



Refereed Journal

Journal of Electronic Systems and Programming

Issue No:2 June 2019

Journal of Electronic Systems and Programming

Electronic Systems and Programming Center

Journal
of
Electronic Systems
and Programming

Issue No: 2 1 June 2019

Editorial Board	
Mr. Abadul Hakim Alhadi Anaiba	Journal Manager
Dr. Khari A. Armih	Editor-in-Chief
Dr. Mustafa Kh. Aswad	Member
Mr. Nagib K. Omran	Member
Mr. Abodalrazag Khalifa	Member
Mr. Mohamed Wnis Huson	Member

Editorial:

Second Issue – Journal of Electronic System and Programming

Welcome to the second issue of the Journal of Electronic Systems and Programming (JESP). As anticipated, we are pleased to report that our first issue received a wonderful reception. It is indication that JESP is a step closer to being recognized as the one of the most useful journal for dissemination of high-quality research within the academic community. With this in mind, it is our pleasure to announce the publication of the second issue of the JESP.


As the official publication of Electronic Systems and Programming Center (EPC), the JESP aims to provide a platform for both researchers from inside and outside the center to share their novel results and latest developments, and also to seek better understanding of the current situation on research related to the area of contributions of EPC.

Finally, we thank our editorial board team, reviewers and authors for their fundamental contribution to the second release of the Journal. We still hope authors could consider JESP to be a place where to publish their work.

Dr. Khari A. Armih
Editor-in-Chief

Table of Contents

1	Investigating the Effects of Neural Network Parameters on Railway Wheel Wear Prediction	1-19
2	Feasibility Study of Photovoltaic Power Plant in Libya; Location, Technology and Economics	21-38
3	Novel Architecture of Self-organized Wireless Sensor Network	39-59
4	Drift Region Voltage Drop in SiC VDMOS - The Influence of Anisotropy	61-79
5	Assessment of different Growth Techniques of Strained Germanium Heterostructures for Electronic and spintronic Devices	61-100
6	Prediction of Wheel and Rail Roughness Parameters Using Artificial Neural Network	101-121
7	Performance Comparison of (AODV and DSDV) Routing Protocols in Mobile Ad Hoc Networks	123-142



Investigating the Effects of Neural Network Parameters on Railway Wheel Wear Prediction

1

Investigating the Effects of Neural Network Parameters on Railway Wheel Wear Prediction

A. Shebani

Institute of Railway Research, University of Huddersfield, Huddersfield, UK

S. Iwnicki

Institute of Railway Research, University of Huddersfield, Huddersfield, UK

Abstract

The prediction of wheel wear is still a great challenge for railway systems. This work examines the effect of radial basis function neural network (RBFNN) parameters such as spread, goal, maximum number of neurons, and number of neurons to add between displays on wheel wear prediction. VAMPIRE vehicle dynamic software was used to produce the vehicle performance data to train, validate, and test the neural network. The wheel wear was calculated using an energy dissipation approach and contact position on straight track. The neural network simulation results were implemented using the Matlab program. The percentage error for wheel wear prediction was calculated. Also, the accuracy of wheel wear prediction using the neural network was investigated and assessed in terms of mean absolute percentage error (MAPE). The results reveal that the railway wheel wear prediction using neural network is dependent on the correct selection of the neural network parameters.

Keywords wheel wear, wear prediction, railway systems, radial basis function neural network, Matlab, Vampire.

1. Introduction

Artificial neural network is currently used to solve a wide range of complex engineering problems. It has the ability to learn by example, consequently, it is a very useful for simulation of any correlation that is difficult to describe with physical models or other mathematical approaches [1]. Though perfect prediction is seldom possible, neural networks can be used to make reasonably good predictions in a number of cases. In particular, feedforward neural networks have been used frequently in this respect [2].

The prediction of wheel wear is a significant issue in railway vehicles. The aim of this work is to investigate the effect of RBFNN parameters on wheel wear prediction.

2. Radial Basis Function Neural Network

Radial basis function neural network (RBFNN) has an input, hidden, and output layer such as in Figure 1.

