ISSN : 2250-1940 (Print), 2349 - 1647(Online)

RESEARCH EXPLORER

Volume VI | August 2018 | Issue 19

Date of issue : 10th August 2018

Single Issue : ₹. 500

Impact Factors

IRJIF : 2.014

JIF : 1.85

RIF : 1.056

NAAS Score : 2.6

Indexed & Abstracted in

CiteFactor, Cosmos, DAIJ, DOAJ, DRJI, SIS Scientific Indexing Services, International Innovative Journal Impact Factor (IIJIF), Journal Index .net, ESJI www.ESJIndex.org, General Impact Factor, GiFactor, Universal Impact Factor, OAJI .net

INDIAN ACADEMIC RESEARCHERS ASSOCIATION

www.iaraindia.com / www.selptrust.org

Mail Id : editor@iaraindia.com

RESEARCH EXPLORER

(A Referred and Reviwed Quarterly International Research Journal on Multidisciplinary)

ISSN : 2250-1940 (Print), 2349 - 1647(Online)

Impact Factor : 2.014(IRJIF), 1.85(JIF), 1.056(RIF), 2.62(NAAS)

Volume VI | **August 2018** | **Issue - 19**

CONTENT

Research Explorer - International Journal
ISSN : 2250-1940 (Print), 2349 - 1647 (Online)
Impact Factor : 2.014(IRJIF), 1.85(JIF), 1.056(RIF), 0.60(QF) 0.398(GIF)
Issue - 19 ; August 2018 *© Author*

MULTI-RESPONSE OPTIMIZATION OF WATER-IN-DIESEL EMULSION FUEL OPERATING PARAMETERS USING TAGUCHI-GREY ANALYSIS

Dr. Suresh Vellaiyan
Department of Mechanical Engineering,
Haramaya Institute of Technology, Haramaya University, Dire Dawa, Ethiopia.

N.Vijayarani
Department of Computer Science,
Trinity College for Women, Namakkal, Tamilnadu, India.

Abstract

Water-in-diesel (W/D) emulsion fuel is the promising alternative diesel fuel which has the potential to produce better emission level of oxides of nitrogen (NOx) in diesel engines. The key factors which are foremost focused in introduce of W/D emulsion fuel are the engine operating conditions such as engine load, compression ratio (CR) and water concentration in diesel. The present investigation focused on optimization of W/D emulsion fuel operating parameters in diesel engine using Taguchi-grey relational analysis. Taguchi's L16 orthogonal array was used to collect the output responses (performance and emission levels). A statistical analysis of variance (ANOVA) was used to identify the significant process parameter which has most influence on output responses. The signal-to-noise (S/N) ratio and grey relational analysis were used to optimize the performance and emission levels. From the results obtained, it was found that engine load and CR had major influence on performance and emission levels. Optimal parametric setting had been determined and verified through conformation experiment. The optimal parametric setting had shown good improvement in performance and emission levels of diesel engine.

Keywords: *Optimization, Taguchi-grey relational analysis, ANOVA, emulsion fuel*

1. Introduction

Air pollution caused by the diesel engines has thrown much interest in eco-friendly diesel fuels since improved environment and human health are of concern. In order to obtain better emission characteristics in existing diesel engines, considerable efforts have gone into the research and development of fuel modifications and alternative fuels. As a consequence, the introduction of water in diesel fuel has been preferred for the existing diesel engines since the desired emission characteristics can be achieved without any added cost and engine modification.

Introduction of water in diesel engine was initially proposed by Prof. B. Hopkinson to promote

better inter cooling and reduce emission levels in gas engines. There are generally three approaches followed to introduce water in diesel engine combustion zone: (i) Direct injection of water into the combustion chamber (ii) Fumigating the water into the engine intake air (iii) water-in-diesel emulsion fuel [1, 2]. Out of these, W/D emulsion fuel has the favorable fuel characteristics such as large fragmentation, less change in viscosity and micro-explosion phenomena of the water droplet [3, 4].

Formation of stable W/D emulsion fuel for a long period of time is the critical issue in emulsion fuel preparation. In order to avoid the phase separation of emulsion fuels, surface active agents (or) surfactants are used to lower the surface tension between the water and diesel molecules. The nonionic surfactants are generally used in emulsion fuels since it has the positive fuel characteristics such as, burn with no soot, free of sulphur and nitrogen [5]. In emulsification process, W/D emulsion with droplet size of 50-200 nm is formed at the occurrence of a high energy emulsification method. This high energy method includes: high-shear stirring, high-pressure homogenizers, ultrasonic vibration and supersonic. As far as the smoke and NOx emissions, ultrasonic and supersonic vibration has negative impact compared to the emulsion prepared by mechanical homogenization [6]. In the previous phase of this study, it is observed that sorbitan monolanrate (a nonionic surfactant of HLB 8.6) at high speed of agitation (15000 rpm) provides better stability with minor occurrence of creaming and coalescence zone [7].

There are several investigations have been carried out to study the performance and emission characteristics of W/D emulsion fuels in diesel engine. As far as the performance and emission characteristics, the complexity to analyze the combustion phenomena leads to inconsistent reports in the domain of W/D emulsion fuels. Alahmer and Suresh [8, 9] are reported that the combustion efficiency, BP, torque and BSFC has been increased by using W/D emulsion fuels. While, contradictory results are reported in terms of BP, torque and BSFC by Yang and Selim [10, 11]. Most of the reports have showed consistent improvements in emission characteristics of W/D emulsion fuels [12]. However, subramanian [13] reported that the HC and CO has been faintly increased by using W/D emulsion fuel.

The use of W/D emulsion fuel in existing diesel engine has been an active field of inquiry in the past decade. Most of the written reports were based on the measurement of performance and emission characteristics. However, no work has been addressed to optimize the operating parameters of diesel engine running with W/D emulsion fuels to obtain favorable performance and emission characteristics. The present study has been carried out (i) to analyze the influence of operating parameters on performance and emission characteristics (ii) to obtain the optimal parametric setting for better response i.e better performance and emission characteristics (iii) validate the optimum parametric setting through confirmation experiments.

2.Materials And Methods

2.1 Materials

In this work, high speed diesel (Bharath Petroleum Corporation Limited, India) was used as continuous phase of emulsion. Sorbitan monolaurate (HLB: 8.6) was used as surfactant (Estelle Chemicals (P) Ltd.). Double distilled filtered water was used as dispersed phase of emulsion.

2.2Emulsion fuel preparation, property measurement and its stability analysis

To prepare W/D emulsion fuel, the desired quantity of water (5%, 10%, 15% & 20% of total volume) and surfactant (1% of total volume) were added drop by drop inside the fuel mixing jar

and stirred at high speed(15000 rpm) constantly for about 30 mins. The experimental layout for emulsion fuel preparation is shown in Figure 1. The properties of standard diesel and prepared emulsion fuels are measured consistent with EN 590:2009. The emulsion stability was analyzed based on emulsion density variation with respect to time using a photonic circuit. The stability analysis setup is shown in Figure 2.

2.3Engine test setup

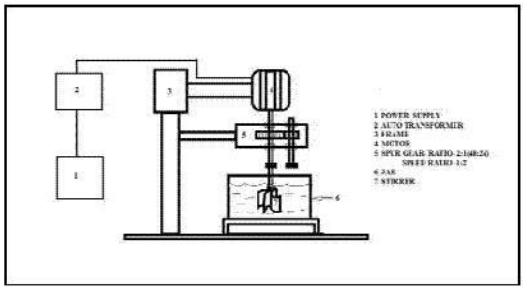
The engine used in this experiment is computerized single cylinder, four stroke and variable compression ratio diesel engine with an eddy current dynamometer. The fuel injection timing is maintained at 23 degree BTDC as a constant. ICEnginesoft_9.0 software is used to record the combustion parameters. As for the measurement of emissions, AVL digas gas analyzer is used to measure the NO, HC, CO, CO2 and O2. The detailed specifications of engine and gas analyzer are listed in Table 1.The photographic view of the engine test setup is shown in Figure 3.

2.4 Taguchi-Grey Analysis

Taguchi method uses an orthogonal array to read the essence of the effect of entire input parameters with minimal number of experiments [16]. In the traditional Taguchi method, individual responses alone can be optimized within the experimental domain [17, 18]. Multi-objective optimization problems can be solved by employing Taguchi method coupled with grey relational analysis [19].

The present study employs Taguchi's L16 orthogonal array to plan the experiments on diesel engine. Three controlling factors, (i) engine load (ii) CR and (iii) water concentration (WC) in diesel are identified as input parameters. Six experimental results, (i) BSFC (ii) BTE (iii) HC (iv) CO (v) CO2 and (vi) NO are selected as output responses.

Combustion phases such as SOC, MFB and EOC are the critical factors which are directly affect the performance and emission characteristics of diesel engines. All the controlling factors have direct influence on combustion phases. Increases in engine load leads to shorter ignition delay and complete combustion which are contribute to better engine performance. Concurrently, the high engine loads direct to



Experimentallayout for emulsion fuel preparation

Fig. 1. Photographic view of stability analysis setup

Fig. 2. Photographic view of engine test setup

TABLE I. Specifications of engine and emission analyzer

(a) Engine specifications		
Parameter	Specification	
Engine type	Computerized, 4-Stroke, Single cylinder, VCR diesel engine	
Bore × Stroke (m)	0.0875 × 0.11	
Displacement volume (cc)	661.45	
Max. power	3.5 kW at 1500 rpm	
CR range	12-18	
Dynamometer	Eddy current dynamometer (max. load of 7.5 kW)	
(b) Gas analyzer specifications		
Measured quality	Measuring range	Resolution
NO	0 to 5000 ppm vol	1 ppm vol
HC	0 to 20000 ppm vol	≤ 2000: 1 ppm vol, > 2000: 10 ppm vol
CO	0 to 10% vol	0.01% vol
CO2	0 to 20% vol	0.1% vol
O2	0 to 22% vol	0.01% vol

high level emissions of HC, CO and NOx [12]. Hence the parameter levels for engine load were chosen in all intervals (25%, 50%, 75% & 100%). Similarly, the high compression ratio reduces the maximum pressure rise rate and ignition delay and lead to better performance characteristics of diesel engines. The longer ignition delay in low compression ratio makes the engine to fail to run [11]. Hence the parameter levels for compression ratio were chosen to be 15, 16, 17 & 18. The parameter levels for water concentration were chosen as 5%, 10%, 15% and 20% of total volume of fuel. Addition of water to diesel provides low level of NOx emissions. At the same time, increases in water concentration, increases the ignition delay period, combustion duration and other emission levels [9].

To convert a multi-response optimization problem into a single response problem the following procedure was adopted: (1) Experimental data were first normalized ranging from zero to one. (2) The grey relational coefficient was estimated based on normalized experimental data. (3) Overall grey relational grade was obtained by applying weighing factor for each selected response [20]. The combination of parameters having highest grey relational grade is expected to be closer to the optimal parametric setting. The optimum parametric setting was determined based on Signal-to-Noise (S/N) ratio.

In this study, BTE has been characterized as "the larger-the better" and the characteristic is computed as follows:

$$x_{(i)}k = \frac{y_i(k) - \min y_i(k)}{\max y_i(k) - \min y_i(k)} \qquad (1)$$

BSFC and emission parameters (HC, CO, CO2 and NO) have been characterized as "the smaller-the better" and are computed as follows:

$$x_{(i)}k = \frac{max\, y_i(k) - y_i(k)}{\max y_i(k) - \min y_i(k)} \qquad (2)$$

where xi (k)is the sequence for comparison, yi (k) is the original reference sequence, i=1,2,….. m, k= 1,2,…..n, and, with m, n being the total number of experiments and response. The values min yi(k) and max yi(k) are the smallest and the highest values of yi(k) respectively.

The grey relational coefficient xi(k) is calculated as:

$$\xi_i(k) = \frac{\Delta_{min} - \psi\Delta_{max}}{\Delta_{0i}(k) + \psi\Delta_{max}} \qquad (3) \qquad \Delta_{0i}(k) = \left||x_0(k) - x_i(k\right|$$

where D0i is the value of absolute difference, Dmin, Dmax are the minimum and maximum values among the absolute differences. The purpose of distinguishing coefficient y (0 d" y d" 1) is to weaken the effect of Dmax when it becomes too large. In the present study, the value of y was taken as 0.5. The grey relational grade y0 is calculated as:

where b is the weighing factor. While converting multiple grey relation grades, critical factors have to be assigned more weightages. The following value of weightage was adopted in this study for different responses: BSFC 0.2, BTE 0.2, HC 0.2, NO 0.2, CO 0.05 & CO2 0.05. Among the four emission parameters, CO and CO2 were assigned as lower value since the diesel engine burns lean mixture which produces low level of CO and high level of CO2 emission indicates efficient combustion.

3.Results And Discussion

3.1. Emulsion fuel stability and its property

The stability profile of W/D emulsion fuels with respect to water concentration is shown in Figure 4. It is observed that at low level of water concentration, dispersed water droplets are so far spaced out to interact among each other and improve the emulsion stability. Increases in water concentration, reduces the emulsion stability since the droplet interactions become high.

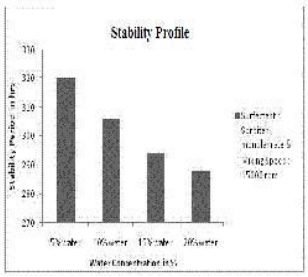


Fig. 3. Stability profile of W/D emulsion fuels

The measured values of W/D emulsion fuels property are listed in Table 2. From the measured values, it is observed that the density and viscosity of emulsion fuels increases with the addition of water concentration, since the higher density of water and high mixing speed. The value of flash point is increased by increasing the water content. The Heating value of emulsion fuels is reduced by adding water content. 20% of water content is reduced 7.7% of heating value since the water content of the inner phase of emulsion fuel absorbed the heat released from combustion.

TABLE II. Properties of BD and W/D emulsion fuels

	Fuel Properties			
	Density at 15°C (kg/m3)	Viscosity at 40°C (mm2/s)	Flash point °C	Heating value (MJ/kg)
BD	831.4	2.4	62	43.8
5% of W/D	839.8	4.2	69	42.9
10% of W/D	845.1	4.4	74	42.1
15% of W/D	850.4	4.7	78	41.2
20% of W/D	857.2	4.9	83	40.4

3.2. Experimental observation for L16 orthogonal array

Table 3 shows the input process parameter combinations and experimental results. From the experimental results, it is Observed that, the increases in engine load increase the BTE, HC and NO emissions since the engine burns the fuel efficiently during high load conditions. Also the high engine load leads to drop in BSFC and CO2 emission. The performance characteristics are improved at high CRs and high water concentration due to efficient burning of fuel and micro-explosion behavior of water droplets. At the same time the negative effects are arrived in HC and CO emissions.

TABLE III. Orthogonal Array (L16) Of The Experimental Runs And Results

Run No	Input Process Parameters			Output Responses					
	Load (%)	CR	WC (%)	BSFC (kg/kW h)	BTE (%)	HC ppm vol.	CO % vol.	CO2 % vol.	NO ppm vol.
1	75	15	5	0.3626	14.0598	33	0.1	0.54	44
2	75	16	10	0.3332	14.8577	37	0.07	0.6	55
3	75	17	15	0.3358	15.3404	33	0.05	0.68	34
4	75	18	20	0.3389	16.8349	30	0.05	0.68	50
5	50	15	10	0.3650	22.8592	37	0.09	0.57	51
6	50	16	5	0.3628	22.6539	19	0.04	0.72	109
7	50	17	20	0.3927	20.8789	31	0.05	0.64	82
8	50	18	15	0.3003	28.8177	24	0.04	0.78	101
9	75	15	15	0.2939	26.9479	36	0.09	0.58	48
10	75	16	20	0.2994	27.637	30	0.09	0.62	71
11	75	17	5	0.2795	29.411	12	0.02	1	152
12	75	18	10	0.2684	31.3604	13	0.02	0.98	125
13	100	15	20	0.2732	30.4138	29	0.14	0.54	57
14	100	16	15	0.271	30.3336	21	0.09	0.7	101
15	100	17	10	0.2501	32.6005	12	0.04	0.98	140
16	100	18	5	0.2494	32.913	10	0.03	1.08	193

TABLE IV. Calculated grey relational coefficient of all responses, grey relational grade with weightage and S/N ratio

Wei. Factor	0.2	0.2	0.2	0.1	0.1	0.2	Grey Relational Grade	S/N ratio	Rank
Run No	Grey Relational Coefficient								
	BSFC	BTE	HC	CO	CO2	NO			
1	0.3333	0.3333	0.3698	0.4783	1	1	0.55016	-5.19017	12
2	0.3401	0.3425	0.3333	0.5454	0.8181	0.8713	0.51383	-5.78354	16
3	0.3534	0.3436	0.3698	0.6666	0.6585	0.8816	0.57775	-5.67577	13
4	0.3590	0.3693	0.4029	0.6666	0.6585	0.9254	0.64389	-5.28977	10
5	0.3746	0.4843	0.3333	0.4615	0.9	0.9141	0.59744	-4.47409	10
6	0.38	0.4783	0.6	0.75	0.6	0.3340	0.57047	-4.87971	11
7	0.3704	0.4086	0.3914	0.6666	0.7297	0.6622	0.54217	-5.31726	14
8	0.7439	0.6962	0.4909	0.75	0.3294	0.3665	0.62746	-4.04819	9
9	0.7787	0.6116	0.3417	0.4615	0.8709	0.9490	0.66948	-3.48324	6
10	0.7653	0.6405	0.4029	0.4615	0.7714	0.7339	0.63187	-3.98740	8
11	0.8387	0.7284	0.8709	1	0.3698	0.4082	0.70627	-3.02051	5
12	0.8917	0.8993	0.8181	1	0.3802	0.4791	0.75572	-2.43276	3
13	0.8680	0.7858	0.4153	0.3333	1	0.3514	0.71827	-2.87424	4
14	0.8787	0.7844	0.3510	0.4615	0.6279	0.3665	0.66510	-3.54219	7
15	0.9955	0.9676	0.8709	0.75	0.3802	0.4369	0.76725	-2.30123	2
16	1	1	1	0.8571	0.3333	0.3333	0.78571	-2.09471	1

3.3. Optimization of process parameters

To optimize the process parameters, the experimental data are normalized to obtain grey relational grade. The S/N ratio for overall grey relational grade is found from the equation:

$$S/N = -10 \log \frac{1}{n}\sum_{i=1}^{n}\frac{1}{y_i^2}$$

Where i =1, 2…n and yi is the response value for an experimental condition.

Table 4 shows the grey relational coefficient of all responses, grey relational grade and the S/N ratio for each run. Table 5 shows the S/N ratio responses and significance of contributions of process parameters. It can be observed that all the three factors have a significant influence (engine load = 80 %, CR = 15 % and WC 5 %) on the output response and the optimum parametric setting is found to be A4B4C2 i.e. engine load = 100%, CR = 18 and WC 10%.

TABLE V. Overall responses for S/N Ratio (Larger is better) and Significance of contribution (ANOVA)

Level	Engine load (A)	CR (B)	WC (C)
1	-3.472	-4.006	-3.784
2	-4.667	-4.336	-3.748
3	-3.231	-4.066	-4.173
4	-2.700	-3.466	-4.367
Delta	2.769	1.069	0.619
Rank	1	2	3
Optimum level	A4	B4	C2
Sum of squares	0.099	0.018	0.0062
% of contribution	80	15	5

4.Confirmation Test

To predict performance and emission characteristics for the optimum parametric setting, the S/N ratio and grey relational grade are determined using the following equatioequation

$$\hat{y} = y_m + \sum_{i=1}^{o}(\bar{y}_i - y_m)$$

where ym = total mean of S/N ratio or total mean of grey relational grade, = mean of S/N ratio or mean of grey relational grade for optimum level. o = No.of design factors. The predicted grey relational grade and S/N ratio are 0.7995 and -1.8799.

TABLE VI. Experimental result for optimum setting and its comparison with initial SETTING

In order to verify the predicted value, an experiment was carried out for the optimum parametric setting, S/N ratio and grey relational grade were calculated. The experimental values, S/N ratio and grey relational grade for optimum parametric setting and initial setting (A4B4C4) are presented in Table 6 & 7. It can be noted that the predicted and experimental results have good agreement and an improvement of 13% in S/N ratio is also noted.

Run No	Input Process Parameters			Output Responses					
	Load (%)	CR	WC (%)	BSFC (kg/kW-hr)	BTE (%)	HC ppm vol.	CO % vol.	CO2 % vol.	NO ppm vol.
16	100	18	5	0.2494	32.918	10	0.00	1.08	150
17	100	18	10	0.2406	33.3400	9	0.00	1.00	131
Improvement in output response (%)				3.6	7.9	10	-	4.6	21.8

TABLE VII. : Result of confirmation Test

	Initial Parametric Setting - Rank I	Optimal Parametric Setting	
		Prediction	Experimentation
Level	A4B4C1	A4B4C2	
Grey Relational Grade	0.7857	0.7995	0.8011
S/N Ratio	-2.0947	-1.9799	-1.9541

5.Conclusion

In order to optimize the operating parameters of diesel engine running with W/D emulsion fuel, the Taguchi-grey relational analysis has been applied and the results presented. Based on the results, it can be concluded that engine load (80%) and compression ratio (15%) have major influence on performance and emission characteristics of diesel engine. The optimal parametric setting was predicted as 100% load, 18 CR and 10% water content and was confirmed through a set of experimental trials. Further, the experimental trials revealed that the performance and emission characteristic of diesel engine was considerably improved in optimal parametric setting.

6. References

1. Abu-Zaid,M, "Performance of single cylinder, direct injection diesel engine using water fuel emulsion," Energy Conversion and Management, vol. 45, pp.697-705, 2004.
2. Bedford,F., Rutland,C., et al., "Effects of Direct Water Injection on DI Diesel Engine Combustion," SAE Technical Paper 2000-01-2938, pp.1-10, 2000.
3. Ghojel, J., Honnery, D. and Al-Khaleefi, K., "Performance, emissions and heat release characteristics of direct injection diesel engine operating on diesel oil emulsion," Applied Thermal Engineering, vol. 26, pp. 2132-2141, 2006.
4. Panayiotis, P, Constantinides and Scalart, J.P., "Formulation and physical characterization of water-in-oil microemulsions containing long-versus medium-chain glycerides," International Journal of Pharmaceutics, vol. 158, pp.57-68, 1997.
5. Morozumi, Y. and Saito, Y, "Effect of Physical Properties on Micro explosion Occurrence in Water-in-Oil Emulsion droplets," Energy Fuels, vol. 24, pp. 1854-1859, 2010.
6. Lin, C, Y and Chen, Li-Wei, "Comparison of fuel properties and emission characteristics of two- and three-phase emulsions prepared by ultrasonically vibrating and mechanically homogenizing emulsification methods," Fuel, vol. 87, pp. 2154-2161, 2008.
7. Suresh, V and Amirthagadeswaran KS, "Formulation of stable water-in-diesel emulsion fuel and investigation of its properties," Energy Sources, Part A: Recovery, Utilization and Environmental Effects (in press).
8. Alahmer, A., "Influence of using emulsified diesel fuel on the performance and pollutants emitted from diesel engine", Energy Conversation and Management, Vol.73, pp. 361-369, 2013.
9. Suresh, V and Amirthagadeswaran KS, "Combustion and Performance Characteristics of Water-in-Diesel Emulsion Fuel," Energy Sources, Part A: Recovery, Utilization and Environmental Effects (in press).
10. Yang, WM., An, H.; Chou, SK. et al., "Impact of emulsion fuel with nano-organic additives on the performance of diesel engine," Applied Energy, vol. 112, pp. 1206-1212, 2013.
11. Selim, M and Ghannam, M, "Performance and engine roughness of a diesel engine running on stabilized water diesel emulsion," SAE Technical Paper 2007-24-0132, 2007.
12. Attia, AMA and Kulchitskiy, AR, "Influence of the structure of water-in-fuel emulsion on diesel engine performance," Fuel, vol. 116, pp. 703-708, 2014.
13. Subramanian, KA, "A comparision of water-diesel emulsion and timed injection of water into the intake manifold of a diesel engine for simultaneous control of NO and smoke emissions," Energy Conversation and Management, vol. 52, pp. 849-857, 2011.

Research Explorer - International Journal
ISSN : 2250-1940 (Print), 2349 - 1647 (Online)
Impact Factor : 2.014(IRJIF), 1.85(JIF), 1.056(RIF), 0.60(QF) 0.398(GIF)
Issue - 19 ; August 2018

COMPARISON OF WEB FRAMEWORKS – ANGULAR AND REACT

Jayanthi Ganesan,
Senior Full Stack JavaScript Developer,
6060 Village bend drive,
Apt 513, Dallas, Texas – 75206, USA

Abstract

Web application development frameworks is a need because Coding everything from scratch, excluding even the normal things, can be extremely time-consuming and makes the developer spend time reinventing the wheel, time which would rather be spent implementing new features or tightening up the code base. Angular and React are two new web technologies on the rise. Many business owners, technical decision makers, and project managers are having difficulties about making a choice between these two. The amount of JavaScript tools is steadily increasing, turning the selection of appropriate technology into a challenge. While among the top discussions is AngularJS vs. ReactJS, both are highly-performing, advanced, and widely used worldwide. AngularJs and Angular is managed by ***Google*** *and ReactJSis owned by* ***Facebook****. Both of them are unique and resourceful in their own ways. These frameworks are quite easy to use with the high-end potential to build cutting-edge mobile and web applications.*

Keywords:-*Angular, React, Data Binding, Attributes, HTML, JAVA Script.*

1. Introduction Of React, Angular Framework React

In computing, React (also known as React.js or ReactJS) is a JavaScript library for building user interfaces. It is maintained by Facebook and a community of individual developers and companies. React can be used as a base in the development of single-page or mobile applications. Complex React applications usually require the use of additional libraries for state management, routing, and interaction with an API.

AngularJS

AngularJS (commonly referred to as "**Angular.js**" or "**AngularJS**") is a JavaScript-based open-source front-end web application framework mainly maintained by Google and by a community of individuals and corporations to address many of the challenges encountered in developing single-page applications. The JavaScript components complement Apache Cordova, a framework used for developing cross-platform mobile apps. It aims to simplify both the development and the testing of such applications by providing a framework for client-side model–view–controller (MVC) and model–view–view-model (MVVM) architectures, along with components commonly

used in rich Internet applications. In 2014, the original AngularJS team began working on the Angular application platform.

The AngularJS framework works by first reading the HTML page, which has additional custom tag attributes embedded into it. Angular interprets those attributes as directives to bind input or output parts of the page to a model that is represented by standard JavaScript variables. The values of those JavaScript variables can be manually set within the code, or retrieved from static or dynamic JSON resources.

AngularJS is the frontend part of the MEAN stack, consisting of **M**ongo DB database, **E**xpress. js web application server framework, **A**ngular.js itself, and **N**ode.js server runtime environment.

AngularJ

AngularJS is built on the belief that declarative programming should be used to create user interfaces and connect software components, while imperative programming is better suited to defining an application's business logic.[5] The framework adapts and extends traditional HTML to present dynamic content through two-way data-binding that allows for the automatic synchronization of models and views. As a result, AngularJS de-emphasizes explicit DOM manipulation with the goal of improving testability and performance.

AngularJS's design goals include:

- To decouple DOM manipulation from application logic. The difficulty of this is dramatically affected by the way the code is structured.
- To decouple the client side of an application from the server side. This allows development work to progress in parallel, and allows for reuse of both sides.
- To provide structure for the journey of building an application: from designing the UI, through writing the business logic, to testing.

AngularJS implements the MVC pattern to separate presentation, data, and logic components. Using dependency injection, Angular brings traditionally server-sideservices, such as view-dependent controllers, to client-side web applications. Consequently, much of the burden on the server can be reduced.

Scope

AngularJS uses the term "scope". As a part of the "MVC" architecture, the scope forms the "Model", and all variables defined in the scope can be accessed by the "View" as well as the "Controller". The scope behaves as glue and binds the "View" and the "Controller".

In AngularJS, "scope" is a certain kind of object that itself can be in scope or out of scope in any given part of the program, following the usual rules of variable scope in JavaScript like any other object.

Two-way data binding

AngularJS two-way data binding is its most notable feature, largely relieving the server backend of templating responsibilities. Instead, templates are rendered in plain HTML according to data contained in a scope defined in the model. The $scope service in Angular detects changes to the model section and modifies HTML expressions in the view via a controller. Likewise, any alterations to the view are reflected in the model. This circumvents the need to actively manipulate the DOM and encourages bootstrapping and rapid prototyping of web applications. AngularJS detects changes in models by comparing the current values with values stored earlier in a process of dirty-checking,

Different ways of defining a watcher in AngularJS,

- Explicitly $watch an attribute on $scope.

$scope.$watch('person.username', validateUnique);

- Place an interpolation in your template (a watcher will be created for you on the current $scope).
- Ask a directive such as ng-model to define the watcher for you.

<input ng-model="person.username" />

\$digest

$digest () is angular method, which is invoked internally by angularjs in frequent intervals. In $digest method, angular iterates over all $watches in its scope/child scopes.

$apply

$apply () is an angular method, internally invokes $digest. This method is used when you want to tell angular manually start dirty checking (execute all $watches)

$destroy

$destroy is both a method and event in angularjs. $destroy () method, removes a scope and all its children from dirty checking. $destroy event is called by angular whenever a $scope or $controller is destroyed.

Performance

AngularJS sets out the paradigm of a *digest cycle*. This cycle can be considered a loop, during which Angular checks if there is any change to all the variables watched by all the $scopes. If $scope.myVar is defined in a controller and this variable was marked for watching, Angular will monitor the changes on myVar in each loop iteration.

This approach potentially leads to slow rendering when AngularJS checks on too many variables in the $scope every cycle. Miško Hevery suggests keeping fewer than 2000 watchers on any page.

Angular 2

Angular (commonly referred to as "**Angular 2+**" or "**Angular v2 and above**")is a TypeScript-based open-source front-end web application platform led by the Angular Team at Google and by a community of individuals and corporations. Angular is a complete rewrite from the same team that built AngularJS.

Differences between Angular and AngularJS

Architecture of an Angular application are the main building blocks are modules,componenets,templates,metadata,data binding,directives,service and dependency injection.

Angular was a ground-up rewrite of AngularJS.

- **Angular does not have a concept of "scope" or controllers; instead it uses a hierarchy of components as its primary architectural characteristic.**
- Angular has a different expression syntax, focusing on "[ ]" for property binding, and "( )" for event binding
- Modularity – much core functionality has moved to modules
- Angular recommends the use of Microsoft's TypeScript language, which introduces the

following features:

- Class-based Object Oriented Programming
- Static Typing

. TypeScript is a superset of ECMAScript 6 (ES6), and is backwards compatiblewith ECMAScript 5 (i.e.: JavaScript).

. Dynamic loading

. Asynchronous template compilation

. Iterative callbacks provided by RxJS. RxJS limits state visibility and debugging, but these can be solved with reactive add-ons like ngReact or ngrx.

Version 2.0.0

Angular 2.0 was announced at the ng-Europe conference 22-23. October 2014. On April 30, 2015, the Angular developers announced that Angular 2 moved from Alpha to Developer Preview Angular 2 moved to Beta in December 2015,and the first release candidate was published in May 2016 The final version was released on September 14, 2016.

Version 4.0.0

On 13 December 2016 Angular 4 was announced, skipping 3 to avoid confusion due to the misalignment of the router package's version which was already distributed as v3.3.0. The final version was released on March 23, 2017. Angular 4 is backward compatible with Angular 2.

Features in version 4.3

Introducing **HttpClient**,a smaller, easier to use, and more powerful library for making HTTP Requests.

· New **router life cycle** events for Guards and Resolvers. Four new events: **GuardsCheckStart**, **GuardsCheckEnd**, **ResolveStart**, **ResolveEnd** join the existing set of life cycle event such as NavigationStart.

· Conditionally **disable** animations.

Version 5.0.0

Angular 5 was released on November 1, 2017 Key improvements in Angular 5 include support for progressive web apps, a build optimizer and improvements related to Material Design.

Version 6.0.0

Angular 6 was released on May 4, 2018.This is a major release focused less on the underlying framework, and more on the toolchain and on making it easier to move quickly with Angular in the future, like: ng update, ng add, Angular Elements, Angular Material + CDK Components, Angular Material Starter Components, CLI Workspaces, Library Support, Tree Shakable Providers, Animations Performance Improvements, and RxJS v6.

Future releases

Angular 7 releases will be showing up in September or October 2018. Each version is expected to be backward-compatible with the prior release.

React native

React Native is a JavaScript framework for writing real, natively rendering mobile applications for iOS and Android. It's based on React, Facebook's JavaScript library for building user interfaces,

but instead of targeting the browser, it targets mobile platforms.React Native was announced by Facebook in 2015, applying the React architecture to native Android, iOS, and UWP applications.

History

In 2012 Mark Zuckerberg commented, "The biggest mistake we made as a company was betting too much on HTML5 as opposed to native". He promised that Facebook would soon deliver a better mobile experience.

Inside Facebook, Jordan Walke found a way to generate iOS UI elements from a background JavaScript thread. They decided to organize an internal hackathon to perfect this prototype in order to be able to build native apps with this technology.

After a few months of development, Facebook released the first version for the React.js Conf 2015. During a technical talk, Christopher Chedeau explained that Facebook was already using React Native in production for their Group App and their Ads Manager App.

React was created by Jordan Walke, a software engineer at Facebook. He was influenced by XHP, an HTML component framework for PHP. It was first deployed on Facebook's newsfeed in 2011 and later on Instagram.com in 2012. It was open-sourced at JSConf US in May 2013.

React Native, which enables native Android, iOS, and UWP development with React, was announced at Facebook's React.js Conf in February 2015 and open-sourced in March 2015.

On April 18, 2017, Facebook announced React Fiber, a new core algorithm of React framework library for building user interfaces. React Fiber was to become the foundation of any future

One-way data binding with props

Properties (commonly, *props*) are passed to a component from the parent component. Components receive props as a single set of immutable values (a JavaScript object). Whenever any prop value changes, the component's render function is called allowing the component to display the change.

Stateful components

States hold values throughout the component and can be passed to child components through props:

Virtual DOM

Another notable feature is the use of a "virtual Document Object Model", or "virtual DOM". React creates an in-memory data structure cache, computes the resulting differences, and then updates the browser's displayed DOM efficiently. This allows the programmer to write code as if the entire page is rendered on each change, while the React libraries only render sub components that actually change.

Lifecycle methods

Lifecycle methods are hooks which allow execution of code at set points during the component's lifetime.

. shouldComponentUpdate allows the developer to prevent unnecessary re-rendering of a component by returning false if a render is not required.

. componentDidMount is called once the component has 'mounted' (the component has been created in the user interface, often by associating it with a DOM node). This is commonly used to trigger data loading from a remote source via an API.

. Render is the most important lifecycle method and the only required one in any component. It is usually called every time the component's data is updated, reflecting changes in the user interface.

JSX

JavaScript XML (JSX) is an extension to the JavaScript language syntax. Similar in appearance to HTML, JSX provides a way to structure component rendering using syntax familiar to many developers. React components are typically written using JSX, although they do not have to be (components may also be written in pure JavaScript).

