/or Hall-effect current sensors. Sensed signals are also used to monitor, measure, and record various performance indices, such as THD, power factor, active and reactive power, crest factor, etc. Signal conditioning is required to make sensed signal compactable or readable by the platform on which control scheme has to implement. The analog or digital filters with low-pass, high-pass or band-pass characteristics are used for signal conditioning.

4.2 Estimation of Compensating Signals

The important aspect of the compensation process is the designing of the control algorithm for the APFs to generate required reference and later compensating signal. The control algorithm should be simple, robust, and

accurate, and it should give its best performance not only in ideal voltage condition but also in distorted and/or distorted and unbalanced voltage condition which are normally present in the electrical distribution system. Proper estimation of compensating signal for achieving particular compensation objectives also affects the steady-state and transient performance of the APFs. Broadly, control strategies for compensating signals generation are based on frequency-domain or time-domain techniques. Fig.6. illustrates the general classification of available techniques.

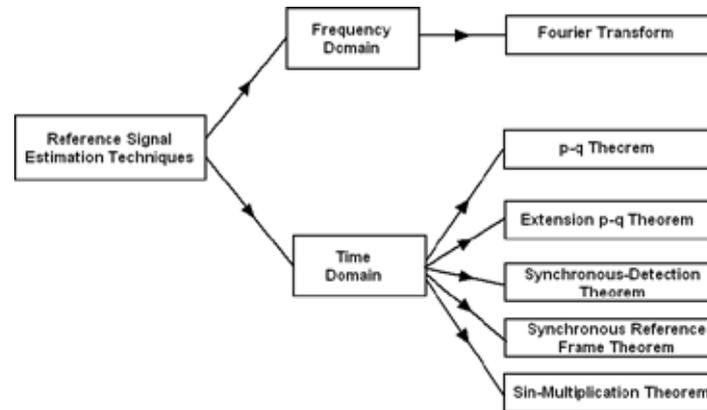


Fig. 6 Classification of Reference Signal Estimation Techniques

Both frequency-domain and time domain techniques have been used with VSI and CSI PWM converters. In recent research new methods based on wavelet transform and Artificial Intelligence (AI) are also reported for extraction of reference signal.

4.2.1 Frequency Domain Approach

In this, reference signal estimation in frequency-domain is suitable for both single and three phase systems. The frequency domain methods are based on Fourier analysis method of discrete signals such as Discrete Fourier Transform (DFT, and Fast Fourier Transform (FFT) and periodicity of distorted voltage and/or current waveforms to be corrected. Its major drawback is the requirement of a window function to analyze the frequency spectrum of the signal; this method also suffers from large memory requirement and large computational power for processor used. To eliminate higher harmonics number of calculations increases, which results in large response time and also not suitable for dynamically varying loads. Although the modified Fourier transform based methods adopting the sliding window show an improved dynamic response. Recent development in the processor technology is helping in reducing the computational time.

4.2.2 Time Domain Approach

This approach is based on instantaneous estimation of reference signal in the form of either voltage or current signal from distorted and harmonic-polluted voltage and current signals. The main advantages of the time domain compensation is its fast response, easy implementation and less computational burden. There is a large number of control methods in the time domain, which are known as instantaneous “p-q” theory, synchronous d-q reference frame method, synchronous detection method, flux-based controller, notch filter method, P-I controller, sliding mode controller, etc.

4.2.3 Control Based on Artificial Intelligence (AI)

AI is a technology to extract information from the process signal by using expert knowledge. Artificial intelligence is popular due to its ability to handle nonlinearity, complex problem without much information about the mathematical model of the system. It can identify the model, if necessary, and give the predicted

performance even with a wide range of parameter variation. Different tools of AI such as fuzzy logic, artificial neural network, genetic algorithm, wavelet theory, etc. are used in various applications of power quality and power electronics for improving the robustness and performance of control algorithm

4.3 Generation of Gating Signal to Control Switches

Once the compensation signals are obtained based on the appropriate control scheme, the next stage of APF control is the generation of switching signals for the switching devices of the PWM converter. These switching signals are obtained by comparing the reference compensating current signals with the actual current in a controller. The switching patterns decide the required compensation of current harmonics. There are different control techniques for generation of gating signals. The performance of an APF is significantly affected by the selection of control techniques.

In linear control technique, voltage or current PWM, sinusoidal internal model control, ramp comparison control, etc. are used for obtaining the PWM signals. Nonlinear current control techniques include hysteresis control and SVM. In this section, a brief description of mostly used PWM techniques and their features have been presented.

4.3.1 Voltage or Current PWM Technique

Linear control technique of switching pulse generation for APF semiconductor switches is accomplished by using a negative feedback system as shown in Fig. 8.

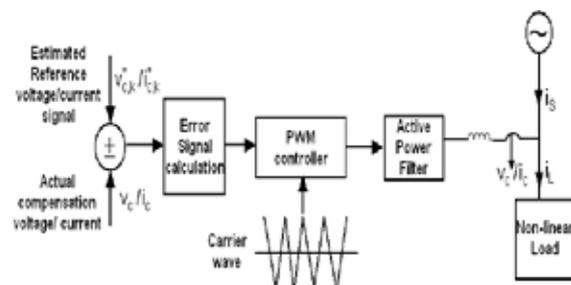


Fig. 8 Block Diagram of Linear Control Technique Linear Control

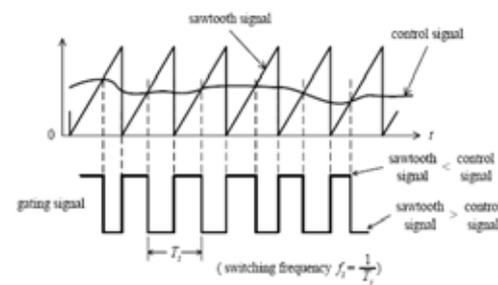


Fig. 9 Generation of Gate Signals Using Technique

voltage v_c signal is compared with its estimated reference signal ($i_{c,ref}$ and $v_{c,ref}$) through the Compensated error amplifier to produce the control signal. The resulting control signal is then compared with a fixed frequency carrier (triangular or ramp) signal. The frequency of the repetitive carrier signal establishes the switching frequency. This frequency is kept constant in linear control technique. As shown in Fig. 9, the gating signal is set high when the control signal has a higher numerical value than the carrier signal and via versa.

4.3.2 Hysteresis Control Technique

This technique forces the compensating current i_c or voltage v_c signal to follow its estimated reference signal ($i_{c,ref}$ or $v_{c,ref}$) within a specified tolerance band, known as hysteresis-band. Switching occurs whenever the error leaves the tolerance band. A basic block diagram of the hysteresis band current control is shown in Fig. 10. H is the hysteresis-band. The APF is switched on in such a way that the peak-to-peak compensation current/voltage signal is limited to a specified band (H) as illustrated in Fig. 11.

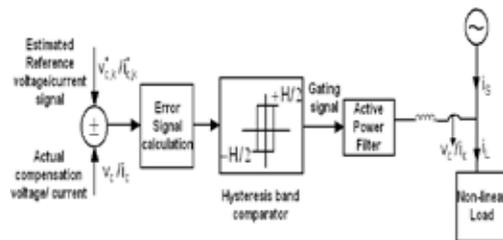


Fig. 10 Block Diagram of Hysteresis Control Technique

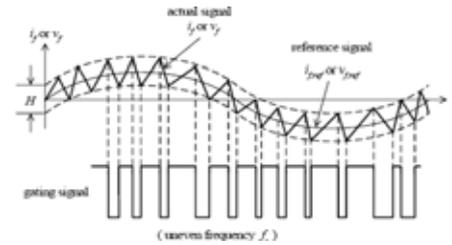


Fig. 11 Gating Signal Generation By Hysteresis Controller

The advantages of using the hysteresis current controller are its excellent dynamic performance and controllability of peak-to-peak current ripple within a specified hysteresis band. The main drawback of this technique is that it produces uneven switching frequency, which affects the APF efficiency and reliability.

4.3.3 Space Vector Modulation (SVM)

This method has several benefits like better voltage utilization, lesser current harmonics, and fixed frequency operation. Although implementation of SVM in digital system is simple, the required calculations and corresponding execution time limits the maximum sampling time and resulting maximum switching frequency and maximum bandwidth. It is reported that hysteresis current control technique is found superior and its performance is almost unaltered by the variation in the firing angle. In direct current control technique APF currents are used as reference compensating current ($\hat{i}_{ca}^*, \hat{i}_{cb}^*, \hat{i}_{cc}^*$) which further compared with actual filter current ($\hat{i}_{ca}, \hat{i}_{cb}, \hat{i}_{cc}$). Reference compensating current are obtained by subtracting load current from the reference supply current and therefore need more number of current sensors. In indirect current control technique the switching signals are obtained by the comparison of estimated reference source current ($\hat{i}_{sa}^*, \hat{i}_{sb}^*, \hat{i}_{sc}^*$) with actual source current ($\hat{i}_{sa}, \hat{i}_{sb}, \hat{i}_{sc}$). It is experimentally verified that indirect current control technique is simpler, requires less hardware and offers better performance. It is also capable of eliminating the harmonics and switching ripples, resulting in sinusoidal supply current.

V. CONCLUSION

After reviewing it is found that extensive efforts are being made to improve the performance of the APFs, with the help of new or improved modifications in topology and control methodologies. Various configurations of APFs have been proposed to compensate aforementioned power quality problems. Complexity of the control algorithm can also be greatly reduced with the involvement of artificial intelligence like fuzzy logic, neural network etc. with the conventional approach to resolve the poor power quality problems not only in steady-state but also in dynamic/transients situation.

REFERENCES

- [1] Akagi H. Kanazawa Y. and Nabae A. ,“Instantaneous reactive power compensators comprising switching devices without energy storage components”, *IEEE Trans. on Industrial Applications*, vol.20 no.3 pp.625-630, 1984.
- [2] Akagi H. and Nabae A., “Control strategy of active power filters using multiple voltage source PWM convertors”, *IEEE Trans. on Industrial Applications*, vol.IA-22, no.3 pp.460-465,May/June 1986

- [3] J. S. Subjak Jr. and J. S. Mcquilkin, "Harmonics-causes, effects, measurements, analysis: An update," *IEEE Trans. Ind. Applicat.*, vol. 26, pp. 1034-1042, Nov./Dec. 1990.
- [4] M. E. Amoli and T. Florence, "Voltage, current harmonic control of a utility system—A summary of 1120 test measurements," *IEEE Trans. Power Delivery*, vol. 5, pp. 1552-1557, July 1990.
- [5] Akagi H. and Fujita H., "A new power line conditioner for harmonic compensation in power systems", *IEEE Trans. on power Delivery*, vol.10 no.3 pp.1570-1575, July 1995.
- [6] H. Akagi, "New trends in active filters for power conditioning," *IEEE Transactions on Industrial Applications*," vol. 32, no. 6, pp. 1312-1322, November/December 1996.
- [7] Aredes M., Hafner J. and Heumann K., "Three phase four wire shunt active filter control strategies", *IEEE Trans. on power Electronics*, vol.12, no.2 pp.311-318, 1997.
- [8] B. Singh, K. Al-Haddad and A. Chandra, "A New control approach to three-phase active filter for harmonics and reactive power compensation" *IEEE Trans. Power Systems*, vol. 13, no. 1, February 1998.
- [9] B. Singh, A. Chandra, and K. Al-Haddad," Performance comparison of two current control technique applied to an active filter," in *Proc. Of 8th international conference on Harmonics and Quality of Power* , vol., 1, pp. 133-138,october 1998.
- [10]B. Singh, K. Al-Haddad and A. Chandra, "A review of active filters for power quality improvement" *IEEE, Trans. Ind. Electron*", vol.46, No. 5, pp. 960-971, 1999.
- [11]B. Singh, A. Chandra, and K. Al-Haddad," Computer-aided modeling and simulation of active power filters," *Taylor & Francis, Inc. Electrical Machine and Power Systems*, vol., 27, pp. 1227-1241,1999.
- [12]B. Singh, A. Chandra, P.Rastgoufard and K. Al-Haddad," DSP based method of active filter:Elimination of switching ripples," in *Proc.of 15th annual IEEE Applied Power Electronics Conference*, vol.,1, pp.427-433,2000.
- [13]Bansal R.C., Bhatti T.S., and Kothari D.P., "Artificial intelligence techniques for monitoring and control of power systems: an overview", in *Proc. of International conference on Control, Instrumentation and information Communication*, Kolkata, India, pp.91-95, 2001.
- [14]D. Rivas, L. Moran, J. Dixon, and J. Espinoza, "A simple control scheme filter for hybrid active power," *IEE Proceedings on Generation, Transmission and Distribution*, vol. 149, no. 4, pp. 485-490, July 2002.
- [15]Jain S.K.,Agarwal P.and Gupta H.O.,"Fuzzy logic controlled shunt active power filter for power quality improvement",in *IEE Proc.on Electric Power Applications*, vol. 149,no.5, pp.317-328,September 2002.
- [16]Aredes M. and Monteiro F.C., "A control strategy for shunt active power filter", in *Proc of IEEE Harmonics and Quality of Power Conference*,vol,2,pp. 472-477,Oct.2002.
- [17]Bansal R.C., Bhatti T.S., and Kothari D.P., "Artificial intelligence techniques for reactive power /voltage control in power systems: a review", in *Proc. of Application Evolutionary Strategies to Power, Signal Processing and Control*, Rourkela, India,pp.57-63,2002.
- [18]Bansal R.C., "Bibliography on the fuzzy set theory applications in power systems (1994-2001)",*IEEE Trans. on Power Systems*, vol.18,no.4,pp.1291-1299,2003.
- [19]Jain S.K.,Agarwal P.and Gupta H.O.,"Design simulation and experimantal investigation on a shunt active power filter for harmonics and reactive power compensation",*Electric Power Components and Systems*, vol. 31,pp.671-692,2003.
- [20]B. Singh, A. Chandra, P.Rastgoufard and K. Al-Haddad," Design simulation and implementation of three pole/four pole topologies for active filters," *IEE Proc. On Electric Power Applications*, vol.151,No. 4, pp.467-476,July 2004.

- [21]Akagi H. "Active harmonic filters", in *Proc. Of the IEEE*, vol.98 no.12 pp. 2128-2141, December 2005.
- [22]Akagi H. "Modern active filters and traditional passive filters", *Bulletin of the Polish Academy of Technical Sciences*, vol.549 no.3 pp.255-269,2006.
- [23]B. Singh and V. Verma, "An indirect current control of hybrid power filter for varying loads," *IEEE Transactions on Power Delivery*, vol. 21, no. 1, pp. 178-184, January 2006.
- [24]Ali Ajami and Hosseini S.H., "Implementation of a novel control strategy for shunt active filter", *ECI Trans. on Electrical Engineering* , vol.4, no.1 pp.40-46,February 2006.
- [25]Akagi H. Watanabe B.H. and Aredes M. "Instantaneous power theory and applications to power conditioning", *IEEE Press Piscataway NJ* 2007.
- [26]B. Singh and V. Verma and J. Solanki, "Neural network based selective compensation of current quality problems in distribution system," *IEEE Trans. on Industrial Electronics*,vol.54, no. 1, pp.53-60,February 2007.
- [27]B. Singh, A. Chandra, P.Rastgoufard and K. Al-Haddad," An improved control algorithm for active filters," *IEEE Trans. Power Delivery*, vol. 22,No. 2, pp. 1009-1020,2007.
- [28]B. Singh and V. Verma, "Selective compensation of power quality problems through active power filter by current decomposition,"*IEEE Trans.on Power Delivery*,vol.23,no.2, pp.792-799,April 2008.
- [29]G. Bhuvaneswari, and M. G. Nair, "Design, simulation, and analog circuit implementation of a three-phase shunt active filter using the $I \cos\Phi$ algorithm," *IEEE Transactions on Power Delivery*, vol. 23, no. 2, pp.1222-1235, April 2008.
- [30]A. Bhattacharya, C. Chakraborty, and S. Bhattacharya, "Current compensation using shunt type active power filters," *IEEE Industrial Electronics Magazine*, March 2009.
- [31]Patidar R.D.and Singh S.P., " Digital signal processor based shunt active filter controller for customer generated harmonics and reactive power compensation", *Electric Power Components and Systems*, vol. 38,pp.937-959,May 2010.

BIOGRAPHICAL NOTES



Ms. Annu Govind is presently pursuing Ph.D. in Electrical Engineering Department of Jamia Millia Islamia, New Delhi. She has done her M.Tech. in Power System and B.E. in Electrical Engineering from Malviya National Institute of Technology (MNIT), Jaipur.



Mr. Omendra K Govind is presently working at Advanced Level Telecom Training Centre (ALTTC), Apex training centre of BSNL. He has done Master of Design in Industrial Design from IIT Delhi and B.Tech. in Electrical Engineering from Jamia Millia Islamia, New Delhi. He has also done MBA in Operation Management. He has delivered training program for Commonwealth Telecommunication Organization (CTO) at South Africa and also for Asia Pacific Telecommununity (APT) at India.

DESIGN ASPECT OF ENERGY EFFICIENT DATA CENTER

Omendra K. Govind¹, Annu Govind²

¹ *Advanced Level Telecom Training Centre (AL TTC), BSNL, Ghaziabad,(India)*

² *Research Scholar, Department of Electrical Engineering, Jamia Millia Islamia, New Delhi, (India)*

ABSTRACT

This paper provides an overview of design aspect of energy efficient data center which covers categories Information Technology (IT) systems and their environmental conditions, data center air management, cooling and electrical systems. IT system energy efficiency and environmental conditions are presented first because measures taken in these areas have a cascading effect of secondary energy savings for the mechanical and electrical systems. This paper concludes with a section on metrics and benchmarking values by which a data center and its systems energy efficiency can be evaluated.

Keywords: Data Center, Design Aspect, Efficiency, Cooling, Benchmarking

I. INTRODUCTION

Data Center is now a day's one of the most critical for successful business operations. Its suitable planning and designing allows sufficient guarantees of quality, efficiency and service continuity, irrespective of its dimension and sector in which it operates. In recent years, data center complexity and criticality has increased with steady growth in capacity and density which strains resources and consequently results in poor performance. The design of energy efficient and reliable data center is the key for business continuity and its success. The data center energy requirements are primarily for running IT and its support systems. Therefore, design aspect of data center spans across IT systems and their environmental conditions, cooling and electrical systems, air flow management and heat recovery, and on-site generations. Energy efficiency of IT systems and their environmental conditions have cascading effect of energy savings for mechanical and electrical systems. This paper also discusses metrics and bench marking values by which a data center and its systems energy efficiency can be evaluated.

II. INFORMATION TECHNOLOGY SYSTEMS

In a typical facility, IT equipment loads can account for over half of the entire facility's energy use. Use of efficient IT equipment with in a data center can significantly reduce these loads, which consequently downsize the equipment required to cool them. The IT equipment loads can be reduced by going for energy-efficient processors, fans, and power supplies, high-efficient network equipment, consolidating storage devices, consolidating power supplies, and implementing virtualization.

2.1 Servers

Servers used in Data Centers are generally rack servers which are main culprit for wasting energy and consumes largest chunk of IT energy load. Though servers are driver of data center operations, still they run below 20%

utilizations most of the times and still draw full power. With the improvement of internal cooling systems and processors, servers are being made to minimize the wasted energy. One of the features available may include variable speed fans which can deliver sufficient cooling while running slower and thereby consuming less energy. It is also to mention that Energy star program recognised high efficiency servers will, on an average, be more efficient by 30% than the standard servers. Another approach will be IT power management by throttling down drive that will put idle servers to sleep while reducing power consumption on idle processors. However, many IT team fear that could negatively impact server reliability but, hardware is designed to handle tens of thousands of on-off cycles. Power drawn by servers can also be regulated by installing “power cycler” software, which on low demand can direct individual devices to power down. But this may cause potential power management risks like slower performance and possible system failure which should be assessed against the potential energy savings.

Energy efficiency can be achieved by using multi-core processors which offer improved performance within same power and cooling load and consolidate shared devices. However, graphic intensive applications and high computing require high speed single core design processor. Energy savings can be further achieved by integration and consolidating IT system redundancies. For example, integrated rack mounted power supplies will operate with higher load factor in comparison with individual power supplies. As shown in Fig.1, increment in load factor improves power supply efficiencies. By sharing other IT resources like CPUs, disk drives, and memory also optimizes electrical usage.

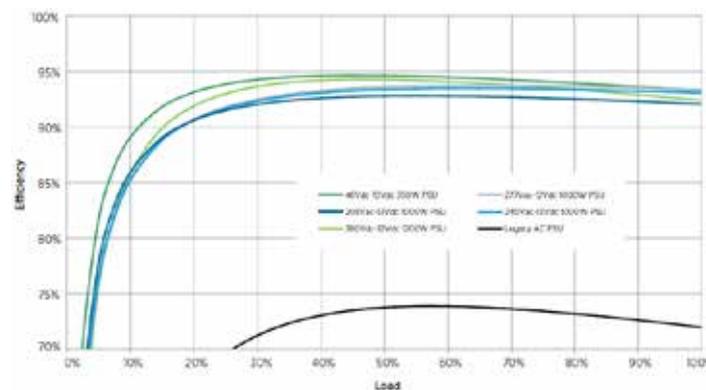


Figure 1 Efficiencies at Varying Load Levels for Typical Power Supplies

(Source: Quantitative Efficiency Analysis of Power Distribution Configurations for Data Centers, The Green Grid)

2.2 Storage Devices

Power consumed by storage devices is roughly linear to the number of storage modules used. The redundancy level for storage needs to be rationalized and right-sized. By using storage consolidating technologies like Storage Area Network (SAN) and Network Attached Storage (NAS) transports data offline which is need not to be accessed readily. By keeping unneeded data offline reduces the data in production environment which results in less storage and CPU requirement that corresponds to lower requirement of cooling and power.

Further, utilization and efficiency can be improved by adopting thin provisioning technology. In this method of maximizing storage capacity utilization, storage is drawn from shared pool on an as-need basis in contrast to traditional storage system where fixed amount of anticipated storage capacity is allotted. This allows addition of extra physical capacity at the later stage when need arises.

2.3 Network Equipment

As newer technologies network equipment can provide more throughput per unit of power consumed, active energy management measures can be adopted to reduce energy usage with the varied network demand. These measures include idle state logic, memory access algorithms, gate count optimization, and Input / Output buffer reduction. Increase of peak data transmission rates requires more power to transmit small amounts of data over time. By quickly switching the speed of the network links to the amount of data that is currently transmitted, efficiency of Ethernet network can be substantially improved.

2.4 Power Supplies

Typical internal or rack mounted alternating current/direct current (AC-DC) power supplies converts AC power to DC power at efficiencies of around 60% to 70%. By using higher efficiency power supplies of 85% - 95%, will directly lower data center's power consumption and indirectly reduce cooling system cost and rack overheating issues. The real operating load should also be considered for selecting power supplies that offers best efficiency at the most frequently expected load level. The optimal power supply load level is typically in the mid-range of its performance curve: around 40% to 60%, as shown in Fig.1.