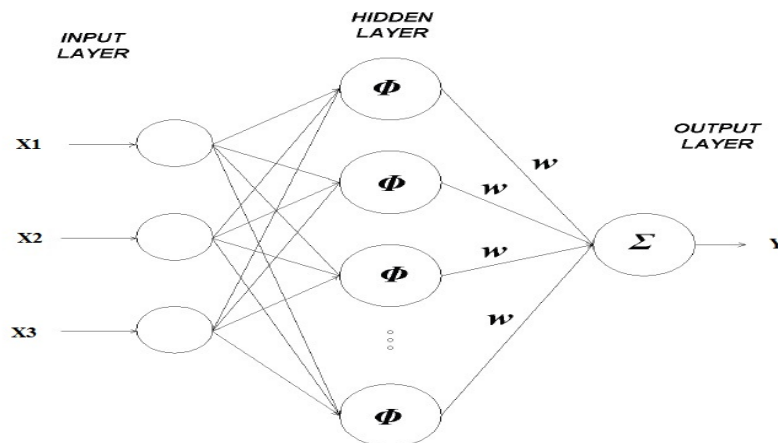


Figure 1: Radial Basis Function Network Architecture [3]-[5]

Where $X_1, X_2, X_3 \dots, X_m$ are the inputs, ϕ is the activation function, and W is the weights [4], [5].

The output of RBFNN can be represented such as shown in the following equation [6], [3], [4]:

$$y = \sum_{j=1}^m W_j \phi_j \quad (1)$$

The common activation function of RBFNN is the Gaussian function (ϕ) [3]-[7]:

$$\phi(x) = \exp\left(\frac{-r^2}{2\sigma^2}\right) \quad (2)$$

$$r = \|x - c\| \quad (3)$$

Where C are the centres, x are the inputs, and σ is the width of activation function.

Euclidean distance method is the most common method which can be used to calculate the width of activation function for RBFNN such as shown in the following equation [8], [7], [9]:

$$E_{\text{dist}} = \sqrt{\sum_{i=1}^n (X_i - c_i)^2}, i = 1, 2, 3, \dots, n \quad (4)$$

Where: X_i are the inputs, c_i are the centres, and n is the vector dimension.

The least mean square algorithm (LMS) is the most common algorithm which can be used for adapting the weights of the output layer for the RBFNN such as shown in the following equation [8], [7]-[10]:

$$W(t + 1) = W(t) + \mu (y(t) - y_m(t))\Phi^T(t) \quad (5)$$

Where $W(t + 1)$ is the updated weights, $W(t)$ is the previous weights originally set to zero, $y(t)$ is the desired output, $y_m(t)$ is the output of

the network, $\Phi^T(t)$ is the hidden layer output (Gaussian output), and μ is the learning factor of the RBFNN. The learning factor is a positive gain factor term that controls the adaptation rate of the algorithm ($0 < \mu \leq 1$).

The mean square error (MSE) is used for measuring the performance of the RBFNN such as shown in the following equation [11], [12], [13]:

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (e_i)^2 = \frac{1}{N} \sum_{i=1}^N (t_i - y_i)^2 \quad (6)$$

Where t_i is the target output and y_i is the estimated output.

In this paper, the “newrb” Matlab command is used to create and train the RBFNN, the design of RBFNN takes six arguments: input vector, target vector, mean square error goal, spread, maximum number of neurons, and the number of neurons to add between displays. These parameters are explicitly set by the user using trial and error.

In this work, the inputs of the neural network are the stiffness parameter, running distance, wheel profile, first derivative of wheel profile, second derivative of wheel profile; while the output of the neural network is the wheel wear.

Radial basis function neural network (RBFNN) is designed in this work using Matlab program using newrb command [14], [15], [16]:

$$\text{Net} = \text{newrb} (\text{P}, \text{T}, \text{goal}, \text{spread}, \text{MN}, \text{DF}) \quad (7)$$

Where:

P: is the input vectors

T: is the target vectors

Goal: is the mean squared error goal (MSE)

Spread: is the spread of radial basis functions

MN: is the maximum number of neurons

DF: is the number of neurons to add between displays

The function `newrb` iteratively creates a radial basis network one neuron at a time. Neurons are added to the network until the sum-squared error falls beneath an error goal or a maximum number of neurons have been reached. The radial basis function neural network architecture is shown in Figure 2.