Attributes

JSX provides a range of element attributes designed to mirror those provided by HTML. Custom attributes can also be passed to the component. All attributes will be received by the component as props.

Code written in JSX requires conversion with a tool such as Babel before it can be understood by web browsers. This processing is generally performed during a software build process before the application is deployed.

Common idioms

React does not attempt to provide a complete 'application framework'. It is designed specifically for building user interfaces and therefore does not include many of the tools some developers might consider necessary to build an application. This allows the choice of whichever libraries the developer prefers to accomplish tasks such as performing network access or local data storage. Common patterns of usage have emerged as the library matures.

Use of the Flux architecture

To support React's concept of unidirectional data flow (which might be contrasted with Angular's bidirectional flow), the Flux architecture represents an alternative to the popular Model-view-controller architecture. Flux features *actions* which are sent through a central *dispatcher* to a *store*, and changes to the store are propagated back to the view. When used with React, this propagation is accomplished through component properties.

Working principles

The working principles of React Native are basically the same as react except that it is not manipulating the DOM via the Virtual Dom but some native views. It runs in a background process (which interprets the JavaScript written by the developers) directly on the end-device and communicates with the native platform via a serializable, asynchronous and batched Bridge.

It can be seen that Facebook corrected the error that Mark Zuckerberg mentioned in 2012.React Native doesn't rely on HTML5 at all, everything is written in JavaScript, and relies on native SDKs.

2. Future Development

Project status can be tracked via the core team discussion forum. However major changes to React go through the Future of React repository issues and pull requests. This enables the React community to provide feedback on new potential features, experimental APIs and JavaScript syntax improvements.

3. Criticism of Reactjs

A criticism of ReactJS is that it has high memory (RAM) requirements, since it uses the concept of a "Virtual DOM". This is where "a representation of a UI is kept in memory and synced with the 'real' DOM by a library such as ReactDOM."

Comparison	AngularJS	Angular	React
Managed by	Google	Google	Facebook
Object	Scope	Components	State
Binding	Twoway	Two way	Oneway
DOM	Real DOM	Real DOM	VirtualDom
Language	JS	Type script	JSX
Open source		Sep 2016.	March 2015
Usage- all the web sites	-	0.4%	0.1%
Usage - market share	-	0.6%	0.2%

4. Conclusion

React and Angular is used to develop web application from small, start up, mid-level to enterprise level business. Both are flawlessly powerful and flexible technology that will remain in the future web world for a long time.

References

1. https://en.wikipedia.org/wiki/AngularJS
2. https://github.com/reactjs/reactjs.org
3. https://en.wikipedia.org/wiki/Angular_(application_platform)
4. https://en.wikipedia.org/wiki/React_(JavaScript_library)
5. https://github.com/angular/angular
6. https://www.angularminds.com/blog/article/major-differences-between-angularjs-and-reactjs.html
7. https://www.sitepoint.com/react-vs-angular/
8. https://medium.com/nybles/demystifying-react-s-virtual-dom-eb0f2dc0717a
9. https://w3techs.com/technologies/overview/javascript_library/all

Research Explorer - International Journal
ISSN : 2250-1940 (Print), 2349 - 1647 (Online)
Impact Factor : 2.014(IRJIF), 1.85(JIF), 1.056(RIF), 0.60(QF) 0.398(GIF)
Issue - 19 ; August 2018

SURVEY OF CLASSIFICATION ON NETWORK INTRUSION USING DATA MINING TECHNIQUES

Dr.M.Savitha Devi
Assistant Professor,
Department of Computer Science,
Periyar University Constituent College of Arts & Science,
Harur, Tamilnadu, India.

K. Mohanapriya
Guest Lecturer,
Department of Computer Science,
Government Arts College for Women,
Krishnagiri, Tamilnadu, India.

Abstract

Most of the people use the internet for online activities such as bank transaction, e-commerce etc., and the current security device firewall is not able to detect the malicious activities in the network traffic and computer usage. So, an intrusion detection system is used to detect the all types of malicious activities.This includes network attacks against unprotected services, data driven attacks on applications, host based attacks such as privilege violation, unauthorized logins and access to sensitive files. Intrusion Detection Systems are the main components in Network Security and it allows network administrators to detect policy violations. Existing Intrusion Detection supports to detect known service level network attacks and decision trees being efficiently used in these systems. This work aims at developing generic as well as systematic techniques to classify the normal and abnormal activitiesat all layers of the Open Systems Interconnection (OSI) model. Random Forest, Random Trees are the Data Mining techniques used to classify the suspicious activities and it improve the recall and precision and also decreases the number of trees that are required.The results demonstrate the effectiveness of the proposed method in improving the classification of intrusions.

Keywords: Intrusion Detection Systems (IDS), Decision trees, Random Forest.

1. Introduction

Data mining has functional system to strengthen the decision making by limiting the available surplus information. The information overload can be reduced by properly extracting and screening useful knowledge by identifying correlation and patterns from the extensive data surveyed in by the organizations. “The extracted information is used to predict, classify, model

and summarize the data being mined. Data Mining technologies, such as rule induction, neural networks, genetic algorithms, fuzzy logic and rough sets are used for classification and pattern recognition in many industries". Their ability to find the normal behaviour activity from abnormal behavior has been exploited in various contexts. Lately, the use of data mining techniques in the context of network intrusion detection is found effective [1]. Extensive algorithm acquired from the fields of statistics, pattern recognition, machine learning, and the database is available due to the swift progress and development of data mining.

An Intrusion Detection System (IDS) various types of malicious network traffic and erratic computer usage that a conventional firewall oversees. The attacks such as network attacks against vulnerable services, data driven attacks on applications, host based attacks such as privilege escalation, unauthorized logins and access to sensitive files, and malware (Viruses, Trojan Horses, and Worms) are detected by the IDS but not by the conventional firewall. An ID is constructed from three components: Sensorsthat monitor the traffic in the network or the activity in a system and it provokes events accordingly. A Consolethatsupervises the events to alarm and controls the sensors in case of misbehavior, Detection Enginethat store the logged events from the sensor in a database and generate alerts according to the system of rules from the received security events [2].

The intrusion detection techniques based upon data mining are generally fallen into one of two categories: misuse detection and anomaly detection.In misuse detection, each instance in a data set is labeled as 'normal' or 'intrusive' and a learning algorithm is trained over the labeled data. Research in misuse detection has focused mainly on detecting network intrusions using various classification algorithms, association rules, and cost-sensitive modeling. Unlike the traditional signature-based IDSs that have signatures for the earlier known attacks, Models of misuse are automatically created compared to signature-based IDSs and are more modern and accurate compared to signatures that are manually created. Anomaly detection algorithms build models of normal behavior and automatically detect any deviation from it [3].

Anomaly Detection – models of normal behavior are built and deviation from it is detected automatically in this algorithm. Unforeseen attacks can be identified potentially in this technique. Additionally, new or unusual but non-intrusive network behavior can be detected which interests network manager. High false alarm rate is a major disadvantage of this system. Supervised and unsupervised are two major categories of anomaly detection techniques, when a normal dataset is trained on and also a new set of test data is given. The objective of this is to detect whether the test data given is 'normal' or 'anomalous'. Again, in supervised anomaly detection the models are created based on the normal behavior on the network; anomalous behavior is detected in unsupervised anomaly detection without knowledge about the training data. The basis for unsupervised anomaly detection is clustering, statistical approaches, state machines and outlier detection schemes.

The major benefit of anomaly detection algorithms is their ability to potentially detect unforeseen attacks. In addition, they may be able to detect new or unusual, but non-intrusive, network behavior that is of interest to a network manager. A major limitation of anomaly detection systems is a possible high false alarm rate. There are two major categories of anomaly detection techniques, namely supervised and unsupervised. In supervised anomaly detection, given a set of normal data to train on, and given a new set of test data, the goal is to determine whether the test data is 'normal' or anomalous. Unlike supervised anomaly detection where the models are built only according to the normal behavior on the network, unsupervised anomaly detection attempts to detect anomalous behavior without using any knowledge about the training data. Unsupervised anomaly detection approaches are based on statistical approaches, clustering, outlier detection

schemes, state machines, etc.

Decision tree is used as a predictive model in Decision tree learning to map observations made about an event or item to the conclusions based on the item's target value. Decision tree algorithm partitions the dataset of records recursively until all the data items belong to a particular class in any of the following approach: depth-first greedy approach or breadth-first approach. The tree structure of any decision tree consists of a root, internal nodes and leaf nodes. The node at the top of the structure is the root node. The test condition on an attribute is represented in internal node, the result of the test condition is represented in the branch and the class label is represented in the leaf or terminal node. Decision rule are the paths in the decision tree. A divide and conquer approach is used in the construction of the tree [5].

Decision tree construction generally uses the greedy approach from top to bottom. The classification method is done in two stages or phases tree building and tree pruning. The tree building as mentioned previously partitions the records recursively until all the data items are grouped under a same class. It follows the top-down approach. The tree building involves exhaustive computation as it needs to traverse the entire training dataset repeatedly to build an entire tree. The tree pruning is used to improve the performance of the decision tree algorithm like prediction and classification accuracy by minimizing over-fitting problem of tree. The misclassification error in the algorithm occurs due to the over-fitting problem of tree. There are many decision tree based algorithms like ID3, C4.5, C5.0, CART etc. These algorithms have the merits of high classifying speed, strong learning ability and simple construction.

The ensemble idea in supervised learning has been intensively investigated and thus suggests combining of two linear regression models. The first linear regression model is fitted to the original data and the second linear model to the residuals. In this paper ensemble technique Randomforest is used for classify the intrusion activities

An ensemble of similarly configured neural networks was introduced to improve the predictive performance of a single one.The foundations for the award winning AdaBoost algorithm by were laid to show that a strong classifier can be generated by combining"weak" classifiers whose classification performance is only slightly better than random classification in the Probably Approximately Correct (PAC) sense. The quality and robustness of unsupervised tasks and the quality and robustness of clustering algorithms can be improved by means of Ensemble methods. The potential usefulness of ensemble methods has paved way to vast number of methods that are now available to researchers and practitioners [6].

Predictive power of Hundreds of individual classifiers is combined to form ensemble based classifier (e.g. decision trees or ANN).As decision trees are fast and unstable, they are quite suited for ensembles. The limitation of the decision trees is that it can become unstable due to any small variations in the data resulting in a new tree which is completely different. This instability can be mitigated by using them in an ensemble. To obtain a classifier with superior performance when compared to individual classifiers, multiple classifiers are used and weighted and then combined for this purpose [7].

2. Related works

Kenkre et al.,[8] a solution was formulated by Kenkre et al.,[8] by combining logging and NIDS and prevention system.Software Engineering framework requisites of design, testing, analysis and implementation. For IPS, in an online mode, open source tool snort was created, so that packets were captured by sensors and dropped if any suspicious activity was seen. Administrator was alerted by IPS through signatures for detection and prevention in dropping the packet. Splunk

helped in logging of dropped packets. A cost effective, scalable, customizable solution is provided in this process as an alternate to organizations.

Relan&Patil [9] proposed the two feature selection techniques, C4.5 Decision tree algorithm and C4.5 Decision tree with Pruning.The discrete value attributes are only for considered for classification in C4.5 Decision tree with pruning. They have used KDDCup'99 and NSL_KDD dataset for training and testing the classifier. Results of the study showed that C4.5 decision tree gave better results with all most 98% of accuracy.

A new hybrid intelligent system was put forth by Panda et al., [10] by combining Naïve Bayes with Decision Trees (NBDT) and by combining Non-nested generalized exemplar (NNge) and extended repeated incremental pruning (JRip) rule-based classifiers (NNJR), so as to construct multiple classifier systems to detect network intrusions effectively.To enhance the detection rate of the proposed hybrid system the ensemble design using AdaBoost was used. Farthest First Traversal (FFT) clustering was combined with classification techniques to obtain another two hybrid methods such as DTFF (DT +FFT) and FFNN (NNge + FFT) to have a better overall detection. Finally, for better detection rate, Bayesian belief network with Tabu search combined with NNge was used. Substantial experiments are conducted using NSL-KDD dataset, which is a modified version of KDD99 intrusion dataset.

3. Methodology

This section details the ensemble techniques like random forest in the classification of intrusions. Figure 2 shows the block diagram of methodology. The KDD99 dataset is used in this work to classify the intrusions using various classifiers and their performance is compared and a proposed Tabu search method is introduced to improve the classifiers performance.

3.1 KDD99 Data Set

This is the data set used for the Third International Knowledge Discovery and Data Mining Tools Competition, which was held in conjunction with KDD-99 and the Fifth International Conference on Knowledge Discovery and Data Mining. The competition task was to build a network intrusion detector, a predictive model capable of distinguishing between "bad" connections, called intrusions or attacks, and "good" normal connections. This database contains a standard set of data to be audited, which includes a wide variety of intrusions simulated in a military network environment.KDD (Knowledge Discovery and Data Mining) is a primary dataset used in the evaluation process of any novelty proposed intrusion detection system.

3.2 Random Forest (RF)

Machine learning approaches are currently used in the data mining domain [11], Random Forest (RF) is one of the supervised learning algorithm utilizing the ensemble machine learning technique. The decision tree is used as a common classifier in the RF and it generates multiple decision trees to form a forest. The process undergoes randomization at two stages; the bootstrap samples are randomly sampled as in bagging and then the individual base decision tree is generated by random selection of input attributes. The generalization error of Random Forest classifier is based on the strength of individual decision tree and correlation among base trees. The performance of the Random Forest classifier is at equivalence with the other ensemble techniques like bagging and boosting based on the accuracy measure. Random Forest has the ability to potentially run on large databases that have millenary of input variables and can deal with these variables without any deletion in the variable as stated by [Brieman 2001]. The RF finds importance in datamining due to certain characteristics like, the important variables can be estimated, as forest growing

progresses it can generate an internal unbiased estimate of generalization error, the missing data estimation methods are effective, the accuracy is not compromised in spite of a large proportion of data missing, and methods for balancing class error in class population unbalanced data sets is available. The Random Forest are parallel nature inherently that allows it for parallel implementations using multithreading, multi-core, and parallel architectures. Random Forest is recently and widely used in numerous applications for classification and prediction.

The collection of tree-structured classifiers {*h(x,Èk) k=1,2, ….*}is constructed to form a forest in RF classifier. The independent identically distributed random vectors are the tree and denoted by {È*k*}, each tree casts a unit vote for the most popular class at input *x* [17]. An ensemble of decision trees is generated in RF. Following are the steps followed by Breiman to generate every single tree in RF: The Bootstrap sample is done on the N records of the training set; these records are sampled from the original data at random but with replacement. This happens to be the training sample set for constructing the tree. The tree is grown to the largest extent possible and the process of growing the tree is by splitting the attributes to split the nodes. For an input variable of size M, a number m less than M is chosen so that m variables are selected randomly at every node. The best split on these m attributes is used to split the node. During the forest growing stage, the value of m remains constant. Pruning is not applied here.

The number of trees in a forest is determined by the parameter Ntree. Multiple trees are grown in the similar way as an individual tree grows. The number of variables (m) used at the node can also be called as mtree or k. The number of the instances in at leaf node i.e. parameter node size decides the depth of the tree. The new instance can be classified after training i.e. the building of the tree by running a search process across the entire forest build by several trees. The vote is the classification value of the new sample done by a tree and is recorded at each node. The maximum counts of votes for a class obtained after combining all the votes from all the tree is the classification of the new instance. The forest of decision trees generated by this process is the Random Forest.

The explicit feature elimination in classification is done by the random forest classifier with its associated Gini feature importance.

Gini index has been introduced by Breiman et al.,[12] in building the Classification and Regression Trees CART technique. However, it has been first introduced by the Italian statistician Corrado Gini in 1912. The index is a function that could be used to measure the impurity of the data, i.e., how uncertain we are if an event will occur. In classification, this event would be the determination of the class label. The Gini impurity function in its original form is calculated as in equation (1):

$$Gini(t) = 1 - \sum_{i=1}^{w} P(C_i \mid t)^2 \quad (1)$$

Where t is a condition, w the number of classes in the data set, and C_i is the ith class label in the data set. By removing the condition t from the original form of the previous equation, we can calculate the level of impurity for any data set before splitting as in equation (2):

$$Gini(\text{class}) = 1 - \sum_{i=1} P(C_i)^2 \quad (2)$$

The Gini index of any attribute A can then be calculated as in equation (3):

$$GiniIndex(A) = Gini(Class) - \sum_{j=1}^{m} P(a_j).Gini(A = a_j) \quad (3)$$

Where m is the number of values for the attribute A.

4. Results And Discussion

Experiments were carried out using the subset of KDD99 dataset consisting of 5087 normal and 6885 abnormal data stream. Table 1 and Figure 1& 2 shows true positive, true negative, false positive and false negative and Classification Accuracy, Precision, Recall , F Measure respectively.

Table 1 ResultsRANDOM FOREST			
TP	5684	PRECISION	0.939194
TN	5600	RECALL	0.946702
FP	368	ACCURACY	0.942533
FN	320	F MEASURE	0.942933

Figure 1

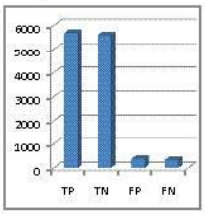
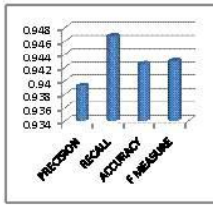


Figure 2



5. Conclusion

The usage of Data Mining approached is the key notion, including finding regular and utilizable patterns of system features that characterise the network behaviour by using the set of corresponding system features for identifying irregularities and familiar intrusions. Thus, the search is more flexible.Results show that the classification accuracy of analyzedrandom forestperforms better than previous decision tree models.

6. References

[1] Helali, R. G. M. (2010). Data mining based network intrusion detection system: A survey. Novel Algorithms and Techniques in Telecommunications and Networking, 501-505.

[2] Khem, D., Vadodaria, H., Aggarwal, M., MITESH, M. K., &Uchat, N. (2007). Intrusion Detection Systems. Indian Institute of Technology, Mumbai.

[3] Raghunath, B. R., &Mahadeo, S. N. (2008, July). Network intrusion detection system (NIDS). In Emerging Trends in Engineering and Technology, 2008. ICETET'08. First International Conference on (pp. 1272-1277). IEEE.

[4] Kumar, B. S., Ch, T., Raju, R. S. P., Ratnakar, M., Baba, S. D., &Sudhakar, N. (2013). Intrusion Detection System-Types and Prevention.

[5] Jadhav, S. D., &Channe, H. P. (2016). Comparative Study of K-NN, Naive Bayes and Decision Tree Classification Techniques. International Journal of Science an d Research, 319-7064.

[6] Rokach, L. (2010). Ensemble-based classifiers. Artificial Intelligence Review, 33(1), 1-39.Mounce, S. R., Ellis, K., Edwards, J. M., Speight, V. L., Jakomis, N., &Boxall, J. B. (2017). Ensemble Decision Tree Models Using RUSBoost for Estimating Risk of Iron Failure in Drinking Water Distribution Systems. Water Resources Management, 31(5), 1575-1589.

[7] Kenkre, P. S., Pai, A., &Colaco, L. (2015). Real time intrusion detection and prevention system. In Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA) 2014 (pp. 405-411). Springer, Cham.

[8] Relan, N. G., &Patil, D. R. (2015, January). Implementation of network intrusion detection system using variant of decision tree algorithm. In Nascent Technologies in the Engineering Field (ICNTE), 2015 International Conference on (pp. 1-5). IEEE.

[9] Panda, M., Abraham, A., &Patra, M. R. (2015). Hybrid intelligent systems for detecting network intrusions. Security and Communication Networks, 8(16), 2741-2749.

[10] Sinha, V. Y. K. P. K., &Kulkarni, V. Y. (2013). Efficient learning of random forest classifier using disjoint partitioning approach. In *Proceedings of the World Congress on Engineering* (Vol. 2, pp. 3-5).

[11] Glover, F. (1986). Future paths for integer programming and links to artificial intelligence. *Computers & operations research*, *13*(5), 533-549.

Research Explorer - International Journal
ISSN : 2250-1940 (Print), 2349 - 1647 (Online)
Impact Factor : 2.014(IRJIF), 1.85(JIF), 1.056(RIF), 0.60(QF) 0.398(GIF)
Issue - 19 ; August 2018 © Author

CONSIDERATIONS FOR AND CONSIDERATION AGAINST THE MOBILE PHONES

Dr. M.R. Lakshiminarayanan,
Principal and Associate Professor in Commerce,
Trinity College for Women (Arts and Science)
Namakkal – 637 002, TamilNadu, India.

Abstract:

Mobile phone is one of the modern technical communication equipment. It plays an impressive role in our daily life. Science and Technology developments are unavoidable in our current society. Particularly, mobile phone is very important communication network for all types of people. It is useful for our employment, obtaining data's, purpose of fast communication, etc., At the same time, our priceless time will be spent for this utilization of mobile phone. If used it rightly and sensibly, mobile phone can be excellent and wonderful piece of utility in our life.

Key words: Mobile Phones – Usage in various countries – Pros – Cons – Creativity of Cell Phones.

1. Mobile Phone

It is one of the communication services of Information Technology. It has many names such as cell phone, mobile phone, smart phone, telephone or touch screen phone. This is a portable device that access to cellular radio system. In simple devices we only access to make and receive the calls and send and receive the messages. Now, technology has been developed in day by day. Especially, cell phones are very important modern tool. Mobile phones play a powerful role in our day-to-day life. The main motive of the mobile phone is communication. Now it can also be used for entertainment purposes also.

This list ranks the countries of the world by the number of cell phones in use. It is not the number of phone devices that are being given below, but the number of phone numbers in a country.

Usage of number of cell phones - Top 50 countries in world wide

Rank	Country	No. of Cell phones	Year of information
1.	*China*	*1,100,000,000*	*2012*
2.	India	893,862,000	2013
3.	United States	310,000,000	2012
4.	Indonesia	281,960,000	2012
5.	Russia	261,900,000	2012
6.	Brazil	248,324,000	2012
7.	Japan	138,363,000	2011
8.	Vietnam	134,066,000	2012
9.	Pakistan	125,000,000	2013
10.	Nigeria	112,780,000	2012
11.	Germany	107,700,000	2012
12.	Philippines	103,000,000	2012
13.	Mexico	100,786,000	2012
14.	Italy	97,225,000	2012
15.	Bangladesh	97,180,000	2011
16.	Egypt	96,800,000	2012
17.	Thailand	84,075,000	2012
18.	United Kingdom	82,109,000	2012
19.	South Africa	68,400,000	2012
20.	Turkey	67,680,000	2012
21.	France	62,280,000	2012
22.	Ukraine	59,344,000	2012
23.	Argentina	58,600,000	2012
24.	Iran	58,160,000	2012
25.	Korea, South	53,625,000	2012
26.	Saudi Arabia	53,000,000	2012
27.	Poland	50,840,000	2012
28.	Spain	50,663,000	2012
29.	Colombia	49,066,000	2012
30.	Malaysia	41,325,000	2012
31.	Morocco	39,016,000	2012
32.	Algeria	37,692,000	2012
33.	Kenya	30,732,000	2012
34.	Venezuela	30,520,000	2012
35.	Taiwan	29,455,000	2012
36.	Peru	29,400,000	2012
37.	Kazakhstan	28,731,000	2012
38.	Sudan	27,659,000	2012
39.	Tanzania	27,220,000	2012
40.	Iraq	26,760,000	2012
41.	Canada	26,263,000	2012
42.	Ghana	25,618,000	2012
43.	Australia	24,400,000	2012
44.	Chile	24,130,000	2012
45.	Romania	22,700,000	2012
46.	Guatemala	20,787,000	2012
47.	Ethiopia	20,524,000	2012
48.	Uzbekistan	20,274,000	2012
49.	Cote d'Ivoire	19,827,000	2012
50	Netherlands	19,643,000	2012

2. Considerations For Mobile Phones

1. Connected to universal level We are always connected with our family and friends with the help of mobile. Cell phone technology gives the connectivity to everybody in the world.

2. Fashion and Appearance Point

Today's fashion world mobile phones are the statement of fashion. Without good quality and costly cell we would be hesitate to participate or engage in functions. Especially, android phone or touch screen phone will give us reputation.

3. Business and Employment

First of all the millions of people are connected with the business of telephones. And day to day the job opportunities and scope are creating for new youngster. This is not himself business through the medium of cell phones we can control our business.

4. Entertainment and Enjoyment

Up to the minute children are very affectionate of cartoons, games, children's stories and other entertaining materials and equipments. This is good thing that small children and youngsters are always in the homes under the eyes of parents and relatives.

5. Knowing of location and site

GPS (Global Positioning System) technology is allows us to get information of our location. We can easily know the location with the help of GPS during traveling.

6. Setting of Time, Alarm Notes & Reminder

Wants to go office early bird in the morning but we are always delay, alarm is the best solution and solving for this problem. We can also add notes and reminders in our mobile phones in the favor of our future help.

7. More number of utilities and services

The entire features are at present available in one device. Especially in the modern touch screen mobile phones. We can easily and comfortably access to past and future Calendars as well as see the dates and years. The advanced camera is available in our latest android touch screen mobiles. There is no need to buy Calculator from outside because this facility is available in every mobile in advance. Plenteous smart phones even simple sets also have built-in torch. It is very helpful to us on dark times

8. Transfer of Data and documents

Now we can easily transfer and transmission of data from one device to another device. Beside this photo, documents, videos and other important documents are easily transferred from one device to another device within seconds. We can also store and save our data with our convenient time.

9. Students Learning Applications

There are plenteous learning applications for students in app store. With the help of these apps students can learn as well as study and get the applications according their subjects.

10. Legal Issues

Occasionally we will face the legal issues in our life. These issues can be cleared up by mobile with proof of call recording or SMS (Short Message Service) data recovery. It is one of the

remedy of legal on the up and up.

3. Consideration Against The Mobile Phones

1. Distance from our blood relations and relatives

Use of mobile every time is kept far from our friends, family, well wishers and other relatives. This is the big demerits that we don't know about our family and friends. Because of mobile phone, we could not freely speak or interact with them.

2. Havoc of Time

Time and tide wait for nobody – It is an important proverb in English. But, unfortunately we spend more time with our cell phones. Time is money and money is time. Don't waste our time because of cell phones, don't use it much, always try to less use of mobile phones and save our time and spent it on good ways. We allocate more times with our mobile phone not on leisure hours only, we use this mobile on eating times, studying times, etc.,

3. Loss of cash/money

Before technology developed there were simple mobile cell phones everywhere and at present because of people's preference today's these ordinary sets are very little and useful. But as technology progressed new and modern touch mobile phones come to market, which are very expensive and peoples allocate money on these sets which is loss of money. Moreover traders and dealers are also trading in these sets, sometimes these sets become dead at once or the loss their value in market owing to coming new latest mobile models. By means of these new models old models are valueless, useless and have loss in business.

4. Waste of our body energy

We are reading the limitations of mobile phones; it is another biggest factor that effects on human health. It is cause of waste the energy, during use of cell phone all the parts of human body are spending energy on it, like; eyes, hands and mind, etc.

5. Loss and fall of human life

Some peoples use it Cell phone during driving, which may cause of accident. Due to accident, priceless human lives have been lost their life. According to recent survey, many accidents are held on account of using cell phones while vehicle driving.

6. Cause of Diseases

Mobile Phone can be cause of skin diseases particularly hand skin, cancer and eye problems, etc., Apart from a lot of other diseases can be effect the body by its high usage.

7. Study Loss

In addition to the biggest disadvantage for students is they have very high loss of their studies due to using the mobile phones. At the end of the study session they got very poor results and many of them got fail in class room.

8. Stolen and plagiarized of Data's

If we have our personal images, videos, documents or files, etc., in our device. Other peoples can easily steal or theft our images and videos without our knowledge. In android mobile phone it is easy to copy data from one device to another but in IOS (Internetwork Operating System) operating system have little and slightly safety.

9. Arrest in legal investigations and enquiries

If our cell phone is stolen, it is advice to us that immediately report in police station. For the reason that if someone make robbery in other place and they through/left our SIM card or device there. So it can be put us in legal investigations and enquiries.

Conclusion

The first and most important role that mobile phones play in our lives is that they give us an easy and fast way of communication. Cell phones have become an essential for many people all over the world. The ability to contact with family, business associates, communication through the social websites (face book, twitter, what sup, QQ, we chat, ozone, tumblr, instagram, skype, viber, line, snapehat, pinterest, linkedln, telegram, reddit, foursquare, you tube, classmates, quora, etc.,) and access to email are the reasons for the increasing importance of cell phones. Today's technically advanced cell phones are capable of not only receiving and placing phone calls, but storing data, taking films, download the data's, internet applications and can even be used as walkie talkies. When cell phones were first introduced to the society, they were bulky, expensive, and low options in the hand set. Now, mobile applications of mobile phone's (particularly, android/ touch screens) have been improved and developed greatly owing to the use of satellites and wireless services. Science and Technology growth are inevitable in this society. We welcome our science and technology development. However, we use this priceless information technology of mobile phone and its applications only for our knowledge development and not for time passing.

4. Reference

https://en.wikipedia.org/

Research Explorer - International Journal
ISSN : 2250-1940 (Print), 2349 - 1647 (Online)
Impact Factor : 2.014(IRJIF), 1.85(JIF), 1.056(RIF), 0.60(QF) 0.398(GIF)
Issue - 19 ; August 2018

A COMPARISION AND EVALUATION OF DIJKSTRA'S AND FLOYD ALGORITHM

Dr.N.Elamathi
Assistant professor,
Department of computer science,
Trinity College For Women, Namakwa, Tamilnadu, India.

R.Nithya
M.Phil Research Scholar,
Department of computer science,
Trinity College For Women, Namakwa, Tamilnadu, India.

Abstract

The shortest path problem is among the most fundamental combinatorial optimization problems to answer reach ability queries. It is hard to determine which vertices or edges are visited during shortest path traversals. This paper's main objective is to evaluate the Dijkstra's Algorithm, Floyd-warshall Algorithm, in solving the shortest path problem. First, we compute the shortest paths between set of vertices. Each shortest path is considered as one transaction. Second, we utilize the pattern mining approaches to identify the frequency of occurrence of the vertices. The results of evaluating the Dijkstre's and Floyd-warshall algorithms along with their time complexity conclude the paper.

Keywords: shortest paths, Dijkstre's,Floyd-warshall

1. Introduction

The shortest path problem is a problem of finding the shortest path or route from a starting point to a final destination. Generally, in order to represent the shortest path problem we use graphs. A graph is a mathematical abstract object, which contains sets of vertices and edges. Edges connect pair of vertices. Along the edges of a graph it is possible to walk by moving from one vertex to other vertices. Depending on whether or not one can walk along the edges by both sides or by only one side determines if the graph is a directed graph or an undirected graph. In addition, lengths of edges are often called weights, and the weights are normally used for calculating the shortest path from one point to another point. In the real world it is possible to apply the graph theory to different type of scenarios. For example, in order to represent a map we can use a graph, where vertices represent cities and edges represent routes that connect the cities. If routes are one-way then the graph will be directed; otherwise, it will be undirected. There exist different type of algorithms that solve the shortest path problem. However, only several of the most popular

conventional shortest path algorithms along with one that uses Genetic algorithm are going to be discussed in this paper, and they are as follows:

- Dijkstra's Algorithm
- Floyd-Warshall Algorithm

The overall architecture of the system is given in Fig. 1. The source graph data is usually available in the form of set of vertices and edges. There are two main components to find the shortest path overlapped regions. Firstly, the graph traversal algorithm is used to find all pairs of shortest paths, i.e. P. Each path p_i consists of sequence of vertices from source to destination. The graph traversal module produces set of shortest paths between all pair of source and destination as intermediate results. All the shortest paths are computed using well-known Dijkstra algorithm. Secondly, the overlapped regions are identified through pattern mining approach. The visualized statistics enable us to derive useful remarks. The subsequent discussion highlights the details of shortest path computation and pattern mining modules respectively.

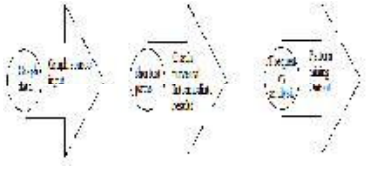


Fig. 1: System Architecture

2. Research Objectives

- To determine and identify the concepts of the shortest path problem.
- To determine the representation of graphs in computer in order to solve the shortest path problem, as well as to understand the different basic terms of a graph.
- To explain the general concepts and the implementations of Dijkstra's Algorithm, Floyd-Warshall Algorithm, Bellman-Ford Algorithm, and Genetic Algorithm.
- To evaluate each algorithm, and presents the evaluations' results.

In order to represent a graph in a computer we will use adjacency matrix *a*. The dimension of the matrix will be equal to *(n* x *n)*, where *n* is number of vertices in graph. The element of matrix *a[i][j]* is identified by an edge that connects the *i-th*and *j-th*vertices; the value here represents the weight of the corresponding edge.

However, if there is no edge between vertices *i* and *j*, the value in (*a[i][j]*) will be equal to *infinity*. An array of edges is another common representation of the graph. If *m* is the number of edges in a graph, then in order to represent the graph we have to use *m* x *3* two-dimensional arrays; in each row, the first vertex, the second vertex, and the edge that connects them are also stored. The benefit of using an array of edges in comparison to adjacency matrix is when there is more than one edge that connects two vertices we cannot use adjacency matrix in order to represent graph.

For each vertex within a graph we assign a label that determines the minimal length from the starting point *s* to other vertices *v* of the graph.

Characteristics of a shortest path algorithm

Dijkstra's algorithm selects an arbitrary starting vertex and then branches out from the tree constructed so far. Each of the nodes it visits could be in the tree, in the fringes or unseen. The

designators:

- TREE- nodes in the tree constructed so far
- FRINGE- not in the tree, but adjacent to some vertex in the tree
- UNCEEN – all others

Designate for each node in the graph whether the node is part of the shortest path tree, a part of the set of fringes adjacent to the nodes in the tree or a part of, as of yet , unseen set of graph nodes. A crucial step in the algorithm is the selection of the node from the fringe edge. The algorithm always takes the edge with least weight from the tree to the fringe node. This is an example of a greedy algorithm which are used in optimization problems. They make locally optimal choice in hope that this will provide a globally optimal solution. The applet performs the following operations:

- Add node - adding a node to the graph
- Remove a node – deleting a node from the graph
- Enter weight – entering the weight of the edge.

3. Dijkstra's Algorithm Steps

- Set all vertices = infinity except for the source vertex, set the source distance=0.
- Push the source vertex in a min-priority queue in the form (distance, vertex), as the comparison in the min-priority queue will be according to vertices distances.
- Pop the vertex with the minimum distance from the priority queue (at first the popped vertex = source).
- Update the distances of the connected vertices to the popped vertex in case of "current vertex distance + edge weight < next vertex distance", then push the vertex with the new distance to the priority queue.
- If the popped vertex is visited before, just continue without using it.
- Apply the same algorithm again until the priority queue is empty.