As efficient power supplies usually have a minimal incremental cost at the server level. There are many certification programs that have standardized power supplies efficiencies. For example, the 80 PLUS program offers power supplies with efficiencies of 80% or greater at 20%, 50%, and 100% of their rated loads with true power factors of 0.9 or greater.

2.5 Consolidation

2.5.1 Hardware Location

The more efficient cooling system performance can be achieved by grouping together equipment with similar heat load densities and temperature requirements. Grouping equipment by their environmental requirements of temperature and humidity allows controlling of cooling system for each location to the least energy-intensive set points. This also applies for consolidating underutilized data center spaces to a centralized location which can improve data center efficiency by condensing the implementation to one location.

2.5.2 Virtualization

Virtualization is a method of running multiple independent virtual operating systems on a single physical system. It allows increase of server utilization. Virtualization allows combining of processing power of individual system onto fewer servers that operate at higher utilization. It will drastically reduce the number of servers in a data center, reducing required server power and consequently the size of the necessary cooling equipment. There are some overhead for implementation of virtualization, but results in higher saving and efficiency.

III. ENVIRONMENTAL CONDITIONS

3.1 Ashrae Guidelines and IT-Reliability

Designing of cooling and air management systems in a data center requires a look at the standardized operating environments for equipment recommended by the American Society of Heating, Refrigerating and Air-

Conditioning Engineers (ASHRAE) or Network Equipment Building System (NEBS). The recommended and allowable condition for Class 1 and 2 data centers are shown in Fig.2 and tabulated in Table 1.

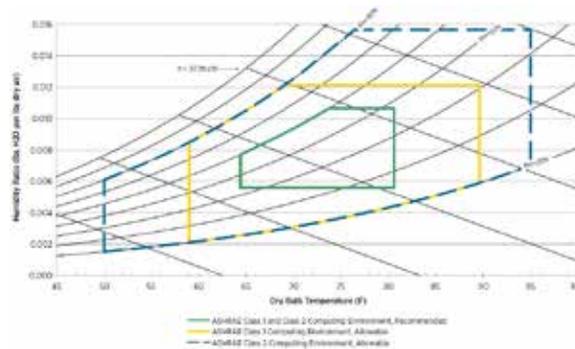


Figure 2 2009 ASHRAE Environmental Envelope for IT Equipment Air Intake Conditions

(Source: Rumsey Engineers)

	Class 1 and Class 2 Recommended Range	Class 1 Allowable Range	Class 2 Allowable Range
Low Temperature Limit	64.4°F DB	59°F DB	50°F DB
High Temperature Limit	80.6°F DB	89.6°F DB	95°F DB
Low Moisture Limit	41.9°F DP	20% RH	20% RH
High Moisture Limit	60% RH & 59°F	DP80% RH & 62.6°F	DP80% RH & 69.8°F DP

Table 1 ASHRAE Recommended and Allowable Inlet Air Conditions for Class 1 and 2 Data Centers

(Source: Rumsey Engineers)

The recommended environmental envelope is guidance for energy-efficient operation of data center while maintaining high reliability while allowable envelope outlines the environmental boundaries tested by equipment manufacturers for equipment functionality, not reliability.

IV. AIR MANAGEMENT

The purpose of design and configuration of air management in a data center is to minimize or eliminate mixing of cooling air supplied to equipment and the hot air rejected from the equipment. Effective air management implementation minimizes the bypass of cooling air around rack intakes and the recirculation of heat exhaust back into rack intakes. When designed correctly, an air management system can reduce operating costs, reduce first cost equipment investment, increase the data center’s power density (Watts/ square foot), and reduce heat related processing interruptions or failures. A few key design issues include the configuration of equipment’s air intake and heat exhaust ports, the location of supply and returns, the large scale airflow patterns in the room, and the temperature set points of the airflow.

4.1 Implement Cable Management

The obstructions in false flooring and false ceiling / overhead often interfere with cooling air distribution. This interference is undesirable as this can significantly reduce air flow being handled by air handlers and negatively affect the air distribution. Cable congestion in raised-floor plenums can sharply reduce the total airflow as well as degrade the airflow distribution through the perforated floor tiles. Both effects promote the development of undesirable hot spots. A minimum effective clear height of 2 feet must be provided for raised floor installations.

Greater under floor clearance can help achieve a more uniform pressure distribution in some cases. A data center should have a cable management strategy to minimize air flow obstructions caused by cables and wiring and should target the entire cooling air flow path, including the rack-level IT equipment air intake and discharge areas as well as under-floor areas. Persistent cable management is a key component of maintaining effective air management.

4.2 Aisle Separation and Containment

When equipment racks and the cooling system are designed to prevent mixing of the hot rack exhaust air and the cool supply air into the racks is called basic hot aisle/cold aisle configuration is created. In this quipment is laid out in rows of racks with alternating cold (rack air intake side) and hot (rack air heat exhaust side) aisles between them. Strict hot aisle/cold aisle configurations can significantly increase the air-side cooling capacity of a data center's cooling system. All equipment rows of racks are placed back-to-back, and holes through the rack (vacant equipment slots) are blocked off on the intake side to create barriers that reduce recirculation, as shown in Figure 3 below. Additionally, cable openings in raised floors and ceilings should be sealed as tightly as possible. With proper isolation, the temperature of the hot aisle no longer impacts the temperature of the racks or the reliable operation of the data center; the hot aisle becomes a heat exhaust. The air-side cooling system is configured to supply cold air exclusively to the cold aisles and pull return air only from the hot aisles.

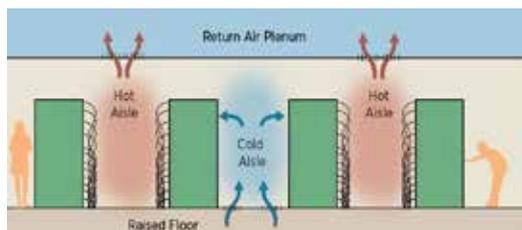


Figure 3 Example of Hot Aisle/Cold Aisle Configuration

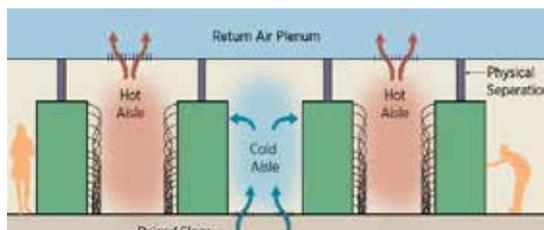


Figure 4 Sealed Hot Aisle/Cold Aisle Configuration

The hot rack exhaust air is not mixed with cooling supply air and, therefore, can be directly returned to the air handler through various collection schemes, returning air at a higher temperature. Further improvements can be achieved by using flexible plastic barriers, such as strip curtains to seal the space between the tops of the rack and air return location while allowing flexibility in accessing, operating, and maintaining the computer equipment below as shown in Fig.4.

4.3 Optimize Supply and Return Air Configuration

Hot aisle/cold aisle configurations can be served by overhead or under-floor air distribution systems. When an overhead system is used, supply outlets that 'dump' the air directly down should be used in place of traditional office diffusers that throw air to the sides, which results in undesirable mixing and recirculation with the hot aisles. The diffusers should be located directly in front of racks, above the cold aisle and temperature monitoring to control the air handlers should be located in areas in front of the computer equipment.

In under-floor air supply systems, under-floor plenum often serves both as a duct and a wiring chase. Coordination throughout design and into construction and operation throughout the life of the center is necessary since paths for airflow can be blocked by electrical or data trays and conduit. The location of supply

tiles needs to be carefully considered to prevent short circuiting of supply air and checked periodically if users are likely to reconfigure them.

4.4 Raising Temperature Set Points

Higher supply air temperature and a higher difference between the return air and supply air temperatures increases the maximum load density possible in the space and can help reduce the size of the air-side cooling equipment required, particularly when lower-cost mass produced package air handling units are used. The lower required supply airflow due to raising the air-side temperature difference provides the opportunity for fan energy savings. Additionally, the lower supply airflow can ease the implementation of an air-side economizer by reducing the sizes of the penetrations required for outside air intake and heat exhaust.

V. COOLING SYSTEMS

When beginning the design process and equipment selections for cooling systems, it is important to always consider initial and future loads.

5.1 Direct Expansion (DX) Systems

Packaged DX air conditioners likely compose the most common type of cooling equipment for smaller data centers. These units are generally available as off-the-shelf equipment from manufacturers, commonly described as CRAC units. An enhancement to the air-cooled condenser is a device which sprays water over the condenser coils. The evaporative cooling provided by the water spray improves the heat rejection efficiency of the DX unit. Additionally, these units are commonly offered with air-side economizers.

5.2 Air Handlers

Better performance can be achieved with specifically designed central air handler systems. A centralized system offers many advantages like they use larger motors and fans that tend to be more efficient. They are also well suited for variable volume operation through the use of VSDs and maximize efficiency at part-loads. It also has maintenance benefits, and the reduced footprint. Implementation of an airside economizer system is simplified with a central air handler system. Optimized air management, such as that provided by hot aisle/cold aisle configurations, is also easily implemented with a ducted central system.

5.3 High-Efficiency Chilled Water Systems

Use efficient water-cooled chillers in a central chilled water plant. A high-efficiency VFD-equipped chiller with an appropriate condenser water reset is typically the most efficient cooling option for large facilities. Chiller part-load efficiency should be considered since data center often operate at less than peak capacity. Chiller part-load efficiencies can be optimized with variable frequency driven compressors, high evaporator temperatures and low entering condenser water temperatures.

5.4 Free Cooling

The cooling load for a data center is independent of the outdoor air temperature. However, a proper engineering evaluation of the local climate conditions must be completed to evaluate whether this is the case for a specific data center. Due to the humidity and contamination concerns associated with data centers, careful control and

design work may be required to ensure that cooling savings are not lost because of excessive humidification and filtration requirements, respectively.

VI. ELECTRICAL SYSTEMS

Similar to cooling systems, it is important to always consider initial and future loads, in particular part- and low-load conditions, when designing and selecting equipment for a data center's electrical system.

6.1 Power Distribution

In Data center, typically have an electrical power distribution path consisting of the utility service, switchboard, switchgear, alternate power sources, paralleling equipment for redundancy, and auxiliary conditioning equipment. These components each have a heat output that is tied directly to the load in the data center. Efficiencies can range widely between manufacturers and variations in how the equipment is designed. However, operating efficiencies can be controlled and optimized through thoughtful selection of :

6.1.2 Uninterruptible Power Supplies

UPS systems provide backup power to data centers, and can be based on battery banks, rotary machines, fuel cells, or other technologies. A portion of all the power supplied to the UPS to operate the data center equipment is lost to inefficiencies in the system. To minimize these losses, first evaluate requires a UPS system. Increasing the UPS system efficiency offers direct, 24-hour-a-day energy savings, both within the UPS itself and indirectly through lower heat loads and even reduced building transformer losses. Evaluate the need for power conditioning. Line interactive systems often provide enough power conditioning for servers at a higher efficiency than typical double conversion UPS systems.

6.1.3 Power Distribution Units

A PDU passes conditioned power to provide reliable power distribution to multiple pieces of equipment. Maintaining a higher voltage in the source power lines fed from a UPS or generator allows for a PDU to be located more centrally within a data center. As a result, the conductor lengths from the PDU to the equipment are reduced and less power is lost in the form of heat.

6.1.4 Distribution Voltage Options

Another source of electrical power loss for both AC and DC distribution is that of the conversions required from the original voltage supplied by the to that of the voltage at each individual device within the data center (usually a low voltage around 120V AC to 240V AC). In order to provide electrical power in the most energy-efficient manner possible:

- Minimize the resistance by increasing the cross-sectional area of the distribution path and making it as short as possible.
- Maintain a higher voltage for as long as possible to minimize the current.
- Use switch-mode transistors for power conditioning.
- Locate all voltage regulators close to the load to minimize distribution losses at lower voltages.

6.1.5 DC Power

In a conventional data center power is supplied from the grid as AC power and distributed throughout the data center infrastructure as AC power. However, most of the electrical components within the data center, as well as the batteries storing the backup power in the UPS system, require DC power. As a result, the power must go through multiple conversions resulting in power loss and wasted energy. One way to reduce the number of times

power needs to be converted is by utilizing a DC power distribution which may involve significantly higher cost.

6.1.6 Lighting

Data center spaces are not uniformly occupied and, therefore, do not require full illumination during all hours of the year. UPS, battery and switch gear rooms are examples of spaces that are infrequently occupied. Therefore, zone based occupancy sensors throughout a data center can have a significant impact on reducing the lighting electrical use. Careful selection of an efficient lighting, lamps and ballasts will also reduce not only the lighting electrical usage but also the load on the cooling system.

VII. DATA CENTER METRICS AND BENCHMARKING

Energy efficiency metrics and benchmarks can be used to track the performance of and identify potential opportunities to reduce energy use in data centers.

7.1 Power Usage Effectiveness (PUE)

PUE is defined as the ratio of the total power to run the data center facility to the total power drawn by all IT equipment:

$$PUE = \frac{\text{Total Facility Power}}{\text{IT Equipment Power}}$$

Standard	Good	Better
2.0	1.4	1.1

An average data center has a PUE of 2.0; however, several recent super-efficient data centers have been known to achieve a PUE as low as 1.1.

7.2 Data Center Infrastructure Efficiency (DCiE)

DCiE is defined as the ratio of the total power drawn by all IT equipment to the total power to run the data center facility, or the inverse of the PUE:

$$PUE = \frac{\text{Total Facility Power}}{\text{IT Equipment Power}}$$

Standard	Good	Better
2.0	1.4	1.1

PUE and DCiE are defined with respect to site power draw. These donot define overall efficiency of an entire data center. These metrics could be alternatively defined using units of average annual power or annual energy (kWh) rather than an instantaneous power draw (kW).

7.3 Source PUE

This is defined with reference to source energy.

$$\text{Source PUE} = \frac{\text{Total Facility Energy (kWh)}}{\text{UPS Energy (kWh)}}$$

7.4 Energy Reuse Effectiveness (ERE)

ERE is defined as the ratio of the total energy to run the data center facility minus the reuse energy to the total energy drawn by all IT equipment:

$$ERE = \frac{\text{Cooling + Power + Lighting + IT - Reuse Energy}}{\text{UPIT Equipment Energy}}$$

An ERE of 0 means that 100% of the energy brought into the data center is reused elsewhere, outside of the data center control volume.

7.5 Rack Cooling Index (RCI) and Return Temperature Index (RTI)

RCI measures how effectively equipment racks are cooled according to equipment intake temperature guidelines established by ASHRAE/NEBS. By using the difference between the allowable and recommended intake temperatures from the ASHRAE Class 1 (2008) guidelines, the maximum (RCI_{HI}) and minimum (RCI_{LO}) limits for the RCI are defined as follows:

$$RCI_{HI} = \left[\frac{\sum (T_x - 80)}{(90 - 80)n} \right] \times 100 [\%] \quad RCI_{LO} = \left[\frac{\sum (65 - T_x)}{(65 - 59)n} \right] \times 100 [\%]$$

where,

T_x = Mean temperature at equipment intake x n = Total number of intakes.

An RCI of 100% represents ideal conditions for the equipment, with no over or under temperatures. An RCI < 90% is often considered to portray poor conditions.

7.6 RTI

RTI evaluates the energy performance of the air management system. RTI is defined as:

$$RTI = \frac{\Delta T_{AHU}}{\Delta T_{EQUIP}} \times 100\%$$

where, ΔT_{AHU} is the typical (airflow weighted) air handler temperature drop ΔT_{EQUIP} is the typical (airflow weighted) IT equipment temperature rise.

The RCI and RTI parameters allow an objective method of measuring the overall performance of a data center air management system. They should be used in tandem to ensure the best possible design.

VIII. CONCLUSION

The Data Center is a complex system which requires thorough knowledge of different aspect of its design. These different areas are interrelated with each other and the impact others performance in reference to energy efficiency. It is also inferred that consideration of load while designing also plays role in overall dimensioning and achievement of required efficiency level.

REFERENCES

1. *Design Recommendations for High Performance Data Centers*, Rocky Mountain Institute, 2003.
2. *High Performance Data Centers – A Design Guidelines Sourcebook*. Pacific Gas and Electric, 2006. http://hightech.lbl.gov/documents/data_centers/06_datacenters-pge.pdf.
3. *Best Practices for Datacom Facility Energy Efficiency*. ASHRAE Datacom Series, 2008.
4. Efficiency of Power Supplies in the Active Mode. EPRI. <http://www.efficientpowersupplies.org>.
5. 80 PLUS® Energy-Efficient Technology Solutions. <http://www.80plus.org>.
6. *Program Requirements for Computer Servers, Version 1.0*. Energy Star, 2009.
7. *Data Processing and Electronic Areas*, Chapter 17. ASHRAE HVAC Applications, 2007.

8. The Green Data Center 2.0, Chapter 2, Energy-Efficient Server Technologies, 2009.
9. The Green Grid, *Quantitative Efficiency Analysis of Power Distribution Configurations for Data Center*. <http://www.thegreengrid.org/en/Global/Content/white-papers/Quantitative-Efficiency-Analysis>.
10. Juniper Networks. *Energy Efficiency for Network Equipment: Two Steps Beyond Greenwashing*.
11. Energy Star. *Enterprise Server and Data Center Energy Efficiency Initiatives*.
12. *Thermal Guidelines for Data Processing Environments*, 2nd Edition, ASHRAE Datacom Series 1, 2009. • *Thermal Guidelines for Data Processing Environments*, TC9.9 Mission Critical Facilities, ASHRAE
13. *Best Practices Guide for Variable Speed Pumping in Data Centers*, Ernest Orlando, Lawrence Berkeley.
14. *Variable-Primary-Flow Systems Revisited*, Schwedler P.E.
15. *Thermal Guidelines for Data Processing Environments*, TC9.9 Mission Critical Facilities, ASHRAE, 2004.
16. Federal Energy Management Program, *Best Practices Guide for Energy-Efficient Data Center Design*, US Department of Energy, Revised 2011.
17. Industrial Technologies Program. Saving Energy in Data Centers. *DC Pro Software Tool Suite*. <http://www1.eere.energy.gov/industry/datacenters/software.html>. Accessed December 3, 2009.
18. *Self-Benchmarking Guide for High-Tech Buildings: Data Centers*, Ernest Orlando Lawrence Berkeley National Laboratory. <http://hightech.lbl.gov/benchmarking-guides/data.html>. Accessed February 4, 2010.

BIOGRAPHICAL NOTES

	Mr. Omendra K Govind is presently working at Advanced Level Telecom Training Centre (ALTTC), Apex training centre of BSNL. He has done Master of Design in Industrial Design from IIT Delhi and B.Tech. in Electrical Engineering from Jamia Millia Islamia, New Delhi. He has also done MBA in Operation Management. He has delivered training program for Commonwealth Telecommunication Organization (CTO) at South Africa and also for Asia Pacific Telecommununity (APT) at India.
	Ms. Annu Govind is presently pursuing Ph.D. in Electrical Engineering Department of Jamia Millia Islamia, New Delhi. She has done her M.Tech. in Power System and B.E. in Electrical Engineering from Malviya National Institute of Technology (MNIT), Jaipur.

EFFICIENT DATA AGGREGATION TECHNIQUE FOR CATTLE MONITORING USING ODMR PROTOCOL

Dr.Sivakumar.D¹, Suriya Prakash.S², Veera.V³,
Richerd Ebimeni Doss.A.T⁴

¹ Professor- Department of IT, Easwari Engineering College, Chennai (India)

^{2,3,4} Student -Department of Information Technology, Easwari Engineering College, Chennai (India)

ABSTRACT

In Cattle monitoring a central hub needs to get update status regards the presence of living inside the region. Frequent data aggregation is necessary in application like Cattle monitoring. All nodes in the coverage area should respond to query raised by central hub rather than particular node. Existing work uses On-demand unicast routing to collect the defect node information where it takes considerable amount of latency. To overcome this issue we introduce On-demand multicast routing to reduce latency of network. This leads to efficient complete path identification within mille seconds and reduce latency with lesser network traffic. The proposed protocol is validated through Proteus and ccs compiler.

Keywords: Data Aggregation, Cattle Monitoring, Efficient Multipath Routing

I. INTRODUCTION

Modern, intensive farms make the farmer totally responsible for all cattle under his control. Over the last three decades farming practice has moved away from self-sustaining mixed livestock enterprises with relatively small numbers of several species, towards large, single species units. Thus information regarding infective cattle must be routed with lesser latency. When latency is concern for this application we discussed two major routing methodology and proposed efficient routing way. One of the basic cattle monitoring tasks is routing between various nodes. It is nothing other than establishing a path between the source and the destination. However in large and complex networks routing is a difficult process because of the possible intermediate hosts it has to cross in reaching its final destination. In order to reduce the complexity, the network is considered as a collection of sub domains and each domain is considered as a separate entity. This helps routing easy. However basically there are three routing protocols in ad hoc networks namely proactive, reactive and hybrid routing protocols. Of these reactive routing protocols establish and maintain routes based on demand. A primary issue in managing multicast group dynamics is the routing path built for data forwarding. Most existing ad hoc multicasting protocols can be classified as tree based or mesh-based. The tree-based protocol, a tree-like data forwarding path is built with the root at the source of the multicast session.

The mesh-based protocol [eg.ODMRP], in contrast, provide multiple routes between any pair of source and destination, intended to enrich the connectivity among group members for better resilience against topology changes.

II. RELATED WORKS

In [1] Wireless sensor network for cattle health monitoring by Ivan Andonovic, Craig Michie, Michael Gilroy, Hock Guan Goh, Kae Hsiang Kwong, Konstanios Sasloglou and Tsungta Wu investigates an adaptation of Wireless sensor networks to cattle health monitoring. The proposed solution facilitates the requirement for continuously assessing the condition of individual animals, aggregating and reporting this data to the farm manager. There are several existing approaches to achieving animal monitoring, ranging from using a store and forward mechanism to employing GSM-based techniques; these approaches only provide sporadic information and introduce a considerable cost in staffing and physical hardware. The core of this solution overcomes the aforementioned drawbacks by using alternative cheap, low power consumption sensor nodes capable of providing real-time communication at a reasonable hardware cost. In this paper, both the hardware and software have been designed to provide real-time data from dairy cattle whilst conforming to the limitations associated with WSNs implementations.