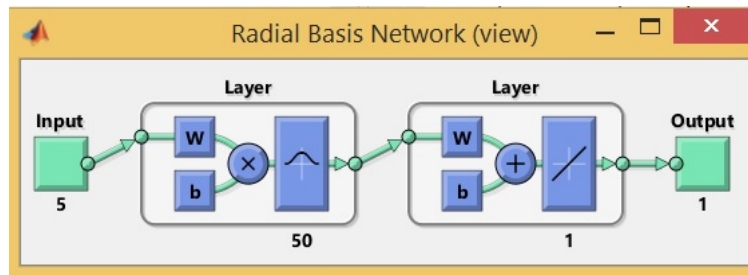


Figure 2: Radial basis function neural network architecture

3. VAMPIRE Vehicle Dynamics Software

VAMPIRE uses a multi-body modelling method that enables the user to assemble a mathematical model of almost any rail vehicle configuration. The VAMPIRE pro. 6.30 was used in this work to perform the simulations. The VAMPIRE GUI is shown in Figure 3. For wheel wear prediction, the transient analysis programme is run and the energy expended per unit distance travelled calculated for each wheel/rail contact. This is the product of creep force and creepage ($T\gamma$), is one of the output types available in the transient programme. Experimental work has demonstrated that the amount of metal removed through wheel is proportional to the energy dissipated in the wheel–rail contact. Therefore, the expected wear of wheel can be studied and predicted by calculating the energy dissipated between wheel and rail ($T\gamma$) [17], [18].

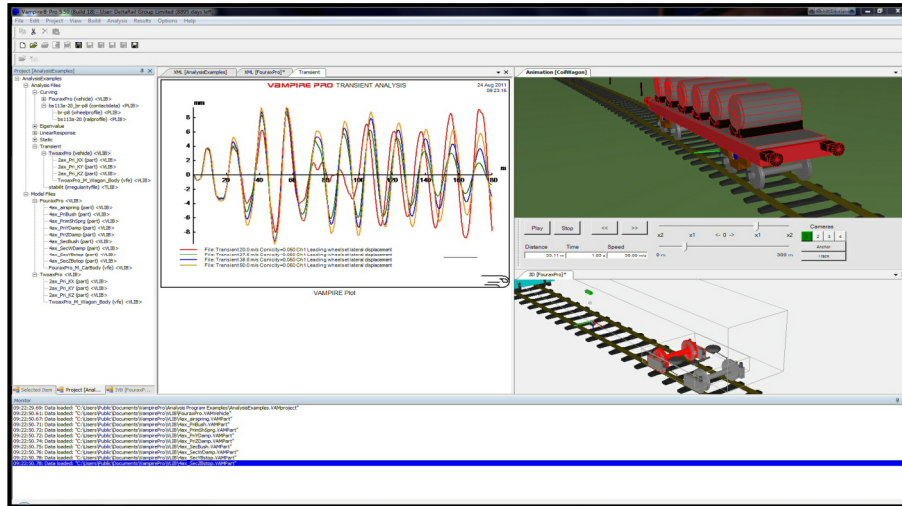


Figure 3: VAMPIRE vehicle dynamics software platform [17], [18]

4. Effects of RBFNN Parameters on Wheel Wear Prediction

In this section, the effect of the RBFNN parameters on wheel wear prediction was investigated. VAMPIRE vehicle dynamics software was used to collect data to train, validate, and test the neural network model.

In this work, an artificial neural network was developed to predict railway wheel wear in case of changing parameters such as vertical bush stiffness, lateral bush stiffness, lateral bush stiffness, and vertical shear stiffness. All results shown in this work are for unseen data.

A. Effect of Spread Parameter on Wheel Wear Prediction

The vertical bush stiffness simulation was used to investigate the effects of the spread parameter of RBFNN on wheel wear prediction. Wheel wear predicted using VAMPIRE, and wheel wear predicted using RBFNN with different values of spread are shown in Figure 4.