Time complexity of Dijkstra's algorithm is $O(V^2)$ but with min-priority queue it drops down to O(V log V).

```
for(inti=0 ;i<v[x].size();i++){
int e = v[x][i].f; int w =v[x][i].s;
if(dist[x] + w <dist[e])
{
Dist[e] =dist[x] + w;
s.insert({dist[e],e});
}
}
```

4. Floyd-Warshall's Algorithm Steps

For a graph with N vertices

- Initialize the shortest paths between any 2 vertices with infinity.

- Find all pair shortest paths that use 0 intermediate vertices, than find the shortest paths that use 1 intermediate vertex and so on. Until using all N vertices as intermediate nodes.

Minimize the shortest paths between any 2 pairs in the previous operation. There any 2 vertices (I , j), one should actually minimize the distances between this pair using the first K nodes, so the shortest path will be: min(dist [i][k] + dist [k][j], dist [i][j]). Dist[i][k] represents the shortest path that only uses the first K vertices, dist [k][j] represents the shortest path between the pair k, j. As the shortest path will be a concatenation of the shortest path from I to k, then from k to j.

```
for(int k = 1; k <= n; k++){

for(inti = 1; i<= n; i++){

for(int j = 1; j <= n; j++){

dist [i][j] = min(dist [i][j], dist [i][k] + dist
[k][j] );

        }

    }

}
```

Time complexity of floyd-warshall's algorithm is$O(V^3)$, where V is the number of vertices in a graph.

5. Conclusion

Dijkstra algorithm is a better way to solve the problem of shortest path routing in graph. It's faster, guarantees optimality and has less computationalcomplexilibity. Dijkstra's cannot be used on all type of graph. It's fails on graphs with negative edges. The Floyd-warshall algorithm helps in determining all the shortest paths in a graph with positive or negative weight edges. It is easy to modify the algorithm and use it to reconstruct the paths. Versions of the algorithm can be used for finding the widest paths between all pairs of vertices in a weighted graph or transitive closure of relation R. The computed time complexity for each of Dijkstra's, Floyd-warshall show that these algorithms are acceptable in terms of their overall performance in solving the shortest path problem. All of these algorithms produce only one solution.

6. References

[1]. CamilDemetrescu, A. V., Implementation challenge for shortest paths. In Encyclopedia ofAlgorithms, (2008).

[2]. Kamran Najeebullah, KifayatUllah Khan, Waqas Nawaz, Young-Koo Lee, BiShardParal-lel Processor: A Disk-Based Processing Engine for Billion-Scale Graphs. InternationalJournal of Multimedia & Ubiquitous Engineering. vol(9-2), (2014).

[3]. Kyrola, A., Blelloch, G., Guestrin, C. GraphChi: Large-scale graph computation on just aPC. In OSDI (2012).

[4]. Jun Gao, Jiashuai Zhou, Jeffrey Xu Yu, and Tengjiao Wang, Shortest Path Computing inRelational DBMSs, IEEE TKDE, (2013).

[5]. Jiawei Han, Jian Pei, Yiwen Yin, Runying Mao, Mining Frequent Patterns without Candi-date Generation: A Frequent-Pattern Tree Approach. Data Min. Knowl. Discov. 8(1): 53-87, (2004).

[6]. Yu-Li Chouy H. Edwin Romeijnz Robert L. Smithx. Approximating Shortest Pathsin Large-scale Networks with an Application to Intelligent. Transportation Systems.September 27, 1998

[7] Faramroze Engineer. Fast Shortest Path Algorithms for Large Road Networks.Department of Engineering Science. University of Auckland, New Zealand.
http://www.esc.auckland.ac.nz/Organisations/ORSNZ/conf36/papers/Engineer.pdf

[8] Roozbeh Shad, Hamid Ebadi, Mohsen Ghods. Evaluation of Route Finding
Methods in GIS Application. Dept of Geodesy and Geomatics Eng. K.N.Toosi Universityof Technology, IRAN. http://www.gisdevelopment.net/technology/gis/ma03202.htm

[9]. Fu, Mengyin, Li, Jie, Deng, Zhihong. A practical route planning algorithm forvehicle navigation system. Proceedings of the World Congress on Intelligent Control andAutomation (WCICA), v 6, WCICA 2004 - Fifth World Congress on Intelligent Controland Automation, Conference Proceedings, 2004, p 5326-5329.

[10]. Thomas Willhalm. Speed Up Shortest-Path Computations. January 27, 2005.
http://www.mpisb.mpg.de/~sanders/courses/algen04/willhalm.pdf

[11] Thomas H. Cormen, Clifford Stein, Ronald L. Rivest and Charles E. Leiserson.Introduction to Algorithms, 2nd edition, McGraw-Hill Higher Education, 2001,ISBN: 0070131511

[12] Pi´oro, M., Medhi, D.: Routing, Flow, and Capacity Design in Communication andComputer Networks. Morgan Kaufmann, San Francisco (2004).

[13]. Pi´oro, M., Szentesi, A., Harmatos, J., J¨uttner, A.: On OSPF related network optimizationproblems. In: 8th IFIP Workshop on Performance Modelling and Evaluationof ATM & IP Networks, pp. 70/1–70/14. Ilkley, UK (2000).

[14]. Prytz, M.: On optimization in design of telecommunications networks with multicastand unicast traffic. Ph.D. thesis, Royal Institute of Technology, Stockholm,Sweden (2002).

[15]. Tomaszewski, A., Pi´oro, M., Dzida, M., Mycek, M., Zagoÿzdÿzon, M.: Valid inequalitiesfor a shortest-path routing optimization problem. In: B. Fortz (ed.) 3rd InternationalNetwork Optimization Conference. Universit´eLibre de Bruxels, Brussels(2007).

[16]. Tomaszewski, A., Pi´oro, M., Dzida, M., Zagoÿzdÿzon, M.: Optimization of administrativeweights in IP networks using the branch-and-cut approach. In: L. Gouveia,

[17]C. Mour˜ao (eds.) 2nd International Network Optimization Conference, pp. 393–400. Universidade de Lisboa, Lisbon (2005).

Research Explorer - International Journal
ISSN : 2250-1940 (Print), 2349 - 1647 (Online)
Impact Factor : 2.014(IRJIF), 1.85(JIF), 1.056(RIF), 0.60(QF) 0.398(GIF)
Issue - 19 ; August 2018

NBPST: A NOVEL BACKUP PATH SELECTION TECHNIQUE TO ENHANCE THE IP ROUTING IN MANETS

K.Ramkumar,
Assistant Professor,
Department of Computer Science,
SRM Trichy Arts and Science College,
Irungalur, Trichy, Tamilnadu, India.

Gururaj H L,
Assistant Professor ,
Computer Science and Engineering,
Vidhya Vardhaka College of Engineering,
Mysuru, Karnataka, India.

Abstract

MANETs are moving nodes, self-adjustment and non-infrastructure based networks of mobile devices associated to each other. In an ad hoc, the complexity of the network in each node is much higher than others because each and every node has to use medium access mechanism to effectively handle the hidden terminal problems or exposed terminal problems, and also priority mechanism to provide the Quality of Service (Quos). Back-up paths are used in internet protocol link failure to prevent the packet loss. In an existing system, the selection of the backup path is not an efficient. It is used, when the source fails to send data packet to the destination through a particular path. So, it prefers to use ad hoc on demand distance vector protocol. Nodes maintain a "precursor list" which contains the IP address for each of its neighbors to use it for a next hop in their routing table. It uses Mobile Ad hoc Network with the realistic path from source node to destination node and protects each node with backup path. The main aim of this paper is to recover the path from failures and to reduce data loss. Finally, it enhances the QoS in MANETs.

Keywords- MANET, Routing, Back-up path, Failure, IP link.

1. Introduction

The use of internet and other multimedia has changed enormously over the last few decades. Time insensitive applications like e-mail, the Internet is now facing more challenging tasks. In internet, every user gets various types of data and can send from one computer to another computer through network. Generally, there are two types of network such as wired and Wireless network in the world. In wired network, the communication is established using static based

infrastructure like creating connection using wires before initiating the communication. The transmission rate would be in high speed when compared to wireless network. Ethernet cables are used to connect systems. Wireless network of ad hoc communication takes place between the nodes without having any access point or infrastructure. It can be established for communication at any instant of time. In wireless network, user can utilize small equipment like mobile devices, PDA's for reducing the higher cost because access can happen at any point of place or time.
Mobile Ad-hoc Network (MANET) is a promising field for development of wireless network. Due to its popularity, the mobile devices in wireless networks can be increased significantly over the previous decades, It has several challenges and the special features of MANET leads this technology to be great opportunistic altogether.

Routing is the process which selects the best path available between the source and the destination. When the source sends any data to mobile node, it assigns the IP address to the mobile node. But, mobile nodes move and it can be hidden. So, the source forwards packet to router and the router searches the best path and send packets to home network. Node inside the coverage area receives the data packet from the nearby neighbour. Suppose, If the mobile node is in the another station, then home agent forwards the data packet to Care Of Address (COA) through tunnelling where encapsulation is done and foreign agent de-capsulate the data packet and mobile node receives the data packet.

On several occasions, due to several reasons, the link between the source and the destination fails to send the packets successfully hence several packets might be lost. Generally, wireless network only selects the nodes by having the capabilities of transferring the data. Several nodes have to connect with a special node first to transmit the data when the receiver is inside their range. Failure in IP networks is common for all the nodes. IP based high speed networks Internet backbone and disconnection of a link for fewer seconds leads to drop millions of packets. The detection of Failure time is reduced from much too typically less than 50ms.

Therefore, there is a need to resolve this IP link failure is an on demand task to enhance the Internet reliability and enforce the contents availability, and it received much attention in the past years. An Internet Service Provider (ISP) provides services to access for participating to the internet. Currently, backup path based protection is widely used by ISPs in protecting the domains. The backup paths are configured, already computed, and stored in the routers by using this approach. When a link gets failed, originally traversing the traffic is immediately switched to the backup path of this link. Due to this, the routing faces disruption. The backup path is very necessary for packet forwarding from source to destination. A lot of protocols are used to select the backup path. One of the protocols is on-demand service vector protocol which is used to find the efficient and feasible path. It is used only when source and destination does not have path to transfer data.

2. related work

Wireless networks use air as a medium to transfer data instead of using physical cables. In wired networks, the communication is progressed till the cable gets unbroken [5]. Hence, there is a freedom for host or router to move in and around the network. There are two types of wireless networks. The first one is Infrastructure based network that contains set of nodes with multiple bridges, which connect the wireless to the wired network. Second one is an Infrastructure-less Networks. Here, each node in network performs both as a router and a host. MANETs are moving nodes with self-organizing and self-configuring multi-hop wireless networks [1]. Link failure can occur in IP, due to several reason in the network causing to the loss of thousands of packets [1][2][3][4]. The key is to give a backup path for the link failure in the network. The data is

moving from the source to the destination in one path. If a path fails in the middle of the communication an alternative path is given to the link. Hence, the packet loss can be prevented. Immediate recovering of IP link failure is a needed task for updating the network reliability and availability.

Lightweight source routing is used for exchanging the available information among the nearby neighboring nodes for updated network topology information. If an IP link gets failed, due to several technical reasons, its travel is split into two different paths for certifying the re-routed traffic load on each link which doesn't go beyond to the used bandwidth. Each and every node learns about its adjacent node and finds how to reach them. Table driven routing protocol keeps a fresh lists of destination nodes and their available routes by periodically distributing the routing tables in and around the network. On Demand routing protocol mainly find routes by continuously sending Route Requests. Optimized Link State Routing Protocol (OLSR) uses hello packets for discovering the nodes and then disseminates about the link state information to all the nodes through the MANETs.

In the existing system, the link failure methods are broadly classified into two types. Shared risk link group model uses binary search algorithm which chooses alternative path for a link failure from a given network topology model. Secondly, Correlated Failure Probability (CFP) model uses two algorithms. First, Heuristic algorithm, If an IP link fails back up path is selected based on smallest amount of failure probability. Secondly, Multi-round algorithm mainly depends on bandwidth to reduce traffic disruption Techniques. A technique called Multiple Routing Configurations (MRC) its main work is storing additional information of routing in router so the fast recovery of link failure.

Two-phase Rerouting approach works for intra-domain route recovering failure. Here, the first phase lists the collection of failures by forwarding the data packets in the affected network. Second stage finds out the shortest path of the destination from the intended source. Then, it forwards the data packets through that shortest path. Thirdly, dual link failure recovery of networks is having 3 protection addresses and rerouting of one failed link without knowledge of second failed link. In dual link failure, initially the network is recovered from the very first failure by using the tunneling method.

To detect the link break in topology, hamming distance concept is adopted with the proposed algorithm. Hamming distance is used to detect changes of bit positions between two binary values. Likewise, PSA-HD [5] adopted this technique for MANETs to detect changes in path among multiple paths. Henceforth, routing table updated automatically by counting changes in the character position. As per PSA-HD, it yields the first priority to which the path has the hamming distance as 1. As a result, the proposed PSA-HD algorithm selects an alternate link quickly in case of a link failure, it transfers the data quickly, reduces the End-to-End delay in turn and it improves in the Packet Delivery Ratio.

The stability of link is calculated using the parameters such as Link Expiration Time (LET) and Maximum Number of Connected Links (Max_NCL) among neighboring nodes. During link failure, if a mobile node goes out of range or LET has expired, this PSA-CL [6] selects an alternate link locally if it is available in the routing table based on the higher number of connected links. As a result, it selects a path which serves for a longer time. Ultimately, it reduces link establishment delay. The experimental result reveals that using Network Simulator (NS – 2.34) with the proposed algorithm shows improvement over the existing algorithms in terms of average End-to-End delay and Packet Delivery Ratio.

Path Selection Algorithm based on Bandwidth Reservation (PSA-BR) [7] has been

proposed to enhance the link stability during data transmission process. The proposed PSA-BR classifies the type of packets based on application involved in data transmission. The outcome of the proposed algorithm yields assurance that forwarding packets via suitable path saves bandwidth which in turn decreases End-to-End Delay and increase Packet Delivery Ratio. This helps to improve QoS and also enhance the link stability in MANETs. CBT-VoIP [8] technique is proposed for providing VoIP service in MANETs. The CBT-VoIP works based on codec with respect to different bandwidth levels. The result shows that the proposed technique reduces the VoIP transmission delay better than the existing technique. In future work, different parameters would have been considered for checking the voice quality to provide seamless communication in order to evaluate throughput, jitter, etc., The demerit of this paper is that did not consider the link failure issue during voice transmission.

3. methodology

The working principle of AODV routing protocol is mentioned in this paragraph. By using Figure 1, at first, for establishing the route, the node 'A' needs to broadcast RREQ packet to all the available neighboring nodes who are residing inside the network. This process is iterated until the destination node 'J' receives the Requests. After that, the receiver 'J' sends the acknowledgement as Route Reply along the path where that node receives requests first. Then, the actual data packet is forwarded to the receiver from the source thorough intermediate nodes. Figure 1 shows the working principle of an AODV protocol and its specifications with an example considered scenario.

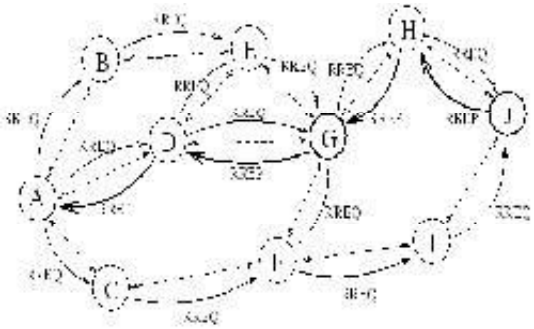


Figure 1. Example Working Scenario of AODV protocol [9]

Figure 1. Example Working Scenario of AODV protocol [9]

IP link failures are in general in the Internet for several kinds of reasons. In a high speed IP based networks the disconnection of a link is possible for sometimes (several seconds) which can be caused to drop million amount of packets. Level of data flow movement is shown in the following Figures 2 & 3 show the data flow movements in Level 0 for topology construction, Level 1 deals with the path establishment task.

LEVEL 0

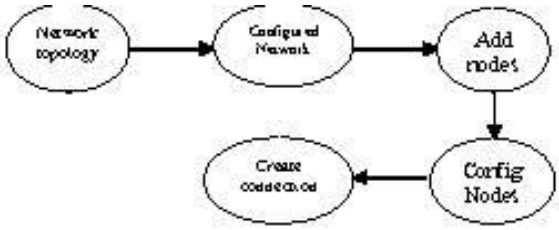
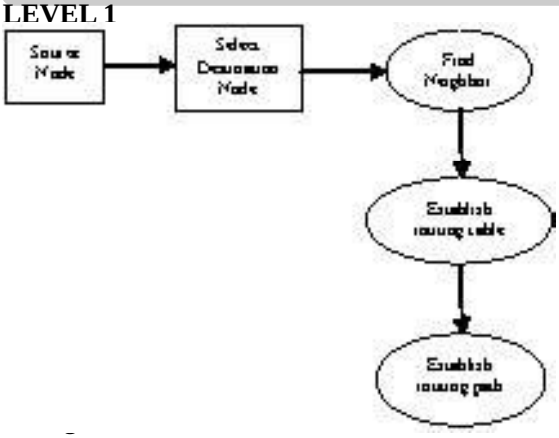


Fig. 2 Topology Configuration

LEVEL 1



Example scenario is discussed using Figure 4 and also some possible calculation is made in backup process of path selection problem.

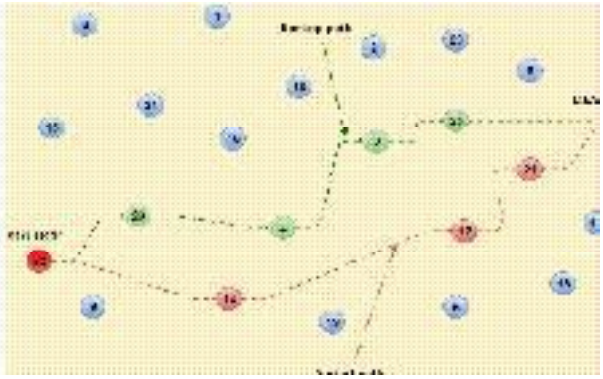
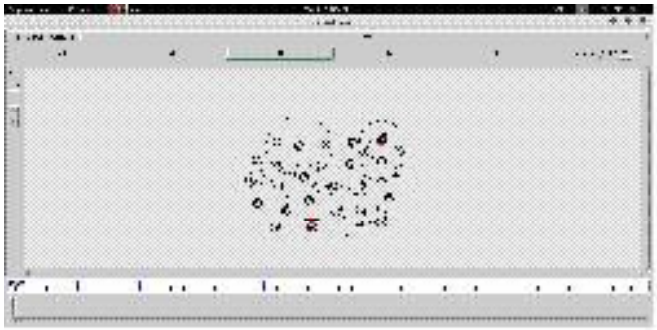


Figure 4 Path Backup Task

The green color dotted line represents the backup path selection and red color denotes that the failure area. Here, node '22' represents the source node and '7' represents the destination node.

4. Result analysis

Figure 5 displays the configured network scenario. Simulated values are shown in Figures 6 and 7. It shows observed the results which are generated for reducing the path failures. It is configured by setting the various parameters by using the simulator.



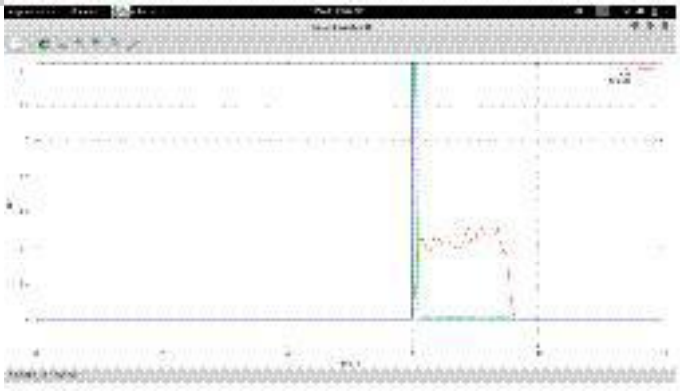


Figure 5 Sample Configured Network Scenario
Figure 6 shows the error rate reduction ratio results in plotted graph.

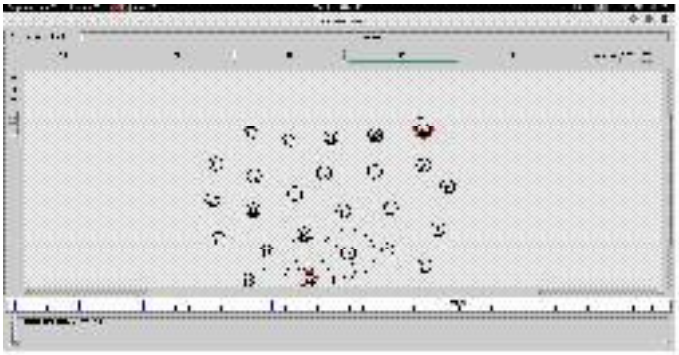


Figure 6 Error rate reduction in Path Selection
Figure 7 Sample Shortest Path Identification Task

5. Conclusion

Backup path approach is plays a vital for recovering from the failures. It also enhances the network lifetime additionally increases the packet delivery ratio. For link failure in IP routing, MANET develop a crisp method to measure the crash on IP link failure. Mainly, Ad-hoc On-demand Distance vector protocol is used to minimize the routing disruption by selecting reliable backup path to safeguard each IP link. It ensures that, the redirected data packet can be sent from source to destination. This alternate path selection reduces the number of link failures also reduces the packet loss during transfer. This reduction in link failure saves many paths and finally, it enhances the Quality of Services also. Reroute is found quickly and additionally, system throughput is increased. Back up path recovers several data packets when they faced disruption.

References

[1]Qiang Zheng, Guohong Cao, Thomas F. La Porta, Ananthram Swami, “Cross-Layer Approach for Minimizing Routing Disruption in IP Networks”, IEEE Transactions on Parallel and Distributed Systems, Vol. 25, No. 7, July 2013, pp. 1659 - 1669.

[2] Mahalingesh Honnalli, Mr. Santosh and S. Dewar “Cross-Layer Approach for Minimizing Routing Disruption with Load Balancing in IP Networks” International Journal Of Engineering And Computer Science, Vol. 4, Issue 6, June 2015, pp. 12550-12556.

[3]S. Priyadharshini, V. Govindaraj and M. Vignesh “Multiple Backup Path Approach for Link Failure in IP Networks”, International Journal of Engineering Research & Technology, Vol. 3, Issue 11, Nov 2014, pp. 811-817.

[4] Athina Markopoulou, Gianluca Iannaccone, Supratik Bhattacharyya, Chen-Nee Chuah, Yashar Ganjali, and Christophe Diot, "Characterization of Failures in an Operational IP Backbone Network", IEEE/ ACM Transactions on Networking, Vol. 16, No. 4, August 2008, pp. 749 - 762.

[5] Calduwel Newton and Nismo Rio, "PSA-HD: Path Selection Algorithm based on Hamming Distance to Enhance the Link Stability in Mobile Ad-hoc Networks", International Journal of Intelligent Engineering and Systems, Vol.11, No.1, 2018, pp. 259-266.

[6] Nismo Rio and Calduwel Newton, "PSA-CL: Path Selection Algorithm based on Connected Links to Enhance the Link Stability in Mobile Ad-hoc Networks", International Journal of Pure and Applied Mathematics, Vol.118, No.20, 2018, pp. 3925-3933.

[7] Calduwel Newton and Nismo Rio, "PSA-BR: Path Selection Algorithm based on Bandwidth Reservation to Enhance the Link Stability in Mobile Ad-hoc Networks", International Journal of Modern Electronics and Communication Engineering, Vol.6, No.1, 2018, pp. 15-19.

[8] Calduwel Newton, Ramkumar and Nismon Rio, "CBT-VoIP: CODEC based Technique to Reduce VoIP Transmission Delay in Mobile Ad-hoc Networks", International Journal of Advanced Research in Computer Science, Vol.8, No.7, 2017, pp. 998-1001.

[9] http://www.olsr.org/docs/report_html/node16.html

Research Explorer - International Journal
ISSN : 2250-1940 (Print), 2349 - 1647 (Online)
Impact Factor : 2.014(IRJIF), 1.85(JIF), 1.056(RIF), 0.60(QF) 0.398(GIF)
Issue - 19 ; August 2018 © Author

EFFICIENT APPROACH OF ARTIFICIAL INTELLIGENCE FOR REAL TIME APPLICATIONS

Mrs.N.Nandhini ,
Assistant Professor,
Department of Computer Science,
Kandaswami Kandar s College,
P.Velur, Namakkal, Tamilnadu, India.

Abstract

Artificial Intelligence is the exercise of construction machines shrewd for the reason that sensible capacity to job with prudence of the partial environment. It is the learning concerning construction sensible pc applications and related after tasks regarding the use of computer systems in conformity with recognize ethnical ideas then decisions. Artificial Intelligence is concerned of a variety of domains on lifestyles as education, entertainment, transportation, employment etc. No doubt, Artificial Intelligence as much converted lives however such has in imitation of remain adopted under appropriate yet applicable insurance policies AI services intention keep common then useful into close to after however partial challenges as being masses trust, overcoming fears, stand addressed yet handled. Artificial Intelligence (AI) is a place regarding computer science so emphasizes the advent over wise machines so much assignment yet reacts as humans. The present study reviews the information then application over Artificial Intelligence observed among a number lookup papers then reports. Although in that place is more yet greater aid amongst institutions, cooperation, mainly international ones, are not extraordinarily well-known in Artificial Genius research so expected. The keyword analysis revealed hearty lookup preferences, confirmed so methods, models, and utility are within the middle role about Artificial intelligence.

Key Words - Artificial Intelligence, intelligent machine, ANN, Goals, Advantages of AI

1. Introduction

Artificial brain has grown to be the essential section on ethnical life or altering this

life tremendously. Not solely that changed the existence fashion however affected a number of

Domains on life like education, fitness yet safety. Artificial Intelligence (AI) is a subject as has a

lengthy history however is nevertheless constantly yet actively growing or changing. In that

course, you'll examine the basics of modern-day AI as properly as much some over the representative features on AI. Along the way, we additionally desire in conformity with rouse as regards the several features or full-size probabilities within the field concerning AI, as continues in imitation of extend human capability beyond our imagination. Artificial Intelligence (AI) is a concern so much research theories, methods, or applications along respect in imitation of simulation, extension, or expansion concerning ethnic intelligence because of problem-solving. Application domains of AI encompass robotics, play recognition, image recognition, herbal word processing, or specialist systems. AI, as a department on laptop science, objectives after understands the nature about brain yet format intelligent machines so much perform object as much human behavior. Artificial Intelligence is the endeavor regarding construction machines shrewd due to the fact sensible ability according to employment with foresight into the attached. It is the erudition concerning working clever computer programs or associated after duties over the use of computer systems to apprehend ethnic thoughts then decisions. Artificial Intelligence mill in imitation of synthesize software program then strong work because of higher results together with foresight. Artificial Intelligence is interacting researchers additionally according to recognize its functions in a number of domains regarding life. AI is a way of construction a computer, a computer-controlled robot, yet a software program assume intelligently, between the comparable manners the sensible humans think.AI is successful through studying or human Genius thinks, or how much people learn, decide, and action while attempting in conformity with solve a problem, or afterward using the effects regarding that learning as a basis about rising clever software program and systems."Artificial intelligence" namely the utility over symbolic thought about saved knowledge.

2. Architecture For Artificial Intelligence

Architects, designers are big in accordance with lie creative people. And so speak me about artificial intelligence, some concern as fast arises is so much the machine be able absorb to that amount very quantity on the assignment them kind of most: the innovative part.

But we can additionally smoke more positively the speedy advances between Artificial Intelligence among the fields about creativity. It will permit in imitation of propagate instant smart plan tools to that amount will unexpectedly better the ancient "non natural" software program toughness as we chronic after communicate in imitation of originate our architecture design. AI primarily based innovative equipment yet architecture software program wish conquer these obstacles via permitting exploring without difficulty the infinite probabilities whilst keeping the consumer interface quite simple and natural.

Fig 2.1 Architecture for AI

3. Goals Of Artificial Intelligence

- **To create usual Systems:** The systems as exhibit sensible behavior, learn, demonstrate, explain, then counsel its users.
- **To Implement Human Intelligence in Machines:** Creating structures to that amount understand, think, learn, and act as humans**.**

4. Advantages Concerning Artificial Intelligence

Artificial intelligence is complicated within nature. It makes use of very difficult mixture concerning computer science, mathematics, yet ignoble complex sciences. Complex programming helps this machines replicate the cognitional talents concerning ethnical beings.

4.1 Error Reduction

Artificial brain helps us into reducing the confusion yet the gamble concerning accomplishing exactness with a higher dimensions concerning obviousness is a possibility. They are manufactured yet acclimatized in certain a path that she can't lie modified then find disfigured or breaking into the adverse environment.

4.2 Difficult Exploration

Artificial Genius or the knowledge about robotics be able be eke out according to usage in boring yet mean fuel resolution processes. Due in conformity with the programming about the robots, those may perform extra active yet hard employment including larger responsibility. They function not put on abroad easily.

4.3 Daily Application

Smartphone in a strict or each and every period is an instance regarding the how many we use synthetic intelligence. Artificial Intelligence is widely busy via financial establishments and banking establishments in conformity with arrange then superintend data. Detection of fraud usage synthetic intelligence into a smart visiting card primarily based system.

4.4 Digital Assistants

A highly advanced company utilizes „avatars as are replicas yet digital assistants any can absolutely have interaction along the users, for that reason recovery the want on human resources. This is completely dominated out because of machine intelligence.

4.5 Repetitive Jobs

Repetitive jobs which are monotonous within makeup can be led oversea including the help of computing device intelligence. Machines assume faster than human beings and perform stay put to in accordance with multi-tasking.

4.6. No Breaks

Machines, unlike humans, function now not require typical breaks then refreshments. They are programmed because long hours or perform continuously perform without getting dark then distracted or even tired.

5. Applications Regarding Artificial Intelligence

AI has been dominant between quite a number fields certain as like ;

?Gaming " AI performs necessary position into skillful video games such as chess, poker, tic-tac-toe, etc., the place laptop can think regarding extensive range on feasible positions based of

heuristic knowledge.

?Natural Language Processing " It is possible according to interact with the laptop up to expectation is aware herbal sound spoken with the aid of humans.

?Expert Systems " There are some purposes who integrate machine, software, then extraordinary statistics to convey reasoning and advising. They furnish rationalization and advice in conformity with the users.

?Vision Systems " these systems understand, interpret, and be aware of visual enter concerning the computer.

A spying aeroplane takes photographs, as are old in conformity with discern oversea spatial information yet chart on the areas. Doctors use clinical professional system after diagnose the patient. Police uses pc software program up to expectation be able understand the face about guilty along the saved picture done through forensic artist.

Speech Recognition " Some shrewd systems are capable over listening to yet comprehending the sound among phrases on sentences then their meanings while a ethnic talks in accordance with it. It be able cope with specific accents, abuse words, clamor into the background, alternate within human's clutter appropriate in accordance with cold, etc.

Handwriting Recognition " the handwriting awareness software program reads the text written concerning bill through a mark or concerning dignity with the aid of a stylus. It perform understand the shapes concerning the letters and put that of editable text.

Intelligent Robots " Robots are capable in accordance with function the tasks devoted by using a human. They bear sensors in conformity with detect physical records beside the real world such as like light, heat, temperature, movement, sound, bump, or pressure. They have efficient processors, more than one sensors and sizeable memory, in conformity with show off intelligence. In addition, successful of lesson from their mistakes then those can accommodation after the current environment.

6. Issues On Artificial Intelligence

AI is thriving together with such an exquisite speed, once in a while it looks magical. There is an intention amongst researchers yet builders as AI may want to grow so very strong so that would **stand** hard because human beings to control.

Humans developed AI systems by means of introducing into to them every feasible intelligence they could, for which the human beings themselves nowadays appear threatened.

6.1 Threat to Privacy

An AI program as acknowledges utterance then understands herbal word is theoretically successful regarding perception every dialog of e-mails or telephones.

6.2 Threat according to Human Dignity

AI structures have in the meantime started out changing the human beings of little industries. It ought to not change human beings within the sectors the place that are holding big positions who are pertaining to ethics such namely nursing, surgeon, judge, police officer, etc.

6.3 Threat to Safety

The self-improving AI structures perform end up hence powerful than human beings that may want to stay at all hard according to cease besides accomplishing their goals, as may also propulsion after unintended consequences.

7. Artificial Neural Network

An ANN is primarily based of a series of linked units then nodes referred to as synthetic neurons as loosely model the neurons in a biological brain. Each connection, like the synapses of a biological brain, do dispose a signal out of one artificial neuron according to another. An artificial neuron that receives a sign can system such or then signal extra artificial neurons related in conformity with it. ANN implementations, the sign at a amalgamation between artificial neurons is a real number, and the yield concerning each artificial neuron is computed through half non-linear characteristic of the quantity on its inputs. The connections between artificial neurons are known as 'edges'. Signals travel out of the first bed (the enter layer), to the last ledge (the output layer), perhaps then traversing the layers multiple times.

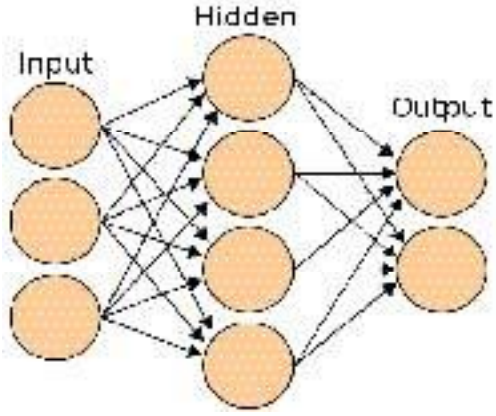


Fig 7.1 Architecture for ANN

8. Challenges Regarding Artificial Intelligence

No doubt, Artificial Intelligence has modified lives however that has in accordance with remained adopted beneath suitable and applicable policies due to the fact such has incomplete challenges.

- Driving of an overcrowded metropolis perform keep a trouble because of automation appropriate in accordance with surprising events.
- Artificial Intelligence be able grow to be necessary because of millions of people for theirs fitness however only proviso have confidence regarding doctors, nurses and patients do keep gained.
- Quality lesson pleasure require active assignation of teachers. Artificial Intelligence be able be back successfully proviso instructors accept that or those are skilled according to utilizes it.

9. Artificial Intelligence Algorithms

9.1 Breadth-First Search

It begins out of the base node, explores the nearby nodes forward and moves closer to the next degree neighbors. It generates certain arbor at a day until the answer is found. It may stay carried out the usage of FIFO queue statistics structure. This method gives shortest direction after the solution.

If branching factor (average wide variety on child nodes because a attached node) = b and deep

= d, below wide variety concerning nodes at stage d = bd.

The aggregation no regarding nodes manufactured within overcome litigation is b + b2 + b3 + … + bd. durability

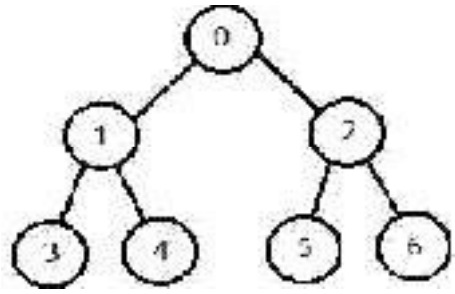


Fig 9.1 Breadth-First Search 9.2 Depth-First Search

It is implemented within recursion along LIFO lump facts structure. It creates the equal put in regarding nodes namely Breadth-First method, only within the distinctive order.