III. OVERVIEW OF ON-DEMAND ROUTING PROTOCOLS

The reactive routing protocols (e.g. AODV) usually use distance-vector routing algorithms that keep only information about next hops to adjacent neighbors and costs for paths to all known destinations. On the other hand hybrid routing protocols combine the advantages of both proactive and reactive protocols. Reliable multicast in mobile network was proposed by Prakash et al. In their solution the multicast message is flooded to all the nodes over reliable channels. The nodes then collectively ensured that all mobile nodes belonging to the multicast group get the message. If a node moves from one cell to another while a multicast is in progress, delivery of the message to the node was guaranteed. Tree-based multicast routing provides fast and most efficient way of routing establishment for the communications of mobile nodes in MANET. The authors described a way to improve the throughput of the system and reduce the control overhead. When network load increased, MAODV ensures network performance and improves protocol robustness. Its PDR was found to be effective with reduced latency and network control overhead. On Demand Multicast Routing Protocol is a multicast routing protocol (ODMRP) designed for ad hoc networks with mobile hosts. Multicast is nothing but Communication between a single sender and multiple receivers on a network and it transmits a single message to a select group of recipients. Multicast is commonly used in streaming video, in which many megabytes of data are sent over the network. The major advantage of multicast is that it saves bandwidth and resources. Moreover multicast data can still be delivered to the destination on alternative paths even when the route breaks. It is an extension to Internet architecture supporting multiple clients at network layers. The fundamental motivation behind IP multicasting is to save network and bandwidth resource via transmitting a single copy of data to reach multiple receivers. Single packets are copied by the network and sent to a specific subset of network addresses.

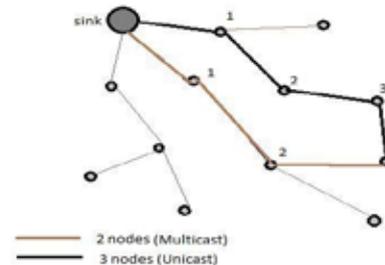
These addresses point to the destination. Protocols allowing point to multipoint efficient distribution of packets are frequently used in access grid applications. It greatly reduces the transmission cost when sending the same packet to multiple destinations.

IV. EFFICIENT DATA AGGREGATION TECHNIQUE

The multicast extensions for the AODV (Ad-hoc On- Demand Distance Vector) routing protocol discover Multicast routes on demand using a broadcast route discovery mechanism. One of the distinctive feature of on demand Unicast (AODV) is unicast reply. It is a tree based hard state. In case of link failure force actions

to repair the state. AODV unicasts the reply back to the source, if an intermediate node on the path moves away, the reply is lost and the route is lost.

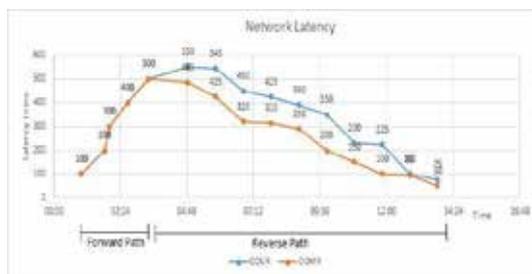
It does not activate multicast route immediately because it's bi-directional. While ODMRP provides alternative paths and a link failure need not trigger the re computation of the mesh. Its immediately activate the multicast route in case of failure. So latency will be reducing and also network traffic slow down greatly.



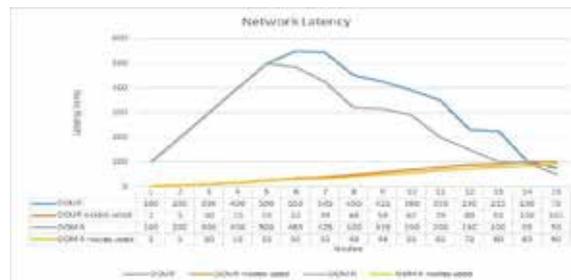
(a).No of nodes used by unicast and multicast routing

V . SIMULATION RESULTS

(a).Network Latency



(i) latency during forward and backward pass



(ii)latency in terms of node used.

From the graph shown above we depreciated real-time reading of our proposed system. It clearly shows that number of nodes used is greatly reduced with reduced latency.

VI. CONCLUSION AND FUTURE WORK

In this paper we proposed an efficient data aggregation technique using On-demand multicast routing intended for cattle monitoring system. In proposed system on-demand multipath technique greatly reduced the latency of network we also study that this methodology reduces energy consumption considerably. Simulation experiments under protoues and ccs compiler to evaluate the proposed system in real time scenario, however this system leads to end-to-end delay. For future work, we plan to validate the proposed schemes on different scenarios with various parameters like throughput and PDR.

REFERENCES

- [1]Wireless sensor network for cattle health monitoring by Ivan Andonovic,Craig Michie,Michael Gilroy,Hock Guan Goh,Kae Hsiang Kwong,Konstanios Sasloglou and Tsungta Wu.
- [2] Rishiwal, S.Verma, and S. K. Bajpai, ``QoS based power aware routing in MANETs," Int. J. Comput. Theory Eng., vol. 1, no. 1, pp. 4754, 2009.
- [3] S. Gupta, C. K. Nagpal, M. Kaur, and B. Bhushan, ``Impact of variable transmission range on MANET performance," Int. J. Ad Hoc, Sensor Ubiquitous Compute., vol. 2, no. 4, pp. 5966, 2011. .
- [4] C. Poongodi and A. M. Natarajan, ``Optimized replication strategy for intermittently connected mobile networks,"Int. J. Business Data Communication .

BLOCKING MISBEHAVING USERS IN UNIDENTIFIED NETWORK

K.Dinesh Raja¹, V. Aravindhana², N. Noor Alleema³

*^{1,2} Scholar, ³Assistant Professor, Department of Information Technology,
SRM University, Chennai (India)*

ABSTRACT

The advent of anonymizing networks assured that users could access internet services with complete privacy avoiding any possible hindrance. This arrangement where series of routers form a network, hide the user's IP address from the server. However malfeasance of few malpractitioners has left this system with a loophole where users make use of this anonymity to deface popular websites. Administrators who cannot practically block a user using IP address are forced to shut all possible nodes that lead to exit. Thus deny access to both behaving and non-behaving users altogether. And so end up blocking users with no compromise to their anonymity. Hence propose a system which is undogmatic with different servers. Thus they aim at giving the administrator the right to block the malicious user without hindering the anonymity of the rest.

Keywords: *Anonym Zing Networks, Blacklisting, Symmetric Cryptography, Tor, Pseudonym, Nymble Ticket, Subnet-Based Blocking, Rate-Limiting, Non-Frame Ability, Anonymous Authentication, Backward Unlinkability, Subjective Blacklisting, Rate-Limited Anonymous Connections, Revocation Auditability..*

I. INTRODUCTION

Propose a system with following features: Anonymous authentication, backward unlink ability, subjective blacklisting, fast authentication speeds, rate-limited anonymous connections, revocation auditability (where users can verify whether they have been blacklisted). In this system aim to generate nymbles, which are not easy to connect, however a stream of these nymbles assure a simulation to anonymous access. Here provide a means where the website administrator can block user without knowing his IP address (ie through pseudonym generated: which is a random secret identity with the pseudonym manager) without hindering the remaining network. User also has his complete privacy without having to compromise until he behaves.

II. RELATED WORKS

Anonymous credential systems like Camenisch and Lysyanskaya's [9, 10] use group signatures for anonymous authentication, wherein individual users are anonymous among a group of registered users. Non-revocable group signatures such as Ring signatures [14] provide no accountability and thus do not satisfy our needs to protect servers from misbehaving users. Basic group signatures [3, 4, 2, 12] allow revocation of anonymity by no one except the group manager. As only the group manager can revoke a user's anonymity, servers have no way of

linking signatures to previous ones and must query the group manager for every signature; this lack of scalability makes it unsuitable for our goals.

Traceable signatures allow the group manager to release a trapdoor that allows all signatures generated by a particular user to be traced; such an approach does not provide the backward anonymity that desire, where a user's accesses before the complaint remain anonymous. Specifically, if the server is interested in blocking only future accesses of bad users, then such reduction of user anonymity is unnecessarily drastic.. And misbehaving users should be blocked from making further connections after a complaint.

In some systems, misbehavior can be defined precisely. For instance, double-spending of an —e-coinl is considered misbehavior in anonymous electronic cash systems . Likewise, compact e-cash , k-times anonymous authentication and periodic n-times anonymous authentication deem a user to be misbehaving if she authenticates—too manyl times. In these cases, convincing evidence of misbehavior is easily collected and fair judgment of misbehavior can be ensured. While such approaches can encourage certain kinds of fair behavior in anonymizing.It is difficult to map more complex notions of misbehavior onto —double spendingl or related approaches. It may be difficult to precisely define what it means to —deface a webpage| and for

It proves to a trusted party that a particular webpage was defaced. How can the user be sure these—proofsl are accurate and fairly judged? Can avoid the problem of judging misbehavior entirely? In this paper answer affirmatively by proposing a system that does not require proof of misbehavior. Websites may complain about users for any reason; our system ensures users are informed of complaints against them, thus —making everybody happy|—except, of course, the misbehaving users, who remain anonymous but are denied access.

II. EXISTING MODEL

Verifier-Local Revocation (VLR): In order to overcome the problem of lack of backward unlinkabilit VLR was introduced in 2004 by —Dan Bonehl and —Hovav Shachaml.This was an approach of membership revocation in group signatures known as verifier-local revocation. In this approach, only verifiers are involved in the revocation process, while there is no involvement of the signers. Thus, since signers have no load, this approach is suitable for mobile environments. This stratagem satisfies backward unlink ability to some extent. The backward unlink ability means that even after a member is revoked, signatures produced by the member before the revocation remains anonymous. Verifier-local revocation requires the server to perform only local updates during revocation. Therefore, there will be a lot of burden on the server. Advantages of existing system are :

- 1)Local updating is possible
- 2)Backwardunlinkability

There are many solutions for the problems and difficulties in anonymous networks. But each method has some limitations and issues. They are like: In pseudonym Systems, every individual will be known to the other user by a pseudonym which is blacklisted if a user misbehaves. But this results in pseudonymity for all users and weakens the anonymity. And,also the users are prevented from sharing their pseudonyms.

Group signature is a method by which a member of a group anonymously signs the message on behalf of the group. Here, the server sends complaints to the Group Manager (GM) if a user misbehaves which lacks scalability. Traceable signatures traces the signatures signed by a single party without opening the signature and revealing the identities of any other users. It does not provide backward unlink ability, wherein the previously collected signatures remain anonymous even after the signer's revocation. Since there is no backward

unlinkability, there will be no subjective blacklisting. Subjective blacklisting is the process by which the server can blacklist the user for whatever reason the server desires. Drawbacks are Heavy computation at the server side

Time sqander. Less secure.Hence, due to the unsatisfied results of the existing systems, have implemented the new Nymble system which can give us the fruitful results which need.

III. PROPOSED MODEL

Present a secure system called Nymble, which provides all the following properties: anonymous authentication, backward unlinkability, subjective blacklisting, fast authentication speeds, rate-limited anonymous connections, revocation auditability Without additional information, these nymbles are computationally hard to link,and hence using the stream of nymbles simulates anonymous access to services.Websites, however, can blacklist users by obtaining a seed for a particular nymble, allowing them to link future nymbles from the same user — those used before the complaint remainunlinkable. Servers can therefore blacklist anonymous users without knowledge of their IP addresses while allowing behaving users to connect anonymously. In fact, any number of anonymizing networks can rely on the same Nymble system, blacklisting anonymous users regardless of their anonymizing network(s) of choice

Blacklisting anonymous users. Provide a means by which servers can blacklist users of an anonymizing network while maintaining their privacy.

Practical performance. Our protocol makes use of inexpensive symmetric cryptographic operations to significantly outperform the alternatives.

Open-source implementation. With the goal of contributing a workable system, have built an open source implementation of Nymble, which is publicly available.I provide performance statistics to show that our system is indeed practical.

Advantages

1. Intends to bind identity of an anonymous user to a pseudonym, generated from user's IP address. This idea enables a server to complain about misbehavior of a user and blacklist his future tickets.
2. Honest users remain anonymous, &blacklist future connections of particular users and their requests remain unlinkable.
3. All connections of a blacklisted user before the complaint will remain anonymous.
4. A user can check whether he is blacklisted or not at the beginning of a connection.
5. Users are aware of their blacklist status before accessing a service.
6. Servers can blacklist users for whatever reason, and the privacy of blacklisted users is maintained.

Pseudonym Creation

Equation:

$$\begin{aligned} \text{(a)} \quad f(x) &= \sum_{i=0}^{i=n} U_i \\ \text{(b)} \quad Ps &= P(f(x)) \end{aligned}$$

Where

- $f(x)$ is function to concatenate all string of the field from the user profile
- U_i -> each profile attributes
- Ps -> Pseudonym
- $P(f(x))$ -> Random Function to calculate Pseudonym

Algorithm:

Input: Set $U = \{u_1, u_2, u_3, \dots, u_n\}$

Output: pseudonym (Ps)

- Step 0: Get the User Profile attribute set U
- Step 1: Convert all the attributes to String type
- Step 2: Concatenate all the String to get a single String
- Step 3: Get the auto incremented User ID as I
- Step 4: $x = ID \bmod 7$
- Step 5: for $i=0$ to String length
- Step 6: Fetch x th character from the String
- Step 7: Continue till 7 characters are selected
- Step 8: concatenate all the 7 characters
- Step 9: return Pseudonym

ATTACKS: $A = \{A_1, \dots, A_n\}$ is Attack Set

A1 Attack

Equation:

$$f(A_1) \Rightarrow UP_{data} > lim$$

Where

- $f(A_1)$ is function to identify uploading excess amount of data attack
- UP_{data} -> Uploading data
- Lim -> Limits

Algorithm:

Input: User Uploading data UPdata, threshold size (lim)

Output: User Blocked State

- Step 0: Get the User data on the web server
- Step 1: Get the Current of the file size as C_{lim}
- Step 2: if($C_{lim} > lim$)
- Step 3: Tag user as misbehavior user
- Step 4: Get Pseudonym
- Step 5: Add Pseudonym in blocked list
- Step 6: Update User's state
- Step 7: return user state

A2 Attack

Equation:

$$f(A_2) \Rightarrow U_{data} \notin U_i$$

Where

- $f(A_2)$ is function to identify DMA attack
- U_{data} -> Users Uploaded data
- U_i -> Respective User

Algorithm:

Input: User accessing data Udata

Output: User Blocked State

- Step 0: Allow User to access data on the web server
- Step 1: Get the user accessed data name as U_{data}
- Step 2: if U_{data} does not belongs to him
- Step 3: Tag user as misbehavior user
- Step 4: Get Pseudonym
- Step 5: Add Pseudonym in blocked list
- Step 6: Update User's state
- Step 7: return user state

A4 Attack

Equation:

$$f(A4) \Rightarrow U_{pwd} \notin U_i$$

Where

- $f(A4)$ is function to identify Password attack
- U_{pwd} -> Users Password
- U_i -> Respective User

A5 Unblocking User

Equation:

$$f(A5) \Rightarrow (t_c - t_b) > T$$

Where

- $f(unb)$ is function to identify unblocking user
- t_c -> Current time
- t_b -> blocked time
- T -> threshold time

Algorithm:

Input: User password U_{pwd} and U_{name}

Output: User Blocked State

Step 0: Allow User to login in his account

Step 1: Get the user's credential like U_{name} and U_{pwd}

Step 2: if U_{pwd} does not belongs to U_{name}

Step 3: then warn user for 3 times

Step 4: Reset Password

Step 5: Mail New Password to Original user

Step 6: Get Pseudonym

Step 7: Add Pseudonym in blocked list

Algorithm:

Input: Blocked time as t_b , Current time as t_c and Threshold time as T

Output: Unblocked Blocked State

Step 0: Get t_b and t_c and T

Step 1: if $(t_c - t_b) > T$

Step 2: Get Pseudonym

Step 3: Add Pseudonym in unblocked list

Step 4: Update User's state

Step 5: return user state

IV. IMPLEMENTED MODULES

4.1 The Nymble Manager

After obtaining a pseudonym from the PM, the user connects to the Nymble Manager (NM) through the anonymizing network, and requests nymbles for access to a particular server (such as Wikipedia). A user's requests to the NM are therefore pseudonymous, and nymbles are generated using the user's pseudonym and the server's identity. These nymbles are thus specific to a particular user-server pair. Nevertheless, as long as the PM and the NM do not collude, the Nymble system cannot identify which user is connecting to what server; the NM knows only the pseudonym-server pair, and the PM knows only the user identity-pseudonym. Servers have the right to blacklist anonymous users without having to know their IP addresses while allowing behaving users to stay intact anonymously. The system ensures the user has complete knowledge about being blacklisted, that he should disconnect immediately if they are blacklisted. Although our work applies to anonymizing networks in general, we consider Tor for purposes of exposition. In fact, any number of anonymizing networks can rely on the same Nymble system, blacklisting anonymous users regardless of their anonymizing network(s) of choice.

4.2 Pseudonym Manager

The user must first contact the Pseudonym Manager (PM) and demonstrate control over a resource; for IP-address blocking, the user must connect to the PM directly (i.e., not through a known anonymizing network), ensuring that the same pseudonym is always issued for the same resource.

4.3 Blacklisting A User

Users who make use of anonymizing networks expect their connections to be anonymous. If a server obtains a seed for that user, however, it can link that user's subsequent connections. It is of utmost importance, then, that users be notified of their blacklist status before they present a nymble ticket to a server. In our system, the user can download the server's blacklist and verify her status. If blacklisted, the user disconnects immediately.

IP-address blocking employed by Internet services. There are, however, some inherent limitations to using IP addresses as the scarce resource. If a user can obtain multiple addresses she can circumvent both nymble-based and regular IP-address blocking. Subnet-based blocking alleviates this problem, and while it is possible to modify our system to support subnet-based blocking, new privacy challenges emerge; a more thorough description is left for future work.

4.4 Nymble-Authenticated Connection

Blacklist ability assures that any honest server can indeed block misbehaving users. Specifically, if an honest server complains about a user that misbehaved in the current linkability window, the complaint will be successful and the user will not be able to —nymble-connect, i.e., establish a Nymble-authenticated connection, to the server successfully in subsequent time periods of that linkability window. Rate-limiting assures any honest server that no user can successfully nymble-connect to it more than once within any single time period. Non-frameability guarantees that any honest user who is legitimate according to an honest server can nymble-connect

to that server. This prevents an attacker from framing a legitimate honest user, e.g., by getting the user blacklisted for someone else’s misbehavior. This property assumes each user has a single unique identity. When IP addresses are used as the identity, it is possible for a user to —frame an honest user who later obtains the same IP address. Non-frameability holds true only against attackers with different identities (IP addresses). A user is legitimate according to a server if she has not been blacklisted by the server, and has not exceeded the rate limit of establishing Nymble-connections. Honest servers must be able to differentiate between legitimate and illegitimate users.

Anonymity protects the anonymity of honest users, regardless of their legitimacy according to the (possibly corrupt) server; the server cannot learn any more information beyond whether the user behind (an attempt to make) a nymble-connection is legitimate or illegitimate

A user is legitimate according to a server if she has not been blacklisted by the server, and has not exceeded the rate limit of establishing Nymble-connections. Honest servers must be able to differentiate between legitimate and illegitimate users.

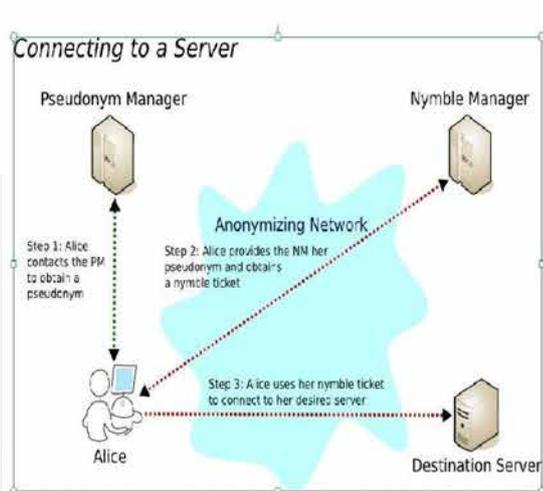


Fig.1 Pseudonym Manager

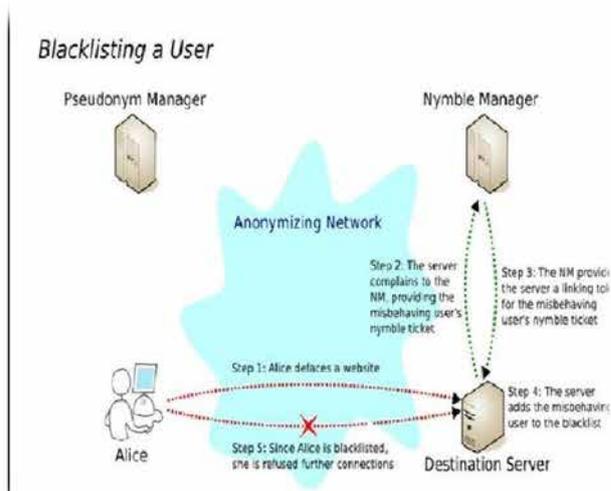


Fig.2 Blacklist a User

V. FEATURES

Anonymous Authentication: Anonymous authentication allows any user to access any public content without providing a user name and password challenge to the client browser. If some content should be viewed only by selected users, It must configure the appropriate permissions to prevent anonymous users from accessing that content. Only registered users to view selected content, configure an authentication method for that content that requires a user name and password.

5.1 Backward Unlinkability

Backward unlinkability means that even after a user is revoked, signatures produced by the user before the revocation remain anonymous. However, all the signatures produced from the revoked user are linkable. This means that the anonymity of signatures produced before the revocation is compromised. In some cases that all

signatures from an illegal person should be traced. The approaches without backward unlinkability need to pay careful attention to when and why a user must have all their connections linked and users must worry about whether their behaviors will be judged fairly.

5.2 Fast Authentication Speed

In Nymble system there is a fast Authentication speed that implies the presence of a database to provide persistent data to be used as a part of the verification process. Database access must be kept to a minimum so that the request/response process remains fast and uninhibited by database overhead.

5.3 Revocation Auditability

In this the user can check whether he is blacklisted or not and if he is blacklisted then the user can be revoke

5.4 Subjective Blacklisting

If the authorized server complains about a user that misbehaved in the current linkability window, the complaint will be successful and the user will not be able to a Nymble-authenticated connection.

5.4 Anonymity

User is legitimate according to a server if user has not been blacklisted by the server, and has not exceeded the rate limit of establishing Nymble connections. Honest servers must be able to differentiate between legitimate and illegitimate users.

5.5 Non-Frameability

It guarantees that any honest user who is legitimate according to an honest server can nymble connect to that server. This prevents an attacker from framing a legitimate honest user.

5.6 Rate Limited Anonymous Connection

Rate-limiting assures any honest server that no user can successfully nymble-connect to it more than once within any single time period.

VI. FUTURE SCOPE

Nymble project can be extended in next version called nymble and also can be developed on android platform. Expecting that our work will increase the mainstream acceptance of anonymizing networks such as Tor, which has thus far been completely blocked by several services because of users who abuse their anonymity. By providing a mechanism for server administrators to block anonymous misbehaving users, to make the use of anonymizing networks such as Tor more acceptable for server administrators everywhere. All users remain anonymous misbehaving users can be blocked without deanonymization, and their activity prior to being blocked remain unlinkable

This work can also be extended into a multiple rounds of pseudonym construction in which the PM participates in multiple rounds of communication with the user. Another enhancement would be is to provide service providers with the ability to detect repeat offenders and revoke these users' access for longer durations of time.