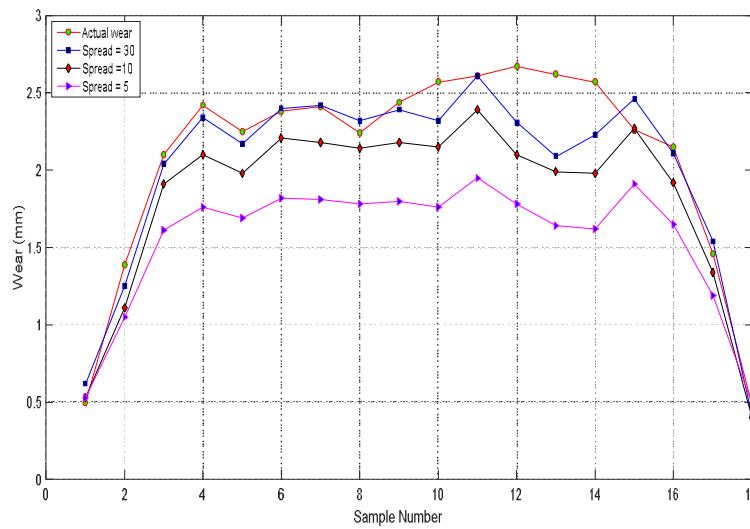


Figure 4: Actual wheel wear and predicted using RBFNN; with different values of spread

Wheel wear predicted using VAMPIRE, wheel wear predicted using RBFNN, and the percentage error are shown in Table 1. The mean absolute percentage error (MAPE) was 10.98% at spread of 30, was 14.70% at spread of 10, and was 32.36% at spread of 5.

Table 1: Actual wheel wear and predicted using RBFNN, and error; with different values of spread

Sample Number	Actual wheel wear (mm)	Wear predicted using RBFNN (mm) spread = 30	Error%	Wear predicted using RBFNN (mm) spread = 10	Error%	Wear predicted using RBFNN (mm) spread = 5	Error%
1	0.50	0.62	19.02	0.53	5.13	0.53	5.45
2	1.39	1.25	11.47	1.11	25.23	1.05	31.72
3	2.10	2.04	3.05	1.91	9.96	1.61	30.56
4	2.42	2.34	3.72	2.10	15.54	1.76	37.50
5	2.25	2.17	3.89	1.98	13.80	1.69	33.27
6	2.38	2.40	0.83	2.21	7.79	1.82	31.07
7	2.41	2.42	0.07	2.18	10.78	1.81	33.27
8	2.24	2.32	3.49	2.14	4.63	1.78	26.03
9	2.44	2.39	2.13	2.18	11.54	1.80	35.09
10	2.57	2.32	10.56	2.15	19.19	1.76	45.74
11	2.61	2.61	0.10	2.39	9.32	1.95	33.65
12	2.67	2.31	15.49	2.10	27.40	1.78	49.74
13	2.62	2.09	25.03	1.99	31.51	1.64	59.86
14	2.57	2.23	15.43	1.98	29.88	1.62	58.24
15	2.26	2.46	8.30	2.27	0.52	1.91	17.99
16	2.15	2.11	2.09	1.92	11.85	1.65	30.79
17	1.46	1.54	5.49	1.34	8.84	1.19	21.96
18	0.48	0.41	18.14	0.40	21.71	0.49	0.63

B. Effect of mn Parameter of RBFNN on Wheel Wear Prediction

The lateral bush stiffness simulation was used to investigate the effects of the mn parameter of RBFNN on wheel wear prediction. Wheel wear predicted using VAMPIRE, and wheel wear predicted using RBFNN with different values of mn are shown in Figure 5. Where the mn is the maximum number of neurons of RBFNN.