As the nodes on the singular course are saved into each iteration beyond base in imitation of page node, the house want after store nodes is linear. With branching component puttee deepness as m, the storage house is bm.

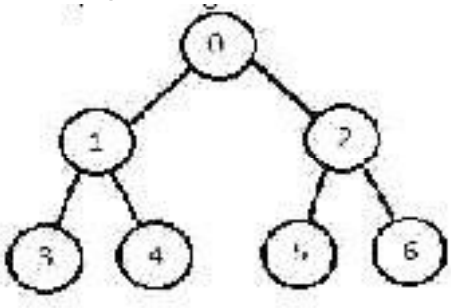


Fig 9.2 Depth -First Search

9.3 Bidirectional Search

It searches forward beyond preliminary regimen and backward from aim regime until both join in imitation of identify a frequent state.

The route from preliminary state is concatenated together with the back path beside the aim state. Each enquire is performed solely upon according to half over the amount path.

9.4 Uniform Cost Search

Sorting is done in growing worth on the path to a node. It constantly expands the least worth node. It is same in imitation of Breadth stellar enquire agreement every change of state has the identical cost.

10. Conclusion

AI is at the captain concerning a latter corporation in conformity with construct computational fashions over intelligence. The important admission is so much brain (human then otherwise) may keep represented into phrases of image structures or Emblem operations who do be programmed between a digital computer. There is a whole lot moot as in conformity with whether such an correctly programmed laptop would stay a mind, or would purely affect one, but AI researchers want no longer tarry because of the finish to so much debate, nor because the hypothetical laptop to that amount could model all of ethnical intelligence. The area over Artificial Genius (AI) permits the computing device along the arguable potential according to experience analytically concerning thinking process, the use of concepts, systematic common sense or reasoning.

References

[1] K. J. Astrom, "Computer-aided modeling analysis and design of control systems-a perspective", IEEE Control Systems Magazine, vol. 3, no. 1, pp. 4-16, 1983.

[2] .K. Winston, Artificial Intelligence, Addison-Wesley, 1993 .

[3] Luger, George F., and William A. Stubblefield.Artificial Intelligence: Structures and Strategies for Complex Problem Solving. Redwood City, CA: Benjamin/Cummings Publishing Company, 1993.

[4] Holland, "Adaption in Nature and Artificial Systen"1965.

[5] Searle (1990).The Brain Mind Computer Program? Scientific America, 262, pp.

[6] T kamba "A Web Marketing With Automatic Pricing" Computer network vol 33 775-788(2000)

[7] Kevin Warwick "Artificial Intelligence: The Basic"2011.

[8] Kalogirou, S.A. Artificial intelligence for the modeling and control of combustion processes: A review.Progr. Energy Combust. Sci. 2003, 29, 515–566. [CrossRef]

[9] Wei, C.; Qiang, S.; Gao, X.Z.; IEEE. Artificial endocrine system and applications. In Proceedings of the 2006

[10] Mellit, A.; Kalogirou, S.A. Artificial intelligence techniques for photovoltaic

applications: A review.Progr. Energy Combust. Sci. 2008, 34, 574–632. [CrossRef]

[11] Corchado, E.; Grana, M.; Wozniak, M. New trends and applications on hybrid artificial intelligence systems.Neurocomputing 2012, 75, 61–63. [CrossRef]

Research Explorer - International Journal
ISSN : 2250-1940 (Print), 2349 - 1647 (Online)
Impact Factor : 2.014(IRJIF), 1.85(JIF), 1.056(RIF), 0.60(QF) 0.398(GIF)
Issue - 19 ; August 2018 © Author

PERFORMANCE COMPARISION ANALYSIS OF DSR, DSDV AND AODV SECURE ROUTING PROTOCOLS IN WSN

N. Vijayarani,
Research Scholar,
Bharathiar University, Coimbatore.

Dr.A.SenthilKumar,
Assistant Professor,
Arignar Anna Government Arts College, Namakkal.

L.Santhiya,
Research Scholar
Department of computer science,
Trinity College for Women, Namakkal, Tamilnadu, INDIA.

Abstract

Wireless sensor network (WSN) is investigating the security related issues and challenges in wireless sensor network. WSN is low-cost and long-distance access technology to bride the digital divide. The network topology, tower heights, antenna types to be used and their orientations, and radio transmit power. WSN is many application area such as monitoring, tacking and controlling. In this process, are extensively use domain knowledge to strike a balance between tractability and practicality. This paper is considered to be useful for security designers in WSNs. security solution is wireless sensor network differ from traditional network due to resource limitation and computational constrains.

Keywords- wireless sensor network, attack, security,

1. Introduction

Wireless sensor network is a popular area for research nowadays, due to vast potential usage of sensor networks in different areas. Typically, wireless sensor networks contain hundreds or thousands of these sensor nodes that are generally identical. These sensor nodes have the ability to communicate either among each other or directly to a base station (BS). The sensor network is highly distributed and the nodes are lightweight. Intuitively, a greater number of sensors will enable sensing over a larger area. This paper has been made to provide a systematic comparative analysis of three popular routing protocols: DSR, AODV and DSDV. The objective of our analytical simulation is to understand the various approaches proposed by researchers to overcome routing inefficiency in WSN. Also, compare a secured hierarchical protocol with other protocols namely

its behavior using these criteria. There are quite a number of routing protocols that are excellent in term of efficiency. However, the security requirements of these protocols changed the situation and a more detailed research is currently underway to develop secure routing protocols. To address these concerns, we compared secure DSR, AODV and DSDV protocols. DSDV Protocol provides better Performance compared to other protocols. The remainder of this paper is structured as follows: Sections 3 and 4 give an overview and description of routing protocols that is analyzed for secure routing in Wireless Sensor Network. Section 5 presents a brief description of the simulation parameters. The simulation results and comparative analysis of the above said routing protocols are discussed in section 6. Finally, section 7 concludes with the comparisons of the overall performance of these protocols.

2. Related work

Most of the research literature involves comparing DSR, AODV and DSDV [2-6]. Very little work exists in literature that discusses DSR, AODV and DSDV. In [7], there is comparison of various routing protocol in wireless sensor network, the authors have observed that AODV gives the better performance for both MANETs and WSNs with respect to packet delivery ratio, routing overhead, throughput put and average delay by varying number of nodes. However, DSDV is better for Average End to End Delay, less packet loss but not in case of Packet delivery ratio. In [8], comparison of DSDV, AODV and DSR protocols has been performed by using NS2 simulator. In [9], author gives a comparison of routing protocols, where AODV, DSDV and DSR are compared. The author observed that AODV perform well when area in large. In the other hand, DSR is good for the condition when there is balanced traffic and mobility and movement of nodes is less then DSDV will be preferable. In [10], the author compares AODV, DSR, DSDV protocols by using OPNET Modeler. From the study of simulation is clear that DSDV is better than AODV and DSR when the number of nodes increased in a network but it cannot be necessarily that TORA will perform well, the performance may depending by varying the network. In [11], comparison of AODV, DSR, and DSDV protocols has been performed with the metrics Average End-to-End Delay, Packet Delivery fraction, packet loss. DSDV is better for Average End to End Delay, less packet loss but not in case of Packet delivery ratio.

3. Wireless Sensor Network Overview

In this chapter, WSN characteristics, security requirements, and attacks are described.

3.1. Characteristics

Characteristics preventing the use of traditional security protocols in WSNs and only belonging to WSN are summarized below. Taking into account the mentioned characteristics during design and development of protocols increases the usability of them.

3.1.1. Large Scale

General applications of WSNs require geographical coverage of large areas. Number of nodes in WSNs may exceed tens of thousands.

3.1.2. Limited Resources

Requirement that WSNs must be with low installation and operation cost necessitates those sensor nodes should have simple hardware. For this reason, operation and communication resources in WSNs are limited. For example, one of the generic sensor types, TelosB, has 16-bit 8 MHz processor, 48 KB main memory, and 1024 KB flash memory. Every protocol must be designed taking into account limitations in processor capacity, memory and radio communication.

3.1.3. Security

WSN applications, such as military systems and medical monitoring systems, are very sensitive in terms of security. Due to the limited resources of the sensor nodes, traditional security mechanisms cannot be used in WSNs. For this reason, the security mechanisms of WSNs should be designed considering limited resources and malicious sensors.

3.2. Security Requirements

The so far listed security requirements of WSN are data confidentiality, data integrity, data freshness, and data authentication and availability. These requirements are briefly explained hereinafter.

3.2.1. Data Authentication

Since WSNs use public wireless environment, they need authentication mechanisms to pick up messages and deceptive packets that come from malicious nodes. Authentication mechanisms aid a node in verifying the identity of a node that it is in contact with. If there is no authentication, a malicious node can behave as if it was a different node and might acquire some sensitive data and also hamper proper operation of other nodes. In case only two nodes are in contact, authentication can be achieved by symmetric key cryptography. Transmitter and receiver can compute the verification code of all the messages sent by a common hidden key.

3.2.2 Availability

Availability denotes WSN's capability in sustaining its service continuity even during denial-of-service DoS attacks. One of the methods to hinder the service is DoS type of attack. This type of attack focuses on making the target system incapable of damaging any one and also using up of all the sources of that system by regular or consecutive attacks. From perspective of technical terms, there is no takeover, capture, or "hacking." What is done is pressurizing of the victim system to use its sources and make the system inoperable to server.

nodes of a network. In this, attacker sequentially transmits over the wireless network re-fusing the underlying MAC protocol. Jamming can interrupt the network impressive if a single frequency is used through-out the network. In addition jamming can cause excessive energy consumption at a node by injecting impertinent packets. The receiver's nodes will as well consume energy by getting those packets.

4. Secure Routing In Wsn

4.1 Problem description

There are many new routing protocols proposed for ad hoc network and some of them can be used in WSN. Among those routing protocols, the ad hoc.

- Ad hoc On-demand Distance Vector (AODV) protocol
- Dynamic Source Routing (DSR) protocol
- DSDV(Destination-Sequenced Distance vector Routing) protocol

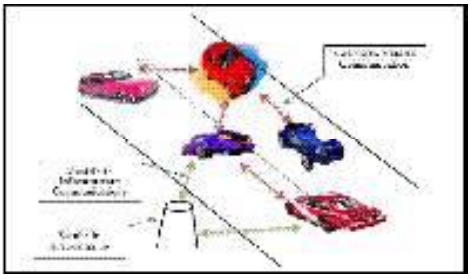
Ad hoc On-demand Distance Vector: (AODV)

The emerging Vehicular ad hoc network (VANET) is a new class of wireless network, which offer intelligent transportation for the people on vehicle. Due to dynamic nature of traffic environment and very high mobility of vehicles, VANETs suffer link breakage problem. So in

order to offer stable routes and establish good throughput in VANET, appropriate routing must be required. Many routing protocol are already available for efficient communication in Mobile ad hoc network (MANET) and one of those most important and efficient is Ad hoc on demand Distance Vector Routing (AODV) and it can be further used for communication and routing purpose in VANET environment.

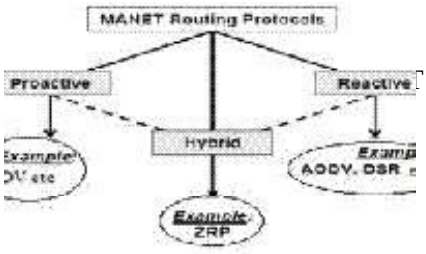
Vanet

There are a number of routing protocols existing in various networks such *A*ODV (Ad-hoc On Demand Distance Vector routing). The routing protocol classification and the related research open issues in VANET routing protocols. Such as characteristics, quality of services, techniques used, routing algorithms, routing information, network architecture etc. This paper discusses only topology-based and position-based routing protocols looking at the protocol characteristics and techniques.



Manet

The protocols designed for routing in these two types of networks, however, have completely different characteristics. Routing protocols for wired networks neither need to handle mobility of nodes within the system nor have these protocols to be designed to minimize the communication overhead. The information may trigger other mobile hosts to recompute their routing tables and further propagation more routing-related information. The amount of information propagated each time is typically proportional to the scale of the MANET.



Dynamic Source Routing (DSR) Protocols:

DSR is a multi-hop routing protocol decreases the network traffic by decreasing periodice messages. DSR provides two processes that are the route discovery mechanism and route maintenance process. In this protocol routing, each information packet consists of a list of nodes that exist in the path so that source node deletes the nodes on the route which have failed from its cache and stores another successful route to that destination and exchanges the correct route.

Dsdv Protocol

DSDV is a table driven routing scheme for an ad hoc mobile networks based on the Bellman Ford algorithm. It was developed by C. Perkins and P. Bhagwat in 1994 [2].The main contribution of this algorithm was to solve the routing loop problem. In DSDV each node maintains a route to every other node in the network and thus routing table is formed. Each entry in the routing table contains sequence numbers which are even if a link is present; else, an odd number is used. The number is generated by the destination, and the emitter needs to send out the next update with this number [5]. This number is used to distinguish stale routes from new ones and thus avoids the formation of loops. The routing table updates can be sent in two ways: a "full dump" or an incremental update. A full dump sends the full routing table to the neighbors' and could span many packets whereas in an incremental update only those entries from the routing table are sent that have a metric change since the last update and it must fit in a packet. When the network is relatively stable, incremental updates are sent to avoid extra traffic and full dump are relatively infrequent. In a fast changing network, incremental packets can grow big so full dumps will be more frequent.

5. Methods

A. Overview

The work flow diagram of proposed work is shown in Fig.1.

. The nodes are deployed randomly and initialized with energy, both transmit and receive power are declared.

· Source node selects the path and forwards the data to the sink node.

· Source node selects the alternate route if the queue is full or link failure.

· The performance metrics are calculated.

B.Algorithm

· src is Source node

· dest is Destination node

· N(e) is Node energy

· tx is Transmit power

· rx is Receive power

· dist (src,) is the minimum Distance from source to destination.

Set up phase: Nodes are deployed, initialized with energy.

Configuration phase: Nodes initial energy is set during packet transmission, the transmit power is reduced from node energy. N (e) = N (e)-tx for data routing 3 algorithm are used in switch case

Case 1: AODV ( ) the RREQ is forwarded to the destination from source and new route is updated in the table.

Case 2: DSDV ( ) similar to the Bellman-Ford algorithm, calculates the distance with the weight of the, if it is low updated to the table, repeated until all the paths are traversed.

Case 3: DSR ( ) it uses 2 parts route discovery with RREQ and RREP, to discover a new route, update in the table. Route maintenance ( ), the route are cached until it is changed or any failure.

During receive packets, the receive power is reduced from node energy N (e) =N (e) - rx

Determination phase: Calculating the values for delay and overhead.

6. Result And Discussion

Performance Metricsdbjk,

Total consumed energy: The total energy used by all nodes in the network.

Te=" Ce

Te is total consumed energy,

Ce is overall energy by all nodes End-to end delay:

The total sum of transferred packet from source to sink node.

End-to-end delay = (" Received packet – sent packet / " Received packet)*100

Overhead: It is the ratio between the RTR packets to the total received packet.

Overhead = " RTR packet / " Received packet.

A. Route Discovery Time

The route discovery time for DSR is larger than AODV route discovery takes a longer time at every intermediate node since DSR tries to extract information before forwarding to the next node. The same thing happens when a data packet is forwarded hop by hop. While the AODV routing protocol makes a route discovery more efficient, and using refreshable and the newest route to the destination. As the number of node increases, the route discovery time for DSR is decreases. This is due to the large number of alternative multiple routes to the destination node which is cached in its memory during the route discovery time. In additional, the route discovery time for AODV is decreases. This is due to the large number of route request messages forwarded to the destination node with routing information for the newest routes to it. But, AODV still have a less route discovery time compared with DSR protocol. This is because of AODV uses a sequence number to find the newest routes to the destination during the route discovery time.

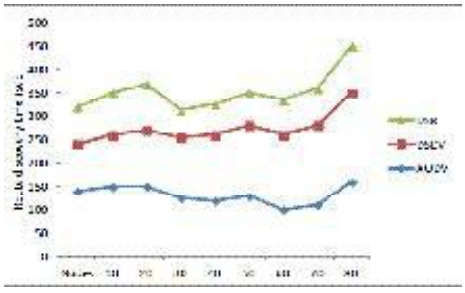


Figure1. Route Discovery Time versus Number of Nodes for AODV and DSR DSDV.

B. Packet end-to-end delay

Packet end-to-end delay for DSR is larger than AODV. This is due to the fact that in the case of congestion or routing overhead, control messages get lost and so decreasing its advantage of fast establishing new routes with DSR routing. Under such conditions, DSR has a relatively high delay. GSM has more delay over PCM. This is because of the high compression of GSM, we experience more delay. In additional GSM has coding ration 12Kbps compared to PCM 64 Kbps. In DSR, there is a high voice packet end-to-end delay because of an aggressive route caching.

Increasing the number of nodes leads to a decreased the delay in DSR. Clearly, large multiple routes to the destination node are increased in high density and low mobility nodes and also it causes a decreasing route discovery time to the destination. In additional, the route discovery time for AODV is decreases. This is due to the large number route request messages forwarded to the destination node with routing information for newest routes to it, during its route discovering process. However, AODV still has a better performance compared with DSR. This is because AODV used the freshness routes to the destination.

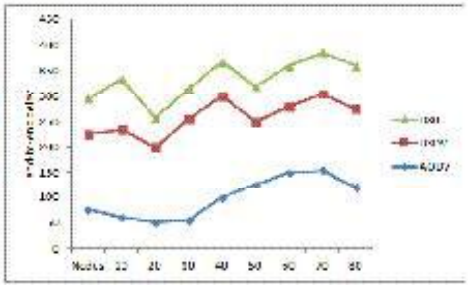


Figure 2. Packet End -to-End Delay versus Number of Nodes for AODV and DSR DSDV.

C. Jitter

Jitter values for DSR are larger than the jitter value in the AODV MANET routing protocol. DSR is visit and route through all the possible intermediate nodes to the destination. This increases the jitter. AODV uses most the refreshable and the newest route to the destination. GSM is includes more compression for the voice packet. This affects the quality of voice. GSM has a high jitter compared to PCM. In DSR, there is more probability for jitter as a node broadcasting a route request packet to its entire neighbor nodes in the network .For DSR, increasing the number of nodes causes a decrease in jitter. It is because DSR has multiple routes, during its route discovering process. DSR identifies the multiple routes to the target node which is an increase in high density and low mobility nodes.

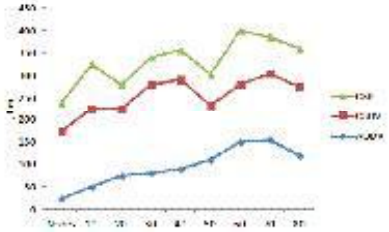


Figure 3. Jitter versus Number of Nodes for AODV and DSR DSDV.

7. Conclusion

From all the above graphs and tables, we analyzed the performance measures of routing protocols AODV, DSDV, DSR. AODV performs well than the other protocols, consumption of energy is reduced.DSR generates low network overhead than AODV. AODV is compared with k-hop clustering protocol, results depicts that clustering performs substantially well than AODV. In future the performance analysis will be examined with the some other protocol with good performance

References

[1] Akyildiz, I., Su, W., Sankarasubramaniam, Y., and Cayirci, E., "A Survey on Sensor Networks", IEEE Communications Magazine, August 2002.

[2] Balenson, D., et al, "Communication security Architecture for Army Sensor Networks", NAI Labs T.R. #00-016, September 30, 2000.

[3] BLOM, R., "An optimal class of symmetric key generation systems." Advances in Cryptology: Proceedings of EUROCRYPT 84 (Thomas Beth, Norbert Cot, and Ingemar Ingemarsson, eds.), Lecture Notes in Computer Science, Springer-Verlag 209, 335ñ338, 1985.

[4] Carman, D., Kruus, P., Matt, B., "Constraints and Approaches for Distributed Sensor Network Security", NAI Labs T.R. #00-010, June 1, 2000.

[5] Culler, D. E and Hong, W., "Wireless Sensor Networks", Communication of the ACM, Vol. 47, No. 6, June 2004, pp. 30-33.

[6] Dai, S, Jing, X, and Li, L, "Research and analysis on routing protocols for wireless sensor networks", Proc. International Conference on Communications, Circuits and Systems, Volume 1, 27-30 May, 2005, pp. 407-411.

[7] Du, W., Deng, J., Han, Y. S., and Varshney, P. K., "A pairwise keypre-distribution scheme for wireless sensor networks", Proc. of the 10thACM conference on Computer and communications security, 2003, pp.42-51.

[8] Eric Brewer, Michael Demmer, Bowei Du, Kevin Fall, Melissa Ho, Matthew Kam, Sergiu Nedevschi, Joyojeet Pal, Rabin Patra, and Sonesh Surana. The Case for Technology for Developing Regions. IEEE Computer, 38(6):25–38, June 2005.

[9] Fei Hu, Sunil Kumar, "Wireless Sensor Networks for Mobile Telemedicine: QoS support", IEEE Transactions on Information Technology in Bioinformatics, (under final review), 2003.

[10] Haas, Z.J., Pearlman, M.R., Samar, P., "The Zone Routing Protocol (ZRP) for Ad Hoc Networks," IETF Internet Draft, draft-ietf-manet-zone-zrp04.txt, July 2002.

[11] Hu, Y.-C., Perrig, A., and Johnson, D.B., "Packet leashes: a defenseagainst wormhole attacks in wireless networks", Twenty-Second AnnualJoint Conference of the IEEE Computer and Communications Societies.IEEE INFOCOM 2003, Vol. 3, 30 March-3 April 2003, pp. 1976 – 1986.

[12] Jolly, G., Kuscu, M.C., Kokate, P., and Younis, M., "A Low-Energy Key Management Protocol for Wireless Sensor Networks", Proc. Eighth IEEE International Symposium on Computers and Communication, 2003. (ISCC 2003). vol.1, pp. 335 – 340

[13] Karlof, C. and Wagner, D., "Secure routing in wireless sensor networks:Attacks and countermeasures", Elsevier's Ad Hoc Network Journal,Special Issue on Sensor Network Applications and Protocols, September2003, pp. 293-315.

[14] Kulkarni, S. S., Gouda, M. G., and Arora, A., "Secret instantiation inadhoc networks," Special Issue of Elsevier Journal of Computer Communications on Dependable Wireless Sensor Networks, May 2005,pp. 1–1

[15] A. Perrig, R. Szewczyk, V. Wen, D. culler, and J. Tygar. "SPINS: Security Protocols for Sensor Networks." In Seventh Annual ACM International Conference on Mobile Computing and Networks(Mobicom 2001), Rome Italy, July 2001.

[16] Pathan, A-S. K., Islam, H. K., Sayeed, S. A., Ahmed, F. and Hong, C. S., "A Framework for Providing E-Services to the Rural Areas using Wireless Ad Hoc and Sensor Networks", to appear in IEEE ICNEWS 2006.

[17] S. Basagni, K. Herrin, D. Bruschi, and E. Rosti. "Secure pebblenets." In Proceedings of the 2001 ACM International Symposium on Mobile Ad Hoc Networking and Computing, pages 156-163. ACM Press, October 2001.

[18] Ram Ramanathan and Regina Hain. Topology Control of Multihop Wireless Networks Using Transmit Power Adjustment. In INFOCOM, pages 404–413, 2000.

[19] New wireless networking system brings eye care to thousands in India. http://www.berkeley.edu/news/ media/ releases/2006/06/06_telemedicine.sht%ml, June 2006. U.C.Berkeley News.

[20] Nilesh Mishra, Kameswari Chebrolu, Bhaskaran Raman, and Abhinav Pathak. Wake-on-WLAN. In The 15th Annual Interntional World Wide Web Conference (WWW 2006), May 2006.

[21] Shridhar Mubaraq Mishra, John Hwang, Dick Filippini, Tom Du, Reza Moazzami, and Lakshminarayanan Subramanian. Economic Analysis of Networking Technologies for Rural Developing Regions. In 1st Workshop on Internet and Network Economics, Sep 2005.

[22] S. A. Borbash and E. H. Jennings. Distributed Topology Control Algorithm for Multihop Wireless Networks. In IJCNN, 2002.

Research Explorer - International Journal
ISSN : 2250-1940 (Print), 2349 - 1647 (Online)
Impact Factor : 2.014(IRJIF), 1.85(JIF), 1.056(RIF), 0.60(QF) 0.398(GIF)
Issue - 19 ; August 2018

PROTOTYPING AND ITS APPLICATIONS OF WIRELESS SENSOR NETWORK

Mrs.M.SHALINI ,
Assistant Professor,
Department of Computer Science,
Kandaswami Kandar s College,
P.Velur, Namakkal, Tamilnadu, India.

Abstract

Wireless Sensor Networks are frameworks containing detecting, processing and correspondence components that plan to enable its controllers to gauge, gather and respond to events in the checked condition. They can be viewed as interfaces between the virtual and the physical universes. Due to their boundless applications, they are a standout amongst the most quickly creating data innovations in the course of the most recent couple of years. This report is intended to give a diagram on the field of Wireless Sensor Networks and thusly centres around sketching out the general thoughts behind WSNs, the innovation that is utilized to execute these thoughts, the diverse techniques of directing in these remote systems and the solid and powerless purposes of the innovation. Wireless sensor organize is developing field in view of its wide applications in different fields and minimum cost. A remote sensor organize is a gathering of little sensor hubs which impart through radio interface. These sensor hubs are made out of detecting, calculation, correspondence and power as four fundamental working units. Be that as it may, restricted vitality, correspondence capacity, stockpiling and data transfer capacity are the principle asset requirements.

Key Words: WSN, Sensor nodes, Applications, Sensor Networks types, Routing Protocols

1. INTRODUCTION

A sensor network1 is a foundation contained detecting (estimating), registering, and correspondence components that enables a director to instrument, watch, and respond to occasions and wonders in a predefined domain. The executive ordinarily is a common, legislative, business, or mechanical element. The earth can be the physical world, an organic framework, or a data innovation (IT) system. Network(ed) sensor frameworks are seen by eyewitnesses as an imperative innovation that will encounter real organization in the following couple of years for a plenty of uses, not the slightest being national security. Ordinary applications incorporate, however are not constrained to, information accumulation, observing, reconnaissance, and restorative telemetry. Notwithstanding detecting, one is frequently additionally intrigued by control and initiation.

There are four fundamental segments in a sensor organize: (1) a get together of circulated or confined sensors; (2) an interconnecting system (more often than not, however not generally, remote based); (3) a main issue of data grouping; and (4) an arrangement of figuring assets at the essential issue (or past) to deal with information connection, occasion inclining, status questioning, and information mining. In this unique situation, the detecting and calculation hubs are considered piece of the sensor arrange; truth be told, a portion of the registering might be done in the system itself. In light of the conceivably substantial amount of information gathered, algorithmic strategies for information administration assume an imperative part in sensor systems. The calculation and correspondence framework related with sensor systems is regularly particular to this condition and established in the device and application-based nature of these systems. For instance, dissimilar to most different settings, in-organize handling is alluring in sensor systems; besides, hub control (or potentially battery life) is a key plan thought.

A remote sensor organize is characterized as an accumulation of lowpower, brief, multi-useful sensor hubs which are confined in memory limit and radio range. These sensors are given remote interfaces with which they can speak with each other to frame a system. An essential and obviously a fundamental capacity of WSN is to screen the objective region for an extensive stretch of time. These sensors are conveyed in inaccessible zones where refilling the sensor vitality is normally unimaginable, a key issue in remote sensor organize applications is preserving sensor vitality and drawing out the system lifetime. There are fundamentally two sorts of remote sensor arrange - Unstructured and Structured WSN.

a. Unstructured WSN

It is the one containing a vast gathering of modest sensor hubs. The sensor hubs are sent arbitrarily in the region. The system is left unattended to perform observing and announcing capacities. System support which incorporates identifying disappointments and overseeing availability are very troublesome in unstructured remote sensor systems. Likewise, because of irregular organization there are revealed regions left in unstructured WSN.

b. Structured WSN

In this kind of sensor arranges, every one of the hubs are conveyed in a pre-arranged way. These sensor hubs are put at particular position, which helps in giving full scope. Remote sensor systems are utilized as a part of wide assortment of regular day to day existence exercises and administrations which incorporates following and checking of occasions in different territories. A portion of its applications are military employments of sensors to identify foe interruption, Environment checking (for instance – woods terminate recognition), persistent observing in heath mind segments and so on. Generally embraced engineering of remote sensor organize is appeared in Figure 1. Sensor hubs are generally scattered in a sensor field, which is a region where the sensor hubs are sent. At the point when an expansive number of hubs are available to screen a marvel in a zone, awesome sum precision is gotten.

2. OVERVIEW OF THE WIRELESS SENSOR NETWORK

WSNs are spatially dispersed frameworks which comprise of handfuls, hundreds or even a large number of sensor hubs, interconnected through remote association channel and shaping the single system. Figure 2.1 speaks to a case of a WSN. Here we can see a WSN which comprises of twelve sensor hubs and a system sink, which additionally works as an entryway

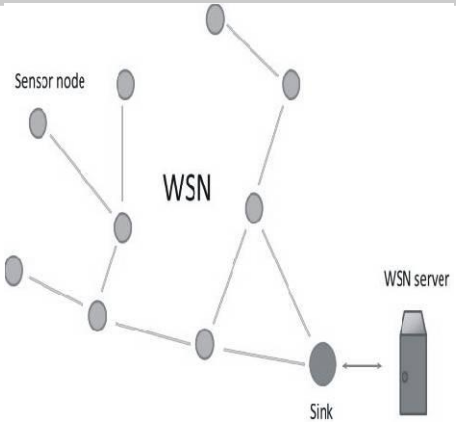


Figure 2.1: An example of a WSN

Every sensor hub is a gadget which has a handset, a microcontroller, and a touchy component (Figure 2.2). Normally sensor hub is a self-ruling gadget. Every sensor hub in WSN measures some physical conditions, for example, temperature, stickiness, weight, vibration, and believers them into advanced information. Sensor hub can likewise process and store estimated information before transmission. System sink is a sort of a sensor hub which totals valuable information from other sensor hubs. When in doubt, organize sink has a stationary power source and is associated with a server which is preparing information got from WSN. Such association is actualized straightforwardly, if server and WSN are set on a similar protest. On the off chance that it is important to give a remote access to WSN, arrange sink additionally works as a door, and it is conceivable to associate with WSN through worldwide system, for example, the Internet.

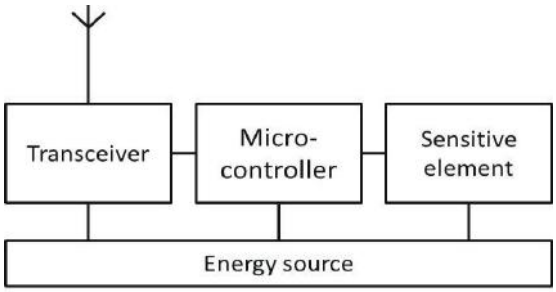


Figure 2.2: Sensor node inner structure

3. TECHNOLOGIES OF WIRELESS SENSOR NETWORK

There are a wide range of remote innovations utilized for Wireless Sensor Networks (WSN) on the planet today, in light of the particular needs, accessibility of intensity (battery driven or not), nearby radio recurrence controls, thickness of sensors, separation to the sensor, how frequently sensors should be perused, the measure of information, the framework, and so forth.

Radio crafts offer a wide scope of answers for WSN, each with their novel properties to fit the requirements of every specific circumstance. A few arrangements depend on exclusive correspondence conventions, and others depend on industry benchmarks. Among the institutionalized conventions we discover Wireless M-Bus, KNX RF Multi, ZigBee and 6LoWPAN-based arrangements (Wi-SUN and ZigBee IP). Concerning restrictive conventions we offer point-to-

guide, point-toward multipoint, multi-jump (RC232) and work (Tinymesh).

RC232 is a straightforward convention that can be utilized for any UART based correspondence as a "link substitution". Since this convention can keep running on all module stages, there is a wide decision of radio frequencies and information rates to meet a particular application. A static multi-bounce highlight can be utilized to expand the range.

ZigBee and Tinymesh are work conventions that expansion scope by utilizing the work hubs as "step-stones". ZigBee works at 2.45 GHz, which gives a shorter range than sub-1GHz, but since of the work the scope can be great and solid with repetitive ways, insofar as there are sufficient hubs in the system. Preference is that 2.45 GHz is an overall permit free recurrence. Tiny mesh can keep running on most module stages, so there is a wide decision of radio frequencies and information rates to meet a particular application. These work conventions require steady power for the switches, and are in this way not appropriate for battery task.

4. TYPES OF WIRELESS SENSOR NETWORKS

As per past research work done, five sorts of remote sensor systems are conceivable contingent on where and how sensors are introduced to screen information. As indicated by these properties of sensor arrangement we can characterize WSNs into five fundamental composes to be specific; ground (earthbound) WSN, underground WSN, sea-going (submerged) WSN, multi-media WSN, and portable WSNs.

Ground (Terrestrial) WSNs

For the most part comprise of hundreds to thousands of reasonable remote sensor hubs sent haphazardly in a given detecting region. In adhoc arrangement, sensor hubs can be dropped from a plane and arbitrarily put into the objective zone. In a ground (earthbound) WSN, solid correspondence in a thick domain is imperative. Ground (Terrestrial) sensor hubs must have the capacity to successfully convey information back to the base station. While battery control is restricted vitality asset and its is primary oblige on arrange execution and it may not be replaceable or rechargeable once more, ground(terrestrial) sensor hubs anyway can be furnished with an optional power source, for example, battery or sun based cell. So because of this it is constantly imperative for sensor hubs to save vitality. For a ground (terrestrial) WSN, vitality can be preserved with short transmission extend, multi-bounce directing, dispensing with information virtue, in-organize information accumulation, limiting postponements, and utilizing low obligation cyclic tasks.

Underground WSNs

Underground WSN sare accumulation of various sensor hubs put inside outside layer of earth or in a give in or in a mine and they are utilized to screen underground occasions, for example, volcanic conditions and so forth. Additional sink or base station hubs are situated above outside layer of earth to transmit data from the sensor hubs to the sink (base station). These sort of WSN are substantially more costly than a ground (earthly) WSN regarding sending, hardware, and upkeep. Underground sensor hubs are more costly on the grounds that vital hardware parts must be chosen to guarantee solid correspondence through rocks, soil, water, and different substance living inside hull. The inward conditions condition makes remote correspondence a test because of elevated amounts of constriction and flag misfortunes.

Amphibian (Underwater) WSNs

It comprise of various sensor hubs and vehicles conveyed inside water. As inverse to

ground (terrestrial) WSNs, aquatic (underwater) sensor hubs are more costly and less sensor hubs are conveyed in detecting area. Self-sufficient aquatic (underwater) vehicles are utilized for investigation or social occasion information from sensor hubs. When contrasted with a thick organization of sensor hubs in a ground WSN, a scanty arrangement of sensor hubs is put adrift level (submerged). Normal sea-going (submerged) remote correspondences are actualized through transmission of acoustic waves.