VII. CONCLUSION

The proposed and built a comprehensive credential system called Nymble, which can be used to add a layer of accountability to any publicly known anonymizing network. Our new design is not only scalable and robust, but also securer under various types of attacks. A new system is proposed that adds an additional layer of security to the anonymous networks.

In Our system they tried to blacklist user's activities, considered several types of attacks. This system is used to block the misbehaving users in anonymizing networks. It automatically finds the misbehaving user and blacklists them without affecting their privacy and anonymity. This adds one more layer of security to the system. The proposed method motivates the need for dynamic forgiveness and security in anonymous networks and this system will increase the acceptance of anonymous networks that is blocked by several services because of users who misuse their anonymity

REFERENCES

- [1] Stefan Brands. Untraceable off-line cash in wallets with observers (extended abstract). In Douglas R. Stinson, editor, CRYPTO, volume 773 of LNCS, pages 302–318. Springer, 1993
- [2] Mihir Bellare, Haixia Shi, and Chong Zhang. Foundations of group signatures: The case of dynamic groups. In Alfred Menezes, editor, CT-RSA, volume 3376 of LNCS, pages 136–153. Springer, 2005.
- [3] Giuseppe Ateniese, Jan Camenisch, Marc Joye, and Gene Tsudik. A practical and provably secure coalition-resistant group signature scheme. In Mihir Bellare, editor, CRYPTO, volume 1880 of LNCS, pages 255–270. Springer, 2000.
- [4] Mihir Bellare, Daniele Micciancio, and Bogdan Warinschi. Foundations of group signatures: Formal definitions, simplified requirements, and a construction based on general assumptions. In Eli Biham, editor, EUROCRYPT, volume 2656 of LNCS, pages 614–629. Springer, 2003.
- [5] Jan Camenisch, Susan Hohenberger, Markulf Kohlweiss, Anna Lysyanskaya, and Mira Meyerovich. How to win the clonewars: efficient periodic n-times anonymous authentication. In Ari Juels, Rebecca N. Wright, and Sabrina De Capitani di Vimercati, editors, ACM Conference on Computer and Communications Security, pages 201–210. ACM, 2006.
- [6] Jan Camenisch, Susan Hohenberger, and Anna Lysyanskaya. Compact e-cash. In Ronald Cramer, editor, EUROCRYPT, volume 3494 of LNCS, pages 302–321. Springer, 2005.
- [7] Paul F. Syverson, Stuart G. Stubblebine, and David M. Goldschlag. Unlinkable serial transactions. In Rafael Hirschfeld, editor, Financial Cryptography, volume 1318 of LNCS, pages 39–56. Springer, 1997.
- [8] Isamu Teranishi, Jun Furukawa, and Kazue Sako. k-times anonymous authentication. In Pil Joong Lee, editor, ASIACRYPT, volume 3329 of LNCS, pages 308–322. Springer, 2004.
- [9] Jan Camenisch and Anna Lysyanskaya. An efficient system for non-transferable anonymous credentials with optional anonymity revocation. In Birgit Pfitzmann, editor, EUROCRYPT, volume 2045 of LNCS, pages 93–118. Springer, 2001.
- [10] David Chaum. Blind signatures for untraceable payments. In CRYPTO, pages 199–203, 1982. LNCS,

pages 246–264. Springer, 1990.

[11] Patrick P. Tsang, Apu Kapadia, and Sean W. Smith. Anonymous IP-address blocking in tor with trusted computing (work-in-progress).

[12] In The Second Workshop on Advances in Trusted Computing (WATC '06 Fall), November 2006.

[13] Ronald L. Rivest, Adi Shamir, and Yael Tauman. How to leak a secret. In Colin Boyd, editor, ASIACRYPT, volume 2248 of LNCS, pages 552–565. Springer, 2001.

DESIGN AND IMPLEMENTATION OF TOOL FOR CONVERTING A RELATIONAL DATABASE INTO AN XML DOCUMENT: A REVIEW

Sunayana Kohli

*Masters of Technology, Department of Computer Science, Manav Rachna College of Engineering,
Faridabad,(India)*

ABSTRACT

There are four central problems in data management: capture, storage, retrieval, and exchange of data. XML is a standard that solves the problem of data exchange between various web applications. Electronic data interchange (EDI), the traditional data exchange standard for large organizations, is giving way to XML, which is likely to become the data exchange standard for all organizations.

However despite the big extent of XML language most of the applications store their data in relational databases. So there must be a mechanism to convert a relational database into an XML document. The interoperation of relational databases and xml database involves schema and data translations. Converting the relational database into XML document involves converting an existing Relational Database which is in the form of tables, into XML file format. The proposed approach performs the conversion of database in MS SQL to XML document. First the database connection is established. Then the user is asked to select a table which is in the form of dropdownlist. The selected table is shown as grid view. Thereafter the table is converted to xml document on button click.

Keywords: *Data Exchange, Relational Database, XML*

I. INTRODUCTION

A database is a means of storing information in such a way that information can be retrieved from it. In simplest terms, a relational database is one that presents information in tables with rows and columns. A table is referred to as a relation in the sense that it is a collection of objects of the same type (rows). Data in a table can be related according to common keys or concepts, and the ability to retrieve related data from a table is the basis for the term relational database. A Database Management System (DBMS) handles the way data is stored, maintained, and retrieved. In the case of a relational database, a Relational Database Management System (RDBMS) performs these tasks. ^[9].

Extensible Markup Language (XML) is a simple, very flexible text format derived from SGML (ISO 8879). Originally designed to meet the challenges of large-scale electronic publishing, XML is also playing an increasingly important role in the exchange of a wide variety of data on the Web and elsewhere. XML is a markup language for documents containing structured information. Structured information contains both content (words, pictures, etc.) and some indication of what role that content plays (for example, content in a section heading has a different meaning from content in a footnote.). Almost all documents have some structure. A

markup language is a mechanism to identify structures in a document. The XML specification defines a standard way to add markup to documents^[10].

XML is **not a replacement** for HTML since both these languages were designed for **different goals**.

XML was designed to **describe data** and to focus on **what data is** while HTML was designed to **display data** and to focus on **how data looks**.

The need to convert a Relational database into an XML document emerged due to the fact that whenever two systems having different characteristics(say for example both systems having databases in different formats) had to share data then such a data exchange proved to be complex and costly. Thereafter XML based formats have been used for storing data and for exchanging data between various programs, applications and systems. The success of using XML based formats is accounted by the following factors^[11]:

- Firstly, because the specification of the XML syntax is **publicly available**, any programmer can write software to process a document that is created in accordance with that syntax.
- Secondly, because XML is **text-based** and the specification is explicitly based on other international **standards** for the exchange of text (Unicode, UCS), there is no room for ambiguity in the reading of an XML document. If necessary, an XML document can be viewed using the simple text editors available on all systems - though the creator of an XML document rarely intends that a user views a document in this "native" form.
- Thirdly, because the XML syntax is relatively **simple**, it is quite easy to write such programs ("parsers"), and those parsers are usually quite small and efficient. Software developers building XML-based applications do not have to concern themselves with developing the code to read or write the XML syntax; they can reuse an existing parser (probably one of a range available) and concentrate their efforts on the characteristics of the data.
- Finally, XML's capacity for describing data structure is sufficiently **flexible** to incorporate a wide range of structures. XML's approach to structure is discussed further below.

So, the owner or manager of the information can have confidence that the data they create can be read on a remote system (or indeed on their own system after a change of software). And the software developer is freed from at least some of the work in having to read or write many different (and rapidly changing) file formats. Many major commercial vendors have an interest in facilitating such data exchange, and several of the largest contributed to the development of the XML specification and have subsequently invested heavily in software initiatives that build on XML^[11].

II. RELATED WORK

2.1. On XML Technology^[5]

Fong et al. ^[6]outline steps for converting relational database into XML document. These steps show how to translate relational schema into XML schema, and mapping relational data to an XML document^[5].

2.2 On data conversion^[5]

Ref^[3] categorizes the earlier translation methods (Relational database to XML translation) into four categories:

- **Flat Translation** - Converts relations to XML without using nesting.

The simplest method for translating relational data to XML is Flat translation^[2] where each relation is translated to an element E, and each attribute of the relation is either translated to a sub element (element approach) or attribute (attribute approach) of E.

- **2. Query-based Translation** - Conversion occurs by using a query extraction language that is an extension of SQL or a new XML query language. Methods for defining an XML document by specifying a query or extraction rules allow user control of the conversion process, but require complete specification of the mapping. That is, the entire target XML schema is defined during the mapping.
- **Model-based Translation** - Converts the relational schema to an intermediate model which then is mapped to XML using conversion rules. Most of these methods are more concerned with designing good XML modelling languages than performing the translation, although translation can be performed as long as the relational model can be transformed into the intermediate model.
- **Dependency-based Translation** - Converts the relational schema using dependency information.

Dependency-based translation uses functional, multi-valued, and inclusion dependencies for translation^[3].

Wang Xiaoling^[4] proposed a method based on genetic algorithm for XML document storage to seek the optimum mapping relational between XML Schema and relational database table^[1].

Ref^[7] analyses the different variants of implementation for publishing the relational data as XML documents. To be able to analyse the alternatives for publishing relational data as XML documents the main differences between the relational tables and the XML documents were taken into account which means that XML documents, unlike the relational tables, have tags and nested structure. Therefore, somewhere along the process of converting relational tables to XML documents, the tags and the structure must be added. So there existed several possibilities of approaching i.e. the adding of tags during the final step of the query processing (late tagging) or by doing this at the beginning of the process (early-tagging). Similarly, the structure adding can be done as the final step of processing (late structuring) or it can be done earlier (early-structuring)^[7].

III. PROPOSED METHODOLOGY

This approach converts an existing relational database which is in the form of tables into an XML document. The database which we are using is MS SQL2008. Suppose there are two organizations which want to share data. But both these organizations have different database formats. It becomes very complex and time consuming to convert one database format to another. Hence an easier solution is to convert the database into an XML file format which can be easily transmitted over the web to the other organization. Then the intermediate XML file format could be converted to the respective database format used by the organization. This has an additional benefit that the native relational database remains unaltered.

The steps to be followed for converting the relational Database to an XML document are listed below:

- Firstly a database connection is established. We have made the database in MS SQL 2008.
- Now the user can select a table from the dropdownlist. The selected table is shown as a grid view.
- Thereafter the selected table is converted to an XML document on button click. A message is displayed to the user regarding the location where the XML document is created.

3.1 Implementation

The tool that we have designed converts a table of Relational database into an XML document. The input to the tool is a relational database table. Then the processing of the input takes place. Finally the user gets an XML document as an output. The graphical user interface is designed so that the user can select a table which he wants to convert into an XML document. On button click the XML document is created and a message is shown to the user where he can find the created XML document.

3.2 Technology Used

The GUI is designed using the visual studio 2010. The language used is C#(C sharp).The database is designed using the MS SQL 2008.

3.3 Screenshots of the Tool

3.3.1 Homepage Design of the Proposed Tool

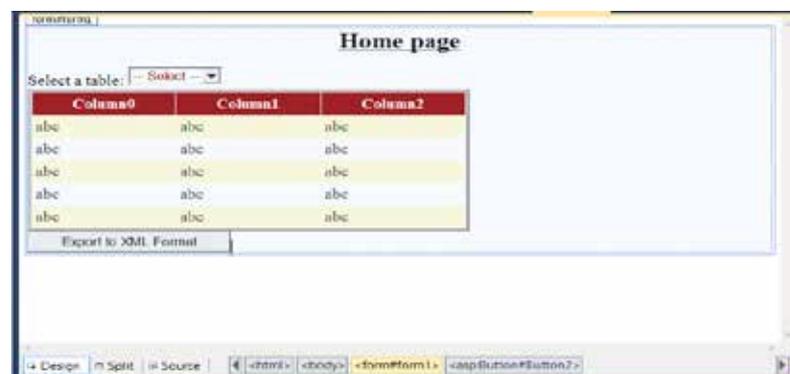


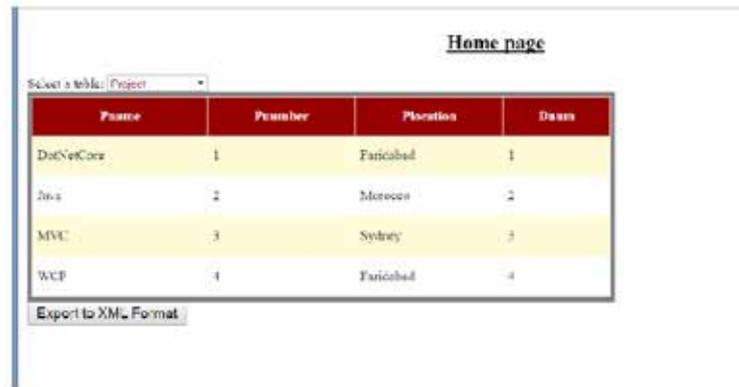
Figure 3.3.1: Homepage Design View

3.3.2 The user is asked to select a table. The table is a Relational database table which serves as an input to the tool.



Figure 3.3.2: The user is prompted to select a table

3.3.3 The table is displayed in Grid view.

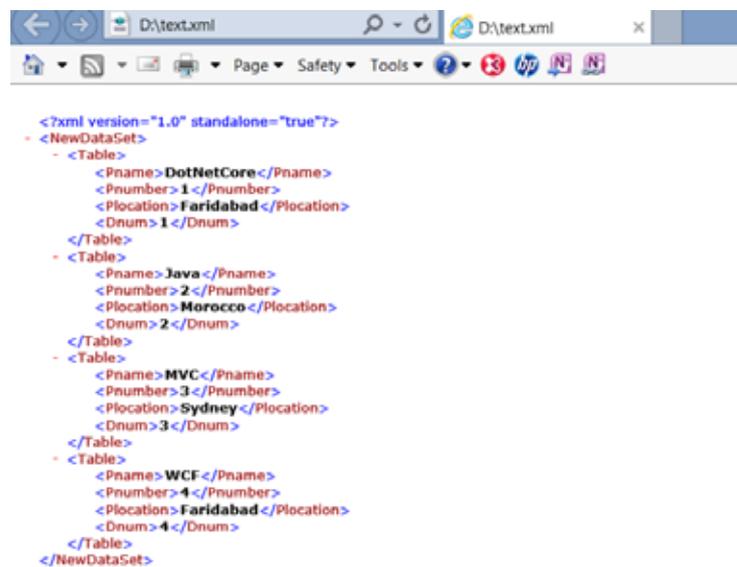


Name	Pnumber	Plocation	Dnum
DotNetCore	1	Faridabad	1
Java	2	Morocco	2
MVC	3	Sydney	3
WCF	4	Faridabad	4

Export to XML Format

Figure 3.3.3 Grid view display of the selected table

3.3.4 XML document is created on button click



```
<?xml version="1.0" standalone="true"?>
<NewDataSet>
  <Table>
    <Pname>DotNetCore</Pname>
    <Pnumber>1</Pnumber>
    <Plocation>Faridabad</Plocation>
    <Dnum>1</Dnum>
  </Table>
  <Table>
    <Pname>Java</Pname>
    <Pnumber>2</Pnumber>
    <Plocation>Morocco</Plocation>
    <Dnum>2</Dnum>
  </Table>
  <Table>
    <Pname>MVC</Pname>
    <Pnumber>3</Pnumber>
    <Plocation>Sydney</Plocation>
    <Dnum>3</Dnum>
  </Table>
  <Table>
    <Pname>WCF</Pname>
    <Pnumber>4</Pnumber>
    <Plocation>Faridabad</Plocation>
    <Dnum>4</Dnum>
  </Table>
</NewDataSet>
```

Figure 3.3.4 : XML document

3.5 Advantages of Xml

The XML is a text formatted flat file, so this method has an ability to store any kind of data from the application system^[1]. XML will improve the efficiency of data exchange in several important ways, which include^[8]:

- **Write once and format many times:** Once an XML file is created it can be presented in multiple ways by applying different XML style sheets. For instance, the information might be displayed on a web page or printed in a book.
- **Hardware and software independence:** XML files are standard text files, which means they can be read by any application.
- **Write once and exchange many times:** Once an industry agrees on a XML standard for data exchange, data can be readily exchanged between all members using that standard.
- **Faster and more precise web searching:** When the meaning of information can be determined by a computer (by reading the tags), web searching will be enhanced. For example, if you are looking for a specific book title, it is far more efficient for a computer to search for text between the pair of tags

<booktitle> and </booktitle> than search an entire file looking for the title. Furthermore, spurious results should be eliminated.

- **Data validation** XML allows data validation using XSD or DTD which is a contractual agreement between two interacting parties.

IV. CONCLUSIONS AND FUTURE WORK

XML is here to stay. There is no doubt that virtually every branch of the computing industry has embraced the XML standard in one form or another. The proposed method converts the relational database table into an XML document. The tool can be extended to capture all the constraints imposed on a Relational database. Such constraints should also be taken into account while converting the Relational database into XML document so that all the semantics are preserved. These constraints are not covered by the tool we have designed.

V. ACKNOWLEDGEMENTS

I owe my deepest gratitude to Assistant Professor, Ms GunjanChandwani, the Guide of the project and also the Head of the Department, Mr Sachin Lakra who helped and supported me all the way for the successful completion of this paper.

REFERENCES

- [1] Kanagaraj.S and Dr.SunithaAbburu, Converting Relational Database Into Xml Document, *IJCSI International Journal of Computer Science Issues*, Vol. 9, Issue 2, No 1, March 2012
- [2] D. Lee, M. Mani, F. Chiu, and W. Chu. NeT&CoT: Translating relational schemas to XML schemas using semantic constraints. In Proceedings of the 11th CIKM, pages 282{291, 2002.
- [3] A Cost-Based Approach for Converting Relational Schemas to XML, University of Iowa Technical Report 05-02, Ramon Lawrence,IDEA Lab, University of Iowa,Iowa City, IA, USA, 52242
- [4] WangXiaoling, LanJinfeng, Dong Yisheng, Basing on GA to Store XML data in RDBMS[j]. *Journal ofComputer Research and Development*,2003,40(7):1110-1116.
- [5] Converting relational database into XML documents with DOM, J. Fong, H.K. Wong, Z. Cheng, *Information and Software Technology* 45 (2003) 335–355
- [6] J. Fong, H.K. Wong, A. Fong, Performance analysis between XML-enabled database and native XML database, in: A. Chaudhri (Ed.), XML Data Management, Addison-Wesley, Reading, MA, 2003.
- [7] From Relational Databases to XML Documents: Efficient alternatives for Publishing,Mihai Stancu, *International Journal of Digital Information and Wireless Communications (IJDWC)* 1(2): 545-553, *The Society of Digital Information and Wireless Communications*, 2011(ISSN 2225-658X)
- [8] http://en.wikibooks.org/wiki/XML_-_Managing_Data_Exchange/Introduction_to_XML
- [9] <http://docs.oracle.com/javase/tutorial/jdbc/overview/database.html>
- [10] <http://www.xml.com/pub/a/98/10/guide0.html?page=2>
- [11] <http://www.ukoln.ac.uk/nof/support/help/papers/metaxml/>

IMPACT OF EDGE DETECTION ALGORITHM IN IMAGE PROCESSING

C.A.Md. Zameer¹, A.Viswanthan², A.A.Khadar Maideen³

^{1,2} Assistant Professor, Department of Computer Science, Islamiah College [Autonomus],
Thiruvalluvar University, vaniyambadi, (India)

³ Research Guide & Associate Professor, Department of Computer Science, Islamiah College
[Autonomus], Thiruvalluvar University, vaniyambadi, (India)

ABSTRACT

The basic pitfall in Image analysis is Edge detection. Edge characteristics and boundaries are useful information in Image segmentation. Edges play an important role in many applications of image processing, in particular for machine vision systems that analyze scenes of man-made objects under controlled illumination conditions. Image edge detection is a process of locating the edge of an image which is important in finding the approximate absolute gradient magnitude at each point of an input grayscale image. The problem of getting an appropriate absolute gradient magnitude for edges lies in the method used. This paper aims at throwing light on the various edge detection methods. The various Edge detection methods are applied on the sample data. The pros and cons of them are summarized.

Keywords: Image analysis, Edge Detection operator, Gray scale input.

I. INTRODUCTION

Image processing is important in modern data storage and data transmission especially in progressive transmission of images, video coding (teleconferencing), digital libraries, and image database, remote sensing. It has to do with manipulation of images done by algorithm to produce desired images. Digital Signal Processing (DSP) improve the quality of images taken under extremely unfavorable conditions in several ways: brightness and contrast adjustment, edge detection, noise reduction, focus adjustment, motion blur reduction etc. The advantage is that image processing allows much wider range of algorithms to be applied to the input data in order to avoid problems such as the build-up of noise and signal distortion during processing.[1,2,3]. With the fast computers and signal processors available in the 2000's, digital image processing became the most common form of image processing and is general used because it is not only the most versatile method but also the cheapest. The process allows the use of much more complex algorithms for image processing and hence can offer both more sophisticated performances at simple tasks. Detection of edges in an image is a very important step towards understanding image features. Edges consist of meaningful features and contained significant information. It's reduce significantly the amount of the image size and filters out information that may be regarded as less relevant, preserving the important structural properties of an image.

II. TYPES OF EDGE DETECTION

An edge operator is a neighborhood operation which determines the extent to which each pixel's neighborhood can be partitioned by a simple arc passing through the pixel where pixels in the

neighborhood on one side of the arc have one predominant value and pixels in the neighborhood on the other side of the arc have a different predominant value. Usually gradient operators, Laplacian operators, zero-crossing operators are used for edge detection. The gradient operators compute some quantity related to the magnitude of the slope of the underlying image gray tone intensity surface of which the observed image pixel values are noisy discretized sample. The Laplacian operators compute some quantity related to the Laplacian of the underlying image gray tone intensity surface. The zero-crossing operators determine whether or not the digital Laplacian or the estimated second direction derivative has a zero-crossing within the pixel. Gradient based Roberts, Sobel and Prewitt edge detection operators, Laplacian based edge detector and Canny edge detector.

The common methods that are used for edge detection are:

- Gaussian Method
- Gradient Method (First Order Derivate)
- Laplacian Method(Second Order Derivative)

2.1 Gradient Method

In this edge detection method the assumption is that edges are the pixels with a high gradient. A fast rate of change of intensity at some direction given by the angle of the gradient vector is observed at edge pixels. An ideal edge pixel and the corresponding gradient vector is shown. At the pixel, the intensity changes from 0 to 255 at the direction of the gradient. The magnitude of the gradient indicates the strength of the edge. If we calculate the gradient at uniform regions we end up with a zero vector which means that there is no edge pixel.

In natural images we usually do not have the ideal discontinuity or the uniform regions as in the figure and we process the magnitude of the gradient to make a decision to detect the edge pixels. The simplest processing is applying a threshold. If the gradient magnitude is larger than the threshold we decide that the corresponding pixel is an edge pixel. An edge pixel is described using two important features:

- Edge strength, which is equal to the magnitude of the gradient.
- Edge direction, which is equal to the angle of the gradient.