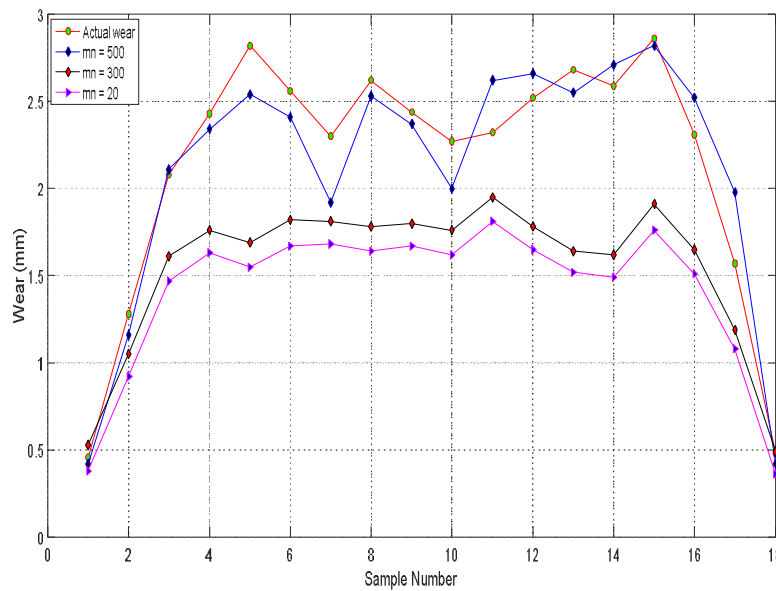


Figure 5: Actual wheel wear and predicted using RBFNN; with different values of mn

Wheel wear predicted using VAMPIRE, and wheel wear predicted using RBFNN, and the percentage error are shown in Table 2. The mean absolute percentage error was 9.11% when mn was 500, was 36.39 % when mn was 300, and was 49.31% when mn was 20.

Table 2: Actual wheel wear and predicted using RBFNN, and error; with different values of mn

Sample Number	Actual wheel wear (mm)	Wear predicted using RBFNN (mm) mn =500	Error%	Wear predicted using RBFNN (mm) mn =300	Error%	Wear predicted using RBFNN (mm) mn =20	Error%
1	0.456	0.42	7.51	0.53	14.39	0.38	17.87
2	1.28	1.16	10.40	1.05	21.46	0.92	38.48
3	2.08	2.11	1.60	1.61	29.21	1.47	40.78
4	2.43	2.34	3.75	1.76	37.96	1.63	49.26
5	2.82	2.54	10.86	1.69	66.94	1.55	81.24
6	2.56	2.41	6.45	1.82	40.98	1.67	52.93
7	2.30	1.92	19.75	1.81	26.92	1.68	37.08
8	2.62	2.53	3.24	1.78	46.89	1.64	59.41
9	2.44	2.37	3.08	1.80	35.17	1.67	45.48
10	2.27	2.00	13.46	1.76	29.11	1.62	40.23
11	2.32	2.62	11.29	1.95	18.97	1.81	28.17
12	2.52	2.66	5.15	1.78	41.32	1.65	52.79
13	2.68	2.55	5.21	1.64	63.68	1.52	76.35
14	2.59	2.71	4.17	1.62	59.69	1.49	73.33
15	2.86	2.82	1.14	1.91	49.31	1.76	61.89
16	2.31	2.52	8.04	1.65	40.43	1.51	52.74
17	1.57	1.98	20.45	1.19	31.60	1.08	45.94
18	0.48	0.42	14.02	0.49	0.95	0.36	33.63

C. Effect of Goal Parameter of RBFNN on Wheel Wear Prediction

The lateral shear stiffness simulation was used to investigate the effects of the goal parameter of RBFNN on wheel wear prediction. Wheel wear predicted using VAMPIRE, and wheel wear predicted using RBFNN with different values of goal are shown in Figure 6. Where the goal is denotes the mean squared error goal.