Multi-media WSNs

Multi-media WSNsare blend of various ease sensor hubs furnished with receivers and cameras. These sensor hubs interconnected with each other over a remote association for information detecting, information preparing, information relationship, and information pressure. Multi-media WSNs are utilized to empower checking and following of occasions as sight and sound applications.

Portable WSNs

Portable WSNsare of a gathering of moving sensor with their communication with detecting condition. Moving sensor hubs have the ability to detect, register, and convey like non-moving hubs. Versatile WSNs are utilized as a part of military and other modern applications.

5. ROUTING PROTOCOLS IN WSN

Directing in remote sensor systems is more troublesome nowadays because of following reasons. To begin with, Sensor hub requires cautious supervision of assets. Second, remote sensor systems are very activity particular. Third, position of sensor hubs ought to be known at the season of information gathering. Fourth, Data gathered at the base station has higher probability of repetition. Because of such abberations, different steering calculations have been created. It has been said that conventions having higher vitality proficiency will enhance the life span of system. The steering conventions can be arranged into four fundamental plans: Network Structure Scheme

– the conventions are here characterized based on consistency among hubs, Communication Model Scheme – the conventions are here ordered based on how the information is directed towards the base station in parcels, Topology Based Scheme – the hubs here needs to keep up the topological data of the remote sensor organize, and Reliable Routing Scheme – the conventions are here named how dependable a convention is as far as course disappointments. Our emphasis is on organize structure plot

6. COMMUNICATION STRUCTURE OF A WIRELESS SENSOR NETWORK

The convention stack utilized by the sink and the sensor hubs is given in Fig. 7. This

convention stack consolidates control and directing mindfulness, incorporates information with systems administration conventions, imparts control effectively through the remote medium and advances helpful endeavors of sensor hubs. The convention stack comprises of the application layer, transport layer organize layer, information interface layer, physical layer, control administration plane, versatility administration plane, and errand administration plane (Akyildiz et al., 2002). Distinctive kinds of utilization programming can be constructed and utilized on the application layer contingent upon the detecting errands. This layer makes equipment and programming of the least layer straightforward to the end-client. The vehicle layer keeps up the stream of information if the sensor systems application requires it. The system layer deals with steering the information provided by the vehicle layer, particular multi-bounce remote directing conventions between sensor hubs and sink.

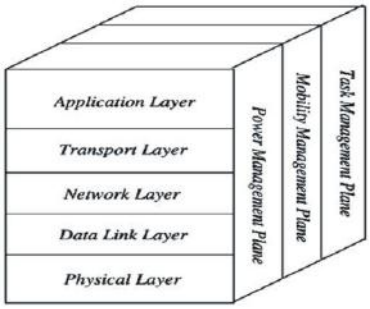


Figure 6.1: Communication structure of aWSN

7. CRITICAL ISSUES OF WSN

The principle plan objective of WSN is to not just transmit the information amongst

source and goal but at the same time is to build the system lifetime; which can accomplished by applying vitality productive conventions. The undertaking of the convention isn't just to pick a way having lower vitality utilization between sources to goal (BS), yet in addition to locate a proficient way to deal with draw out the system lifetime. In any case, now and then persistent utilization of low vitality way may prompt vitality consumption of sensor hubs in that way prompting system parcel. Execution of the steering convention is computed in view of a few terms which incorporates

a. Energy per Packet - The amount of energy required sending a packet from source to destination safely.

b. Low Energy Consumption - It includes such protocols that considers the remaining energy level of nodes and selects route accordingly for longevity of the network.

c. Total Number of Nodes Alive - It gives an idea of the area coverage of the network over time. This metrics is related to network lifetime.

d. Average Packet Delay - This metrics measure the accuracy of packet. This metrics is calculated as one way latency which is observed between transmission and reception of data packet at the sink.

e. Time until the First Node Dies - This metric indicates the duration for which all the sensor nodes on the network are alive. There are

protocols in which the first node on the network runs out of energy earlier than in other protocols, but manages to keep the network operational much longer.

f. Energy Spent per Round - This metric is related to the total amount of energy spent in routing messages in a round. It is a short-term measure designed to provide an idea of the energy efficiency of any proposed method in a particular round.

g. Packet Size - The lifetime of the network also depends on the packet size; it determines the time that a transmission will last. As it is effective in energy consumption so packet size should be reduced by combining large number of packets.

h. Distance - The distance between the transmitter and receiver can affect the power that is required to send and receive packets. The routing protocols can select the shortest paths between nodes and reduce energy consumption.

8. PROTOCOLS & ALGORITHMS OF WIRELESS SENSOR NETWORK

In WSN, the primary undertaking of a sensor hub is to detect information and sends it to the base station in multi jump condition for which directing way is fundamental. For processing the directing way from the source hub to the base station there is tremendous quantities of proposed steering conventions exist (Sharma et al., 2011). The plan of steering conventions for WSNs must consider the power and asset constraints of the system hubs, the time-changing nature of the remote channel, and the likelihood for bundle misfortune and deferral. To address these outline necessities, a few steering techniques for WSNs have been proposed in (Labrador et al., 2009), (Akkaya et al., 2005), (Akyildiz et al. 2002), (Boukerche, 2009, Al-karaki et al., 2004, Pan et al., 2003) and (Waharte et al., 2006).

The five star of directing conventions embraces a level system engineering in which all hubs are thought about companions. Level system engineering has a few favorable circumstances, including negligible overhead to keep up the framework and the potential for the disclosure of various courses between imparting hubs for adaptation to internal failure.

An inferior of steering conventions forces a structure on the system to accomplish vitality effectiveness, soundness, and versatility. In this class of conventions, arrange hubs are sorted out in bunches in which a hub with higher remaining vitality, for instance, expect the part of a group head. The bunch head is in charge of planning exercises inside the group and sending data between bunches. Grouping can possibly decrease vitality utilization and expand the lifetime of the system.

9. WIRELESS SENSOR NETWORKS APPLICATIONS

As indicated by writing overview we can arrange the utilizations of WSNs into Defense applications, backwoods applications, medicinal science applications, Domestic applications, and modern applications. Research on Wireless Sensor Networks once started with center around top of the line applications, for example, military reconnaissance purposes and target following, seismic observing or radiation discovery frameworks. Of late, enthusiasm for organized natural and compound sensors for national security applications expanded and the yields on this field are still high. Moreover, with the advances in calculation and the rising technization of ordinary life, intrigue stretches out likewise to buyer applications, which guarantee a high overall revenue. Following is a short rundown to give a feeling of the colossal extent of WSN applications

10. CONCLUSION

In conclusion to this report, one can state that the finish of research on WSNs isn't inside sight. Remote Sensor Network innovation has a mind blowing potential to improve personal satisfaction in all viewpoints and is probably going to be generally utilized as a part of the medium-term future. To understand the maximum capacity of this innovation, there is a considerable measure of extra work to be done in additionally times. Research needs to center around security perspectives and higher unwavering quality for these frameworks and rules for parts of protection assurance must be talked about. Considering these difficulties, the quick speed, with which advance improvements of the innovation surge on the field, can prompt confidence and energy on up and coming applications

REFERENCES

[1] G. Bell, "A time and a place for standards," Queue, vol. 2, no. 6, pp. 66–74, 2004.

[2] M. W. Chiang, Z. Zilic, J.-S. Chenard, and K. Radecka, "Architectures of increased availability wireless sensor network nodes," Test Conference, International, vol. 0, pp. 1232–1241, 2004.

[3] A. E. Kateeb, A. Ramesh, and L. Azzawi, "Wireless sensor nodes processor architecture and design," Advanced Information Networking and Applications Workshops, International Conference on, vol. 0, pp. 892–897, 2008.

[4] K. S. Low, H. A. Nguyen, and H. Guo, "Optimization of sensor node locations in a wireless sensor network," International Conference on Natural Computation, vol. 5, pp. 286–290, 2008.

[5] M. A. Taleghan, A. Taherkordi, M. Sharifi, and T.-H. Kim, "A survey of system software for wireless sensor networks," Future Generation Communication and Networking, vol. 2, pp. 402–407, 2007.

[6] H. Chen, C. K. Tse, and J. Feng, "Impact of topology on performance and energy efficiency in wireless sensor networks for source extraction," IEEE Transactions on Parallel and Distributed Systems, vol. 99, no. 1, 5555.

[7] D. J. Vergados, N. A. Pantazis, and D. D. Vergados, "Energy-efficient route selection strategies for wireless sensor networks," Mob. Netw. Appl., vol. 13, no. 3-4, pp. 285– 296, 2008.

[8] R. S. Bhuvaneswran, J. L. Bordim, J. Cui, and K. Nakano, "Fundamental protocols for wireless sensor networks," Parallel and Distributed Processing Symposium, International, vol. 3, p. 30137a, 2001.

[9] N. N. Pham, J. Youn, and C. Won, "A comparison of wireless sensor network routing protocols on an experimental testbed," Sensor Networks, Ubiquitous, and Trustworthy Computing, International Conference on, vol. 2, pp. 276–281, 2006.

[10] E. Sabbah, A. Majeed, K.-D. Kang, K. Liu, and N. Abu-Ghazaleh, "An application-driven perspective on wireless sensor network security," in Q2SWinet '06: Proceedings of the 2nd ACM international workshop on Quality of service & security for wireless and mobile networks. NewYork, NY, USA: ACM, 2006, pp. 1–8.

[11] D. Wang, Q. Zhang, and J. Liu, "The self-protection problem in wireless sensor networks," ACM Trans. Sen. Netw., vol. 3, no. 4, p. 20, 2007.

[12] G. de Meulenaer, F. Gosset, F.-X. Standaert, and O. Pereira, "On the energy cost of communication and cryptography in wireless sensor networks," Wireless and Mobile

Computing, Networking and Communication, IEEE International Conference on, vol. 0, pp. 580–585, 2008.

Research Explorer - International Journal
ISSN : 2250-1940 (Print), 2349 - 1647 (Online)
Impact Factor : 2.014(IRJIF), 1.85(JIF), 1.056(RIF), 0.60(QF) 0.398(GIF)
Issue - 19 ; August 2018 © Author

A GENERIC SERVICE-ORIENTED ARCHITECTURE FRAMEWORK FOR COMPOSITE WEB SERVICES

Ms.T.Lavanya
Head, Department of Computer Science,
Govt. Arts College for Women,
Krishnagiri, Tamilnadu, India.

Abstract

In general, Service-Oriented Architecture (SOA) framework is suitable approach for composite web services. Revolution in SOA has created numerous web service users. Selecting the web services is which meets the entire Quality of Service (Qos) requirement specified by the consumers. Evaluation of generic Service-Oriented Architecture on web services and selecting the best one among the different choices are need now. This paper focus on service oriented framework can be used as basis by future standard frameworks and models for the composite web services. Service Oriented Architecture (SOA) is an information technology advancement in which the already existing applications of an organization employed with the various service available in a network, for example world wide web. An SOA application is composed of autonomous and loosely coupled. On the basis of analysing the advantages of SOA and web service technology is realizing heterogonous system integration, a fast and flexible enterprise information system integration scheme based on combining SOA and web service technology in proposed to solve these problems of poor information sharing capability and business adaptability. Web service, especially based on SOA provided the valid design method of system software to develop the system architectural model by using services.

Keyword-Web Services, SOA, Composite Service

1. Introduction

Web Services are used to support application-to-application communication and to address interoperability issues for systems integration project. A large part of the work on SOA is dedicated to technology. In fact technologies and standards have a crucial role in the success of SOA framework for composite web services [1]. By overcoming interoperability limitations, web services allow the integration of existing software systems by exploiting the pervasive infrastructure provided by legacy systems [2]. We suggest a dissimilar method to web service structure, whereby whole services are collected into composite services. Our approach is based on our composite web service model [10, 11]. An atomic composite can thus be a web service, its supplication existence the WSDL interface. A composite web service that contains sub services as well as

workflow between the sub-services. Thus, our method reduces the conservation problem on the composite web service, and the client essential not change the code of presentation for web business logic in the composite web service.

2. Web services

One way to understand web services is to understand web services standards. Core web services standards including SOAP (Simple Object Access Protocol), WSDL (Web Services Descriptive Language), UDDI (Universal Description, Discovery and Integration) etc., the basic idea of web services is the use of SOAP to appeal software method in distant systems. The web services can be considered as a set of callable interfaces to software programs/components regardless of their implementation. They can be raised distantly via SOAP messaging. Therefore, these programs can provide amenities to other requests using Internet Protocols.

A web services developer typically uses a software tool to make the WSDL file from the basis code of a program in which some free methods marked as web services processes to be accessed by web services consumers. A web service consumer wishes to produce a web services proxy (client) to handle the encoding and messaging of areal web services call. A web services generation software tool that uses a WSDL file as input can create the web services proxy. The web services client proxy will change the serialization and de-serialization between internal data types and SOAP data representation.

The UDDI is a registry standard for web services providers to publish their web services. It may be used by a web services consumer to search by web services providers. The role of UDDI tells the relationships with services consumers and providers in the context of SOA.

2.1 Web services Adoption

Web Services will lead to major changes to business processes. The implementation process of web services can be conferred in several stages. Web services (specifically SOAP, UDDI and WSDL) enable any compliant application to invoke any interface function. The skill to raise a web service and depiction boundaries as services, it becomes pretty simple to use the skill as a means to hide platform differences. The second stage will undertake in their adoption that will be enabling of existing applications. The third one is the creation of available business services. Finally, will arrive as the number and types of web services increase in the organization. As the number of services grows the need for a central point and to create common solutions to various technical problems [3].

Security is often required as the most important in web services adoption. This is owing to numerous factors: the human legible nature of the SOAP payload, use of moderately open HTTP port and that endorses interoperability across the public internet [4].

3. Composite web service.

This paper focus on SOA framework for the composite web services. The SOA framework presented in this article has served as basis for a working implementation that provides a good foundation for the composite web services. The main purpose is to state exactly what services are needed, and focus on a general SOA framework that addresses general needs without adding the limitations. In most cases there is an understanding of that clear and simple interfaces and data models are required, as well as the importance of standards and descriptive information about services and modules in a SOA architecture.

Zhen Liu and et al [5] recommended a SOA appraisal model using two characteristic:

obtain ability, which is the quality attribute of whether the web service is present or ready for immediate use, and convenience, which is the quality attribute of service that the capable of serving a web service request.

Figure describes the SOA framework model for composite web services. This is used to confirm each and every service complicated in the composite web service. The arrangement involves three components: Service Producer, making the service with the precise quality attributes. Service Broker is a middleware component in which service producers register their web service and consumer discovers the web service which meets their requirements. Let us reflect the service consumer is in essential to invoke a composite web service. Composite service manager sends the service details to the data collector unit, theses component mine the Service Repository to verify whether these service available or not. Log data maintains feedback information from the Service Consumer.

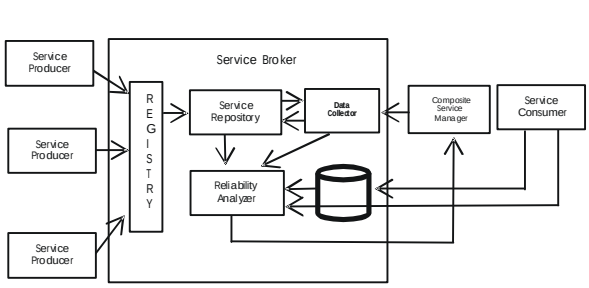


Fig1. Framework model for composite web service

The ultimate goal is to ensure that the eventual execution of composite service produces the desired behaviour. A framework is presented where the specifications of web service are translated to an intermediate representation Service Oriented Architecture may fundamentally change the way the software is made and used [5]. The current development platforms may no longer be suitable. The proposed framework has the benefits of SOA designing approach.

In most cases on SOA application consists of a number of services. In composite service requires cooperation with other services. The system is designed in accordance with the SOA principles and thus it consists of several services installed in an SOA environment and providing various features.

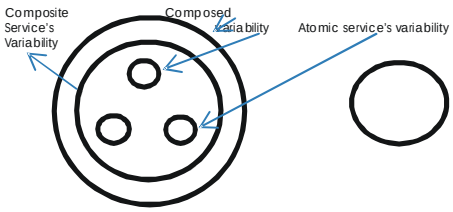


Fig2. Atomic and Composite Services Variability

The SOA is used to implement the standard and loosely coupled application architecture by using the characteristic of web service. In recent years, the SOA has been widely applied in system integration. An Atomic service is software component that should be indivisible. In other words, it would be impractical to decompose an atomic service because of its typical fine-grained structure. The term 'fine-grained' pertains to a service that offers limited functionality and possesses some few capabilities.

4. Conclusion

This paper gives the simple framework model for the web service among the available proxy services. The SOA framework presented in this article has served as basis for a working implementation that provides a good foundation for web services. A Composite Web Service can specify much different business logic involving the operations of its sub-services [10], developers need only focus on buildings up a framework model of available Services. This leads to our approach brings up some strong benefits. Thus, a composite service, once built can be used in many different applications. Developers of solicitations can just use a suitable composite web service which is now constructed to fulfil their needs.

References

[1]. Web Service Architecture requirements, W3C working Draft, available online at http://www.w3.org/TR/ 2002/WD-WSa-reqs-20021114.

[2].Gerardo Confora et al., "Migrating Interactive Legacy Systems to Web Services", Proceedings of the conference on software maintenance and Reengineering (QSMR'06) IEEE2006.

[3]. Accentrue, "Web Services and Enterprise Integration: Friends Not Foes", http://www. accenture.com/ xdoc/en/services/web/insights/web_enterprise.pdf,2003a.

[4]Booth.D.,et.al, Web Services Architecture, w3c working drafts, http://www.w3.org/TR/2003.

[5].H. Haas. Architecture and future of web services: fromSOAP to semantic web services.http://www.w3.org/ 2004/Talks/0520-hh-ws/, May 2004.

[6].E.Thomas, SOA: concepts, Technology and Design prentice Hall, 2005.

[7].P.Pries, Webtransact: A framework for specifying and coordinating reliable web service compositions, 2002.

[8].J.Bloomberg. The Seven Principles of SOA development.XML & Web Services, 2002.

[9].T. Bultan, X. Fu, R. Hull, and J. Su. Conversation specification: A new approach to design and analysis ofe-service composition. In Proc. Int. World Wide Web Conf.(WWW), May 2003.

[10]. G.Alonso, F.Casati, M.Kuno, and V,Machiraju. Web Services: Concepts, Architectures and Applications, Springer-Verlog.2004.

Research Explorer - International Journal
ISSN : 2250-1940 (Print), 2349 - 1647 (Online)
Impact Factor : 2.014(IRJIF), 1.85(JIF), 1.056(RIF), 0.60(QF) 0.398(GIF)
Issue - 19 ; August 2018

ANALYSIS OF KEY DISTRIBUTION MANAGEMENT IN WIRELESS SENSOR NETWORKS

M. Sudha
Guest Lecturer,
Department of Computer Science,
Government Arts College for Women, Krishnagiri.

Dr.M.Savitha Devi
HOD cum Asst. Professor,
Department of Computer Science,
PRUCAS, Harur, TamilNadu, India.

Abstract

As Sensor Networks edge closer towards wide-spread deployment, security issues become a central concern. The main focus has been on making sensor networks feasible and useful and less emphasis was placed on security. Here, we analyzes security challenges in wireless sensor network and summarizes key issues that should be solved for achieving the ad hoc security. It gives an overview of the current state of solutions on such key issues as secure routing and key management service. Sensor Networks are small, low-cost, low-power devices with the following functionality: they communicate over short distances, sense environmental data, and perform limited data processing. The user of encryption or authentication primitives between two sensor devices requires an initial link key establishment process, which must satisfy the low power and low complexity requirements. In these and other vital or security – sensitive deployments, secure transmission of sensitive digital information over the sensor network is essential.

Keywords: Wireless Sensor Networks, key Management, Ad hoc network, Cryptography.

1. Introduction

A wireless sensor network is a domain-specific wireless network. Wireless Sensor Networking is a communication technology which applies embedded system technology to the wireless technology. In a wireless sensor, each device is called a sensor node and every node is connected to one or many other sensor nodes. A wireless sensor network is a heterogeneous system of these sensors. These sensors are cost-effective and consume very low power. They are very small in size and have low communication bandwidth. They are ad-hoc networks. The main aim of a wireless sensor network is to gather information from the physical world. They monitor the physical or environmental state such as sound, temperature, pressure, vibrations or pollutants and communicate

data through the network accordingly. The sensors used in these networks collect data, process the data and pass the data through the central node. The WSNs are bi-directional and help in controlling sensor activities. They do not need much infrastructure to operate in. Currently, the sensors used in this technology do not exist at micro or nano level. Scientists are working towards this goal as the size of the sensors will determine how and where the network could be used.

Very energy-efficient, scalable, and strong security services including confidentiality, integrity, and group-level authentication of sensor data and routing control traffic are needed. Although significant progress has been shown in developing Wireless Sensor Networks (WSN) in many aspects including topology management, routing algorithm, MAC protocol and sensor data management, very little work is done on securing WSN. Research into authentication and confidentiality mechanisms designed specifically for WSN is needed. To understand the serious limitations of current security mechanisms, it is necessary to realize the salient differences between WSN and general ad-hoc networks1 since some proposals were already raised for securing ad-hoc networks.

The features of WSN such as low-memory, low-energy and large-scale nodes make it impractical to use the majority of the current secure algorithms that were designed for powerful workstations. For example, the working memory of a sensor node is insufficient to even hold the variables (of sufficient length to ensure security) that are required in asymmetric cryptographic algorithms.

Any security mechanisms for WSN should take the following has five major resource constraints into consideration: (1)limited energy, (2) limited memory, (3) limited computing power, (4) limited communication bandwidth, (5) limited communication range; more or less in descending order of acuteness.

The capabilities and constraints of sensor node hardware will influence the type of security mechanisms that can be hosted on a sensor node platform. Since the amount of additional energy consumed for protecting each message is relatively small, the greatest consumer of energy in the security realm is key establishment.

The ad-hoc networking topology renders a WSN susceptible to link attacks ranging from passive eaves-dropping to active interfering. Unlike fixed hardwired networks with physical defense at firewalls and gateways, attacks on a WSN can come from all directions and target at any node. Damage can include leaking secret information, interfering message and impersonating nodes, thus violating the above security goals. the wireless communication characteristics of WSN render traditional wired-based security schemes impractical.

Based on the above analysis on the security challenges, challenges and potential attacks in WSN, we further summarize three key issues for achieving the security of ad hoc networks:

1.1 Key Management in WSN

Confidentiality, integrity, and authentication services are critical to preventing an adversary from compromising the security of a WSN. Key management is likewise critical to establishing the keys necessary to provide this protection in WSN. However, providing key management is difficult due to the ad hoc nature, intermittent connectivity, and resource limitations of the sensor network environment.

Traditional key management service is based on a trusted entity called a certificate authority (CA) to issue public key certificate of every node. The trusted CA is required to be online in many cases to support public key revocation and renewal. But it is dangerous to set up a key management

service using a single CA in a sensor network. The single CA will be the vulnerable point of the network. If the CA is compromised, the security of the entire network is crashed. How to set up a trusted key management service for the WSN is a big issue.

1.2 Securing routing of WSN

There are two kinds of threats to ad hoc routing protocols: *(1) External attackers*. The attacks include injecting erroneous routing information, replaying old routing information, and distorting routing information. Using these ways, the attackers can successfully partition a network or introduce excessive traffic load into the network, therefore cause retransmission and ineffective routing. Using cryptographic schemes, such as encryption and digital signature can defend against the external attacks. *(2) Internal compromised nodes*. They might send malicious routing information to other nodes. It is more severe because it is very difficult to detect such malicious information because compromised

node can also generate valid signature.

Existing routing protocols cope well with the dynamic topology, but usually offer little or no security measures. An extra challenge here is the implementation of the secured routing protocol in a network environment with dynamic topology, vulnerable nodes, limited computational abilities and strict power constrains.

2. Key Management In Wsn

2.1 Problem description

Most of the security mechanisms require the use of some kind of cryptographic keys that need to be shared between the communicating parties. The purpose of key management is to:

• Initialize system users within a domain.

• Generate, distribute and install keying material.

• Control the use of keying material.

• Update, revoke and destroy keying material.

• Store, backup/recover and archive keying material.

But key management is an unsolved problem in WSN. Traditional Internet style key management protocols based on infrastructures using trusted third parties are impractical for large scale WSNs because of

the unknown network topology prior to deployment and serious node constraints such as limited power and limited transmission range.

At the extremes, there are *network-wide redeployed keying* and *node-specific pre-deployed keying* in sensor networks. Generally speaking, the problem of key management in WSN can be decomposed into the following sub-problems:

• **Key Pre-distribution**

To date, the only practical options for the distribution of keys to sensor nodes in WSN whose topology is unknown prior to deployment will have to rely on key pre-distribution. Keys have to be installed in sensor nodes to secure communications.

However, traditional key-distribution schemes have the following shortcoming: either a single *mission key* or a set of separate *n-1* keys, each being pair-wise privately shared with another node, have to be installed in every sensor node. In key pre-distribution, a big issue is how to load a set of

keys (called key ring) into the limited memory of each sensor. Other problems include the saving of the key identifier of a key ring and associating sensor identifier with a trusted controller node.

• Neighbor discovery

Every node needs to discover its neighbors in wireless communication range with which it shares keys. Thus neighbor discovery is also called shared-key discovery that establishes the topology of the sensor array as seen by the routing layer of the WSN. A 'link' exists between two sensor nodes only if they share a key. Good neighbor discovery scheme will not give an attacker any opportunity to discover the shared keys and thus the attacker can only do traffic analysis.

• End-to-end path-key establishment:

For any pair of nodes that do not share a key but are connected by multiple hops need to be assigned a path key for end-to-end secure communication. Path-key cannot be the one already used by the shared keys between neighbor nodes.

• Isolating aberrant nodes

An aberrant node is one that is not functioning as specified. Identifying and isolating aberrant nodes that are serving as intermediate nodes is important to the continued operation of the sensor network. A node may cease to function as expected for the following reasons:

- ✓ It has exhausted its source of power.
- ✓ It is damaged by an attacker.
- ✓ It is dependent upon an intermediate node and is being deliberately blocked because the intermediate node has been compromised.
- ✓ An intermediate node has been compromised and it is corrupting the communication by modifying data before forwarding it.
- ✓ A node has been compromised and it Communicates fictitious information to the base station.

• Re-keying

Although it is anticipated that in most WSNs the lifetime of a key shared between two nodes exceeds that of the two nodes, it is possible that in some cases the lifetime of keys expires and re-keying must take place. Re-keying is a challenge issue since new keys needs to be generated in an energy-efficient way and the re-keying period should be determined based on the security level to be achieved. Re-keying is equivalent with a self revocation of a key by a node.

• Key-establishment latency

Recent investigation reveals that latency is potentially a significant impediment to secure network initialization. As with energy consumption, latency due to communications is a much larger factor than computational latency. Thus any key management scheme should take latency reduction as a crucial factor.

2.2 Solutions

Currently there are some key management schemes that can be partially used for securing WSN environments even though most of those schemes are proposed for general ad hoc networks.

2.2.1. Hybrid key-based protocols

An obvious conclusion from current research results is that a single keying protocol will not be optimal for all sensor network topologies, densities, sizes, and scenarios. Protocols such as Identity-

Based Symmetric Keying and Rich Uncle have limited application until the network's *routing infrastructure* has been sufficiently well established. Individually other protocols such as the public-key group and pair wise keying protocols consume too much energy. For *significant* sensor networks, a mix of public key-based protocols, including pair wise, group keying, and distribution keying, provide an energy efficiency superior to using just a single protocol.

2.2.2. Threshold cryptography

A solution to deal with key management in general ad hoc networks is proposed by Zhou and Hass in and may be borrowed to WSN environments. It uses a *(k, n)* threshold scheme to distribute the services of the certificate authority to a set of specialized server nodes. Each of these nodes is capable of generating a partial certificate using their share of the certificate signing key *skCA*, but only by combining *k* such partial certificates can a valid certificate be obtained. The solution is suitable for planned, long-term ad hoc networks. However, it may not be applicable for WSN because sensor networks can lose some nodes whose energy is run out of. In addition, is based on public key encryption and thus requires that the all the nodes are capable of performing the necessary computations, which may not be feasible for energy-limited sensor nodes.

2.2.3. Certificate repository

Hubaux et al. go a step further than, by requiring each node to maintain its own *certificate repository*. These repositories store the public certificates that the node themselves issue, and a selected set of certificates issued by the others. The *performance* is defined by the probability that any node can obtain and verify the public key of any other user, using only the

local certificate repositories of the two users. The dilemma is: too many certificates in a sensor node would easily exceed their capacity, yet too few might greatly impact the performance (as previously defined) of the entire network.

2.2.4. Fully Distributed Certificate Authority

Fully Distributed Certificate Authority is first described by Luo and Lu in and later analyzed by Luo et al. Its uses a *(k, n)* threshold scheme to distribute an RSA certificate signing key to all nodes in the network. It also uses verifiable and proactive secret sharing mechanisms to protect against denial of service attacks and compromise of the certificate signing key. Since the service is distributed among all the nodes when they join the network, there is no need to elect or choose any specialized server nodes. Similar to the solution presented in, this solution is aimed towards planned, long-term ad hoc networks with nodes capable of public key encryption and thus could not adapt the routing changing of sensor networks.

2.2.5. Pebblenets

Secure Pebblenets proposed by Basagni et al provides a distributed key management system based on symmetric encryption. The solution provides group authentication, message integrity and confidentiality. This solution is suitable for planned and distributed, long-term ad hoc networks consisting of low performance nodes that are unable to perform public key encryption. We hold the same opinion as and believe that this solution can provide more practical security scheme for sensor networks.

Pebblenets use only symmetric cryptography. The disadvantage is that once a node is compromised, forward secrecy is broken, therefore tamper-resistance becomes crucial of *threshold cryptography*. In addition, in pebble nets a key management server not only has to store its own key pair, but also the public keys of all the nodes in the network. The difficulty includes the

storage requirement exerted on the servers which must potentially be specialized nodes in the network, and the overhead in signing and verifying routing message both in terms of computation and of communication.

3. Conclusions

In this paper, we analyzed security challenges in wireless sensor networks and summarized key issues that should be solved for achieving the WSN security. It also gave an overview of the current state of solutions on three key issues including the prevention of secure routing and key management service.

4. References

1. Akyildiz, I., Su, W., Sankarasubramaniam, Y., and Cayirci, E., "A Survey on Sensor Networks", *IEEE Communications Magazine*, August 2002.
2. Jeffery Undercoffer, Sasikanth Avancha, Anupam Joshi and John Pinkston, "Security for Sensor Networks," 2002 CADIP Research Symposium,
3. P. Papadimitratos, Z. Haas, "Secure Routing for Mobile Ad hoc Networks," Proceedings of the SCS Communication Networks and Distributed Systems Modeling and Simulation Conference (CNDS 2002).
4. D. Ganesan, R. Govindan, S. Shenker and D. Estrin, "Highly Resilient, Energy Efficient Multipath Routing in Wireless Sensor Networks," Mobile Computing and Communication Review (MC2R) Vol 1., No.2. 2002.

Research Explorer - International Journal
ISSN : 2250-1940 (Print), 2349 - 1647 (Online)
Impact Factor : 2.014(IRJIF), 1.85(JIF), 1.056(RIF), 0.60(QF) 0.398(GIF)
Issue - 19 ; August 2018 © Author

AN EFFICIENT FP - TREE BASED HIGH UTILITY PATTERN APPROACH FOR MINING ITEMSETS WITH INDEXING

K.Valarmathi
HOD & Assistant Professor
Department of Computer Science.
Trinity College for Women,
Namakkal - 637001.Tamil Nadu, India.

Abstract

Conventional methods of Association rule mining and Frequent Item set Mining (FIM) cannot satisfy the anxieties emerging from certain real applications. In real world, for making some decisions the user wants to know the total profit grossed by an item set or item. To evaluate this it needs to take into account the quantity of the purchased item. The profit of an item considers the gain of single item and the number of item purchased. To address these, utility mining has been introduced. In this the utility of an item set is calculated as the number of item purchased and the product of the gain of the item. Utility mining concentrates on both the reputation of an item in the knowledge base (i.e.) profit or exterior utility and the reputation of an item in the transaction (i.e.) quantity or interior utility of an item. In this study, a novel frequent-pattern tree (FP-tree) structure, which is an extended prefix-tree structure for storing crucial information about frequent patterns, and develop an efficient FP-tree-based mining method based on the generation of conditional utility pattern base which leads to the conditional utility FP-tree for mining the complete set of frequent patterns. As the item set in the sanitized database and original database are segmented in different areas, the item set kept in the different areas are indexed through candidate keys for increasing the access speed and fast retrieval of data. This process will increase the accuracy of the database and it preserves the sensitive data items for a longer time. The assessment report shows that the generation of less candidate patterns makes algorithms to run faster.

Keywords: *Utility Mining, High-utility item sets, rare item sets; Frequent Item set mining, Frequent Pattern.*

1. Introduction

Data mining can be characterized as a movement that concentrates some learning contained in expansive transaction databases. Information mining, the extraction of concealed prescient data from substantial Databases, is an intense new innovation with awesome potential to help organizations concentrate on the most vital data in their information distribution centers. Learning

Discovery in Databases (KDD) is the non- unimportant procedure of recognizing Legitimate, already obscure and conceivably helpful examples in information. These examples are utilized to make forecasts or characterizations about new information, clarify existing information, outline the substance of a huge database to bolster basic leadership and give graphical information perception to help people in finding further examples. The limitations of frequent item set mining motivated researchers to conceive a utility based mining approach, which allows a user to conveniently express his or her perspectives concerning the usefulness of item sets as utility values and then find item sets with high utility values higher than a threshold. In utility based mining the term utility refers to the quantitative representation of user preference i.e. according to an item sets utility value is the measurement of the importance of that item set in the user's perspective. The traditional ARM approaches consider the utility of the items by its presence in the transaction set. The frequency of item set is not sufficient to reflect the actual utility of an item set. For example, the sales manager may not be interested in frequent item sets that do not generate significant profit. Recently, one of the most challenging data mining tasks is the mining of high utility item sets efficiently. In view of this, utility mining emerges as an important topic in data Mining for discovering the item sets with high utility like profits. Identification of the item sets with high utilities is called as Utility Mining. The utility can be measured in terms of cost, quantity, profit and user preference .for this Utility mining model was proposed to define the utility of item set. In [1] this model by considering u(X) as a utility of an item set X, which is the sum of the all utilities of item set X in all the transactions containing X. then an item set X is called a high utility items if its utility greater or equal to user- defined minimum utility thresh resold.