The gradient for the ideal continuous image is estimated using some operators. Among these operators "Roberts, Sobel and Prewitt" are the commonly used.

2.2 Roberts Operator

The Roberts cross operator provides a simple approximation to the gradient magnitude:

$$G[f[i,j]] = |f[i,j] - f[i+1,j+1]| + |f[i+1,j] - f[i,j+1]|$$

Using convolution masks, this becomes:

$$G[f[i,j]] = |G_x| + |G_y|$$

$$G_x = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad G_y = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

Fig. 1 Roberts Operator

where G_x and G_y are calculated using the following mask.

As with the previous 2×2 gradient operator, the differences are computed at the interpolated point $[i + 1/2, j + 1/2]$. The Roberts operator is an approximation to the continuous gradient at this interpolated point and not at the point $[i, j]$ as might be expected[5].

2.3 Sobel Operator

A way to avoid having the gradient calculated about an interpolated point between pixels is to use a 3×3 neighborhood for the gradient calculations. Consider the arrangement of pixels about the pixel $[i, j]$ The Sobel operator is the magnitude of the gradient computed by:

where the partial derivatives are computed by:

$$M = \sqrt{s_x^2 + s_y^2}$$

$$s_x = (a_2 + ca_3 + a_4) - (a_0 + ca_7 + a_6)$$

$$s_y = (a_0 + ca_1 + a_2) - (a_6 + ca_5 + a_4)$$

with the constant $c = 2$. Like the other gradient operators, S_x and S_y can be implemented using convolution masks:



Fig. 2: Sobel Operator

this operator places an emphasis on pixels that are closer to the center of the mask. The Sobel operator is one of the most commonly used edge detectors. The labeling of neighborhood pixels used to explain the Sobel and Prewitt operator.

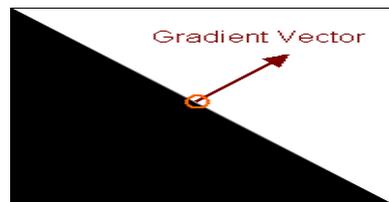


Fig.3: Gradient Vector

2.4 Prewitt Operator

The Prewitt operator uses the same equations as the Sobel operator, except that the constant $c = 1$. Therefore:

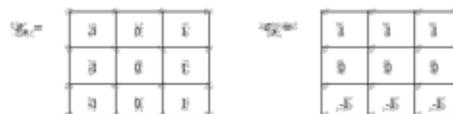


Fig. 4: Prewitt Operator

Unlike the Sobel operator, this operator does not place any emphasis on pixels that are closer to the center of the masks

2.5 Laplacian Based Edge Detection

The edge points of an image can be detected by finding the zero crossings of the second derivative of the image intensity. The idea is illustrated for a 1D signal in Figure 5.3.1. However, calculating 2nd derivative

is very sensitive to noise. This noise should be filtered out before edge detection. To achieve this, "Laplacian of Gaussian" is used. This method combines Gaussian filtering with the Laplacian for edge detection. In Laplacian of Gaussian edge detection there are mainly three steps:

- Filtering,
- Enhancement,
- Detection.

Gaussian filter is used for smoothing and the second derivative of which is used for the enhancement step. The detection criterion is the presence of a zero crossing in the second derivative with the corresponding large peak in the first derivative.

In this approach, firstly noise is reduced by convoluting the image with a Gaussian filter. Isolated noise points and small structures are filtered out. With smoothing; however; edges are spread. Those pixels, that have locally maximum gradient, are considered as edges by the edge detector in which zero crossings of the second derivative are used. To avoid detection of insignificant edges, only the zero crossings, whose corresponding first derivative is above some threshold, are selected as edge point. The edge direction is obtained using the direction in which zero crossing occurs[7,8].

The output of the Laplacian of Gaussian (LoG) operator; $h(x,y)$; is obtained by the convolution operation:

$$\begin{aligned} h(x,y) &= \Delta^2 [g(x,y) * f(x,y)] \\ &= [\Delta^2 g(x,y)] * f(x,y) \text{ where} \\ \Delta^2 g(x,y) &= \left(\frac{x^2 + y^2 - 2\sigma^2}{\sigma^4} \right)^{- (x^2 + y^2) / 2\sigma^2} \end{aligned}$$

is commonly called the Mexican hat operator.

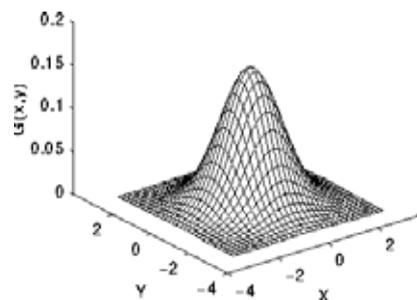


Fig.5: Mexican hat operator

In the LoG there are two methods which are mathematically equivalent:

- Convolve the image with a gaussian smoothing filter and compute the Laplacian of the result,
- Convolve the image with the linear filter that is the Laplacian of the Gaussian filter.

This is also the case in the LoG. Smoothing (filtering) is performed with a Gaussian filter, enhancement is done by transforming edges into zero crossings and detection is done by detecting the zero crossings.

III.ALGORITHM IMPLEMENTATION OF VARIOUS METHODS

Pseudo code for Laplacian Edge detection Algorithm

Step 1: Start with an image of a good looking team member. Since no such images were available, we used the image shown to the right.

Step 2: Blur the image to "smooth" it using a general low pass filter has ripples.

Step 3: Perform the laplacian on this blurred image. A one dimensional image signal, with an edge as highlighted below.

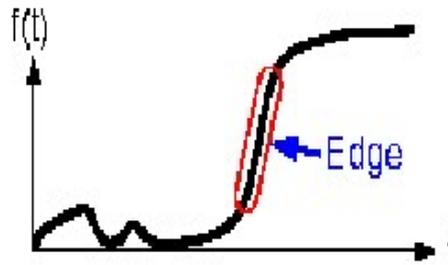


Fig. 6: Edge

we take the gradient of this signal (which, in one dimension, is just the first derivative with respect to t) we get the following:

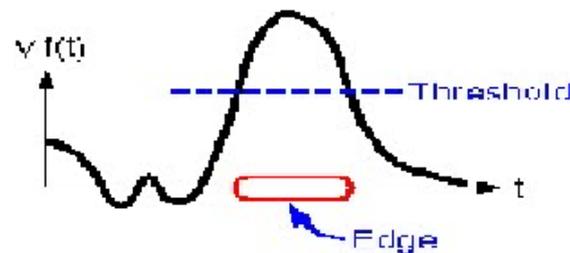


Fig. 7: Signal Edge

Clearly, the gradient has a large peak centered around the edge. By comparing the gradient to a threshold, we can detect an edge whenever the threshold is exceeded. we can localize it by computing the Laplacian (in one dimension, the second derivative with respect to t) and finding the zero crossings.

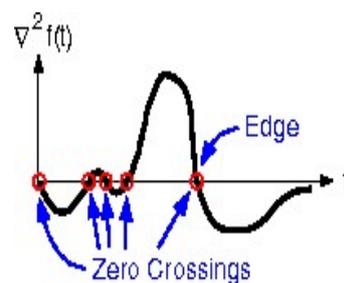


Fig. 8: Zero Crossings Edge

The above figure shows the laplacian of our one- dimensional signal. As expected, our edge corresponds to a zero crossing, but we also see other zero crossings which correspond to small ripples in the original signal. When we apply the laplacian to our test image, we get the following:

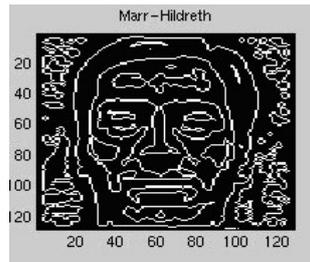


Fig. 9: Test image of laplaican

Step 4: Find the zero crossing of the laplacian and compare the local variance at this point to a threshold. If the threshold is exceeded, declare an edge.

Step 5: Median Filter the image. We apply a median filter because it removes the spot noise while preserving the edges

Pseudo code for Canny's Edge detection Algorithm

The Canny Edge Detection Algorithm has the following steps:

- Smooth the image with a Gaussian filter,
- Compute the gradient magnitude and orientation using finite difference approximations for the partial derivatives,
- Apply non-maxima suppression to the gradient magnitude,
- Use the double thresholding algorithm to detect and link edges.

Canny edge detector approximates the operator that optimizes the product of signal-to-noise ratio and localization. It is generally the first derivative of a Gaussian

Smoothing : Let $I[i, j]$ denote the image. The result of convolution of $I[i, j]$ with gives an array of smoothed data as:

$$G[i, j, \sigma]$$

$$S[i, j] = G[i, j, \sigma] * I[i, j]$$

Gradient Calculation: Firstly, the gradient of the smoothed array $s[i, j]$ is used to produce the x and y partial derivatives $p[t, f]$ and $Q[i, j]$ respectively as:

$$P[i, j] \approx (S[i, j+1] - S[i, j] + S[i+1, j+1] - S[i+1, j]) / 2$$

$$Q[i, j] \approx (S[i, j] - S[i+1, j] + S[i, j+1] - S[i+1, j+1]) / 2$$

The x and y partial derivatives are computed with averaging the finite differences over the 2x2 square. From the standard formulas for rectangular-to-polar conversion, the magnitude and orientation of the gradient can be computed as:

$$M[i, j] = \sqrt{P[i, j]^2 + Q[i, j]^2}$$

$$\theta[i, j] = \arctan(Q[i, j], P[i, j])$$

Here the arctan(x,y) function takes two arguments and generates an angle.

Non-maxima Suppression : Given the being the magnitude image array one can apply the thresholding operation in the gradient-based method and end up with ridges of edge pixel. But canny has a more sophisticated approach to the problem. In this approach an edge point is defined to be a point whose strength is locally maximum in the direction of the gradient. This is a stronger constraint to satisfy and is used to thin the ridges found by thresholding. This process, which results in one pixel wide ridges, is called "Non-maxima Suppression".

After nonmaxima suppression one ends up with an image $N[x, y] = \text{nonmax}(M[x, y], G[x, y])$ which is zero everywhere except the local maxima points. At the local maxima points the value of is preserved.

Thresholding: In spite of the smoothing performed as the first step in edge detection, the non-maxima suppressed magnitude image $N[x, y]$ will contain many false edge fragments caused by noise and fine texture. The contrast of the false edge fragments is small.

These false edge fragments in the non-maxima- suppressed gradient magnitude should be reduced. One typical procedure is to apply a threshold to $N[x, y]$. All values below the threshold are set to zero. After the application of threshold to the non-maxima-suppressed magnitude, an array E(i,j) containing the edges detected in the image $E[x, y]$ is obtained. However; in this method applying the proper threshold value is difficult and involves trial and error. Because of this difficulty, in the array E(i,j) there may still be some false edges if the threshold is too low or some edges may be missing if the threshold is too high. A more effective thresholding scheme uses two thresholds[6,7].

The algorithm performs edge linking as a by-product of thresholding and resolves some of the problems with choosing a threshold.

Pseudo code for Sobel Edge detection Algorithm

- Accept the input image.
- Apply mask S_x, S_y to the input image
- Apply sobel edge detection algorithm and the gradient
- Masks manipulation of S_x, S_y separately on the input image
- Results are combined to find the absolute magnitude of the gradient
- The absolute magnitude is the output edges
- Apply noise smoothing to the original image

$$M = \sqrt{S_x^2 + S_y^2}$$

Smoothing: Filter the image using S_x and S_y to obtain I_1 and I_2 Estimate the gradient magnitude at each pixel as

$$S(i,j) = \sqrt{(I_1)^2 + (I_2)^2} \text{ Marking the pixel as edge points if } S(i,j) > \beta .$$

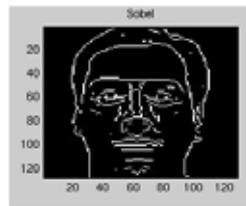


Fig. 10: Pixel Edge Point

IV.RESULT ANALYSIS

The merits and the demerits of the different methods is analyzed and summarized.

Type	Merit	Demerit
Gradient (all 3 Methods)	1. Simple 2. Detection of edges and their orientations.	1. Sensitive to noise 2. inaccurate

Gaussian	<ol style="list-style-type: none"> Using probability for finding errorate Localization and response Improving signal to noise ratio. Better detection specially in noise conditions. 	<ol style="list-style-type: none"> Complex Computations False zero crossing Time consuming
Laplacian	<ol style="list-style-type: none"> Finding the place of edges Testing the area around pixel 	<ol style="list-style-type: none"> Malfunctioning at corners, curves and where the gray level intensity function varies. Not finding the orientation of edge because of using the Laplacian filter

V.CONCLUSION

The following advantages of Sobel edge detector justify its superiority over other edge detection techniques.

(i.) **Edge Orientation:** The geometry of the operator determines a characteristic direction in which it is most sensitive to edges. Operators can be optimized to look for horizontal, vertical, or diagonal edges.

(ii.) **Noise Environment:** Edge detection is difficult in noisy images, since both the noise and the edges contain high-frequency content. Attempts to reduce the noise result in blurred and distorted edges. Operators used on noisy images are typically larger in scope, so they can average enough data to discount localized noisy pixels. This results in less accurate localization of the detected edges.

(iii.) **Edge Structure:** Not all edges involve a step change in intensity. Effects such as refraction or poor focus can result in objects with boundaries defined by a gradual change in intensity. The operator is chosen to be responsive to such a gradual change in those cases. Newer wavelet- based techniques actually characterize the nature of the transition for each edge in order to distinguish, for example, edges associated with hair from edges associated with a face. Detecting edges of an image represents significantly reduction the amount of data and filters out useless information, while preserving the important structural properties in an image. Hence, edge detection is a form of knowledge management.

REFERENCES

- [1] Canny, John, "A Computational Approach to Edge Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-8, No. 6, 1986, pp. 679-698.
- [2] Lim, Jae S., *Two-Dimensional Signal and Image Processing*, Englewood Cliffs, NJ, Prentice Hall, 1990, pp. 478-488.
- [3] Parker, James R., *Algorithms for Image Processing and Computer Vision*, New York, John Wiley & Sons, Inc., 1997, pp. 23-29.

- [4] Gonzalez, R. C., and Woods, R. E. *Digital Image Processing* (Reading, MA: Addison-Wesley, 1992
- [5] Marr, D., and Hildreth, E. "Theory of Edge Detection," *Proceedings of the Royal Society London* 207 (1980) 187-217.
- [6] Haralick, R. M., and Shapiro, L. G. *Computer and Robot Vision*, vol.1 (Reading, MA: Addison-Wesley, 1
- [7] Jain, Anil k., *Fundamentals of Digital Image Processing*, prentice Hall,1989
- [8] www.av.com

PERFORMANCE ANALYSIS OF OLSR ROUTING PROTOCOL IN MOBILE AD HOC NETWORK

Uma Dagar

Research Scholar, School of Engineering & Technology, Poornima University, Jaipur, Rajasthan, (India)

ABSTRACT

Wireless communication is the transfer of information between two or more points that are not connected by an electrical conductor. Due to their infrastructure less and fast deployment practices, they are vulnerable to many severe attacks. we address so many attacks as Denial of service attack , black hole attack, man-in-middle attacks, sinkhole attacks , jamming attacks and flooding attacks etc. Routing is a crucial factor of security in MANET and hence the focus of this thesis along with the performance analysis of routing protocols. The comparison analysis will be carrying out about these protocols using Simulator NS 3. In this paper we find the solution approaches which are Result Verification Mechanism for OLSR Routing Protocol in MANETs. Wireless Network is based on demand routing protocol to improve the performance and provide secure communication between two devices.

Key Words: MANET, OLSR, EE-OLSR, Denial Of Service At Attack, Black Hole Attack, Simulator NS 3

I. INTRODUCTION

A Mobile ad hoc network is a group of wireless mobile computers (or nodes) in which nodes collaborate by forwarding packets for each other to allow them to communicate outside range of direct wireless transmission. Ad hoc networks require no centralized administration or fixed network infrastructure such as base stations or access points, and can be quickly and inexpensively set up as needed. A MANET is an autonomous group of mobile users that communicate over reasonably slow wireless links. The network topology may vary rapidly and unpredictably over time, because the nodes are mobile. The network is decentralized, where all network activity, including discovering the topology and delivering messages must be executed by the nodes themselves [17]. A mobile ad hoc network is collection of self-configuring and adaption of wireless link between communicating devices (mobile devices) to form an arbitrary topology without the use of existing infrastructure. In wireless network technology, simulative analysis is a significant method to understand the performance of routing protocol. In this paper an attempt has been made to analysis performance of OLSR routing protocol for mobile ad hoc networks: The Proactive OLSR Routing Protocol offers quick adaptation to dynamic link conditions, low processing and memory overhead, low network utilization, and determines unicast routs to destinations within the ad hoc network. The performance differentials are analyzed using varying network load, mobility, and network size [8].

II. PERFORMANCE ISSUES IN MANET

Performance in Mobile Ad-Hoc Network (MANET) is the most important concern for the basic functionality of

network. To judge the quality of a protocol one needs to test them on the basis of metrics i.e. both qualitative and quantitative. These metrics are used to measure the suitability and performance of the different protocols. The metrics should be chosen carefully and should be independent of any routing protocol. The following is a list of all desirable qualitative properties of MANET as follows:-

- **Security:** - A MANET routing protocol is vulnerable to many forms of attack. They are more prone to security replay transmission, do spoofing threats than other general wired networks because the network structure is not strictly defined. Also a number of nodes keep on getting added as well as deleted from the network making it very easy for a malicious node to enter a network. Then it will be relatively easy for that node to snoop on network traffic, redirect traffic and flood the entire network. Security is very important to stop any kind of disruption of the network. There are three Attacks in security as follows:
- Denial of Service Attack Detection in MANET
- Black Hole Attack Detection in MANET
- Distributed Denial of Service Attack Detection in MANET
- Routing :-

§ OLSR Routing Protocol

III.PERFORMANCE EVALUATION MATRICES

Throughput-Throughput is the average rate of successful message delivery over a communication channel. Throughput is usually measured in bits per second (bits/sec), and sometimes in data packets per second or data packets per time slot. High throughput is always desirable in a communication system.

Packet Delivery Ratio: The ratio between the numbers of packets received by the TCP sink at the final destination and the number of packets originated by the “application layer” sources. It is a measure of efficiency of the protocol.

End to End Delay- End-to-end delay refers to the time taken for a packet to be transmitted across a network from source to destination. A data packet may take longer time to reach to the destination due to queuing and different routing paths.

Average Jitter- Jitter is the variation in delay by different data packets that reached the destination and can seriously affect the quality of audio/video and thus an unwanted parameter.

Average Queue Length- It is FIFO Queue Size (bytes) in MAC layers. The length of Queue depends on congestion and route discovery.

III SOLUTION APPROACHES AND RESULTS

OLSR Routing Based Analysis	
<i>Solutions</i>	<i>Result</i>
Mobility and Load Aware Algorithm using OLSR protocol	Packet delivery ratio is improved 10%.
EA-Willingness Setting mechanism usind OLSR protocol	More packets are delivered
profile based protection	Provides 100% performance

scheme (PPS)	compare to normal routing behavior of network
BAMBi technique	BAMBi achieve greater than 99% successful packet delivery rate
AODV routing protocol	Packet delivery ratio is improved and performance is evaluated over throughput
Multi-metric efficient routing scheme for the OLSR protocol	PDR has an improvement of up to 10% in the case of the modified OLSR
ML-OLSR Protocol	ML-OLSR improves the PDR by 8% comparing to original OLSR protocol
EE- OLSR protocol	0.097 msec Delay is less in EE-OLSR as compared to OLSR

IV.COMMON FINDINGS

- OLSR Routing Protocol have the advantage of finding a route between two nodes in the network in a very short time.
- EE-OLSR saves battery energy in a dense mobile network with high traffic loads.
- Mobility and Load aware OLSR (ML-OLSR) protocol achieves better performance over the OLSR protocol in terms of packet delivery ratio and average end-to-end delay where packet delivery ratio is 95%.
- When FAP takes effect at 600s, the performance becomes better and packet delivery rate keep up about 80%.
- Multi-metric energy efficient routing scheme for the OLSR protocol leads to a significant increase (5%-20%) in network lifetime compared to standard OLSR and slightly better performance in terms of Packet Delivery Ratio (PDR).
- Hierarchical OLSR (HOLSR) greatly reduces the required protocol overhead, which improves protocol scalability in large size heterogeneous networks.
- PA-OLSR protocol reflects more superior performance, and it can improve the success rate more than 5% highest
- There are no requirement of any special hardware like directional antennas, time synchronization, and any unwanted assumptions and any complex calculation in the AODV routing protocol.
- The packet delivery ratio of the MPOLSR-Fuzzy-WRR version is higher compared to original MP-OLSR and improvement of MP-OLSR-Fuzzy-WRR is up to 13.15% compared to original MP-OLSR.

- BAMBi technique can achieve more than 99% packet delivery success and suffers from very little false positives.

4.1 Performance Analysis of Olsr Routing In Mobile Ad Hoc Network

Objectives can be outlined as follows:

- To Design Network scenario for implementing existing OLSR Routing Protocol.
- To make certain modifications in OLSR Routing Protocol.
- To decide input/output parameters.
- To carry out Performance Analysis of Existing and modified routing protocol.

Selection of simulation parameter:

S no.	Parameter	Value
1	Network Size	1000*1000
2	Simulator	NS-3.11
3	Number of Nodes	25
4	Packet Size	1000 Bytes
5	Source Node	24
6	Sink Node	0
7	pkt interval	1/sec
8	Simulation time	100 sec
9	Mobility Model	Constant Position Mobility Model
10	Routing protocol	OLSR Routing
11	Traffic type	CBR

An experiment is performed to analsis the performance of OLSR routing protocol in mobile ad hoc network using NS3. Network Parameters are Taken that shows above table. The Performance analysis of OLSR routing is represented as follows.



I.

5.1 OLSR Routing Protocol in MANET



5.2 OLSR Routing Protocol in MANET

V RESULT

OLS Routing protocol is the representative of proactive protocol in MANET and it has become hot spot in recent years. In order to improve the performance of the network which has implemented OLSR protocol on it, we use NS3 simulation tool to simulate the network and compare the analysis result then the packet dropped Ratio is 90% in the Mobile Ad hoc Network using OLSR Routing Protocol and packet are more reliably send from source to destination node.

VI. CONCLUSION

The review of 20 research papers has been carried out in the area of Simulation Based Performance Analysis of OLSR Routing Protocol in Mobile ad hoc Network and find out current challenges and scope of work in the area. After the review, one issue was found which should be given proper concern, during the designing and implementation of performance analysis mechanism of routing protocols. The solution approaches under particular issues were studied in depth and were analyzed on the basis of various findings, which helped to understand the strengths and weaknesses of the solution approaches. OLSR routing protocol plays a key role to measure the performance of a MANET.

As initial objectives, some more specific research papers related to analysis performance of OLSR routing protocol were reviewed. We analyzed the performance investigation of performance of routing such as throughput, end to end delay and PDR.