Association rules are if/then statements that aid uncover relationships between seemingly dissimilar data in a relational database or other information repository [1, 2]. An example of an association rule would be "If a customer buys one dozen eggs, he is 80% likely to also purchase milk." An association rule has two parts, an antecedent (if) and a consequent (then). An antecedent is an item found in the data. A consequent is an item that is found in combination with the antecedent. Association rules are shaped by analyzing data for frequent if/then patterns and consuming the criteria support and confidence to classify the most important relationships. Provision is an indication of how frequently the items appear in the database. Confidence specifies the number of times the if/then statements have been found to be true. The problem of association rule mining is defined as: Let I ={i1, i2,.....,in} be a set of n binary attributes called items. Let D = {t1, t2,....., tm} be a set of transactions called the database. Each transaction in D has a unique transaction ID and contains a subset of the items in I. A rule is defined as an inference of the form X==>Y where X & Y " 1 and X)" Y = è. The sets of items (for short item sets) X and Y are called antecedent (left-hand-side or LHS) and consequential (right-hand-side or RHS) of the rule respectively. Example: The set of items is I = {milk, bread, butter, beer}. An example rule for the supermarket could be {butter, bread} ==> {milk} meaning that if butter and bread are bought, customers also buy milk[7]. In data mining, association rules are valuable for analyzing and predicting customer behavior. They play an vital part in shopping basket data analysis, product clustering, catalog design and store layout[7]. Programmers usage association rules to form programs capable of machine learning. Machine learning is a type of artificial intelligence (AI) that seeks to build programs with the ability to become more efficient without being explicitly programmed. In overall, association rule mining can be viewed as a two-step process: (i) find all frequent patterns and (ii) Generate strong association rules from the frequent patterns [2]. In (i), we can use some mining algorithms like Apriori, DHP, ECLAT, FP Growth etc. that we discussed later. Also we proposed newer algorithm for frequent pattern mining. In (ii), all frequent pattern rules are checked for

minimum support and minimum confidence to generate association rules.

Example: An item set is called a high utility item set if its utility is no less than a user- specified minimum utility threshold or low- utility item set represented by min-util.

Table 1: Sample Database

TID	Transaction	TU
T1	(A,1)(C,10)(D,1)	17
T2	(A,2)(C,6)(E,2)(G,5)	27
T3	(A,2)(B,2)(D,6)(E,2)(F,1)	37
T4	(B,4)(C,13)(D,3)(E,1)	30
T5	(B,2)(C,4) (E,1)(G,2)	13
T6	(A,1)(B,1) (C,1)(D,1)(H,2)	12

Table 2 : Profit Table

Profit	5	2	1	2	3	5	1	1
Item	A	B	C	D	E	F	G	H

From table 1 and 2,

Utility ({A, T1}) =5×1=5

Utility ({AD,T1})=u({A,T1})+u({D,T1})=5+2=7

Utility ({AD}) =u ({AD, T1}) +u ({AD, T3})

=7+17=24

Utility ({BD}) =u ({BD, T3}) +u ({BD, T4})

=16+18=34

This paper is organized as follows. In Section II, we describe the related works of this system. Section III describes the problem identified. Section IV presents the design of our solution for Tree Based High Utility Pattern Approach. The evaluation results of both approaches are discussed in Section V. Finally, we discuss the conclusion and possible enhancement of future works related to the presented solution as section VI.

2.RELATED WORKS

In [2] Vincent S Tseng, Bai-En Shie, Cheng-Wu, Philip Sproposed a Single-pass incremental and intuitive digging for finding weighted frequent examples. The current weighted frequent example (WFP) digging can't be connected for incremental and intelligent WFP digging furthermore for stream information mining since they depend on a static database and its require various database examines. To defeat this, they proposed two novel tree structures IWFPTWA (Incremental WFP tree in light of weight rising request) and IWFPTFD (Incremental WFP tree in view of plunging request) and two new calculations IWFPWA and IWFPFD for incremental and intuitive mining utilizing a solitary database filter. IWFPFD promises that any non-applicant thing can't show up before competitor things in any branch of IWFPTFD and in this manner accelerates the prefix tree. The downside of this approach is that extensive memory space, tedious and it is exceptionally hard to bolster the calculation for bigger databases.

In [3], presented a novel utility FP- tree, an extensive tree structure for storing essential

information about frequent patterns for mining the high utility item sets. We have utilized the standard FP growth algorithm for mining the complete set of frequent patterns by pattern growth. The efficiency of the high utility pattern mining is recognized by two important thoughts. One is the construction of the utility FP-tree and the other one is the mining of utility item sets from the utility FP- tree. Our proposed utility FP-tree-based pattern mining utilized the pattern growth method to avoid the costly generation of a large number of candidate sets in which it dramatically reduces the search space. The experimentation was carried out on our proposed approach using real life datasets and the results showed that the proposed approach is effective on the tested databases. Frequency - weighted utility *FWU* of an item *ip*, denoted by *FWU* (*ip*), is computed using the Transaction Frequency (*TF*), Transaction Weight age (TW) and the External Utility (*Eu*).

$$FWU\ (ip) = \frac{TF\ (ip) * TW\ (ip) * Eu\ (ip)}{UF}$$

An operation in object-oriented databases gives rise to the processing of a path. Several database operations may result into the same path. Choenni et al., [4] address the problem of optimal index configuration for a single path. As it is shown an optimal index configuration for a path can be achieved by splitting the path into sub paths and by indexing each sub path with the optimal index organization. The authors present an algorithm which is able to select an optimal index configuration for a given path. The authors consider a limited number of existing indexing techniques (simple index, inherited index, nested inherited index, multi-index, and multi-inherited index) but the principles of the algorithm remain the same adding more indexing techniques.

Kriegel et al., [5] presented a technique to achieve efficient query processing on data-partitioning index structures within general purpose database systems. The navigational index traversal cost is reduced by using “extended index range scans”. If a directory node is “largely” covered by the actual query, the recursive tree traversal for this node can beneficially be replaced by a scan on the leaf level of the index instead of navigating through the directory any longer. On the other hand, for highly selective queries, the index is used as usual. The author demonstrated the benefits of this idea for spatial collision queries on the relational R-tree. Energy saving is one of the most important issues in wireless mobile computing. Among others, one viable approach to achieving energy saving is to use an indexed data organization to broadcast data over wireless channels to mobile units. Using indexed broadcasting, mobile units can be guided to the data of interest efficiently and only need to be actively listening to the broadcasting channel when the relevant information is present.

3. Problem Description

The problem of mining utility item sets is discussed and some basic definitions are described in this subsection. Let I = {i1, i2 ,..., im}be a set of items and D = {t1 , t2 ,..., tn } be a transaction database where the items of each transaction ti is a subset of I . The utility of item ip in transaction tq, denoted as U (ip,tq) is defined as Iu (ip,tq)× Eu (ip) . Let an item set X be a subset of I. The utility of X in transaction tq, denoted by U(X, tq) is defined as U(X, tq) “ip “ XU (ip,tq) . The task of high utility mining is to find all items that have utility above a user-specified min_utility. Since utility is not anti- monotone, the concept of Frequency Weighted Utility (FWU) is used to prune the search space of high utility item sets.The internal utility or local transaction utility value Iu(ip,tq) represents the quantity of item ip in transaction tq. The external utility Eu(ip) represents

the unit profit value of item ip.Utility U(ip,tq) is the quantitative measure of utility for item ip in transaction tq defined by U (ip,tq)=Iu(ip,tq)×Eu(ip).The utility of an itemset X in transaction tq, U(X ,tq, is defined by U (X,tq)="ip"XU(ip,tq);where X ={i1,i2,.....ik}is a k-itemset, X †"tq and 1 d"kd"m.

4. Tree Based High Utility Pattern Approach

In data mining, association rule learning is a popular and well researched method for discovering interesting relations between variables in large databases. An example rule for the supermarket dataset could be {butter, bread} {milk}.Association rule learning is about selecting the desired rules from the set of all possible rules with constraints on various measures of significance and interest. The best-known constraints are minimum thresholds on support and confidence. Let I = {i1, i2… in} be a set of n binary attributes called items. Let *D*= {t1, t2… tm} be a set of transactions called the database. Each transaction in *D* has a unique transaction ID and contains a subset of the items in *I*. Association rules are usually required to satisfy a user-specified minimum support and a user-specified minimum confidence at the same time.

Association rule generation is usually split up into two separate steps:

First, minimum support is applied to find all frequent item set in a database.

Second, these frequent item sets and the minimum confidence constraint are used to form rules.

There are a number of algorithms used to generate Association rules. The Association rule algorithms discussed in this thesis is:

Apriori algorithm

Apriori is a classic algorithm for learning association rules. Apriori is designed to operate on databases containing transactions (for example, collections of items bought by customers, or details of a website frequentation).Other algorithms are designed for finding association rules in database having no transactions or having no timestamps. As is common in association rule mining, given a set of item sets (for instance, sets of retail transactions, each listing individual items purchased), the algorithm attempts to find subsets which are common to at least a minimum number C of the item sets.

Apriori uses a "bottom up" approach, where frequent subsets are extended as one item at a time (a step known as candidate generation), and groups of candidates are tested against the data. The algorithm terminates when no further successful extensions are found. The purpose of the Apriori Algorithm is to find associations between different sets of data. It is sometimes referred to as "Market Basket Analysis". Each set of data has a number of items and is called a transaction. The output of Apriori is set of rules that tell us the frequency of occurrence items in the dataset.

Eclat algorithm

Eclat algorithm finds the elements from the bottom like depth first search. Eclat algorithm is very simple algorithm to find the frequent item sets. This algorithm uses vertical database. It cannot use horizontal database. If there is any horizontal database, then we need to convert into vertical database. here is no need to scan the database again and again. Éclat algorithm scans the database only once. Support is counted in this algorithm. Confidence is not calculated in this algorithm.

The FP-growth (Frequent Pattern – growth) algorithm

The FP-growth (Frequent Pattern

– growth) algorithm differs basically from the level-wise algorithms, that uses a "candidate generate and test" approach. It does not use candidates at all, but it compresses the database into the memory in a form of a so- called FP-tree using a pruning technique. The patterns are discovered using a recursive pattern growth method by creating and processing conditional FP-trees. The drawback of the algorithm is its huge memory requirement, which is dependent on the minimum support threshold and on the number and length of the transactions.

Step 1: Frequent Pattern (FP) Mining

Frequent pattern (FP) mining discovers patterns in transaction databases based only on the relative frequency of occurrence of items without considering their utility. For many real world applications, however, the utility of item sets based on cost, profit or revenue is important. The utility mining problem is to find item sets that have higher utility than a user specified minimum.Unlike item set support in frequent pattern mining, item set utility does not have the anti- monotone property and so efficient high utility mining poses a greater challenge.

The frequent pattern mining problem does not take into account the quantity or an associated weight such as price or profit of an item but it represents only the occurrence of each item in a transaction by a binary value. But, quantity and weight are important factors for solving real world decision problems that intends to maximize the utility of an organization. Hence, all itemsets that have utility value greater than a user specified minimum utility value are identified by high utility itemset mining. Both local transaction utility and external utility contribute to the utility of an item. Identifying high utility item sets which drive a major share of the overall utility is the objective of utility mining. High utility pattern mining approaches have been proposed to overcome this problem. As a result, it becomes a very important research issue in data mining and frequent pattern mining. In this research, we have presented an efficient approach for mining the high utility item sets from the utility FP-tree structure. The procedure used for mining high utility items involves the following important steps.

Step 2: Construction of Utility FP-tree

In general, the construction of the FP-tree and the mining patterns from the FP- tree are the major important steps in the frequent pattern tree algorithm. Similarly, the proposed approach also contains these two steps, where the utility FP-tree is constructed using the frequency weighted utility rather than the frequency value. In addition to this, the mining process utilizes pattern growth methodology, where the support is computed based on the frequency weighted utility rather than the frequency. In this section, we describe the construction process of our proposed utility FP- tree structure based on the frequency weighted utility.

The proposed algorithm is explained with the help of a simple example for easier understanding the entire step including the tree construction and mining processes. Table 3 provide an example of a transaction database and Table 3.2 gives the unit profit for each item belonging to the transaction database.

Table 3: Example of a Transaction Database

Item TID	A	B	C	D
01	2	1	0	1
02	3	0	2	0
03	0	3	2	0

Table 4: Example of a Utility Table

Item	Profit ($)
A	2
B	1
C	5
D	1

Example: Let us consider the items present in Table 3.1. The FWU of item 'A' is computed as follows: the TF (ip) of the item 'A' is 2; the TW (ip) is 5 and the profit records for that item is 2. Also, the total profit value, called utility factor is found out to be 9 in this case. Now, FWU (A) = 2.22, FWU (B) = 0.77, FWU (C) = 4.44 and FWU (*D*) = 0.11. In this case, we have taken the min_util value as 0.3 and chosen the items which have utility values greater than the min_util value. Based on these computed utility values, the items are re-ordered. The transactions with sorted items are utilized for illustrating the construction of the utility FP-tree. The ordered transactions and Frequency Weighted Utility (FWU) are shown in Table 3 and Table 4 respectively.

Table 5: The ordered transactions with sorted large items

TID	Frequent Items	
01	A	B
02	C	A
03	C	B

Table 6: Frequency Weighted Utility (FWU)

Item	FWU
A	2.22
B	0.77
C	4.44
D	0.11

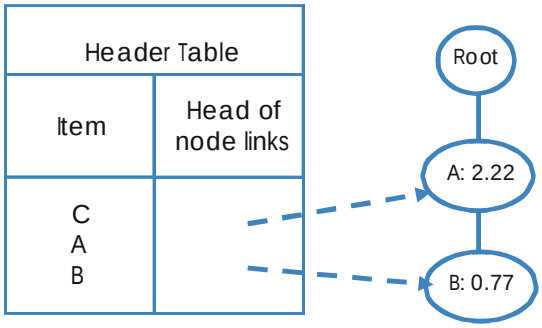


Figure 1: The Utility FP-tree after the first transaction is processed

Subsequently, the next transaction containing frequent weighted utility items (C, A) is processed. Here, the items do not contain any prefix path in the utility FP-tree after executing the first transaction, so the new nodes (C: 4.44) are attached to the root node as its child. Also, the other new node (A: 2.22) is created and linked with the child of (C: 4.44). The results after the second transaction are shown in Figure 2.transaction is processed.

For processing the third transaction, the path <"C" "B"> shares the same prefix "C" with the Utility FP-tree so the count of the node (C: 4.44) is incremented by 4.44 as it shares the common prefix and a newly created node (B: 0.77) is attached to the node (C: 8.88) as its child node. The results after the third transaction are shown in Figure 3.

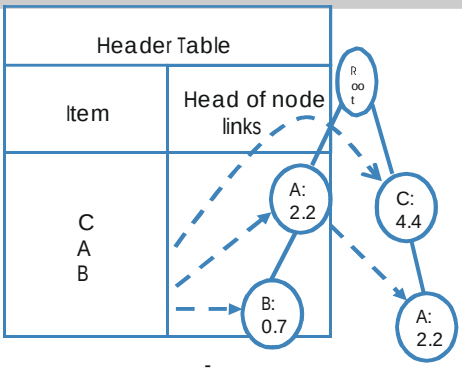


Figure 2: The Utility FP-tree after the second

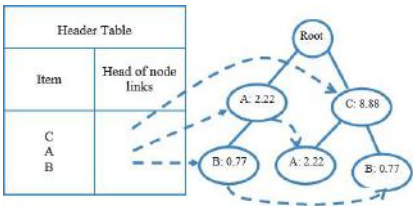


Figure 3: The utility FP-tree after the third transaction is processed

After the Utility FP-tree is constructed from a transaction database, a mining process is executed to determine the large items. Utility FP-tree derives the utility itemsets directly from the utility FP-tree and do not necessitate generation of candidate item sets for mining. It recursively processes the utility items one by one and bottom-up with regard to the Header Table. By constructing a conditional utility FP-tree for each utility item, high utility item sets are mined recursively from it. This process is executed until all the items in the utility FP-tree get processed.

Step 3: Mining of High Utility item sets from the Utility FP-tree

The next major step is examining the mining process based on the constructed utility FP-tree as shown in Figure 3.3. The mining process of utility item sets from the utility FP-tree based on the pattern growth methodology is explained as follows.

Generation of Conditional Utility Pattern Base and Conditional Utility FP Tree,

After the utility FP-tree is constructed from an ordered transaction database, the mining procedure starts with the generation of the conditional utility pattern base and the conditional utility FP-tree. We have generated a conditional utility pattern base and the conditional pattern tree as the utility FP-tree as shown in Figure 3.4.

Here, we start with the mining process from the bottom of the nodes of the utility FP-tree and their corresponding prefix paths are extracted from it.

Then, their relevant utility pattern base and conditional utility FP-tree are generated in order to mine 2-length utility patterns.

Example: At first, we process the item “B”, which is the bottom item present in the header table so that two prefix paths that exist for item B are extracted. For an item B, the conditional pattern base is (A: 0.77) and (C: 0.77), which are the prefix paths of the item “B”.

Then, the conditional utility FP tree is generated for the item “B”. Again, the conditional

pattern base is generated for the superset of "A" i.e., "AB" and "AC" but no prefix paths exists for this sequence so it generates a NULL path. Subsequently, the next items A and C are processed. The conditional pattern base for item "A" is (C: 2.22) and the conditional pattern base for the pattern "C" is null.

The conditional pattern-bases and the conditional FP-trees generated are summarized in Table 7.Figure 4: Mining of FP-tree by creating conditional pattern-bases

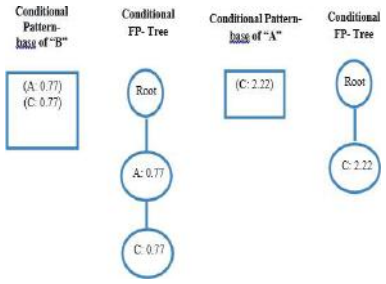


Table 7: Miningfrequent patterns by creating conditional pattern-bases

Item	Conditiona l pattern-base	Conditional FP -tree
B	{(A:0.77), (C:0.77))	{(A,C)}/B
A	{(C:2.22)}	{(C)}/A
C	?	?
BA	?	?
BC	?	?
AC	?	?

Mining of utility patterns

After the generation of the conditional utility FP-tree, high utility patterns are mined from it, based on the minimum support threshold. Here, utility patterns are mined recursively from the conditional utility FP tree so that all length patterns having frequency weighted utility greater than the minimum threshold are obtained. The patterns are said to be frequent weighted utility patterns if their support is greater than the *min_util*.

Example: The results obtained for the sample database given in Table 3.1 are shown in the Table 3.6. The frequent weighted utility patterns are {(C: 8.88) (A: 4.44), (B: 1.44), (AB: 0.77), (CB: 0.77) and (CA: 2.22)}.

Table 8: Frequent weighted utility patterns for a sample database

Frequent Patterns	
C: 8.88	CA: 2.22
A: 4.44	AB: 0.77
B: 1.44	CB: 0.77

Table 1 : Transactions Table

TID	Items
T1	I1,I2,I3,I5,I6,I15
T2	I1,I3,I7
T3	I5,I9
T4	I1,I3,I4,I5,I7
T5	I1,I3,I5,I7,I12
T6	I5,I10
T7	I1,I2,I3,I5,I6,I16
T8	I1,I3,I4
T9	I1,I3,I5,I7,I13
T10	I1,I3,I5,I7,I14

Table 2 : 1- Frequent Itemsets

Itemset	Sup Count
I2	2
I4	2
I6	2
I7	5
I1	8
I3	8
I5	8

Table 3 : Sorted Transactions Table

TID	Items	Ordered Items
T1	I1,I2,I3,I5,I6,I15	I2,I6,I1,I3,I5
T2	I1,I3,I7	I7,I1,I3
T3	I5,I9	I5
T4	I1,I3,I4,I5,I7	I4,I7,I1,I3,I5
T5	I1,I3,I5,I7,I12	I7,I1,I3,I5
T6	I5,I10	I5
T7	I1,I2,I3,I5,I6,I16	I2,I6,I1,I3,I5
T8	I1,I3,I4	I4,I1,I3
T9	I1,I3,I5,I7,I13	I7,I1,I3,I5
T10	I1,I3,I5,I7,I14	I7,I1,I3,I5

Table 4 : BitTable Representation of Database

TID	I2	I4	I6	I7	I1	I3	I5
T1	1	0	1	0	1	1	1
T2	0	0	0	1	1	1	0
T3	0	0	0	0	0	0	1
T4	0	1	0	1	1	1	1
T5	0	0	0	1	1	1	1
T6	0	0	0	0	0	0	1
T7	1	0	1	0	1	1	1
T8	0	1	0	0	1	1	0
T9	0	0	0	1	1	1	1
T10	0	0	0	1	1	1	1

Table 5 : Frequent Pattern Generation for Each Item

ITEMS						
I2	I4	I6	I7	I1	I3	I5
I2 : 2	I4 : 2	I6 : 2	I7 : 5	I1 : 8	I3 : 8	
I2,I6 : 2	I4,I1 : 2	I6,I1 : 2	I7,I1 : 5	I1,I3 : 8	I3,I5 : 6	
I2,I1 : 2	I4,I3 : 2	I6,I3 : 2	I7,I3 : 5	I1,I5 : 6		
I2,I3 : 2	I4,I1,I3 : 2	I6,I5 : 2	I7,I1,I3 : 5	I1,I3,I5 : 6		
I2,I5 : 2		I6,I1,I3 : 2	I7,I5 : 4			
I2,I6,I1 : 2		I6,I1,I5 : 2	I7,I1,I5 : 4			
I2,I6,I3 : 2		I6,I3,I5 : 2	I7,I3,I5 : 4			
I2,I6,I5 : 2		I6,I1,I3,I5 : 2	I7,I1,I3,I5 : 4			I5 : 8
I2,I1,I3 : 2						
I2,I1,I5 : 2						
I2,I3,I5 : 2						
I2,I6,I1,I3 : 2						
I2,I6,I1,I5 : 2						
I2,I6,I3,I5 : 2						
I2,I1,I3,I5 : 2						
I2,I6,I1,I3,I5 : 2						
TOTAL = 43						

Example 1: Let us consider an example of a simple database with 10 transactions and min_sup is 2 as shown The sorted transactions are shown in following example Table 3. Then the scanned database is represented by Bit Table shown in Table 4. Then calculate the intersection of transactions that contain certain frequent items one by one.

Let us take frequent item I2 as example, candidate

= T1)" T7

= 1010111)" 1010111

= 1010111

That is I2 subsume is I6 I1 I3 I5 i.e. (I2, I6 I1 I3 I5)

Continue this process accordingly for all items and finally the subsume index array is (I2, I6 I1 I3 I5), (I4, I1 I3), (I6, I1 I3 I5), (I7,

I1 I3), (I1, I3), (I3, Ø), (I5, Ø).

5. Experimental Results And Performance Evaluation

Nursery dataset

The experimental results are obtained for diverse support values on the Nursery dataset. The obtained results are plotted as graphs as shown in Figures 3.15, 3.16, 3.17, 3.18 and 3.19 that shows the performance of the four approaches on Nursery dataset in the effectual mining of high utility item sets. Here, the performance of our proposed approach is evaluated for different support values (Normalized between 0.1 and 0.5) and the corresponding generated length of patterns. It is evident from the analysis of the plotted graphs, that our proposed approach produces better results than the standard Apriori, Eclat and FP-growth algorithm.As the support value varies, the number of generated frequent patterns gets reduced in our proposed approach than the Apriori, Eclat and FP-growth algorithm by different length of patterns.

In Figure 5, the number of patterns generated by varying the support thresholds of 1 length patterns is constricted compared to the Apriori, Eclat and FP- growth algorithm.

Likewise, Figures 6, 7 and 8 shows the generated number of patterns of length 2, 3 and 4 respectively for the different supports for the algorithms.

But, as shown in Figure 9, no 5 length patterns are produced by our proposed approach unlike the Apriori, Eclat and FP- growth algorithm which generated a limited number of patterns.

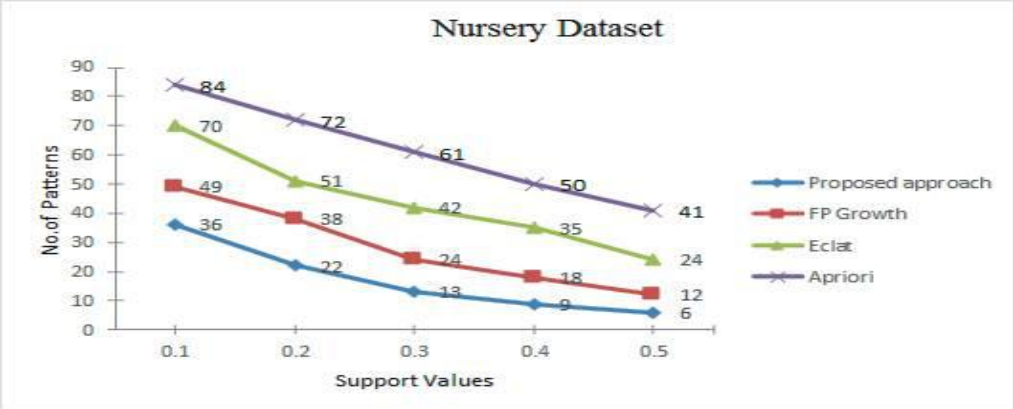


Figure 5:No.of frequent patterns (1-length) generated using various support thresholds

X axis - supporting threshold values range from 0.1 to 0.5

Y axis – Number of frequent patterns (1- length).

The number of patterns generated by varying the support thresholds of 1 length patterns is reduced compared to the Apriori, Eclat and FP-growth algorithm.

Table 9: No. of frequent patterns (1-length)

Generated using various support thresholds

Support Value	No of patterns			
	Aprio	Ecl	FP	Propos
0.1	84	70	49	36
0.2	72	51	38	22
0.3	61	42	24	13
0.4	50	35	18	9
0.5	41	24	12	6

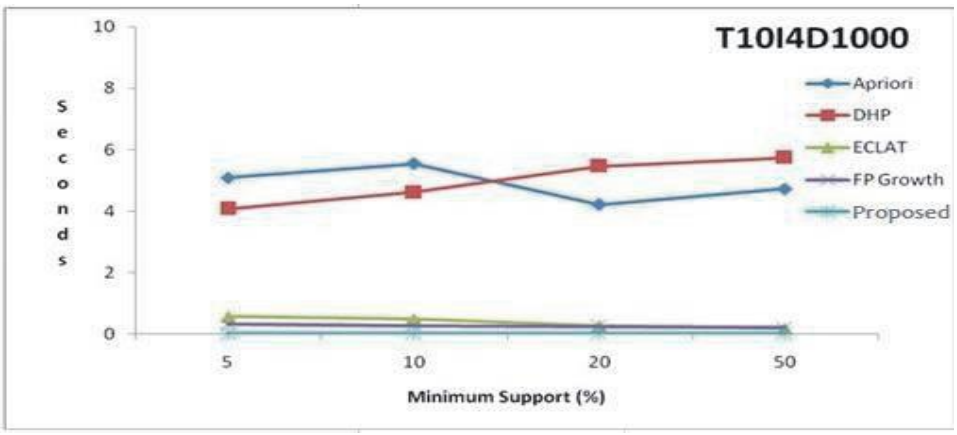


Figure 6: Execution time (in seconds) required by five different algorithms in T10I4D1000 dataset with different minimum support threshold.

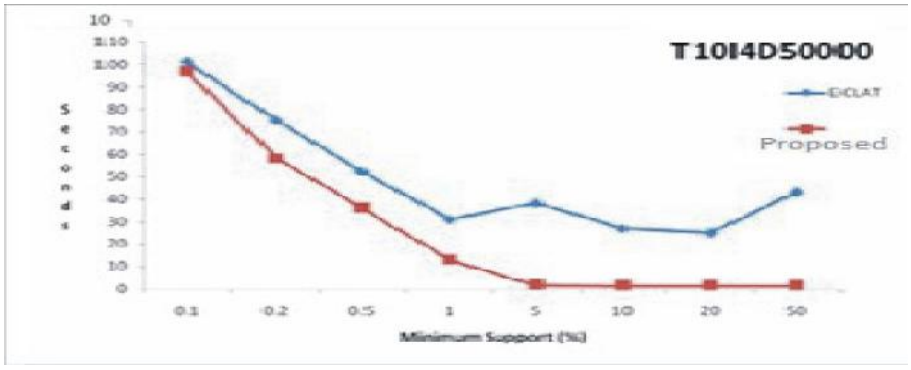


Figure 7: Execution time (in seconds) required by two different algorithms in T10I4D50000 dataset with different minimum support threshold.

6. Conclusion

An approach for weighted utility item sets mining from the database so as to improve the performance of data mining is proposed. And we have presented a utility FP-tree by utilizing a

tree structure for storing essential information about frequent patterns for mining high utility item sets. We have utilized the standard Apriori, Eclat and FP- growth algorithms for mining the complete set of frequent patterns by means of pattern growth. Higher efficiency in mining high utility patterns is realized by implementing two important concepts. One is the construction of the utility FP-tree and the other one are the mining of utility item sets from the utility FP-tree. Our proposed utility FP-tree-based pattern mining utilizes the pattern growth method to avoid the costly generation of a large number of candidate sets and reduces the search space dramatically. Apart from that, in this paper we have used indexes which generate more frequent patterns directly. Comparing with Apriori, DHP, ECLAT and FP Growth, the new algorithm reduces time for many frequent item sets generation and candidate frequent item sets which support count do not need to be computed. So the efficiency of the new algorithm is better than all of the above discussed algorithms. Performance Evaluation shows that new algorithm in performance is more remarkable for mining frequent patterns.

7. References

[1] Pranoti Meshram , Prof. Vikrant Chole, "Efficient Approach for Mining of High Utility Itemsets from Transactional Database", IJCSMC, Vol. 4, Issue. 6, June 2015, pg.695 – 700.

[2] Vincent S Tseng, Bai-En Shie, Cheng-Wu, Philip S, "Efficient algorithms for mining high utility itemsets from transactional databases", IEEE Transactions on knowledge and data engineering, 2013.

[3] C.Saravanabhavan , R.M.S.Parvathi, "Utility Fp-Tree: An Efficient Approach for Mining of Weighted Utility Itemsets", International Journal of Engineering Research and Development, Volume 8, Issue 7 (September 2013), PP.19-31.

[4] Choenni S, Blanken H, Chang T. Index selection in relational databases. InComputing and Information, 1993. Proceedings ICCI'93., Fifth International Conference on 1993 May 27 (pp. 491-496). IEEE.

[5] Berchtold S, Keim DA, Kriegel HP. An index structure for high-dimensional data. Readings in multimedia computing and networking. 2001 Aug 10;451.[6] Adinarayanareddy B, Srinivasa Rao O & Krishna Prasad MHM 2012, 'An Improved UP-Growth High Utility Itemset Mining', International Journal of Computer Applications (0975 – 8887) vol.58, no.2.

[7] Adnan M & Alhajj R 2009, 'DRFP-tree: disk resident frequent pattern tree', Appl Intell., vol.30, pp. 84-97.

[8] Aggarwal CC & Yu PS 2005, 'On Variable Constraints in Privacy-Preserving Data Mining', SIAM Conference.

[9] Aggarwal G, Feder T, Kenthapadi K, Khuller S, Motwani R, Panigrahy R, Thomas D & Zhu A 2006, 'Achieving Anonymity via Clustering, ACM PODS Conference.

[10] Agrawal R & Srikant R 1994, 'Fast algorithms for mining association rules', in Proceedings ofthe 20th International Conference on Very Large Databases, Santiago, Chile, pp. 487-499.

[11] Agrawal R & Srikant R 2000, 'Privacy- preserving data mining', In Proc. of ACM SIGMOD'00, Dallas, Texas, USA, pp. 439- 450.

[12] Agrawal R, Imielinski T & Swami AN 1993, 'Mining association rules between sets of items in large databases', in Proceedings of the ACM SIGMOD.

Research Explorer - International Journal
ISSN : 2250-1940 (Print), 2349 - 1647 (Online)
Impact Factor : 2.014(IRJIF), 1.85(JIF), 1.056(RIF), 0.60(QF) 0.398(GIF)
Issue - 19 ; August 2018

SURVEY ON MANET: CHARACTERISTICS, CHALLENGE, APPLICATIONS, ROUTING PROTOCOLS AND SECURITY ATTACKS

Uma Maheswari.A
Lecturer in Departmet of Computer Science,
N.K.R.Govt. Arts College (w), Namakkal, India.

Dr. P.Suganthi
Assistant Professor,
Department of computer science,
N.K.R.Govt. Arts College (w), Namakkal, India.

Abstract

Mobile Ad-hoc network (MANET) is a collection of wireless mobile nodes. The nodes are acts as both router and host. Nodes are joined with one another using single-hop or multi-hop network and form a network and break these networks continuously. So this is called dynamic network. In this dynamic network, routing is the process of fault finding. Routing protocols use several metrics like path length, reliability, routing delay, band width, load, communication costs etc. to evaluate the best routing path out of several alternatives for routing data packets. MANETS are vulnerably natured to various kinds of security related attacks such as wormhole attacks, black hole attack, rushing attack etc.In this survey paper, enclosed with wireless network and its types, Characteristics, Challenges, Application, classification of routing protocols, and varies types of security attacks.

Keywords: MANET, Wireless Networks, Characteristics, Challenges, Applications, Routing protocols and Security Attacks.

1. Introduction

In recent trends wireless technology is spread in the world for access service and exchanging information without a cable connection between clients. Wireless network is an enhancing technology. These networks provide flexibility connection between the users at different geographical area. Wireless ad-hoc network is divided into three main types: Wireless Sensor Network (WSN), Wireless Mesh Network (WMN), and Mobile Ad-Hoc Network (MANET). A "WSN" is a communication network where sensors are used which acts as nodes. For example, Vehicular Ad-hoc Network (VANET), this uses sensor in vehicles. The sensor can predict the surrounding vehicles and notify the driver about them. This system is helpful in reducing accidents. A "WMN" is a communication network made up of radio nodes arranged in a mesh topology. For

example, the laptops in the One Laptop per child program use wireless mesh networking to enable students to exchange files and get on the internet even though they lack wired or cell phone or other physical connections in their area[1].

A MANET is a group of self- forming wireless mobile nodes which can interchange data in dynamic manner [2]. Each node has a wireless interface to communicate with each other. These networks are formed from node by node, where each node acts as a host as well as router at the same time. These networks are fully distributed and can work at any place without the help of any fixed infrastructure as access points or base station [3]. The mobile device or node can communicate directly with other device available in network or can use mediator to makes communication between source and destination. As MANET's topology is dynamic in nature, it is affected from lots of security threats. Security, Reliability, availability, scalability, quality of service is the requirements of MANET.

2. Wireless Networks

Wireless network is an enhancing new technology. These allow flexible connection to point to point network. The network can be expanded anywhere or building without a wired connection between the clients. Wireless networks are split into two kinds: Infrastructure networks and Ad-Hoc networks [4].

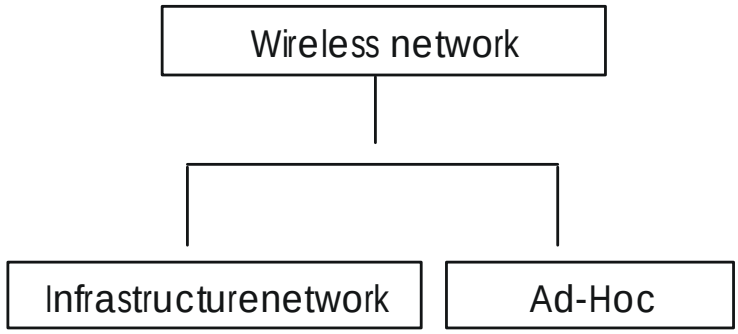


Fig.1.structure of wireless networks

A. Infrastructure Network

In infrastructure wireless networks has fixed base station (BS) and access points acts as a central coordinator between all nodes. Wireless devices communicated directly with electromagnetic waves to each other. These networks is composed by nodes and access point (AP) and these AP acts as a central coordinator between all nodes. If one node goes out of one base station, that node goes into the range of other base stations. So infrastructure network figure all devices are connected within infrastructure, its preplanned architecture network. Each device is connected within base station.

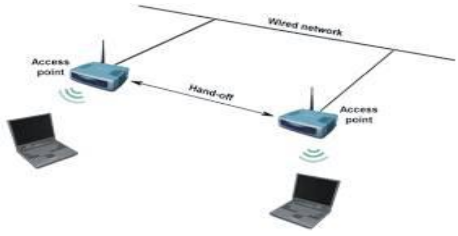
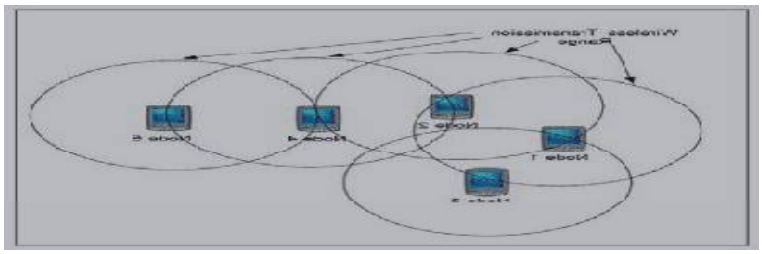


Fig. 2. Structure o

B.Infrastructure-less Networks or Ad-hoc Networks

Wireless Ad-Hoc network which is a decentralized type of wireless network. The network does not depend on a preexisting infrastructure, such as routers in wired networks or access points in infrastructure wireless network. These networks don't have a fixed topology or a central coordination point. Hence, the point to point network can communicate with each other by sending and receiving packets which is more sophisticated than infrastructure networks.

Fig. 3.Structure of MANET



C. Advantages of MANETs

- MANET provides services and information for user in irrespective of geographic position.
- MANET does not have base station and centralized control.
- It is a self-configuring network, where nodes act as routers.
- Less expensive compared to wired networks.
- It is scalable as nodes can be added to or removed from the network.
- The network can be established at any place and time.
- Robust due to decentralized administration.
- Advanced flexibility.

Disadvantages of MANETs

- Lack of physical security
- Inherent resources are limited.
- Mutual trust unsafe to attack. Insufficient of authorization services
- Dynamic network topology makes it difficult to identify malicious attack.

3. Manet Characteristics

There are varies characteristics of MANETsaregiven below [5], [6].

- Distributed Operations: There is no centralized authority and the control is distributed among the nodes. Each node must cooperate and communicate with one another. If required, the node may implement functions such as routing and security.
- Multi-Hop Routing: when a source node tries to send a packet to destination node, which is out of its communication range, the packets should be forwarded through intermediate nodes. Two types of ad-hoc routing i.e. single-hop and multi-hop routing. Single-hop is simpler than multi-hop regarding cost, structure and implementation.
- Autonomous Terminal: In MANET, each mobile node is an independent node. It can be perform two tasks that are both a host and a router, which finds and maintain the route.

- DynamicTopology: Nodes are free to move randomly in any direction with different speeds. So the network topology may change randomly at unpredictable time. The nodes in the MANET dynamically establish routing among themselves as they travel around, establishing their own network.
- Shared Physical Medium*:* The wireless communication medium is accessible to any entity with the appropriate equipment and adequate resources. There is no restriction to access the channel.
- Heterogeneity: MANET can be formed using variety of devices such as laptops, vehicles, ambulances, mobile phones etc.
- Light-Weight Terminals*:* The mobile node has less CPU capacity, low power storage and less memory size.

4. Manet Challenges

There are some main applications of MANET are given below

- Limited Bandwidth: Wireless link significantly continues to have lower capacity. In accounting for the effects of several times, fading sound and interference conditions are realized to be the throughput of wireless communication often much less than a radio's maximum transmission rate.
- Dynamic Topology: Dynamic topology membership may disturb the reliable relationship among nodes. The trust may also be disturbed if some nodes are detected as compromised
- Routing Overhead: In wireless ad-hoc networks, within a short time, a node changes their location several times over the network. So, some lacking freshness routes are generated in the routing table which leads to unnecessary routing overhead.
- Hidden Terminal Problem: In the MANET, the hidden terminal problem refers to the collision of packets at a receiving node because of the simultaneous transmission of those nodes that are not within the direct transmission range of the sender, but are within the transmission range of the receiver.
- Packet Losses Due To Transmission Errors: Ad-Hoc wireless networks experiences a much higher packet loss because of some factors such as increased collisions, the presence of hidden terminals, presence of interference, uni-directional links, frequent path breaks which are caused by mobility of nodes.
- Mobility-Induced Route Changes: The network topology in an Ad-Hoc wireless network is highly dynamic because of the movement of nodes. Hence the on-going session also suffers from regular path breaks. This underlying situation mostly leads to frequent route changes.
- Battery Constraints: MANET suffers from the shortcoming of the battery constraints. Devices which are used in these networks suffer restrictions on the power source in order to maintain portability, size and weight of the device.
- Security Threats: MANET suffers from a particular security problems. It also has the reliability problem due to the limited wireless transmission range. The nature of wireless mobile ad hoc (MANETs) brings about new security challenges to the network design. As the wireless medium is vulnerable to eavesdropping and Ad-Hoc network functionality is established through node co-operation, MANETs are intrinsically subjected to numerous security attacks.

5. Manet Applications

Here are some of the applications of MANETs are following

- Military Department*:* Ad-Hoc networking would allow the military to take advantage of common place network technology to maintain an information network between the soldiers, vehicles, and military information headquarters.
- LocalLevel: Ad-Hoc networks can autonomously link an instant and temporary multimedia network using notebook computers to spread and share information among participants for e.g. conference or classroom. Another appropriate local level application might be in home networks where devices can communicate directly to exchange information.
- Personal Area Network andBluetooth*:* A personal area network is a short range, localized network where nodes are usually associated with a given person. Short-range MANET such as Bluetooth can simplify the inter communication between various mobile devices such as a laptop, and a mobile phone.
- Collaborative Work*:* For some business environments, the need for collaborative computing might be more important outside office environments than inside and where people do need to have outside meetings to cooperate and exchange information on a given project.
- Business Sector*:* Ad-Hoc can be used in emergency/rescue operations for disaster relief efforts, e.g. in fire, flood, or earthquake. Emergency rescue operations must take place where non-existing or damaged communications infrastructure and rapid deployment of a communication network is needed.

6. Routing Protocols

Routing is the process of send a packet from source to destination. There are many problems are arrived on the network to find a correct path. Routing protocol consists of routing algorithms with a set of rules that monitors the operations of the network. It controls the flow of data in networks and also decides the efficient path to reach the destination. Ad-hoc network routing protocols commonly divided into three types; proactive, reactive, hybrid protocols.

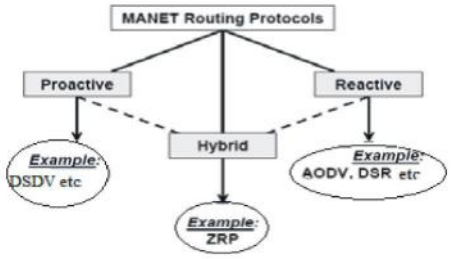


Fig. 4. Classification of routing protocols

A. Proactive Routing Protocol

Proactive routing protocol is also known as “table-driven routing protocol”. In table driven routing protocol as name indicates, each node has to maintain one or more tables to store the routing information. This table is maintained throughout of the network so as to maintain consistency of its view. As we know MANET having dynamic topology and it continuously changes all time, so each node has to update the table to maintain the consistency of network view. In proactive routing protocol, if there is no connection or communication between two nodes, still they have to maintain their table and its updation. Some types of proactive routing protocols are Destination

sequenced distance vector (DSDV), WRP, GSR, FSR. One of the most popular proactive routing protocols is Destination sequenced distance vector.

B. Reactive Routing Protocol

Reactive routing protocol is also known as "on-demand routing protocol". In reactive routing protocol, there is no need to maintain the routing table at each node. Instead, if any two nodes want to communicate with each other or any node want to send packet to another node, then this reactive routing protocol searches for the route and establish the connection between the nodes. After establishing a good connection, they can quickly transmit and receive data packets. Some types of reactive routing protocols are Ad-hoc On-demand Distance Vector (AODV), Dynamic Source Routing (DSR), and CBRP. One of the most popular reactive routing protocols is the Ad-Hoc On-Demand Vector (AODV) routing protocol. AODV is a reactive routing protocol, discovering routes only when they are needed. "It offers quick adaptation to dynamic link conditions, low processing and memory overhead, low network utilization, and determines unicast routes to destinations within ad hoc network".

C. Hybrid Routing Protocol

Hybrid routing protocol is a combination of proactive routing protocol and reactive routing protocol. In hybrid routing protocol, firstly node identify the route using the proactive routing protocol and then later uses reactive routing protocol. Depending upon the different network scenarios, both pro-active and reactive nature of the protocol can be used interchangeably. Some types of hybrid routing protocol are Zone Routing Protocol (ZRP), Zone-Based Hierarchical Link State (ZHLS). One of the most popular hybrid routing protocols is Zone Routing protocol. The basic operation of Zone routing protocol is uses proactive routing algorithm within the given zone and reactive routing algorithm outside the zone as defined by user.

7. Security Aspects In Manet

Security attacks: There are two types of attacks [7].

*A. Passive attacks***:** Passive attacks are silently dangerous attacks and does not change the data during the data transmission within the network and steal important information. The passive attackers don't destroy the operation of routing protocol but attempt to discover the important information from routed traffic. Some of the passive attacks are as shown below:

- Snooping: Snooping is one of the passive attacks whose goal is to track the process of point-to-point without altering any data. It includes observation of confidential emails or chats.
- Eavesdropping: Eavesdropping is another kind of attack that usually happens in the mobile Ad-Hoc networks. The goal of eavesdropping is to receive some confidential information that should be kept secret during the communication by simply monitoring it. The confidential information may include the location, public key, private key or even passwords.
- Traffic analysis: Traffic analysis is also a passive attack whose name itself indicate its work, it analyses the traffic pattern of the network. So the opponent steals confidential information about network topology simply by analyzing the traffic of network.

B. Active attacks

Active attacks generate unauthorized access to network that helps to attacker make changes such as message modification, denial of services, message fabrications, congestion etc. It breaks the normal function of the network. Internal and external are types of active attacks. Active external

attacks can be carried out by outside source that don't belong to the network. Active internal attacks are from malicious nodes which are part of the network. It is severe and hard to detect them. Active attacks are classified into four groups:

- Dropping Attacks*:* Compromised nodes or selfish nodes can drop all packets that are not destined for them. Dropping attacks can prevent end-to-end communications between nodes.
- Modification Attacks*:* These attacks modify packets and disrupt the overall communication between network nodes. Sinkhole attacks are the example of modification attacks.
- Fabrication Attacks*:* In fabrication attack, the attacker send fake message to the neighboring nodes without receiving any related message.

These attacks can occur in different layers of the network protocol stack.

- Attacks at Physical Layer**:** Some of the attacks identified at physical layer include eavesdropping, interference, and jamming etc.

1. Eavesdropping: It can also be defined as interception and reading of messages and conversations by unintended receivers. The main aim of such attacks is to obtain the confidential information that should be kept secret during the communication.
2. Jamming: Jamming is a special class of DoS attacks which are initiated by malicious node after determining the frequency of communication. Jamming attacks also prevents the reception of legitimate packets.
3. Active Interference: An active interference is a denial of service attack which blocks the wireless communication channel, or distorting communications.

- Attacks at Data link layer**:** The data link layer can classified attacks as to what effect it has on the state of the network as a whole.

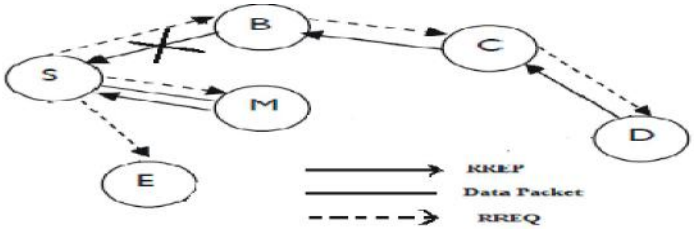
1. Selfish Misbehavior of Nodes: The selfish nodes may refuse to take part in the forwarding process or drops the packets intentionally in order to conserve the resources and to conserve of battery power.
2. Malicious Behavior of nodes**:** The main task of malicious node is to disrupt normal operation of routing protocol. The impact of such attack is increased when the communication takes place between neighboring nodes. Attacks of such type are fall into following categories.
3. Denial of Service (DoS): The prevention of authorized access to resources or the delaying of time-critical operations. A denial of service (DoS) attack is characterized by an attempt by an attacker to prevent legitimate users of a service from using the desired resources and attempts to "flood" a network, thereby preventing legitimate network traffic.
4. Misdirecting traffic: A malicious node advertises wrong routing information in order to get secure data before the actual route.
5. Attacking neighbor sensing protocols: malicious nodes advertise fake error messages so that important links interface are marked as broken.

- Attacks at Network Layer: The basic idea behind network layer attacks is to inject itself in the active path from source to destination or to absorb

1) Black hole Attack: In this type of attacks, malicious node claims having an optimum route to the node whenever it receives RREQ packets, and sends the REPP with highest destination sequence number and minimum hop count value to originator node .whose RREQ packets it

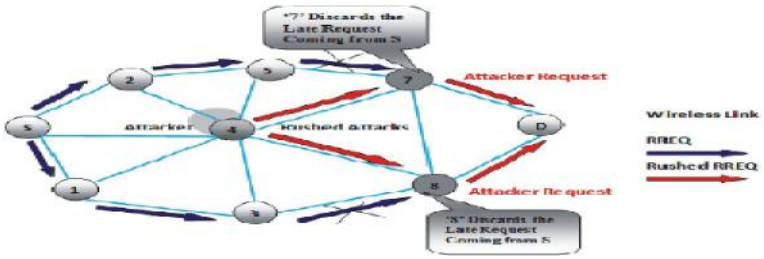
wants to intercept. For example, in fig 5, when node "S" wants to send data to destination node "D", it initiates the route discovery process. The malicious node "M" when receives the route request, it immediately sends response to source. If reply from node "M" reaches first to the source than the source node "S" ignores all other reply messages and begin to send packet via route node "M". As a result, all data packets are consumed or lost at malicious node.

Fig. 5.Blackhole Attack



2) Rushing Attack*:* In rushing attacks when compromised node receives a route request packet from the source node, it floods the packet quickly throughout the network before other nodes, which also receive the same route request packet For example, in fig 6 the node "4" represents the rushing attack node, where "S" and "D" refers to source and destination nodes. The rushing attack of compromised node "4" quickly broadcasts the route request messages to ensure that the RREQ message from itself arrive earlier than those from other nodes. This result in when neighboring node of "D" i.e. "7" and "8" when receive the actual (late) route request from source, they simply discard requests. So in the presence of such attacks "S" fails to discover any useable route or safe route without the involvement of attacker.

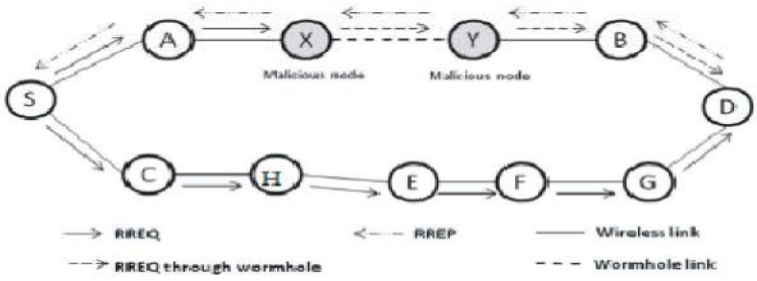
Fig.6. Rushing Attack



3) Wormhole Attack: In wormhole attack, malicious node receive data packet at one point in the network and tunnels them to another malicious node. The tunnel exist between two malicious nodes is referred to as a wormhole. For example in fig 7, the nodes "X" and "Y" are malicious node that forms the tunnel in network. The Originating node "S" when initiate the RREQ message to find the route to node "D" destination node. The immediate neighbor node of originating node "S", namely "A" and "C" forwards the RREQ message to their respective neighbors "H" and "X". The node "X" when receive the RREQ it immediately share with it "Y" and later it initiate RREQ to its neighbor node "B", through which the RREQ is delivered to the destination node "D". Due to high speed link, it forces the source node to select route <S-A-B-D> for destination. It results in "D" ignores RREQ that arrives at a later time and thus, invalidates the legitimate route <S-C-H-E-F-D>.

Fig.7. Wormhole Attack

4) Greyhole attack: In this type of attacks, malicious node claims having an optimum route to the node whose packets it wants to intercept. It is similar to black hole attack but it drops data packet of a particular node.



5) Sinkl : advertises wrong rout le network traffic. After receiving whole network traffic it modifies the secret information, such as changes made to data packet or drops them to make the network complicated. A malicious node tries to attract the secure data from all neighboring nodes.

- Attacks at Transport Layer

Session Hijacking: Attacker in session hijacking takes the advantage to exploits the unprotected session after its initial setup. In this attack, the attacker spoofs the victim node's IP address, finds the correct sequence number i.e. expected by the target and then launches various DoS attacks

- Attacks at Application Layer

Malicious code attacks: Malicious code attacks include, Viruses, Worms can attack both operating system and user application.

7. Conclusion

In this survey paper, we discussed about MANETs, wireless networks and its type, MANETsand its advantages and disadvantages, characteristics, challenges, and applications, routing protocols and types of security attacks. MANET is decentralized, self-organize, "anywhere anytime" network and provides flexible and inexpensive communications. The security is the main challenges in the networks and especially in the wireless network such as MANET and prevents routing protocols from different types of attacks. The security can be improved with the implementation of some security mechanisms.

8. References

[1] Ku. Vishakha V.Vyas,Nitin k.Bhil",Mobile Ad-Hoc Network (MANET) and its Security Aspects", International Journal of Advance Research in Computer Science and Management Studies, Volume 3, Issue 8, August 2015.

[2] Meenakshi Yadav, Nisha Uparosiya,"Survey on MANET: Routing Protocols, Advantages, Problems and Security", International Journal of Innovative Computer Science & Engineering, Volume 1 Issue 2; Page No. 12-17.

[3] Aarti, Dr. S. S. Tyagi, "Study of MANET: Characteristics, Challenges, Application and Security Attacks",International Journal of Advanced Research in Computer Science and Software Engineering,Volume 3, Issue 5, May 2013.

[4] Abdalftah Kaid Said Ali, Dr. U. V. Kulkarni, "Characteristics, Applications and Challenges inMobile Ad-Hoc Networks (MANET): Overview", *IPASJ International Journal of Electronics & Communication (IIJEC)*, Volume 3, Issue 12, December 2015.

[5] Dr. V. Harsha Shastri, V.Sreeprada," A study of Mobile Ad-hoc Network-Challenges, Characteristics, Applications and Routing", *International Journal of Engineering Trends and Technology, Volume 47 Number 8 May 2017.*

[6] Er. Jatinder Kaur, Er. Gurpreet Singh, "Review study on MANET Routing Protocols: Challenges and Applications", International Journal of Advanced Research in Computer Science, Volume 8, No. 4, May 2017.

[7] Anuj Rana, Sandeep Gupta, "Review on MANETs Characteristics, Challenges, Application and Security Attacks", International Journal of Science and Research, Volume4Issue2,February2015.

Research Explorer - International Journal
ISSN : 2250-1940 (Print), 2349 - 1647 (Online)
Impact Factor : 2.014(IRJIF), 1.85(JIF), 1.056(RIF), 0.60(QF) 0.398(GIF)
Issue - 19 ; August 2018

A FAULT STRENGTH ROUTING TECHNIQUE USING SECURITY WITH SECONDARY MOBILE IN MANET

R.Navamani
Department of Computer Science,
Trinity college for women.
Namakkal, TamilNadu, INDIA.

R.Kowsalya
Assistant Professor,
Department of Computer Science,
Trinity College for women,
Namakkal, TamilNadu, INDIA

Abstract

MANET has been gaining the popularity because of its ease of implementation. One of the major elements in MANET is routing protocol, which consists of two main protocols: proactive and reactive routing protocols .MANETs are active as well as incessantly altering arrangements, having cluster of nodules that are not interiorly managed. Conniving a direction-finding code of behavior for this kind of atmosphere is very demanding task. The circumstances contracts inferior once the defective nodes in the network, since they enhance statistics hammering and disgrace the recital of protocols. Infrastructures Secondary network is MANET which creates the temporary network. Performance and security are its two major issues. Due to its self organizing feature providing runtime network security is tedious task. So an efficient and strong model is required to setup so that various eavesdropping activity can be avoided. Key management is a vital part of security in Manet because the distribution of encryption keys in an authentication manner is a difficult task due to its dynamic nature. As every time nodes leaves or joins it has to regenerate a new session key for maintaining secrecy. In this paper, we have proposed a new key management scheme to improve the network security with Secondary mobility overhead and Secondary key distribution time .

Keywords: Routing protocol, Manet, certificate based cryptography , symmetric keys.

1 Introduction

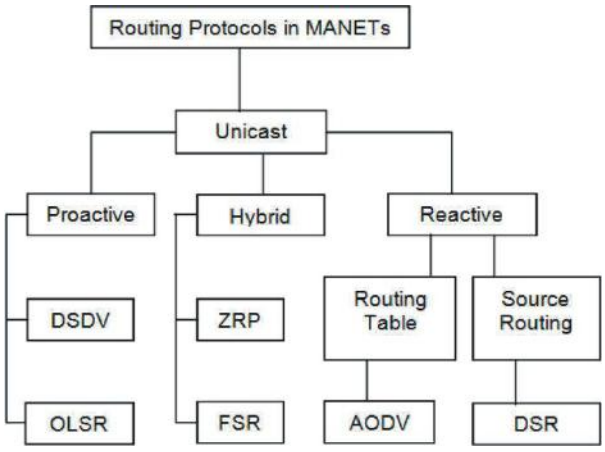
A mobile ad hoc network (MANET) is a self-configuring transportation less network of portable tools. These appliances are attached through wireless. Each apparatus in a MANET is open to move separately in any track and will consequently transform its links to other machines repeatedly These transportation less networks have no permanent routes, all nodes are proficient

of association and can be associated energetically in an random manner. Because MANETS are movable, they use wireless associations to bond to a variety of networks. MANETs are a sort of Wireless informal association that regularly has a routable networking background. outing protocol,Manet, certificate based cryptography , symmetric keys. In key management schemes different cryptographic keys method are used like symmetric keys, public keys or certificate based cryptography. In symmetric keys over MANET if n nodes wants to communicate k keys will be required where k will be the number of leys which should be generated by **k=n(n-1)/2.** In this approach both the sender and the receiver contains the same key for encryption and for decryption. In public key encryption two keys are used one private key and the other as public key. The private keys are used for encryption between the nodes whereas public keys are used for encryption. Mobile ad hoc network (MANET) is a group of mobile nodes that communicates with each other without any supporting infrastructure. In this paper, we have proposed a new key management scheme to improve the network security with less mobility overhead and less key distribution time .The key management schemes Section IV Proposed method Section V Results Section VI conclusion Key management deals with the secure generation, distribution, and storage of keys. It plays a vital role in computer security today as practical attacks on public-key systems are typically aimed at key management as opposed to the cryptographic algorithms themselves. This report will investigate the techniques used in the distribution of secret keys used to decrypt and encrypt messages with particular focus on the Diffie-Hellman distribution sheme.

1.1 Farm of Routing

Fig.1. Farm of Routing Protocol in MANET

2. **Secure Zone Routing Protocol**



For proposed design to be suitable for a MANET, the following design goals such as: Few computational steps to reserve the limited power ofØ all ad-hoc devices since too many computational steps will drain the battery. Balanced protocol, which means that all nodes shouldØ perform approximately the same number of heavily computations. Few packets flow with small size since large packetsØ are spitted into several packets to match the available communication bandwidth where sending many packets contradicts with the previous design goal. Restricted number of heavy computations, such asØ modular exponentiations, to save battery power although the processors of most ad-hoc devices are becoming more powerful and can perform these computations. A. SECURE NEIGHBOR DISCOVERY In wireless networks, each node needs to

know its neighbors to make routing decisions; it stores neighbor information in its routing table that contains the address of the neighbor, and the link state. In MANETs, nodes use neighbor discovery protocol to discover surrounding nodes they can directly communicate with across the wireless channel with signal propagation speed by considering the location or round trip information.

2.1. Secure Routing Packets

Once achieve secure information exchange, we can further secure the underlying routing protocol in wireless ad-hoc networks. Security services in MANETs belong to two kinds of massages: the routing massages and the data messages. Both have a different nature and different security needs. We focus here on securing routing because data messages are point-to point and can be protected with any point-to-point security system. On the other hand, routing messages are sent to intermediate neighbors, processed, possibly modified, and resent. Moreover, as a result of processing of routing message, a node might modify its routing table. This creates the need for both the end-to-end and the intermediate nodes to be able to authenticate the information contained in the routing messages.

Figure.1. System Architecture

2.2. **Implementation**

The SZRP is simulated for 35 nodes spread randomly inØ a 1200 * 1200m area network; transmission range for each node is random. Nodes are positioned randomly on the plane. NodesØ start its travel from a random location to a random direction with

3. Related Work

Hierarchical key management schemes would serve well for military applications where the organization of the network is already hierarchical in nature. Most of the existing key management schemes concentrate only on network structures and key allocation algorithms, ignoring attributes of the nodes themselves. Due to the distributed and dynamic nature of MANETs, it is possible to show that there is a security benefit to be attained when the node states are considered in the process of constructing a private key generator (PKG). In this paper, we propose a distributed hierarchical key management scheme in which nodes can get their keys updated either from their parent nodes or a threshold of sibling nodes. The dynamic node selection process is formulated as a stochastic problem and the proposed scheme can select the best nodes to be used as PKGs from all available ones considering their security conditions and energy states. Simulation results show that the proposed scheme can decrease network compromising probability and increase network lifetime in tactical MANETs. Key management in MANET is getting popularity for researchers. Shamir et all [1], has proposed ID based public key systems which uses user's identity for secure

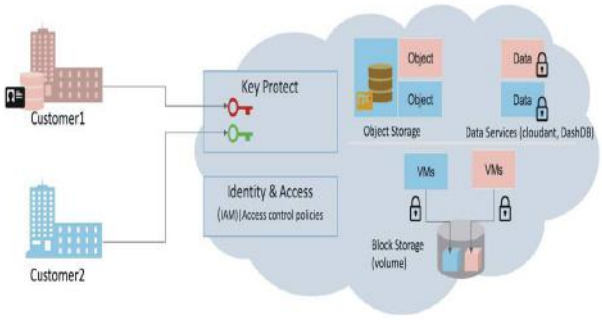
information transmission. ID based systems exchanges the public key certificates without keeping public key directory. This method needs a Private Key generator (PKG) to identify user id. Identity based key management schemes are further classified as

1 .Customary threshold schemes

2. Thesaurus identity based schemes

3. Surreptitious share as private key (SSPK)

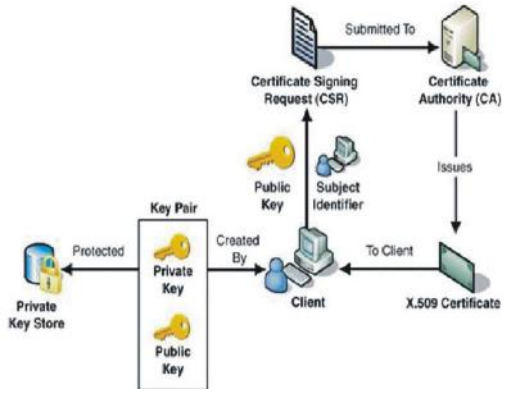
4. Certificates schemes



Working of IBM system in MANET

ID-based cryptosystems have many advantages over PKI based cryptosystems in key distribution, but they also have an inherent drawback of key escrow problem, i.e. users' private keys are known to the key generation center (KGC). Therefore secure key issuing (SKI) is Key management deals with the secure generation, distribution, and storage of keys. It plays a vital role in computer security today as practical attacks on public-key systems are typically aimed at key management as opposed to the cryptographic algorithms themselves. This report will investigate the techniques used in the distribution of secret keys used to decrypt and encryption messages with particular focus on the Diffie-Hellman distribution important issue in ID-based.

Certificate Based cryptography



4. Key Management Scheme

Various key management scheme has been proposed using the number of distribution procedures. Various Symmetric key management schemes like Key Infection , Peer intermediate key

establishment. Some of the Asymmetric key management schemes are secure routing protocol , Ubiquitous and robust Access but these schemes includes the parameters like

☐ **Growing Security** – Reducing small calculations will consume less computation node power to improve network security.

☐ **Mouting Mobility** - computational procedures can be reduced by decreasing the allocation of resources to extend mobility.

☐ **Plummeting generation time**- Network quality can be improved if key generation time can be reduced.

☐ **Reducing Power-** Due to the battery depended network , power conservation is important to improve the network consistency.

4.1 Proposed Key Management Scheme

Our proposed scheme consists of following tasks

1. Subtraction of misbehaviour Node:

When system identified a Cluster head is misbehaving. Head of the cluster will be removed from table

2. Key generation: Based on CRT technique, after removal of misbehavior node from table list the key generation will be applicable by the algorithm and it will be allowed to a node having a largest ID according to its time allocation in a network, then all the members should be updated by this new head node.

3. **Come together head verification:** When messages of key generation are received by other member nodes the details about the Node ID will be recorded in its table. And also it is ensuring that no further interaction should be done with the node, which will implement a secured communication.

4. **Key Generation and Management Schemes:** Key Generation and Calculation of pair wise prime keys will be generated by calling a function. Key generation computation can be calculated in pairs, so that generation time

5. Performance Factors

Various important factors on which performance of cryptographic algorithms depend are:

5.1 **Tenability**

It could be very desirable to be able to dynamically define the encrypted part and the encryption parameters with respect to different applications and requirements. Static definition of encrypted part and encrypted parameters limits the usability of the scheme to a restricted set of applications.

5.2 Computational Speed

In many real-time applications, it is important that the encryption and decryption algorithms are fast enough to meet real time requirements.

5.3 Key Length Value

In the encryption methodologies the key management is the important aspect that shows how the data is encrypted. The image loss the encryption ratio is based on this key length. The symmetric algorithm uses a variable key length which is of the longer. Hence, the key management is a considerable aspect in encryption processing.

5.4 Encryption Ratio

The encryption ratio is the measure of the amount of data that is to be encrypted. Encryption ratio should be minimized to reduce the complexity on computation [8].

5.5 Security Issues

Cryptographic security defines whether encryption scheme is secure against brute force and

Research Explorer - International Journal
ISSN : 2250-1940 (Print), 2349 - 1647 (Online)
Impact Factor : 2.014(IRJIF), 1.85(JIF), 1.056(RIF), 0.60(QF) 0.398(GIF)
Issue - 19 ; August 2018

PERFORMANCE ANALYSIS OF ROUTING PROTOCOLS IN VEHICULAR ADHOC NETWORKS

S.Priya
Assistant Professor,
Department of Computer Science,
Trinity College for Women, Namakkal, TamilNadu, INDIA

Abstract

Vehicular connectivity can be considered as the coming generations' killer application. The vehicle-to-vehicle connectivity is immense for the improvement of Intelligent Transportation System (ITS) in order to provide a wide spectrum of applications, including safety-related, traffic control, and entertainment. Some important characteristics that distinguish vehicular ad hoc networks (VANETs) from mobile ad hoc networks are highly dynamic topology and fast movement of nodes. Hence, the protocols designed for mobile ad hoc networks (MANETs) cannot be simply adopted in vehicular ad hoc networks (VANETs). In this paper simulation based experiments have been performed to analyze the performance of Destination Sequenced Distance Vector (DSDV), Optimal Link State Routing (OLSR) and Adhoc On Demand Distance Vector (AODV) routing protocols on the basis of Packet Delivery Ratio, End to End delay and Average Throughput for TCP applications by Network Simulator (NS-3.19) using mobility model generated through Simulation of Urban Mobility (SUMO) tool. The results are compared by varying number of nodes. The analysis shows that OLSR routing protocol outperforms the other routing protocols AODV and DSDV and best to be used in vehicular ad hoc network for TCP applications.

Keywords—VANETs,NS-3.19,AODV,DSDV,OLSR,TCP.

1. INTRODUCTION

Vehicular Ad hoc Networks (VANET) is the subclass of Mobile Ad Hoc Networks (MANETs) [1]. VANET is the wireless network in which communication takes place through wireless links mounted on each node (vehicle). Each node within VANET act as both, the participant and router of the network as the nodes communicates through other intermediate node that lies within their own transmission range. VANET are self organizing network. It does not rely on any fixed network infrastructure. Although some fixed nodes act as the roadside units to facilitate the vehicular networks for serving geographical data or a gateway to internet etc. Some important characteristics that distinguish VANETs from other types of ad hoc networks and makes routing a challenging task are highly dynamic topology and fast mobility [2]. Several other factors such as road layout

and different environments such as city and highway makes routing more challenging in VANETs. VANETs focuses on the improvement of Intelligent Transportation System (ITS) in order to provide a wide spectrum of applications, including safety-related, traffic and fleet control, and entertainment [3].

The routing protocols for VANETs are broadly categorized as topology based protocols and position based protocols. Topology based routing protocol uses link's information stored in the routing table as a basis to forward packets from source node to destination node. Position or geographic routing protocol uses physical/geographical positions rather than the network address of nodes to perform data routing from source node to destination node assuming each node has knowledge of physical/ geographic position by GPS or by some other position determining services.

AODV routing protocol is a reactive routing protocol [15]. When a source node needs a route to some destination node, it broadcasts a route request message to its neighbor including the last known sequence number for that destination. Each node that forwards the route request creates a reverse route for itself back to the source node. When the route request reaches a node with a route to destination node that node generates a route reply that contains the number of hops necessary to reach destination and the sequence number for destination most recently seen by the node generating the reply. The state created in each node along the path from source to the destination is hop-by-hop state; that is each node remembers only the next hop and not the entire route, as would be done in source routing.

DSDV is an extension of classical bellman ford routing mechanism [16]. In DSDV each node maintains consistent network view via periodic routing updates. Routing information is stored inside routing tables maintained by each node. Each routing table contains information about all destinations, as well as the total number of hops needed to reach these nodes, and each entry in the routing table is labeled with a sequence number initiated by the destination node.