Further on, we can evaluate the performance of proposed model of simulation enabled proactive protocols. Beside this, in Future, we will incorporate route break prediction in our proposed different-different routing protocols.

ACKNOWLEDGMENTS

I would like to express my deep gratitude and thanks to **Prof. Mahesh Bundele (Coordinator, Research), Poornima University** for giving me an opportunity to work under his guidance for review of research papers and his consistent motivation & direction in this regard. I would also express my sincere thanks to **Mr. Prashant Hemrajani Soni (Asst.Professor, CE), Poornima University** for their guidance and support.

REFERENCES

- [1] Amar Rasheed and Rabi N. Mahapatra, 2012, "The Three-Tier Security Scheme in Wireless Sensor Networks with Mobile Sinks", In IEEE, IEEE Transactions on Parallel and Distributed Systems, VOL. 23, NO. 5, pp. 958-967, 2012.
- [2] Alexandra Czarlinska, Deepa Kundur, 2006 " Distributed Actuation Attacks in Wireless Sensor Networks: Implications and Countermeasures", Second IEEE Workshop on Dependability and Security in Sensor Networks and Systems, 2006.
- [3] Chan Yan, Song Nan-nan, 2010, "The research on coal mining security monitoring system based on wireless sensor networks", In IEEE, Science and technology in Kushan city of China , Issue on monitoring in wireless sensor networks, 2010.
- [4] Debao Xiao , Chao Chen, Gaolin Chen, 2005, "Intrusion Detection based Security Architecture for Wireless Sensor Networks", In IEEE, Proceedings of ISCIT, 2005.

- [5] Dmitriy Martynov, Jason Roman, Samir Vaidya, and Huirong Fu, "Design and Implementation of an Intrusion Detection System for Wireless Sensor Networks", IEEE EIT 2007 Proceedings, 2007.
- [6] Dusit Niyato, Zhu Han, Walid Saad, and Are Hjorungnes, 2010, "A Controlled Coalitional Game for Wireless Connection Sharing and Bandwidth Allocation in Mobile Social Networks", In IEEE, IEEE Communications Society subject matter experts for publication in the IEEE Globecom proceedings, 2010.
- [7] Erik Kline, Alexander Afanasyev, Peter Reiher, "Shield: DoS Filtering Using Traffic Deflecting", 19th IEEE International Conference on Network Protocols, pp 37-42,2011.
- [8] Larry Korba, 1998, "Security System for Wireless Local Area Networks", In IEEE, National Research Council of Canada, Issue on security of WLAN, 1998, pp1550-1554.
- [9] Ha Cheol Lee, 2003, "The Relation of Cell and Call Connection Level GoS for the Random Traffic in the Wireless ATM Access Transport", In IEEE, Proceedings of ICCT, pp. 886-890, 2003
- [10] Hany Samul, Weihua Zhuang, Bruno Presis, 2008, "Routing over Interconnected Heterogeneous Wireless Network with Intermittent Connections", In IEEE, IEEE Communication Society, Volume 56, routing in Wireless network, pp 2282-2286,2008.
- [11] Harish Sitaraman, Jesse M. Davis, Joseph B. Evans, 2001, "Broadband Wireless Local Loop Evaluation Via an Active TCP Connection Monitoring Methodology", In IEEE, IEEE Communications Magazine, Issue on Monitoring in Wireless network, pp 594-603, 2001.
- [12] Hassen Redwan and Ki-Hyung Kim, 2008, "Survey of Security Requirements, Attacks and Network Integration in Wireless Mesh Networks", In IEEE, Issue on security on Wireless Mesh Networks, Japan-China Joint Workshop on Frontier of Computer Science and Technology, pp. 3-9, 2008.
- [13] Jun-Zhao, 2001, "Mobile Ad Hoc Networking :An Essential Technology for Pervasive Computing", In IEEE, IEEE Journal on Selected Areas in Communications, Issue on routing in MANET, pp 316-321,2001.
- [14] Khusvinder Gill and Shuang-Hua Yang, "A Scheme for Preventing Denial of Service Attacks on Wireless Sensor Networks", In IEEE, 2009.
- [15] Hany Samul, Weihua Zhuang, Bruno Presis, 2008, "Routing over Interconnected Heterogeneous Wireless Network with Intermittent Connections", In IEEE, IEEE Communication Society, Volume 56, routing in Wireless network, pp 2282-2286,2008.
- [16] Lititng Cao, Shriu Zhou, 2013 "Intelligent Security system based on Wireless Sensor Networks", In IEEE, First international Conference on Intelligent Networks and Intelligent Sytems, PP,143-146,2013
- [17] Maneesha V. Ramesh, Aswathy B. Raj and Hemalatha T., 2012, "Wireless Sensor Network Security: Real-Time Detection and Prevention of Attacks", In IEEE, Fourth International Conference on Computational Intelligence and Communication Networks, pp. 783-787, 2012.
- [18] Mercy J, Abinaya K, 2010, "Security Model Evaluation of 3G Wireless Network", IEEE, Volume 5, Issue on Security on 3G Wireless Network, pp 450-466,2010.
- [19] Muhan Shoaib Siddiqui, Syed Obaid Amin, Jin Ho Ki, Choong Seon Hong, 2010, "MHRP : A Secure Multi-path Hybrid Routing Protocol for Wireless Mesh Network", Next Generation Communication and Wireless Mesh Network, In IEEE, Issue on Routing in Wireless Mesh Networks, 2011
- [20] Mohammad Saiful Islam Mamun, A.F.M. Sultanul Kabir, "Hierarchical design based intrusion detection system for wireless adhoc sensor", International Journal of Network Security & Its Applications (IJNSA), Vol.2, No.3, pp. 102-117, July 2010.

ATTACK DETECTION IN WIRELESS NETWORKS

Usha Dagar¹, Uma Dagar²

^{1,2}Research Scholar, School of Engineering & Technology, Poornima University,
Jaipur, Rajasthan, (India)

ABSTRACT

Wireless communication is the transfer of information between two or more points that are not connected by an electrical conductor. A wireless network is any type of computer network that uses wireless data connections for connecting networks nodes and networks are free to organize by themselves . Due to their infrastructure less and fast deployment practices, they are vulnerable to many severe attacks. we address so many attacks as replication attacks (clone attacks) , sybil attacks, man-in-middle attacks, sinkhole attacks , jamming attacks and flooding attacks e.t.c

Therefore need to generate a system which is free from attacks . To learn about the design and implementation aspects of security of wireless networks, a review process involving 3+ stage approaches has been undertaken for 41 research papers which were published in the period of year 1995 to year 2014. a review process involving 2 stage approach has been undertaken for 30 research papers which were published in the period of year 1998 to year 2013. After an exhaustive review process, Three issues were found, which must be addressed while the developing of skrowten sseleriw. Several solution approaches have been found in the 43 papers. The outcome of the review was in the form of various findings, found under various key issues. The findings included algorithms and methodologies used to solve particular research problem, along with their strengths and weaknesses and the scope for the future work in the area.

Key words: Ad-hoc networks, Security, Wormhole attack, Tunnel, Wireless, MANETs, AODV, Sybil attack, Denial of service attacks, Packet Leashes, HEAP, RTT, Cryptography, FMS, Tunnel.

I. INTRODUCTION

This paper describes different techniques addressing this challenge at different levels of abstraction. In the past, the major concerns of Wireless security is the prevention of unauthorized access or damage to computers using wireless networks. To provide security several techniques including cryptographic, encryption, authentication, key management, attacks and secure routing, intrusion detection are discussed .T sledon fo stol neeb evah eresh -ekil gnituor dna ytiruces rof smsinahcem dnaCryptographic algorithms ,MANET Routing But these technologies are not enough to concern about the security issues.Systems are not enough to provide high speed data and security so LEACH-S and RSD algorithm are used prevented the network from a large number of nodes failure caused by Sybil attacks used to detect and remove Sybil attacks TWPM provides an effective and accurate worm propagation model for detecting worm attacks.

II. REVIEW PROCESS ADOPTED

The review process is divided into five stages in order to make the process simple and adaptable by every researcher. As it reflects from the literature that while beginning the finding of research objectives, it is necessary to start with a broader domain of any area / sub area of interest and narrow down to specific issue, the

process described in the diagram includes the narrowing down. We have followed one of the typical processes to make a literature review and frame the objectives of research. The process diagram is shown in Fig. 1, which includes all five stages defined as under:

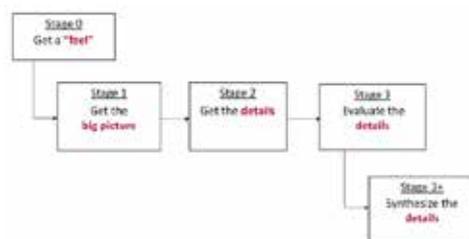


Fig1: Literature Review Process

Stage 0: Get the Feel

Stage 1: Get Big Picture

Stage 2: Get the Details

Stage 3: Evaluate the Details

Stage 3+: Synthesize the Details.

The review process was divided into five stages in order to make the process simple and adaptable.

Stage 0: Get a “feel”: This stage is the beginning of literature review process wherein one has to broadly select his / her area of interest and start searching the scientific research papers from valid sources. I have gone through large number of research papers and selected 25 research papers published in various research journals using this stage.

Stage 1: Get the “big picture”: In order to understand the paper broadly and get an idea whether the paper exactly belong to the research area / sub area selected or it deviates, if deviates how much, these concepts are made clear this stage, known as Get Big Picture. It helps in classifying the research papers according to key issues and narrowing down the survey. Using this stage 25 research papers were clustered according to regions and the type analysis presented by the researchers.

Stage 2: Get the “details”: Stage 2 deals with going in depth of each research paper and understand the details of methodology used to justify the problem. Total 25 paper’s data has been extracted under this process pertaining to my research focus.

Stage 3: “Evaluate the details”: This stage evaluates the details in relation to significance of the problem, Novelty of the problem, significance of the solution, novelty in approach, validity of claims etc. All the selected papers have been processed through this stage.

Stage 3+: “Synthesize the detail”: This stage deals with synthesis of the data, concept & the results presented by the authors. Here imagination of situations different from those presented and expected results has to be predicted. In short it uses interpolation and extrapolation to find out the gaps in the published research. I could exercise this process for all the 26 research papers, of which the final findings and the outcomes have been discussed in the next chapter. This process helped me to easily classify the literature, evaluate, synthesize, summarize and build the problem statement of my research proposal.

III. METHODOLOGIES/SOLUTION APPROACHES

After reviewing the 41 research paper we categories the review in three major issues according to solution approach listed as :

1. Sybil attacks detection
2. Wormhole attacks detection
3. Dos attacks detection

IV. FINDINGS OF SYBIL ATTACK DETECTION

- ✓ CbTM mechanism performs trust management and reputation checker to verify coverage nodes in the participatory sensing.
- ✓ Received Signal Strength Indicator (RSSI) values defines difference between sender within 10 meters in the Identity based attacks.
- ✓ SDTM method does not burden the network with additional communication overhead and excessive computational power.
- ✓ Sybil ACO limits the attack edges between the honest nodes and malicious nodes via ant colony optimization and social network.
- ✓ FMS monitors the features and characteristics of the zones and underlying nodes which can detect the deviations and fixes the malicious node in sequential manner. A LEACH-S mechanism improves the security performance of LEACH protocol against Sybil attacks using the cluster heads number collection.
- ✓ A regional statistics detection scheme (RSDs) against Sybil attacks protocol can prevented the network from a large number of nodes failure caused by Sybil attacks.
- ✓ An identification algorithm based on state information of nodes to detect all Sybil nodes by calculating the distances between the suspected node and its neighbor nodes coordinates.
- ✓ Cross layer detection technique avoids Sybil and collision attack caused in the data link and network layer.
- ✓ Ranging method has costs less, no use for high accurate range.

V. FINDINGS OF WORMHOLE ATTACK DETECTION

- ✓ Signal prints allows to identify the sender based on the received signal characteristics so making it harder to spoof.
- ✓ By cross-layer security optimization DoS attack could be minimized and Dos attack detection rate is increased and network performance also increased.
- ✓ A self –protection system track wireless sensor networks state machine violations and use dynamically updated anomaly detection to detect attacks.
- ✓ Symmetric-key algorithm improves performance instead of Neural Networks for detecting DoS attack.
- ✓ In the cross layer detection of wormhole attack using the Fuzzy Inference System (FIS) in which wormhole attack detected without disturbing any nodes in the network and interfering with the route establishment process.
- ✓ SDTM method does not burden the network with additional communication overhead and excessive computational power.

- ✓ In dynamic wormhole detection method every node responsible to find out there any worm hole between that nodes to it's next to next Node.
- ✓ TWPM provides an effective and accurate worm propagation model for sensor networks and simultaneously captures the time and space dynamics of worms spreading over a sensor network.

VI. FINDINGS OF DOS ATTACK DETECTION

- ✓ By cross-layer security optimization DoS attack could be minimized and Dos attack detection rate is increased and network performance also increased.
- ✓ A self –protection system track wireless sensor networks state machine violations and use dynamically updated anomaly detection to detect attacks.
- ✓ Symmetric-key algorithm improves performance instead of Neural Networks for detecting DoS attack.

VII. ISSUE WISE SOLUTION APPROACHES

Issue	Solutions	Ref
SYBIL ATTACK DETECTION		
	Cooperative RSS-based Sybil Detection(CRSD)	[12]
	A regional statistics detection scheme (RSDs) against Sybil attacks protocol.	[3]
	Ranging based algorithm	[38]
	SDTM (Sybil attack Detection using Traffic Monitoring)	[15]
	Cloud based Trust Management Scheme (CbTMS) framework	[18]
	Ant colony optimiDtion (ACO) algorithm	[11]
	The VaultIC420 security module	[9]
Wormhole Attack Detection		
Wormhole Attack Detection	The Fuzzy Inference System (FIS)	[21]

	A hybrid technology using an divide and conquer technique using AODV Protocol.	[25]
	Wormhole Attack detection in geographic routing protocols using Markov Chain model.	[16]
	Modified wormhole detection AODV protocol	[16]
	AODV protocol	[12]
	AODV routing protocol using HEAP.	[18]
DOS Attack Detection	A cross- layer security technique.	[29]
	Topologically -aware worm propagation model (TWPM)	[12]
	A Clustering-based intrusion detection method.	[18]
	A Wireless Self-organized Networks)WSoN)	[12]
	A method of signal prints to detect identity-based attacks.	[15]

VIII. COMMON FINDINGS

- ✓ Several approaches – LEACH-S and RSD algorithm are used to prevented the network from a large number of nodes failure caused by Sybil attacks using RSSI value.
- ✓ TWPM provides an effective and accurate worm propagation model for detecting worm attacks. Cross layer detection techniques achieve a high accuracy in predicting and defending the network against all Denial of Service attack .
- ✓ FIS system has low false alarm rate most of the wormholes present in the network are detected accurate.
- ✓ By the VaultC420 energy is saved compared to a software based implementation when using ECDSA as signature algorithm. But these technologies are not enough to concern about the security issues. Systems

are not enough to provide high speed data and security. Signal prints technique is particularly suitable for dense networks and requires large networks. The wormhole detection technique having some overhead, control packet is also increases. Underwater security problems explore more efficient counter measures to meet the new challenges in the future.

IX. STRENGTHS

- ✓ FIS system successfully detects the wormhole present in the sensor network using network and physical layer parameters.
- ✓ FMS (Feature monitoring system) approach has no need to maintain any hardware requirements like directional antennas, time synchronization, and any unwanted assumptions.
- ✓ Wormhole detection technique detects the wormhole, and will serve as improved form of existing AODV protocol.
- ✓ Cross layer detection techniques achieves a high accuracy in predicting and defending the network against all Denial of Service attack.
- ✓ The nature of the ACO can limit the number of attack edge in an efficiently and conveniently.
- ✓ Ranging method has less cost, no use for high accurate range but it has extremely precise.
- ✓ Ranging is particularly suitable for the low-cost, lacking resources of wireless sensor network.
- ✓ Advantage of the analytical model is offer great insights for new worm detection techniques.
- ✓ Intrusion Detection System(IDS) provides higher detection rate with a very low false positive rate which generate a realistic traffic pattern for accurate evaluation of protocols.

X. LIMITATIONS

- ✓ Signal prints technique is particularly suitable for dense networks and applications of signal prints faces problem of battery discharge and lack of a central authority.
- ✓ Cross layer detection techniques secure the network leaving room only for attacks by tampering with the sensor physically.
- ✓ In Wireless Self-organized Networks(WSoN) the Intrusion Detection System(IDS) provide very high pulse rate can cause collision at the receiving node, leading to packet drops.
- ✓ Security attacks and security requirements still several types of attacks happens which are needed to resolve.
- ✓ VaultIC420 requires a duty-cycling mechanisms to achieve any savings compared to the software implementation which more expensive.
- ✓ Realization of open wireless architecture system is required. the wormhole detection technique having some overhead, control packet is also increases.
- ✓ AODV protocol doesn't adapt to network which the number of nodes are less than 50.

XI. SCOPE OF WORK

- ✓ CRSD is first step to design an general and effective way to resist against the Sybil attack and more investigations are needed under the general context, e.g., multiple malicious node.

- ✓ The accuracy of the FIS system with 100% detection ratio and improve the security of the network to avoid wormhole attacks in network is next step in process.
- ✓ More duty-cycling MAC layer protocols are required and a study of possible queuing and power saving mechanisms are required to improve the VaultIC420 duty-cycling.
- ✓ In order to detect wormhole proposed technique use larger number of control packet in future we will try negotiates that effect.
- ✓ Based on the performance results, we will continue to study more and more underwater security problems, and explore more efficient counter measures to meet the new challenges in the future.
- ✓ wormhole proposed technique use larger number of control packet in future we will try negotiates that effect.

XII. CONCLUSION

Here after finding the solution approaches I found issues were Sybil attacks, Wormhole attacks and Dos attacks which should be given proper concern, when the enhancement of proper communication of messages from sender to receiver takes place. This paper is a survey of wireless sensor networks security issues that affect the network, Authenticity, Integrity, secure communication. Wireless sensor networks have the potential to bring about the Computation Revolution, in which large-scale computations become universally accessible, affordable and useful. The exhaustive review could finally lead to extract findings in the area, strengths and weaknesses and scope of work during M. Tech thesis.

XIII. ACKNOWLEDGMENTS

We would like to express my deep gratitude and thanks to **Prof. Mahesh Bunde**, Coordinator, M. Tech. and Guide **Asst. prof. Prashant Hemrajani** for giving me an opportunity to work under his guidance for review of research papers and his consistent motivation & direction in this regard. I would also express our sincere thanks to **Mr. Devendra Kumar Somwanshi**, Associate Professor, M Tech, Poornima University for their guidance and support.

REFERENCES

- [1] Larry Korba, 1998, "Security System for Wireless Local Area Networks", In IEEE, National Research Council of Canada, Issue on security of WLAN, 1998, pp 1550-1554.
- [2] Upkar Varshney, 1998, "Design and Performance Evaluation of Connection Routing Schemes for Wireless ATM Environment", In IEEE, International Conference on Universal Personal Communications (ICUPC), pp. 409-415, 1998.
- [3] Upkar Varshney, 1999, "Connection Routing Schemes for Wireless ATM", In IEEE, Proceedings of the 32nd Hawaii International Conference on System Sciences, 1999.
- [4] Jun-Zhao, 2001, "Mobile Ad Hoc Networking : An Essential Technology for Pervasive Computing", In IEEE, IEEE Journal on Selected Areas in Communications, Issue on routing in MANET, pp 316-321, 2001.
- [5] Harish Sitaraman, Jesse M. Davis, Joseph B. Evans, 2001, "Broadband Wireless Local Loop Evaluation Via an Active TCP Connection Monitoring Methodology", In IEEE, IEEE Communications Magazine, Issue on Monitoring in Wireless network, pp 594-603, 2001.

- [6] Prashant Krishnamurthy, Joseph Kabara, Tanapat Anusas-amornkul, 2002, "Security in Wireless Residential Networks", In IEEE, IEEE Transactions on Consumer Electronics, Vol. 48, pp.157-166, 2002.
- [7] Ha Cheol Lee, 2003, "The Relation of Cell and Call Connection Level GoS for the Random Traffic in the Wireless ATM Access Transport", In IEEE, Proceedings of ICCT, pp. 886-890, 2003.
- [8] Debao Xiao , Chao Chen, Gaolin Chen, 2005, "Intrusion Detection based Security Architecture for Wireless Sensor Networks", In IEEE, Proceedings of ISCIT, 2005.
- [9] Shi Zhong, Taghi M. Khoshgoftaar and Shyarn V. Nath , 2005, "A Clustering Approach to Wireless Network Intrusion Detection", In IEEE, Issue on security of wireless networks, Vol.2, No.3, 2005.
- [10] Wider D. YU, Vijaylaxmi Jothiram, 2007, "Security in Wireless mobile Technology for Health care System", In IEEE, Issue on security of wireless mobile technology, 2007, pp,308-311, 2005.
- [11] Hany Samul, Weihua Zhuang, Bruno Presis, 2008, "Routing over Interconnected Heterogeneous Wireless Network with Intermittent Connections ", In IEEE, IEEE Communication Society, Volume 56, routing in Wireless network, pp 2282-2286, 2008.
- [12] Qian Zhang, Ya-Qin Zhang, 2008, " Cross-Layer Design for QoS Support in Multihop Wireless Networks", IEEE, Volume 96, No. 1, Issue on Quality of service in Multihop networks, 2008.
- [13] Hassen Redwan and Ki-Hyung Kim, 2008, "Survey of Security Requirements, Attacks and Network Integration in Wireless Mesh Networks", In IEEE, Issue on security on Wireless Mesh Networks, Japan-China Joint Workshop on Frontier of Computer Science and Technology, pp. 3-9, 2008.
- [14] Nilesh Ananthanarayanan, 2008, "Vehicle Monitoring System in Wireless Networks", In IEEE, Issue on Monitoring of Wireless Networks, Vol. 26 ,no.5, pp. 141 -146, 2008.
- [15] Xiaoyan Hong, Bo Gu and Pu Wang, 2009 "Modeling Worm Propagation through Hidden Wireless Connections", In IEEE, IEEE Communications Society subject matter experts for publication in the IEEE Globecom proceedings, 2009.
- [16] Dusit Niyato, Zhu Han , Walid Saad, and Are Hjørungnes, 2010, "A Controlled Coalitional Game for Wireless Connection Sharing and Bandwidth Allocation in Mobile Social Networks", In IEEE, IEEE Communications Society subject matter experts for publication in the IEEE Globecom proceedings, 2010.
- [17] Mercy J, Abinaya K, 2010, "Security Model Evaluation of 3G Wireless Network", IEEE, Volume 5, Issue on Security on 3G Wireless Network, pp 450-466, 2010.
- [18] Chan Yan, Song Nan-nan, 2010, "The research on coal mining security monitoring system based on wireless sensor networks", In IEEE, Science and technology in Kushan city of China , Issue on monitoring in wireless sensor networks, 2010.
- [19] Muhan Shoaib Siddiqui, Syed Obaid Amin, Jin Ho Ki, Choong Seon Hong, 2010, "MHRP : A Secure Multi-path Hybrid Routing Protocol for Wireless Mesh Network", Next Generation Communication and Wireless Mesh Network, In IEEE, Issue on Routing in Wireless Mesh Networks, 2010
- [20] Mohammad Saiful Islam Mamun, A.F.M. Sultanul Kabir, "Hierarchical design based intrusion detection system for wireless adhoc sensor ", International Journal of Network Security & Its Applications (IJNSA), Vol.2, No.3, pp. 102-117, July 2010.
- [21] Siba K. Udgate, Alefiah Mubeen and Samrat L. Sabat, 2011, "Wireless Sensor Network Security model using Zero Knowledge Protocol", In IEEE, IEEE Communications Society subject matter experts for publication in the IEEE ICC proceedings, 2011.
- [22] Maneesha V. Ramesh, Aswathy B. Raj and Hemalatha T., 2012, "Wireless Sensor Network Security: Real-Time Detection and Prevention of Attacks", In IEEE, Fourth International Conference on Computational Intelligence and Communication Networks, pp. 783-787, 2012.
- [23] Amar Rasheed and Rabi N. Mahapatra, 2012, "The Three-Tier Security Scheme in Wireless Sensor Networks with Mobile Sinks", In IEEE, IEEE Transactions on Parallel and Distributed Systems, VOL. 23, NO. 5, pp. 958-967, 2012.

- [24] Rakesh Bisht, Naveen Chauhan, "A Traffic-Load Adaptive Power-Conserving MAC Protocol for Wireless Sensor Network", IEEE2012, Volume 6, Issue on Quality of Service in Wireless Sensor Network, 2012.
- [25] Okoli. Adaobi, Ejiro Igbesoko, Mona Ghassemain, "Evaluation of Security Problems and Instruction Detection System for Routing Attacks in Wireless Self-Organised Wireless Network", In IEEE, International Journal of Communications and Information Technology, Volume1, Issue on Security in Wireless Network, pp 16-20, 2012.
- [26] Rashween Kaur Saluja, Rajesh Shrivastava, 2012, "A Scenario Based Approach for Gateway Discovery using Manet Routing Protocol", In IEEE, International Conference on Computer Communication and Informatics (ICCI-2012), issue on Routing in Mobile Ad Hoc Network, 2012.
- [27] Yan Yu, Yubao Yao, "Improved AODV Routing Protocol for Wireless Sensor Network and Implementation using OPNET", In IEEE, Third International Conference on Intelligent Control and Information Processing July 15-17, 2012-Dailan China, Issue on Routing in Wireless Sensor Network, pp 709-713, 2012.
- [28] Shi-Zong, Shyam V. Nath, 2012 "A Clustering Approach based to Wireless network intrusion detection system", Issue on security on wireless networks, IEEE, vol. 35, pp.99-101, 2012.
- [29] Woosik Lee, Namgi Kim, 2013, "A study of Transmission Power Control Algorithms in wireless body sensor networks", In IEEE, Soft tech 2013, ASTL, Vol. 19, pp. 31-34, 2013.
- [30] Liting Cao, Shriu Zhou, 2013 "Intelligent Security system based on Wireless Sensor Networks", In IEEE, First international Conference on Intelligent Networks and Intelligent Systems, PP, 143-146, 2013.
- [31] Chan Yan, Song Nan-nan, 2010, "The research on coal mining security monitoring system based on wireless sensor networks", In IEEE, Science and technology in Kushan city of China, Issue on monitoring in wireless sensor networks, 2010.
- [32] Muhan Shoaib Siddiqui, Syed Obaid Amin, Jin Ho Ki, Choong Seon Hong, 2010, "MHRP : A Secure Multi-path Hybrid Routing Protocol for Wireless Mesh Network", Next Generation Communication and Wireless Mesh Network, In IEEE, Issue on Routing in Wireless Mesh Networks, 2010
- [33] Mohammad Saiful Islam Mamun, A.F.M. Sultanul Kabir, "Hierarchical design based intrusion detection system for wireless adhoc sensor", International Journal of Network Security & Its Applications (IJNSA), Vol.2, No.3, pp. 102-117, July 2010.
- [34] Siba K. Udgate, Alefiah Mubeen and Samrat L. Sabat, 2011, "Wireless Sensor Network Security model using Zero Knowledge Protocol", In IEEE, IEEE Communications Society subject matter experts for publication in the IEEE ICC proceedings, 2011.
- [35] Maneesha V. Ramesh, Aswathy B. Raj and Hemalatha T., 2012, "Wireless Sensor Network Security: Real-Time Detection and Prevention of Attacks", In IEEE, Fourth International Conference on Computational Intelligence and Communication Networks, pp. 783-787, 2012.
- [36] Amar Rasheed and Rabi N. Mahapatra, 2012, "The Three-Tier Security Scheme in Wireless Sensor Networks with Mobile Sinks", In IEEE, IEEE Transactions on Parallel and Distributed Systems, VOL. 23, NO. 5, pp. 958-967, 2012.
- [37] Rakesh Bisht, Naveen Chauhan, "A Traffic-Load Adaptive Power-Conserving MAC Protocol for Wireless Sensor Network", IEEE2012, Volume 6, Issue on Quality of Service in Wireless Sensor Network, 2012.
- [38] Okoli. Adaobi, Ejiro Igbesoko, Mona Ghassemain, "Evaluation of Security Problems and Instruction Detection System for Routing Attacks in Wireless Self-Organised Wireless Network", In IEEE, International Journal of Communications and Information Technology, Volume1, Issue on Security in Wireless Network, pp 16-20, 2012.
- [39] Rashween Kaur Saluja, Rajesh Shrivastava, 2012, "A Scenario Based Approach for Gateway Discovery using Manet Routing Protocol", In IEEE, International Conference on Computer Communication and Informatics (ICCI-2012), issue on Routing in Mobile Ad Hoc Network, 2012.
- [40] Yan Yu, Yubao Yao, "Improved AODV Routing Protocol for Wireless Sensor Network and Implementation using OPNET", In IEEE, Third International Conference on Intelligent Control and Information Processing July 15-17, 2012-Dailan China, Issue on Routing in Wireless Sensor Network, pp 709-713, 2012.

WEB SEARCH RESULTS USING QUERY-SPECIFIC KEYWORDS

Ms.M.A.Gunavathie (Assistant Professor)

Jayashree K¹, Jayashree R², Mahalakshmi J³, SaiJanaranjani C⁴

^{1,2,3,4}Information Technology, Panimalar Engineering College, Chennai (India)

ABSTRACT

To solve automated Question Answer system we propose preprocessing algorithm, stemming algorithm and partial match retrieval algorithm. In existing system we search randomly by using user keyword, and then it achieves the results. The result Classified as normally text, image and video results randomly based on the frequent mining. The proposed scheme to find the answer Using Google API technology. For different type of questions we answer only using text, here in our proposed scheme, the user inputs query which is analyzed and processed using algorithm. This is to enrich the answer with multimedia content like images and videos then the user will understand the answer easily. We are using the Google search engine as a tool so the answer will be obtained from web data and we need to analyze and re-rank the search results. That results as classified into multimedia, text, image and video results separately based on the proposed algorithm such stemming and partial match retrieval algorithm to achieve the better results.

Key Words: Data preprocessing, Multimedia Question Answers (MMQA), Partial match retrieval, Stemming, Web results reranking.

I. INTRODUCTION

Web mining - is an act of applying data techniques to discover patterns from the Web. Based on analysis targets, web mining can be classified into three unique types, which are Web usage mining, Web content mining and Web structure mining. Web usage mining is the process of getting useful information from server logs e.g. use Web usage mining is the act of finding out what users are looking for on the Internet. Some of the users will be looking at only textual data, whereas other users will desire multimedia data. Web Usage Mining is the act of applying data mining techniques to discover attractive usage patterns from Web data for better understanding and serve the needs of Web-based applications. Usage data gets the origin of Web users along with the user's browsing behavior at a Web site. Web usage mining can be arranged in similar groups depending on the kind of usage data considered. Web Server Data is the one in which the user logs are collected by the Web server. The data contains IP address, page reference and time of access. Application Server Data: Commercial application servers have important features to enable e-commerce applications to be built on top of them with minimal effort. It is used in confirming the data and validity verification, data integrity and building classifications, content management, content generation and opinion mining.

II. APPROACH OVERVIEW

QUESTION-ANSWERING (QA) is the method for answering a query in an English language. When compared to other systems we are trying to communicate easily between users and computers. It avoids the data contents that are vast in quantity which are displayed as links in search engines instead of getting the exact answers. For complicated questions we are providing the idea to generate the answers. In many cases the results are good which are selected by users. The information gainers are gaining the information for certain specific questions in any topic and get answers.

2.1. Need for the project

The Google search engines return the vast quantity of information the users can easily become overwhelmed. Question-answering (QA) research attempts to tackle this information which is overloaded with problem. Instead of returning a whole set of documents or links, the current search engines return only the exact answers for that query. For returning a perfect answer the search engines make use of algorithm. However, the QA research has largely focused on text. The vast amount of information on the Web is mostly available in multimedia form. We can also change the text based QA to multimedia QA. Some image based answers are available for some question. The three main components of QA framework are document analyzing, question identification, and answer achieving techniques. Research is done based on the three components and has achieved success. Yahoo answers have become a powerful way of analyzing answers. The important need of this project is to get an exact answer that is in single search. Getting answers is based on three methods. First one is based on text, second one is based on video, and the last one is based on text+video. In this project the answer will be provided based on the user's need.

2.2. Objective and Scope of the Project

Question-Answering is a technique for automatically answering a question fixed in usual language. Compared to keyword-based search systems, it greatly enhances the communication between humans and computers by naturally stating user's intention in plain sentences. It also avoids the painstaking browsing of a vast quantity of information contents returned by search engines for the correct answers. Fully automated question answering still faces a challenge that is not easy to tackle, as the deep understanding of complex questions and the sophisticated syntactic, semantic and contextual processing to produce answers. In this system, we use MMQA that automatically determines which type of media information should be added for textual answer by collecting data from web to enrich the answer. In this process extracting the keyword from the given searched query is the first step for identifying the input question type. In other cases, clear words that indicate the question type directly, i.e. "who", "where", or "how many", these words tell the system that the answer should be of type "person", "location", "Number" respectively. Once the system identifies which type of the question, an informative retrieval system is used to find a set of documents containing the correct key words. In this project each question will be answered with not only with text but also with corresponding images and videos if required which provides better understanding to the user. The main objective of the system is to get an enriched textual answer. User will get the information in an easily understandable manner.



Fig 3.1: Traditional QA framework

III. PROPOSED WORK

3.1. Block diagram

The block diagram of the entire project is shown in fig. 4.1 It is composed of three major techniques.

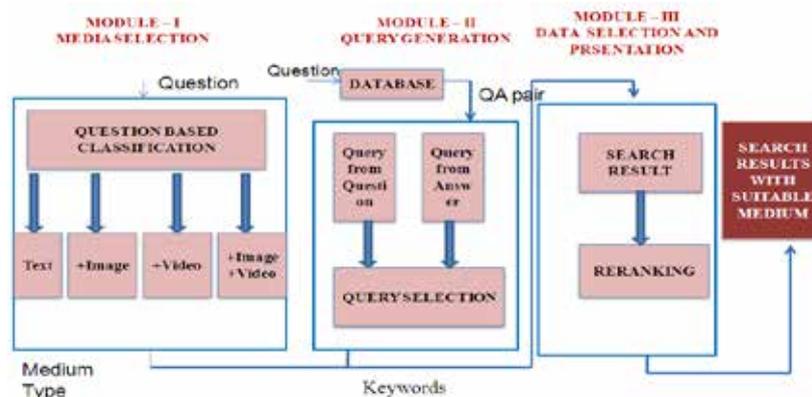


Fig 3.1: Proposed system Block diagram

3.1.1. User Interface

This module is used to get the input query from the user. The user has to interact with the server with his queries and the server will opt for the optimal solution with the user understandability. The user who wants the multimedia enriched output he will enter his query in this search space. All search queries pattern will be stored with the output patterns in the database for further process

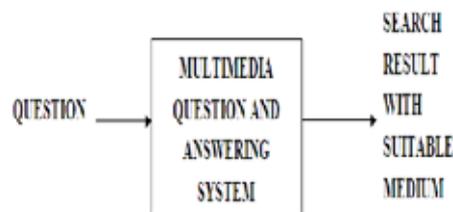


Fig 3.2 Proposed System Input and Output

3.1.2. Query Analysis

Data pre-process: In data pre process module we are going to exclude stop words and other unused data from the search query.

Answered Query: To reduce the server process if the current user search query matches with the existing search query pattern then the response will come from the database result set. Questions are categorized into

following types, yes/no class, choice class, quantity class, enumeration class, description class and store the values in the database.

3.1.3. Pattern Selection

Text: For the “yes/no”, “choice” and “quantity” questions, we classify them into the class of answering with only text. Therefore, given a question, we first judge whether it should use only textual answer based on the interrogative word.

Text & image: If the question is classified into “enumeration” and “description” class then the answer medium will be selected as “text + image” when we search answers in web we will add text search answers and the image search answers for the input query.

Text & Video: The verbs in an answer will be very useful for determining whether the answer can be enriched with video content. In certain cases, if a textual answer contains many complex verbs, it is more likely to explain a changing process and thus it has high probability to be satisfactorily answered by videos. Therefore, verb can be a significant clue.

Text & Image & Video: The verbs in an answer will be useful for judging whether the answer can be enriched with stream of images as well as video content. If the questions contains following string then the answer medium is selected as “text+image+videos” they are president, king, prime minister, kill, issue, nuclear, earth quake, singer, battle, event, war, happened.

3.1.4. Web Search Navigator

Search is useful to visitors who know particularly what they’re trying to locate or discover. But having a search option as a part isn’t an exemption to ignore good information architecture. It’s still of great significance to make sure that your content is locatable for visitors who might not know particularly what they’re looking for or are browsing to discover potentially desirable content.

3.1.5. Query Result (using Google API)

A search engine results page (SERP) is the listing of results returned by a search engine as a result a keyword query. A SERP may refer to only one page of links obtained, or to the set of all links obtained for a search query.

3.2. Used Algorithms and Techniques

3.2.1. Stemming Algorithm

A stemming algorithm is a process related to language normalisation, in which the different forms of a word are reduced to a simple form. It is significant to acknowledge that we use stemming with the purpose of rising the performance of IR In fact from an historical or grammatical viewpoint, a stemming algorithm is likely to make many errors. In these languages words tend to be constant at the front, and are varying at the end.

Management

Manager

Managing ---> manage

Develop the algorithm up bit by bit, trying out limited number of ending removals at a time. For each new ending plus rule added, determine whether, on average, the stemming process is raised or degraded. The words of each entering query would then be stemmed in the same way. When the index terms were viewed by the user, for example during query expansion, they would be visible in their stemmed form. A better approach is to keep each word, X, and its stemmed form, s(X), as a two-way relation in the IR

system. X is held in the index with its own posting list. s(X) could have its separate posting list, but this would be derivable from the class of words that stem to s(X). The important thing is to have the $X \leftrightarrow s(X)$ relation. From X we can derive s(X), the stemmed form. From a stemmed form s(X) we can derive X plus the other words in the IR system which stem to s(X).

Stop words:

The stop word list connects in several ways with the stemming algorithm. The stemming algorithm can itself be used to identify and remove stop words. One would add into the improper forms table something like this,

```
"" , /* null string */
"am/is/are/be/being/been/" /* BE */
"have/has/having/had/" /* HAD */
"do/does/doing/did/" /* DID */
... /* multi-line string */
```

so that the words 'am', 'is' etc. map to the null string. Alternatively, stop words could be eliminated before the stemming algorithm is applied, or after the stemming algorithm is applied. In this latter case, the words to be discarded must themselves have gone through the stemmer, and the number of well defined forms will be greatly reduced as a result.

3.2.2. Naïve Bayes Classifier

Naive Bayes classifiers can manage an absolute number of independent variables whether continuous or categorical. Given a set of variables, $y = \{y_1, y_2, y_3, \dots, y_n\}$, we want to build the posterior probability for the event F_j among a set of possible outcomes $F = \{f_1, f_2, f_3, \dots, f_n\}$. In a more familiar language, Y is the predictors and F is the set of categorical levels present in the dependent variable. Using Bayes' rule:

$$p(F_j | y_1, y_2, \dots, y_n) \propto p(y_1, y_2, \dots, y_n | F_j) p(F_j) \quad (1)$$

Where $p(F_j | y_1, y_2, \dots, y_n)$ is the posterior probability of class membership, i.e., the probability that Y belongs to F_j . Since Naive Bayes supposes that the conditional probabilities of the independent variables are statistically independent we can breakdown the likelihood to a product of terms:

$$p(Y + F_j) \propto \prod_{k=1}^n p(y_k | F_j) \quad (2)$$

And rewrite the posterior as:

$$p(F_j | Y) \propto p(F_j) \prod_{k=1}^n p(y_k | F_j) \quad (3)$$

Using Bayes' rule above, we label a new case Y with a class level F_j that obtains the highest posterior probability. Although the supposition that the predictor (independent) variables are autonomous is not always exact, it does simplify the classification task impressively, since it allows the class conditional densities $p(y_k | F_j)$ to be calculated individually for each variable, i.e., it simplifies a multidimensional task to a number of one-dimensional ones. In effect, Naive Bayes minimizes a high-dimensional density estimation task to a one-dimensional kernel density estimation.

3.2.3. Data preprocessing algorithm

Data pre-processing is significant step in the data mining process. Data-gathering methods are often slackly contained, resulting in out-of-range values, impossible data combinations for instance Sex: Male, Pregnant: Yes), missing values, etc. Analyzing data that has not been cautiously checked for such difficulties can produce misleading outcomes. Thus, the representation and quality of data is the first step before starting an analysis.

If there is much unsuitable and unnecessary information present or noisy and uncertain data, then knowledge discovery during the training phase is more problematic. Data readying and filtering steps can take sizeable amount of processing time. Cleaning, normalization, transformation, feature extraction and selection. The result of data pre-processing technique is the final training set.

Variable-by-variable data cleaning is straightforward filtering approach. Table 1 shows examples of how this metadata can help on detecting a number of possible data quality problems. Partial data is an inevitable problem in dealing with most of the data sources. Generally, there are some significant causes to be taken into account when processing unidentified feature values. One of the most significant ones is the source of 'unknowingness': (i) a value is absent because it was disregarded or lost; (ii) a certain feature is not relevant for a given Instance; (iii) for a given measurement, the designer of a training set does not care about the value of a certain feature.

Discretization: Discretization should minimize the number of likely values of the continuous feature since large number of likely feature values leads to slow and unproductive process. The simplest discretization method is an unattended straightforward method named equal size discretization. It counts the number of values we have from the feature that we are trying to discretize and divides it into intervals consisting of the same number of instances.

Normalization: Normalization is a "scaling down" alteration of the features. Within a feature there is mostly a large variation between the maximum and minimum values. When normalization is performed the value magnitudes are scaled to considerably low values. This is significant for many neural network and k-Nearest Neighborhood algorithms. The two most general methods for this scope are:

- min-max normalization:

$$u' = \frac{u - \min_A}{\max_A - \min_A} (\text{new}_{\max_A} - \text{new}_{\min_A}) + \text{new}_{\min_A}$$

- Z-score normalization:

$$u' = \frac{u - \text{mean}_A}{\text{stand_dev}_A} \quad (4)$$

Where u is the old feature value and u' the new one

3.2.4. Partial match retrieval algorithm

Partial matching enables users to define a portion of the term they are searching for and find any words that include that fragment. Partial-match retrieval is involved with accessing those records of a file which match the user's query, although his query may be only partially specified or defined.

1. A group of information called a file and a separate unit of a file is called a record.
2. A storage or recording method by which to express a file on a physical medium for further reference. This process is called as encoding a file.
3. A method which helps to access and read an encoded file.
4. A user of the system, who has one or more queries to present to the system. The result of a query is assumed to be a subset.
5. A search algorithm. This is a method for accessing and reading part of the encoded file in order to generate the response to a user's query.

IV. CONCLUSION

In this paper, we explain the need and development of MMQA, and it is examined that the existing methods mainly focus on limited area. Targeting at a more native approach, we suggest a new scheme to answer questions using multimedia data by enriching textual answers. For a given QA pair, our scheme first anticipates which type of medium is suitable for enriching the original textual answer. After that, it automatically produces a query based on the QA knowledge and then execute multimedia search with the query. Finally, query adaptive reranking and redundancy removal are done to obtain a set of images and videos along with the original textual answer. Different from the usual MMQA research that aims to automatically produces multimedia answers with given questions, our approach is constructed on the support of community contributed answers, and it can thus deal with more universal questions and achieve better performance and efficiency.

In our future work, we will further improve the scheme, such as creating better query generation method and investigating the suitable segments from a video. We can perform ranking in text and image but in videos we cannot perform the ranking process.

REFERENCE

- [1] R. Datta, D. Joshi, and J.Z. Wang, "Image Retrieval: Ideas, Influences, and Trends of the New Age," *ACM Computing Surveys*, vol. 40, article 5, 2007.
- [2] Q. Yin, X. Tang, and J. Sun, "An associate-predict model for face recognition," In *Proc. CVPR*, 2011.
- [3] A.Tamura, H.Takamura, and M.Okumura, "Classification of Multiple-Sentence Questions," In *Proceedings of (IJCNLP-05)*, pages 426--437, Jeju Island, Korea, October 2005.
- [4] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-Based Image Retrieval," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 12, pp. 1349-1380, Dec. 2000.

BIOGRAPHICAL NOTES

Jayashree K is presently pursuing B.Tech. final year in Information Technology Department from Panimalar Engineering College, Poonamalle, Chennai.

Jayashree R is presently pursuing B.Tech. final year in Information Technology Department from Panimalar Engineering College, Poonamalle, Chennai.

Mahalakshmi J is presently pursuing B.Tech. final year in Information Technology Department from Panimalar Engineering College, Poonamalle, Chennai.

SaiJanaranjani C is presently pursuing B.Tech. final year in Information Technology Department from Panimalar Engineering College, Poonamalle, Chennai.

MULTISTAGE EPICYCLIC LUG WRENCH

¹Shinu Baby, ² Nevin G Ninan, ³ Nithin George, ⁴ Salu Zachariah,
⁵ Aju Zachariah Mani

^{1,2,3,4} Scholar, ⁵ Assistant Professor, Department of Mechanical Engineering,
Saintgits College of Engineering (India)

ABSTRACT

Each and everyone have some kind of vehicle for their daily usage and those who don't have one, use any of the public transport. A common problem associated with all these vehicles is the removal and replacement of their wheels once they get damaged or flat. The lug nuts of heavy vehicles are removed by connecting a socket spanner to a long handle and then blowing this handle manually or with hammers. This takes a lot of human effort and may distort the fasteners and nuts. Although there are various mechanisms for removing lug nuts like lug wrenches, ratcheting socket wrench, impact wrenches, etc. but they are either time consuming, difficult to handle, non portable, power consuming or requires a lot of manual effort.