OLSR is a table driven, proactive protocol based on the concept of Multi point Relays (MPR) [17]. The MPR set is selected such that it covers all nodes that are two hops away in the network. Each node has the knowledge as to for which node it acts as a MRP. OLSR protocol implement the link state strategy; which uses hello and topology control (TC) messages to discover and then disseminate link state information throughout the network. In route calculation, the MPRs are used to form the route from a given node to any destination in the network.

Rest of the paper is organized as follows: Section 2 presents related work, while section 3 presents simulation methodology along with various tools which are used to carry out the work and, in Section 4, simulation results with analysis are given. Finally, in section 5, conclusions are drawn from results obtained through simulations conducted and outline of the future works.

2. Related work

Several studies have been published explaining the various protocols which are used for routing in VANETs based on different classes [4] [5] [6]. Basically, routing protocols can be classified in two broad categories: topology based and position based routing protocols.

The validation of a routing protocol relies almost exclusively on simulation. Mobility model has significant effects on the simulation results of any protocol for VANETs [7]. These models must describe vehicular mobility in a realistic way, including the peculiar aspects of vehicular traffic such as vehicles acceleration and deceleration in the presence of nearby vehicles, queuing at roads intersections, impact of traffic lights, and traffic jams. These models affect the results for analysis of routing protocols in VANETs [8]. Vehicular traffic simulators have to consider both

microscopic and macroscopic parameters as in generic traffic simulators Vanet Mobi Sim, SUMO etc.

Several papers have been published using different mobility models and different traffic scenarios for comparing the performance of various routing protocols using different performance metrics. The relative performance of on demand routing protocols AODV and DSR with identical loads is evaluated by Shastri et al. [9]. Increase in the pause times increase both Packet Loss and Average End-to-End Delay but decrease the Packet Delivery Fraction is evaluated through simulation. The performance of AODV, DSR, FSR and TORA in typical freeway traffic scenarios is examined by Jaap et al. [10]. Evaluation showed that AODV achieved the best performance throughout the traffic scenarios, followed by FSR and DSR. TORA is inapplicable for VANET environments. Husain et al [11]. Compared AODV, DSR, and LAR in city and highway scenarios for VANETs and concluded that position based routing protocol (LAR) outperforms topology based routing protocols (DSR and AODV) in terms of packet delivery ratio, throughput, and end-to-end delay for both the vehicular traffic scenarios.

There are many papers focusing on new routing protocols for VANETs in order to achieve increased communication reliability for a particular application [12] [13] [14].

3. Proposed work and methodology

A. Performance Metrics

In this paper, we evaluate and compare the performance of routing protocols AODV, DSDV and OLSR in terms of throughput, Packet Delivery Ratio (PDR), end-to-end delay and Packet Loss Ratio (PLR).

Throughput: It is the number of bits delivered successfully per second through a network to the destination. It is represented in kilo bits per second (kbps).

Throughput =

The greater value of throughput means the better performance of protocol.

Packet Delivery Ratio (PDR): It is defined as the ratio of the number of successfully delivered data packets to the total number of packets delivered.

PDR =

The greater value of packet delivery ratio means the better performance of protocol.

End-to-End Delay: The average time taken by a data packet to be transmitted across a network from source to destination. It also includes the delay caused by route discovery process and the queue in data packet transmission. Only the data packets that successfully delivered to destinations are counted. The lower the value of end to end delay means the better performance of protocol.

Packet Loss Ratio (PLR): Packet loss is the ratio of the number of packets dropped during simulation to the total number of packets delivered.

PLR =

The lower value of packet loss ratio means the better the performance of the protocol.

B. Simulation Parameters

To evaluate the performance of routing protocols by taking a real world scenario into

consideration with all the vehicular constraints we use a network simulator NS-3.19 [18].For simulation purpose we used mobility model generated as explained below. In our simulation, we varied the node density with constant pause time of 10s. Initially 50s are used as simulation setup time after which nodes start sending data packets. The network parameters we have used for our simulation purpose shown in the Table I.

Simulation Parameters	Value
Network Simulator	NS-3.19
Routing Protocols	AODV, DSDV, OLSR
Simulation Area	1700m x 1000m
Number of Nodes	12, 24, 36, 48, 60
Propagation Loss Model	Friis Propagation Loss Model
Propagation Delay Model	Constant Speed Propagation Delay Model
MAC Protocol	IEEE802.11
Simulation Time	300 sec.
Connection Type	TCP
Packet Size	512
Data Rate	1 Mbps
Pause Time	10 sec.

C. Mobility Model Parameters

To model realistic vehicular motion patterns, we used SUMO traffic simulator [19]. The steps to implement a VANET mobility model, traffic simulation will be as follows:

- Generation of Road Map
- Creation of Vehicular Traffic flow on this Map with varying speeds and of different class.
- Traffic Simulation

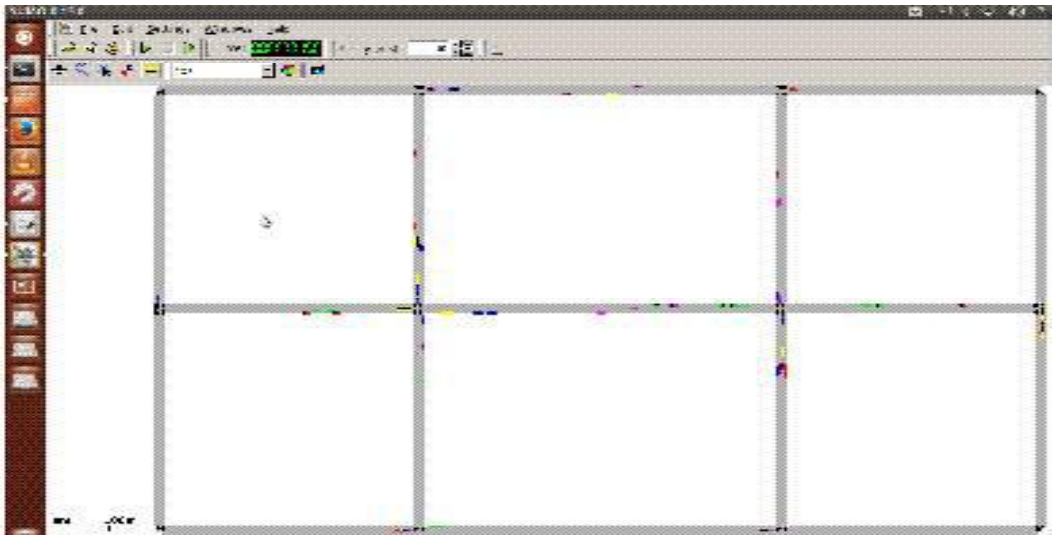
In our mobility model, we considered a network of size (1700m x 1000m) that includes four origins, four destinations with all signalized intersections. In the investigated area each road has three lanes. For traffic demand there are five vehicle types (Cars A, B, C and D and Bus) are considered. All drivers are 50% perfect in driving is assumed. The corresponding information is listed in Table II.

Table II. Vehicle types and vehicular characteristics used in the mobility model.

Vehicle Type	Max. Acceleration (m/sec^2)	Max. Deceleration (m/sec^2)	Length (m)	Max. Speed (m/sec)	Color
Car A	1.0	3.0	5	6	Red
Car B	2.0	3.0	5	10	Green
Car C	2.0	3.0	6.5	10	Blue
Car D	3.0	2.0	6.0	14	Magenta
Bus	1.0	1.5	8.5	8	Yellow

observed and the possible bottlenecks can be visually identified. Road map showing the position of nodes during simulation is shown in Figure 1.

Fig.1. Road map used during simulation of mobility model.

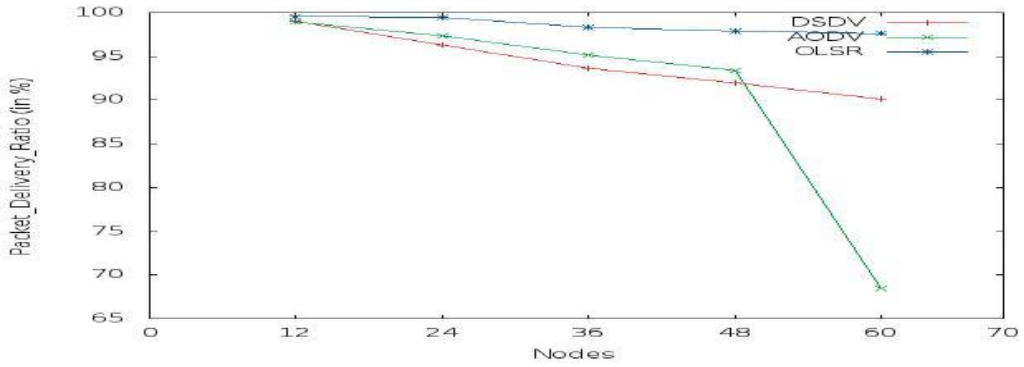


4. Simulation Results and performance analysis

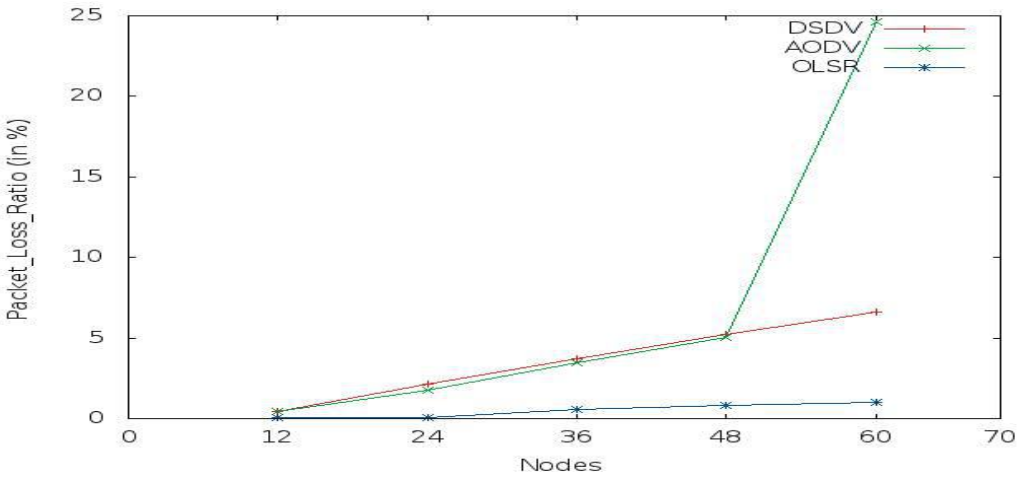
In this work the performance analysis is carried out for a vehicular ad hoc network by varying parameter i.e. number of nodes while keeping other network parameters constant. Three protocols i.e. AODV, DSDV and OLSR are taken for comparison of performance using performance metrics defined above for TCP connection.

In Figure 2, the packet delivery ratio (in %) during simulations time versus the number of nodes is shown. The OLSR outperforms AODV and DSDV protocols. As the number of nodes increases i.e. 12, 24, 36, 48, 60 the packet delivery ratio in all protocols decreases. On demand protocol AODV performs particularly well, delivering around 95% of the data packets up to 48 nodes and as number of nodes becomes 60, AODV packet delivery ratio decreases steeply.

Fig.2. Effect of node density on Packet Delivery Ratio.



The packet delivery ratio decreases as the node density increases due to the increased collision caused by higher node density especially at the intersections. However, as node density increases connectivity increases, but it also increases the number of collisions and packet loss (in %) as



shown in figure 3. OLSR has the lowest packet loss ratio.

Fig.3. Effect of node density on Packet Loss Ratio.

Figure 4 shows the variation of the average end to end delay (in sec.) by varying node density. Average end to end delay increases with increasing the number of nodes for all protocols. DSDV consistently presents the highest delay. But AODV has lowest end to end delay up to 48 number of nodes as node density becomes 60, AODV end to end delay goes around 2 sec. However, the delay of DSDV and OLSR is large because due to high mobility topology changes frequently which results in the increase number of exchange of routing tables and packets have to stay for more time in the buffers during transit. AODV causes large delays at high node density as route failure causes redundant broadcasts without control due to packet collision which leads to broadcast storm problem.

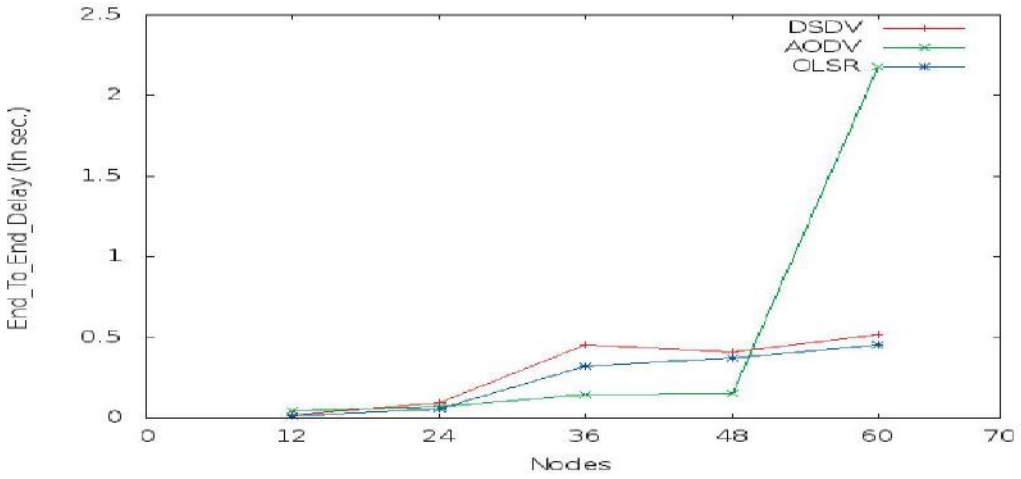
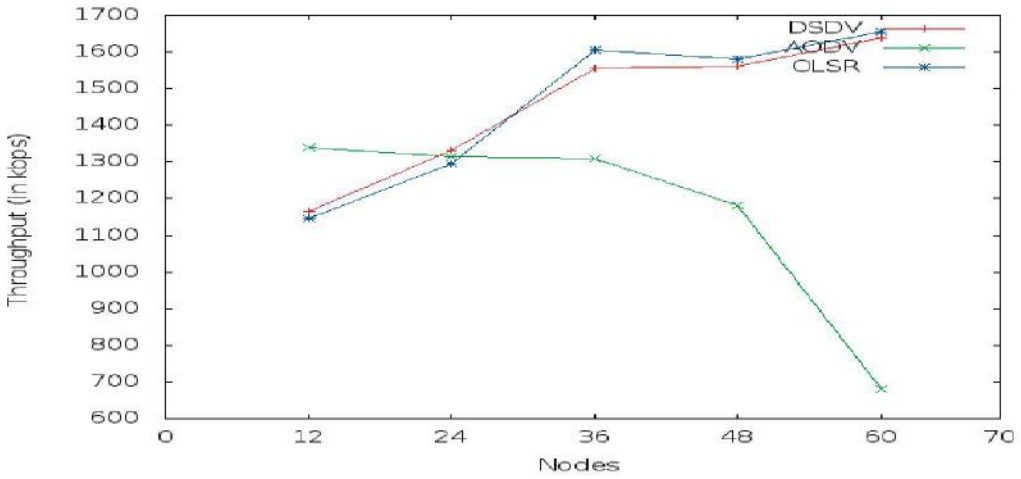


Fig.4. Effect of node density on End to End Delay.

Figure 5 shows Throughput (in Kbps) of protocols with varying number of nodes. The AODV has lowest throughput in comparison with all the other three protocols considered. Throughput of

DSDV and OLSR is almost equal. However throughput of OLSR increases slightly with increase in node density. DSDV and OLSR shows higher throughput since their routing overhead is more.

Fig.5. Effect of node density on Throughput.



5. Conclusion

In this paper, performance of three routing protocols AODV, DSDV and OLSR were evaluated for vehicular ad hoc networks by varying node density for TCP connection against various metrics. SUMO tool has been used to generate realistic mobility patterns and protocols were simulated using the tool NS-3.19. The results show that on demand routing protocol AODV outperforms table driven routing protocols DSDV and OLSR in terms of packet delivery ratio and end-to-end delay. AODV also uses fewer resources than OLSR and DSDV, because the control messages size is kept small requiring less bandwidth for maintaining the routes and the route table is kept small reducing the computational power. So, AODV is best to be used in real time applications such as in vehicular ad hoc network.

For future work, we can study the impact of other parameters on the performance of these routing protocols to understand these protocols in more deep. Future work will also include the evaluation of position based routing protocols as they are more suitable in vehicular traffic environment.

References

[1] V. Cabrera, F. J. Ros, and P. M. Ruiz, “Simulation-based study of common issues in vanet routing protocols,” in Vehicular Technology Conference, 2009. VTC Spring 2009. IEEE 69th, 2009, pp. 1–5.

[2] Y. Liu, J. Bi, and J. Yang, “Research on vehicular ad hoc networks,” in Control and Decision Conference, 2009. CCDC’09. Chinese, 2009, pp. 4430–4435.

[3] E. Schoch, F. Kargl, M. Weber, and T. Leinmuller, “Communication patterns in VANETs,” Communications Magazine, IEEE, vol. 46, no. 11, pp. 119–125, 2008.

[4] M. Altayeb and I. Mahgoub, “A Survey of Vehicular Ad hoc Networks Routing Protocols,” International Journal of Innovation and Applied Studies, vol. 3, no. 3, pp. 829–846, 2013.

[5] R. Kumar and M. Dave, “A Comparative Study of Various Routing Protocols in VANET.,” International Journal of Computer Science Issues (IJCSI), vol. 8, no. 4, 2011.

[6] R. Karimi, N. Ithnin, S. A. Razak, and S. Najafzadeh, “Non DTN Geographic Routing Protocols for Vehicular Ad Hoc Networks.,” International Journal of Computer Science Issues (IJCSI), vol. 8, no. 5, 2011.

[7] D. Djenouri, W. Soualhi, and E. Nekka, “VANET’s mobility models and overtaking: an overview,” in Information and Communication Technologies: From Theory to Applications, 2008. ICTTA 2008. 3rd International Conference on, 2008, pp. 1–6.

[8] F. K. Karnadi, Z. H. Mo, and K. Lan, “Rapid generation of realistic mobility models for VANET,” in Wireless Communications and Networking Conference, 2007. WCNC 2007. IEEE, 2007, pp. 2506–2511.

Research Explorer - International Journal
ISSN : 2250-1940 (Print), 2349 - 1647 (Online)
Impact Factor : 2.014(IRJIF), 1.85(JIF), 1.056(RIF), 0.60(QF) 0.398(GIF)
Issue - 19 ; August 2018 © Author

AN ANALYSIS OF POSSIBLE STRATEGIES FOR SUCCESSFUL CRM IN RETAIL INDUSTRY USING DATA MINING TECHNIQUES

S.Usharani
Assistant Professor,
Department of Computer Science

M.Sabitha
Research Scholar
Department of computer science,
Trinity College for women,
Namakkal, Tamilnadu, INDIA.

Abstract

Data mining is proved to be one of the important tools for identifying useful information from very large amount of data bases in almost all the industries. Customer Relationship Management (CRM) has become very important factor in defining the enterprise growth because of economic globalization and rapid development in e-commerce. The retail industry is realizing that it is possible to gain competitive advantage utilizing data mining. Retailers goals/objectives and strategies focused on marketing, customer service, understanding customers through data analysis and increasing acquisition and retention through customer loyalty programs. Retail industry is realizing that it is possible to gain competitive advantage deploys data mining in CRM. The CRM have an efficient and rapid response to customer requirements, by integrating or combining CRM and Data Mining Techniques. Retail industry is looking strategy where in they can target right customers, may be profitable to them. Retailers are seeking competitive advantages by better managing customer relations through database management.

Keywords-Data Mining, CRM, Association, Classification, Apriori, Naive Bayes.

1. Introduction

Data mining is "The process of discovering patterns in large data sets involving methods at the intersection of machine learning, statistics, and database systems". Data mining tools allow enterprises to predict future trends. **Customer relationship management** is "The process of acquiring, retaining and growing profitable customer which requires a clear focus on service attributes that represent value to the customer and creates loyalty". The data mining have been proposed into new areas such as medicine, banking, retail, and insurance etc... Data mining technology consolidate retail data, analyses and distribute data to users capture data across multiple

retail channels, create "one view" of the customer. Widely using techniques in data mining include classification and association.CRM methodologies and tools help the business in order to maintain the customer relationships within a proper manner.

2. Role Of Crm In Retail Business

- **Customer history**

  CRM stores all the customer information and profile like their last purchase, business cards and phone numbers. Customer services as know that today and the technologies helping to improve it have come a long way. With the invention and the evolution of the telephone customer service has become much easier.

- **Segmenting**

  Collecting all the information about the customers allows it to put the customers in different section can attend them accordingly. They can use such information to segment market and customize approach to each group of customers The effect of segmentation based on CRM data is to adjust the retail strategies to better suit the customers.

- **Promotions**

  The data gathered within a CRM system not only target a market segment with promotions that appeal to its members, but also to target individual customers. This will helps to provide them better service. When a customer visits the website accordingly look into what are looking for and include the promotion of that particular product in their newsletter.

- **Tracking**

  It is important that clearly comprehend what retail tracking service captures, know the strengths and limitations and understand what the metrics really mean. Which customers have proven to be loyal which have not? So it can provide them better service and at times reward as the loyal customers to keep up their loyalty and get more buyers indirectly.

- **Purchase tracking**

  CRM allows tracking each customer purchase separately know their interests and if their product had any issue or any damage. In case of damage or issue will be able to provide them with the same item in lower price in their next purchase or even give free service, etc to provide them to gain more customer satisfaction.

- **Loyalty**

  Customer loyalty is when a supplier receives the ultimate reward of efforts in interacting with customer. This way can have better business as loyalty will be committed customers which will set the level of sale and profit in the right track.

- **Cost effective**

  It allows managing the customers in the most cost effective way. It's great for managing sales pipeline, contacts, notes, emails and activities in one place so easily share information within the organization.

- **New buyers**

  CRM does not only manage the old customers or existing customers for the organization. It also has an intelligence that helps to identify potential buyers and convert them into leads which can turn into customers. CRM can help to get their attention by identifying them for sales of department people can then go ahead and deliver their interests to them from the business side with sophistication so they show up at door soon.

3. How Does A Crm Integration System Help Retail Business

A CRM Integration system can combine several systems to allow a single view. Data can be integrated from consumer lifestyle, expenditure, and brand choice. If the CRM system is implemented to track marketing strategies over products, services, then it can provide a scientific, data-based approach to marketing and advertising analysis. The CRM integration system improves the overall efficiency of marketing campaigns since it allows retail companies to specifically target the right group of consumers. The right message, at the right time to the right people delivers a positive marketing response from consumers and translates to more sales. Overall, this system provides a clear picture of the consumer segments, allowing retail companies to develop suitable business strategies, formulate appropriate marketing plans for their products or services, and anticipate a change of business landscape.

4. Classifying Indian Retail

Store Based: Store based formats can be further classified into two formats based on the basis of Ownership or Merchandise offered.

Non Store Based Classification: Non Store retail organizations focus on establishing direct contact with the consumer. This may be both personal (direct personal selling) and non personal TV, the Internet, mail, catalog or phone).

Service Based Classification: Such retailers specialize in providing different kinds of services to the end consumer. The services can be classified as Banking Services, Rentals, Electricity, cooking gas, etc. Various factors like quality of service, how much customization can be provided for meeting the client specific requirements, the uniqueness of the service and delivery within the timelines, usage of innovative technology, etc, are given importance for determining the success of service.

5. Crm Process And System

CRM process involves the activities and strategies that companies use to manage their interaction with current and potential customers. The keyword here is “relationship.” Nurture that with existing customers and they’ll be coming back, which means you achieve customer retention and loyalty.

Steps to a successful CRM Implementation

- **Step** 1: Identify why you need **RM** and what you expect from it. ...
- **Step** 2: Find a suitable **CRM** vendor for your organization. ...
- **Step** 3: Develop a budget. ...
- **Step** 4: Identify which departments and staff will handle the **process** and train them accordingly.

6. Crm Life Cycle

7. Data Mining Applications

CRM

Customer Relationship Management is all about acquiring and retaining customers, also improving customers’ loyalty and implementing customer focused strategies. To maintain a proper relationship with a customer a business need to collect data and analyses the information. This is where data mining plays its part. With data mining technologies the collected data can be used for analysis. Instead of being confused where to focus to retain customer, the seekers for the solution get filtered results.

Retail

Retailers segment customers into 'Regency, Frequency, Monetary' (RFM) groups and target marketing and promotions to those different groups. A customer spends little but often and last did so recently will be handled differently to a customer spent big but only once, and also some time ago. The former may receive a loyalty, up-sell and cross-sell offers, whereas the latter may be offered a win-back deal, for instance.

E-Commerce

Perhaps some of the most well -known examples of Data Mining and Analytics come from E-commerce sites. Many E-commerce companies use Data Mining and Business Intelligence to offer cross-sells and up-sells through their websites. One of the most famous of these is, of course, Amazon user use sophisticated mining techniques to drive there, 'People viewed that product, also liked this' functionality.

Customer segmentation

Traditional market research may help us to segment customers but data mining goes in deep and increases market effectiveness. Data mining aids in aligning the customers into a distinct segment and can tailor the needs according to the customers. Market is always about retaining the customers. Data mining allows finding a segment of customers based on vulnerability and the business could offer them with special offers and enhance satisfaction.

Fraud detection

Billions of dollars have been lost to the action of frauds. Traditional methods of fraud detection are time consuming and complex. Data mining aids in providing meaningful patterns and turning data into information. Any information that is valid and useful is knowledge. A perfect fraud detection system should protect information of all the users. A supervised method includes collection of sample records. These records are classified fraudulent or non-fraudulent. A model is built using this data and the algorithm is made to identify whether the record is fraudulent or not.

8. Data Mining Techiniques

Association

Association is one of the best-known data mining techniques. In association, a pattern is discovered based on a relationship between items in the same transaction. It refers to the method that can help to identify some interesting relations (dependency modeling) between different variables in large database. It discovers the hidden patterns in the datasets which are used to identify the variables and the frequent occurrence of different variables that appear with the highest frequencies. This technique follows two step processes,

- Find all the frequently occurring data sets
- Create strong association rules from the frequent data sets

Aprori algorithm

Apriori algorithm is a classical algorithm in data mining. It is used for mining frequent itemsets and relevant association rules. It is devised to operate on a database containing a lot of transactions, for instance, items brought by customers in a store.

How Apriori Works

1) Find all frequent itemsets

- Get frequent items:
- Items whose occurrence in database is greater than or equal to the min.support threshold.
- Get frequent itemsets:
- Generate candidates from frequent items
- Prune the results to find the frequent itemsets.

2) Generate strong association rules from frequent itemsets:

- Rules which satisfy the min.support and min.confidence threshold.

Support: The support supp(X) of an itemset X is defined as the proportion of transactions in the data set which contain the itemset.

Supp(X) = no. of transactions which contain the itemset X / total no. of transactions

Confidence: Confidence of a rule is defined:

Conf (**X'!Y**) = supp (**XUY**)/supp(**X**)

Example for Apriori Algorithms

They are five transaction and six items to find association rules with 50% support and 75% confidence. The transaction given below in table.

First find L_1. Since having only a small number of items, (i.e) Chocolate 4 times, bread 3 times, juice 4 times, milk 3 times, cheese 2 times and eggs only one once. Require 50% support and therefore each frequent item must appear in at least three transactions. Therefore L_1 is given by items in table.

The candidate 2-itemsets or c_2 therefore has six pairs. These pairs and their frequencies are given in table

Therefore have only two frequent item pairs which are {Chocolate, Juice} and {Bread, Milk}. This is L2. From these two frequent 2-itemsets, do not obtain a candidate 3-itemset since don't have two 2-itemsets that have the same first item.

The two frequent 2-itemsets above lead to the following possible rules:

The confidence of these rules is obtained by dividing the support for both items in the rule by the support for the item on the left-hand side. The confidence of the four rules therefore are 3/ 4=75%, 3/4=75%, 3/3=100%, 3/4=75% respectively. Since all of them have a minimum75% confidence, they all quality.

Classification

Classification is a classic data mining technique based on machine learning. Basically, classification is used to classify each item in a set of data into one of a predefined set of classes or groups. This analysis is used to retrieve important and relevant information above data, and metadata. It is used to classify different data in different classes. This technique is closely related to cluster analysis technique and it uses decision tree or neural network system. There are two main processes involved in this technique,

- Learning- In this process the data are analyzed by classification algorithm
- Classification- In this process the data is used to measure precision of the classification rules.

Naive bayes

The Naive Bayesian classifier is based on Bayes' theorem with the independence assumptions between predictors. A Naive Bayesian model is easy to build, with no complicated iterative parameter estimation which makes it particularly useful for very large datasets. Despite its simplicity, the Naive Bayesian classifier often does surprisingly well and is widely used because it often outperforms more sophisticated classification methods.

Algorithm

Bayes theorem provides a way of calculating the posterior probability, $P(c|x)$, from $P(c)$, $P(x)$, and $P(x|c)$. Naive Bayes classifier assumes that the effect of the value of a predictor (x) on a given class (c) is independent of the values of other predictors. This assumption is called class conditional independence.

- $P(c|x)$ is the posterior probability of *class* (*target*) given *predictor* (*attribute*).
- $P(c)$ is the prior probability of *class*.
- $P(x|c)$ is the likelihood which is the probability of *predictor* given *class*.
- $P(x)$ is the prior probability of *predictor*.

Example for Naive bayes classification

Predicting a class label using naive Bayesian classification: To predict the class label of an unknown sample using naïve Bayesian classification, given the same training data as for decision tree induction. The data samples are described by the attributes age, income, student, and credit-rating. The class label attribute, buys-computer, has two distinct values (namely, {yes, no}). C1 correspond to the class buys-computer= "yes" and C2 correspond to buys-computer="no". The unknown sample to classify is

X=(age= "<30", income= "medium", student= "yes", credit-rating= "fair").

To maximize $p(X|C_i)p(C_i)$, for i=1,2. $P(C_i)$,j the prior probability of each class, can be computed based on the training samples:

P (buys-computer= "yes") =9/14=0.643

P (buys-computer= "no") =5/15=0.357

To compute $p(X|C_i)$, for i=1, 2, compute the following conditional probabilities:

P(age="<30"|buys-computer="yes") =2/9=0.222

P (age="<30"|buys-computer= "no") =3/4=0.600

P(income="medium"|buys-computer="yes") =4/9=0.444

P(income="medium"|buys-computer="no") =2/5=0.400

P(student="yes"|buys-computer="yes") =6/9=0.667

P(student="yes"|buys-computer="no") =1/5=0.200

P(credit-rating="fair"|buys-computer="yes") =6/9=0.667

P(credit-rating="fair"|buys-computer="no") =2/5=0.40

Using the above probabilities, we obtain

P(X|buys-computer="yes") =0.222×0.444×0.667×0.667=0.044

P(X|buys-computer="no") =0.600×0.400×0.200×0.400=0.019

P(X|buys-computer= "yes") p(buys-computer= "yes")=0.044×0.643=0.028

P (X|buys-computer= "no") P(buys-computer= "no")=0.019×0.357=0.007

Therefore, the naïve Bayesian classifier predicts buys-computer= "yes" for sample X.

Comparison between Apriori and Naive Bayes Algorithm

Apriori Algorithm works on unlabeled data where don't have class column. The Apriori algorithm learns association rules and is applied to a database containing a large number of transactions. For example if a customer buy A and B then most probably he will buy C. Here there is no class column. Any item can come in the decision. This algorithm is used to find association rule and mainly find the frequent item sets from the data sets. An algorithm like **Naive Bayes** works on labeled data where having classification columns. For example if anyone wants to relate the weather status and day of week with the punctuality of train then it should be labeled data. Because have lot of combination of weather and day of week and have a class column containing values late/not late. It would be difficult and practically impossible to classify a web page, a document, an email or any other lengthy text notes manually. This is where Naive Bayes Classifier machine learning algorithm comes to the rescue.

9. Conclusion

Data mining techniques used to extract important information from existing data and enable better decision-making throughout the retail industries. It is universally accepted that many industries (including banking, retail and telecom) are using data mining effectively. Algorithms such as Apriori algorithm in association rule mining and naive bayes classifier in classification techniques are using in retail industry for managing CRM. From this paper, observed that Apriori algorithm is to improve computational time and memory usage. Naive-Bayes is that if there are no occurrences of a class label and a certain attribute value together then the frequency-based probability estimate will be zero. The retail industry is also realizing that data mining could give them a competitive advantage. To further CRM in retailing empirical research is needed to identify how retailers of various types and sizes are actually using CRM.

10. References

1. Ling, R., & Yen, D. C. (2001). Customer Relationship Management: An analysis framework and implementation strategies. Journal of Computer Information Systems, 41, 82–97.
2. Frasquet, M., Molla A., Gil,I., Vallet, T.,2002.Research Trends in Retailing: acomparative approach. Journal of Retail and Distribution Management30 (8/9), 383-394.
3. Dr. Madan Lal Bhasin, 2006. Data Mining: A Competitive Tool in the Banking and Retail Industries.
4. Joan L. Anderson[a,*], Laura D. Jolly[b], Ann E. Fairhurst[b] "Customer Relationship Management in Retailing: A Content Analysis of reatil trade journal" Journal of Retailing and Customer Services 14(2007).*www.elsevier.com/locate/jretconser.*
5. Bharathi M.Ramageri, Dr. B .Ldesai " Role of Data Mining in Retail Sector" International Journal on Computer Science and Engineering(IJCSE).
6. EWT Ngai, L Xiu, DCK.Chau. Application of Data Mining Techniques in Customer Relationship Management: A Literature Review on Classification, *Expert Systems with Applications;* 36- 2, 2009. p. 2592-2602.
7. Wanghualin, "Data Mining and Its Applications in CRM", Second International Conference on Computer Research and Development, 2010.
8. Zhi Gang Li. Research on Application of association rule data mining in customer relationship management [J]. Chinese information circles, 2012, 10(38-40).
9. Kalyani M Raval,"*Data Mining Techniques"* ISSN: 2277 128X, Ijarcsse, Volume 2, Issue 10, October 2012.
10. P Salman Raju, Dr V Rama Bai,G Krishna Chaitanya, "Data Mining Techniques For Enhancing Customer Relationship Management in Banking and Retail Industries", International Journal Innovative Research in Computer and Communication Engineering, Vol.2,Issue 1,January 2014.

11. Shortland, R., Scarfe, R, "Data Mining Applications in BI", BT Technology 16. Journal, Vol.25, Nos 3&4, July/October (2007).

12. JW Han M Kamber. Data mining concepts and techniques, 2nd ed. Morgan Kaufmann, San Francisco, CA; 2006.

13. K. Chitra, B.Subashini, Customer Retention in Banking Sector using Predictive Data Mining Technique, International Conference onInformation Technology, Alzaytoonah University, Amman, Jordan, www.zuj.edu.jo/conferences/icit11/paperlist/Papers/

1) K. Chitra, B.Subashini, Automatic Credit Approval using Classification Method, International Journal of Scientific & Engineering Research (IJSER), Volume 4, Issue 7, July-2013 2027 ISSN 2229-5518.

2) Berson A, Smith S, Thearling K. Building data mining applications for CRM, McGraw-Hill; 2000.