Our aim through this project is to reduce the human effort in unscrewing the lug nut. The huddle lies in the fact that the equipment to be designed must be light weight, portable and lower in power and time consumption. For this, in the present project a multistage epicyclic gear train device is designed to unscrew the lug nuts of vehicles. The device is very compact, simple in construction, easy to handle, reduces the unscrewing time to a considerable extend, portable and is user friendly.

Keywords: Case Hardened Alloy Steel, Gear Design, Gears, Multistage Gears, Torque

I. INTRODUCTION

Nowadays automobiles are an essential part of human life. We cannot imagine a world without them. Each and everyone have some kind of vehicle for their daily usage (whether it is a car or a bike) and those who don't have one use any of the public transport. Heavy vehicles like trucks are used to transport goods and other items. A common problem associated with all these vehicles is the removal and replacement of their wheels once they get damaged or flat. The tool set-up for each vehicle is a T-nut wrench and screw jack which is hard to use for a woman or teen to open their vehicle's lug nuts. Although there are various other mechanisms to unscrew lug nuts, they are either time consuming, non-portable or need a lot of human effort.

A lug nut or wheel nut is a fastener, specifically a nut, used to secure a wheel on a vehicle. Commonly used tool for lug nut removal is lug wrench. Lug wrenches may be L-shaped, or X-shaped. The form commonly found in car trunks is an L-shaped metal rod with a socket wrench on the bent end and a prying tip on the other end. Lug wrenches are much less expensive because they lack the ability to measure or limit the force used. Installing a wheel with a lug wrench thus requires a bit of rough guessing about proper tightness. Excessive force can strip threads or make the nuts very difficult to remove. Also, uneven torque between the various lug nuts, or excessive torque, can

lead to warping of the brake rotor if the car is equipped with disc brakes. An improved form of the lug wrench is the ratcheting socket wrench, often called a ratchet.

There are also power tool versions of "air" ratchets which use compressed air power to drive air powered socket wrenches which tighten or loosen nuts or bolts. A second major variety of compressed air powered tools are impact wrenches which are used for common tasks such as lug nuts on wheels. Electric powered impact wrenches for the same tasks are not uncommon. Small cordless 12 Volt and 18 Volt impact drivers are often used today as powered ratchets to remove and install nuts and bolts. Hydraulic motor ratchets with their characteristic higher torque characteristics are rare outside of heavy industry.

An impact wrench is a socket wrench power tool designed to deliver high torque output with minimal exertion by the user, by storing energy in a rotating mass, then delivering it suddenly to the output shaft. Compressed air is the most common power source for impact wrenches, providing a low-cost design with the best power-to-weight ratio. A simple vane motor is almost always used, usually with four to seven vanes, and various lubrication systems, the most common of which uses *oiled air*, while others may include special oil passages routed to the parts that need it and a separate, sealed oil system for the hammer assembly. Most impact wrenches drive the hammer directly from the motor, giving it fast action when the fastener requires only low torque. Electric impact wrenches are available, either mains powered, or for automotive use, 12-volt, 18-volt or 24-volt DC-powered. Recently, cordless electric impact wrenches have become common, although typically their power outputs are significantly lower than corded electric or air-powered equivalents. Some industrial tools are hydraulically powered, using high-speed hydraulic motors, and are used in some heavy equipment repair shops, large construction sites, and other areas where a suitable hydraulic supply is available. Hydraulic impact wrenches have the advantage of high power-to-weight ratio. But these are not portable. Some of the drawbacks of pneumatic power systems include high cost, require large size compressors to generate high torque and large power consumption. Disadvantages of electrical power systems include inability to operate at low speeds, physically large, expensive to produce and high maintenance cost.

Hence an attempt was made to design a simple device to unscrew the lug nuts of heavy vehicles using epicyclic gear trains which can be hand operated, portable, less expensive and low weight. Objective of the current work also included the development of a solid model of the device in Solid Works as per the design and to conduct structural analysis of the developed model in ANSYS.

In the current work we use gears to design a system that reduces the manual effort. Gears were invented by the Greek mechanics of Alexandria in the third century B.C., were considerably developed by the great Archimedes, and saw wide use in the Roman world. They found two main applications: in heavy-duty machines such as mills and irrigation wheels, where they transmitted considerable power, and in small-scale water-clocks, calendrical instruments and automata which could be of extraordinary sophistication, incorporating the differential and perhaps the hypoid gear [1]. In the current work our objective was to design a multistage speed reducer so that initial input torque given is minimum. A detailed overview of the design of a new two-stage cycloidal speed reducer with tooth modifications can be found in [2]. The effects of the design parameters of involute gears generated by rack-cutters and also a general algorithm for the kinematic synthesis of spur and helical gears can be found in [3].

Planetary gear sets possess numerous advantages over their parallel-axis counterparts in terms of their power density, tolerance insensitivity and noise attributes in addition to their kinematic flexibility. One potential disadvantage of planetary gear sets is power losses due to multiple planet branches, resulting from an increased number of gear meshes and bearings. The power losses of a planetary gear set can be grouped in two categories based on their dependence on load. Load-dependent (mechanical) power losses are induced by friction in external and internal gear mesh contact interfaces as well as at planet bearings while load-independent (spin) losses are associated with drag of the carrier assembly and gears, bearing viscous losses and oil-air pocketing at gear mesh interfaces. With the assumption that power losses of these components are independent of each other, a methodology that implements a family of models to predict total power loss of planetary gear sets including primary mechanical and spin loss components is proposed in [4].

II. EXPERIMENTAL METHODOLOGY

We had to design an epicyclic gear train device which reduces the mechanical leverage in unscrewing lug nuts. It should be compact, easy to handle, have low weight and should be able to produce the desired output torque with least human effort. For heavy vehicles like buses and trucks maximum torque required to unscrew lug nuts is 1000Nm. So if we use a single stage epicyclic gear train, the system will be of large size, heavy and difficult to handle. So we have to use a multistage epicyclic gear system for torque multiplication.

In order to construct a multistage epicyclic gear train, initially we need to find the output of a single stage. So to find the number of rotations of the output shaft for a given input rpm, output torque and the number of stages required to produce the required torque, motion analysis and torque analysis was done [5]. After finding the number of stages required and the output of each stage, we had to find the spur gear data required for the design calculations. So formulas for the dimensional calculation of spur gear in terms of diametral pitch (P) and number of teeth (N) were used.

These spur gear data were then used for the design of the system. For design three different materials were considered. These include case hardened alloy steel, cast iron and bronze. These materials were then checked for dynamic, static and wear tooth loads. For safe design static and wear tooth loads should be more than dynamic tooth load [6]. Next, we had to create a model of the device. So, using Solid Works a model of the required device was created. First, different components including sun gear, planet gears, ring gear, shaft, connecting bar, handle, etc. were generated as different parts and then assembled to have the required device.

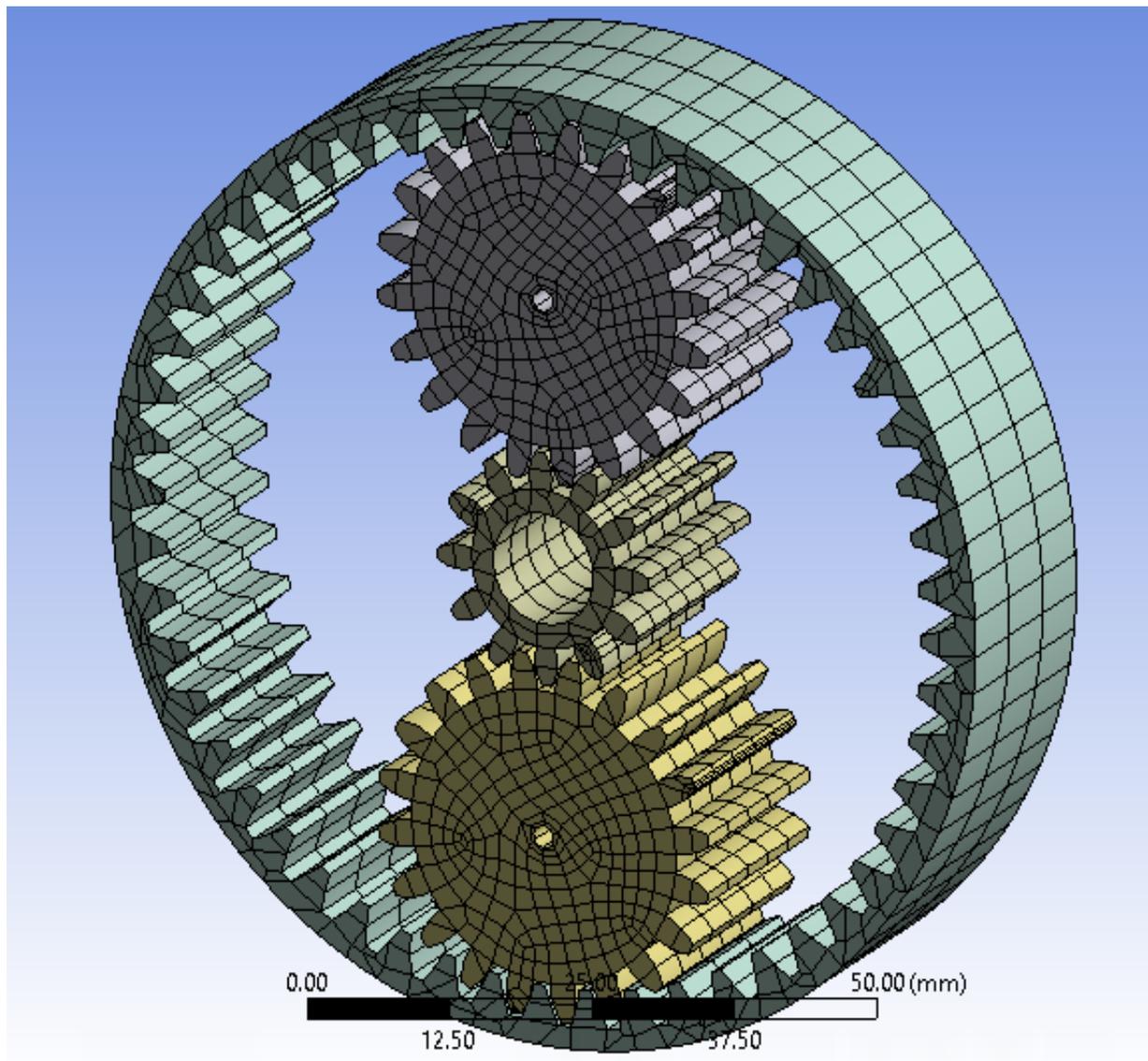


Fig. 1: Meshed view of section of the device being analyzed

Finally, static structural analysis of the created model was done in ANSYS for case hardened alloy steel, cast iron and bronze to find the equivalent stress, equivalent strain and total deformation of the device for various input conditions. Fig. 1 shows a meshed view of the section being analyzed. This was done to find whether the material used for design was safe.

III. RESULTS AND DISCUSSIONS

3.1 Velocity of Arm & Number of Stages

Data obtained from the motion analysis and torque analysis of the system is show in the Table 1. From the table it is clear that for an input velocity of 20rpm the velocity of the arm reduces to 3.75rpm in a single stage. So the velocity

ratio of the system is found to be 0.1875. Results of torque analysis shows that three stages of epicyclic gear train are required to produce the desired torque output. Gear ratio of the system is 5.33. Table 2 shows the input and output velocities and torques for each of the three stages. For an input torque of 10Nm and 20Nm, the output at the end of third stage is 1517.22Nm and 3034.44Nm respectively.

Table 1: Results of velocity and torque analysis

Input velocity = 20rpm		Input torque = 10Nm	
Parameters		Magnitude	
No. of stages		3	
Output velocity at the end of stage one		3.75rpm	
Output velocity at the end of stage three		0.13182rpm	
Maximum output torque		1517.22Nm	
Velocity ratio		0.1875	
Gear ratio		5.33	

Table 2: Input and output velocities and torques for each stage

		Stage		
		1	2	3
Velocity (rpm)	Input	20	3.75	0.703
	Output	3.75	0.703	0.13182
Torque (Nm)	Input	10	53.33	284.495
	Output	53.33	284.495	1517.22
	Input	20	106.667	568.99
	Output	106.667	568.99	3034.44

3.2 Spur Gear Design Data

In the present project, design of spur gear using three different materials was analysed for satisfying different design requirements. The materials used include case hardened alloy steel of BHN 650, cast iron of BHN 225 and bronze of BHN 80. Gears were tested for dynamic, static and wear tooth loads. Table 3 shows results of spur gear design using different materials. From the Table.3 it is clear that the static and wear tooth loads for bronze and cast iron are less than the dynamic tooth load. But for the design to be safe, static and wear tooth loads should be more than the dynamic tooth load. So it is not safe to use bronze and cast iron as gear material in the design. But in the case of case

hardened alloy steel the static and wear tooth loads are more than the dynamic tooth load. So case hardened alloy steel is safe for design.

Table 3: Dynamic, static and wear tooth loads for different materials

Material	Dynamic tooth load (N)	Static tooth load (N)	Wear tooth load (N)
Bronze	2118.749	437.310	142.439
Cast iron-Grade 35	2118.750	359.015	449.092
Case hardened alloy steel	2118.774	2197.370	2428.725

3.3 Solid Model

Fig. 2 shows the model of the multistage epicyclic gear train device design to unscrew the lug nut of vehicles. The device works on hand power. Whenever a lug nut is to be removed, a socket is attached to the lug nut and the output shaft of the device is connected to the socket. Then the input shaft is rotated by means of a handle. The input torque now gets multiplied in each of the three stages, thus providing the necessary torque at the output shaft required to unscrew the lug nut with least human effort. The outer ring serves as the casing for the device, so no additional casing is required. Since the device is compact and is of less weight it can be easily handled and is portable.



Fig. 2: Multistage epicyclic lug wrench

3.4 Structural Analysis Data

The model of the device created was structurally analysed in ANSYS for equivalent stress, equivalent strain and deformation for three different materials (i.e. case hardened alloy steel cast iron and bronze). Table 4 shows the maximum and minimum values of stress, strain and deformation for each of the different materials. From Table.4 it

is clear that the equivalent strain, equivalent stress and total deformation for case hardened alloy steel is much lower than cast iron and bronze. This may be due to the better material properties and strength of case hardened alloy steel compare to the other two. Alloy steel is case hardened by carburizing, quenching and tempering, which increases their fatigue resistance, toughness, hardness and wear resistance. The maximum equivalent stress of alloy steel is less than its ultimate stress, but it is not so in the case of bronze and cast iron. So alloy steel is structurally safe for design, which is in agreement with numerical calculation.

Table 4: Maximum and minimum values of stress, strain and deformation

Parameter		Material		
		Bronze	Cast iron	Alloy steel
Equivalent Elastic Strain	Minimum	1.2389×10^{-15}	7.286×10^{-16}	1.3391×10^{-16}
	Maximum	2.3401×10^{-2}	2.2215×10^{-2}	4.0830×10^{-3}
Equivalent(von-Mises) Stress (MPa)	Minimum	7.037×10^{-11}	3.2785×10^{-11}	1.1503×10^{-11}
	Maximum	2.375×10^3	2.3898×10^3	8.3853×10^2
Total Deformation (mm)	Minimum	0	0	0
	Maximum	2.2537×10^{-1}	2.1230×10^{-1}	3.9019×10^{-2}

Equivalent stress distribution obtained from static structural analysis of the section of the device for case hardened alloy steel, bronze and cast iron are show in figures 3, 4 and 5 respectively. From figures 3, 4 and 5 it is clear that the maximum stress is minimum in the case of case hardened alloy steel. Also, for alloy steel the maximum stress is below its ultimate stress, but it is not so in the case of bronze and cast iron. So alloy steel is safe for design.

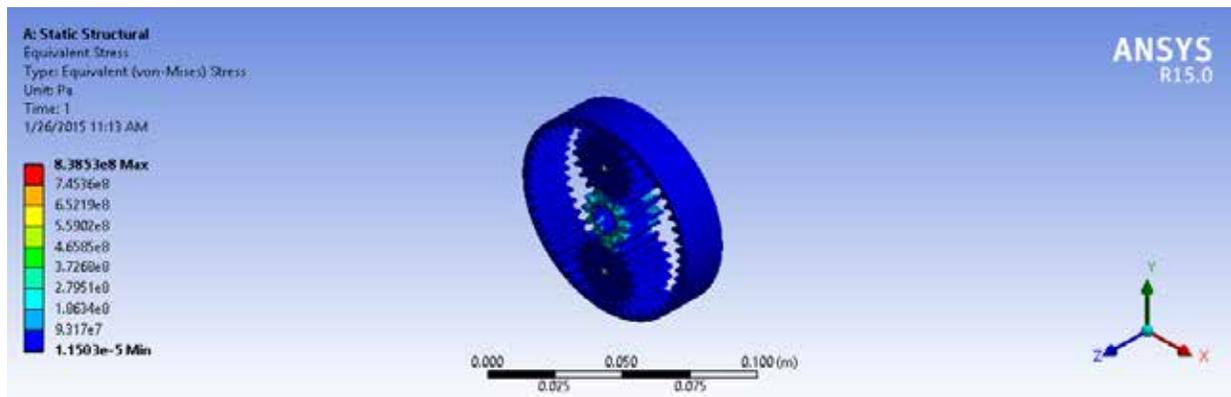


Fig. 3: Equivalent stress distribution for case hardened alloy steel

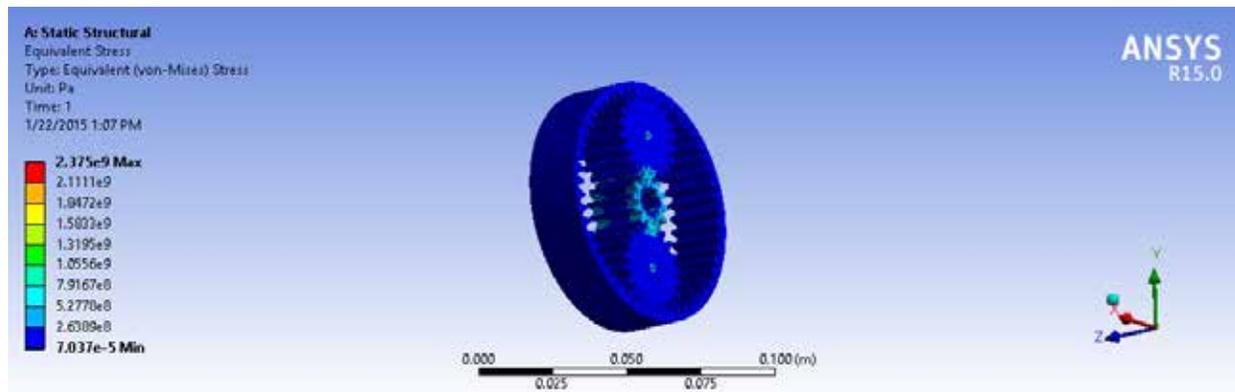


Fig. 4: Equivalent stress distribution for bronze

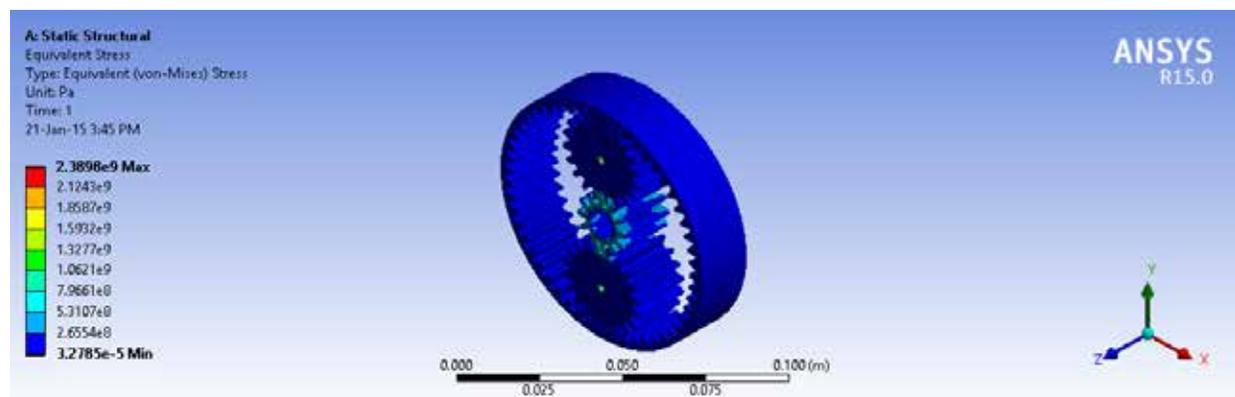


Fig. 5: Equivalent stress distribution for cast iron

IV. CONCLUSIONS

A mechanical device for unscrewing the lug nuts of vehicles was design using multistage epicyclic gear trains. The device offers better advantage over pneumatic, electric and hydraulic impact wrenches as it is hand operated and does not require any external power other than a little human effort. The device is compact, portable and is of less weight, so it can be easily handled. The use of the device can be extended to unscrewing nuts and bolts which are difficult to remove other than the lug nuts.

A model of the device was created in the Solid Works and it was analyzed in ANSYS to find its equivalent stress, equivalent strain and total deformation.

REFERENCES

- [1] M.J.T Lewis., "Gearing in the ancient world," *Endeavour*, Vol. 17, 1993, Issue 3, pp. 110-115.
- [2] Wan-Sung Lin., Yi-Pei Shih, Jyh-Jone Lee, "Design of a two-stage cycloidal gear reducer with tooth modifications," *Mechanism And Machine Theory*, Vol. 79, 2014, pp. 184-197.

- [3] G.Figliolini, P.Rea., "Effects of design parameters of involute gears generated by rack-cutters," *International Gear Conference*, Lyon, 2014, pp. 294-302.
- [4] D.Talbot, A.Kahraman., "A methodology to predict power losses of planetary gear sets," *International Gear Conference*, Lyon, 2014, pp. 625-635.
- [5] S.S. Rattan., *Theory of machines*, 3rd ed., Mc Graw Hill Publications, 2009.
- [6] R.S. Khurmi, J.K. Gupta., *Textbook of machine design*, 14th ed., S Chand Publications, 2005.

Biographical Notes

Mr. Shinu Baby is presently pursuing B. Tech final year in Mechanical Engineering Department from Saintgits College of Engineering, Pathamuttom, Kottayam, India.

Mr. Nevin G Ninan is presently pursuing B. Tech final year in Mechanical Engineering Department from Saintgits College of Engineering, Pathamuttom, Kottayam, India.

Mr. Nithin George is presently pursuing B. Tech final year in Mechanical Engineering Department from Saintgits College of Engineering, Pathamuttom, Kottayam, India.

Mr. Salu Zachariah is presently pursuing B. Tech final year in Mechanical Engineering Department from Saintgits College of Engineering, Pathamuttom, Kottayam, India.

Mr. Aju Zachariah Mani is working as an Assistant Professor in Mechanical Engineering Department, Saintgits College of Engineering, Pathamuttom, Kottayam, India.