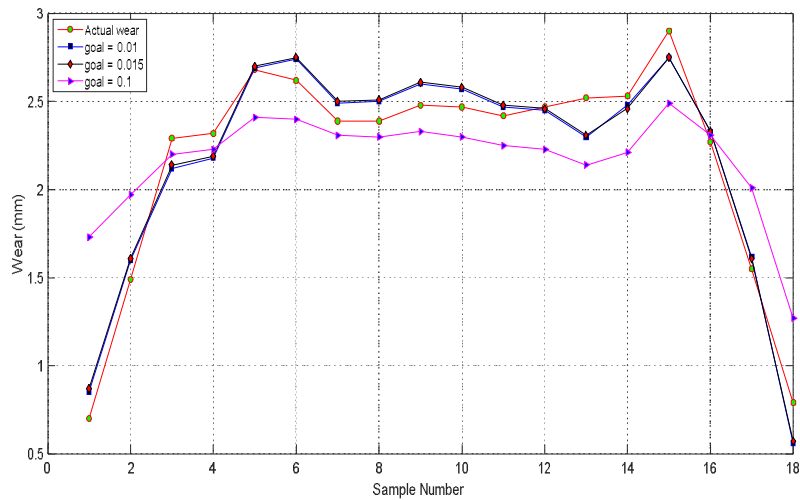


Figure 6: Actual wheel wear and predicted using RBFNN with different values of goal

Wheel wear predicted using VAMPIRE, and wheel wear predicted using RBFNN, and the percentage error are shown in Table 3. The mean absolute percentage error was 9.01% when goal was 0.01, was 9.45% when goal was 0.015, and was 14.55% when goal 0.1.

Table 3: Actual wheel wear and predicted using RBFNN, and error; with different values of goal

Sample Number	Actual wheel wear (mm)	Wear predicted using RBFNN (mm) Goal = 0.01	Error%	Wear predicted using RBFNN (mm) Goal = 0.015	Error%	Wear predicted using RBFNN (mm) Goal = 0.1	Error%
1	0.70	0.85	16.84	0.87	19.31	1.73	59.26
2	1.49	1.60	6.94	1.61	7.40	1.97	24.46
3	2.29	2.12	8.06	2.14	7.09	2.20	4.01
4	2.32	2.18	6.66	2.19	6.22	2.23	4.33
5	2.68	2.69	0.20	2.70	0.57	2.41	11.30
6	2.62	2.74	4.28	2.75	4.79	2.40	9.09
7	2.39	2.49	4.10	2.50	4.61	2.31	3.55
8	2.39	2.50	4.18	2.51	4.57	2.30	3.90
9	2.48	2.60	4.65	2.61	5.03	2.33	6.52
10	2.47	2.57	3.96	2.58	4.38	2.30	7.22
11	2.42	2.47	1.88	2.48	2.25	2.25	7.76
12	2.47	2.45	0.75	2.46	0.43	2.23	10.69
13	2.52	2.30	9.47	2.31	8.79	2.14	17.43
14	2.53	2.48	1.93	2.46	2.56	2.21	13.99
15	2.90	2.75	5.32	2.75	5.35	2.49	16.45
16	2.27	2.33	2.49	2.33	2.64	2.31	1.44
17	1.55	1.62	4.25	1.61	3.91	2.01	23.02
18	0.79	0.56	41.98	0.57	39.58	1.27	37.45

D. Effect of df Parameter of RBFNN on Wheel Wear Prediction

The vertical shear stiffness simulation was used to investigate the effects of the df parameter of RBFNN on wheel wear prediction. Wheel wear predicted using VAMPIRE, and wheel wear predicted using RBFNN with different values of df are shown in Figure 7. Where df represents the number of neurons to add between displays.

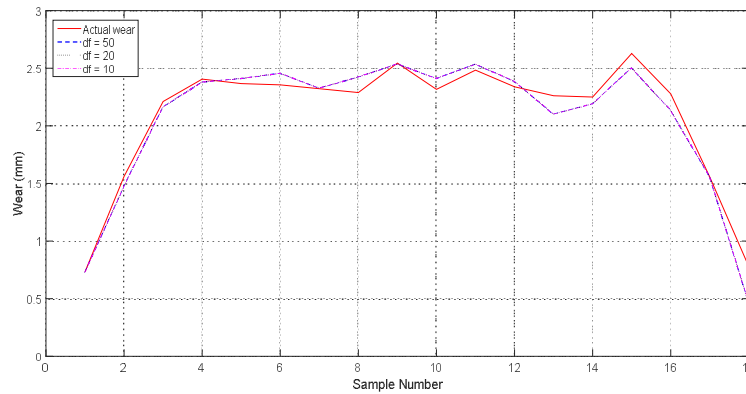


Figure 7: Actual wheel wear and predicted using RBFNN with different values of df

Wheel wear predicted using VAMPIRE, wheel wear predicted using RBFNN, and the percentage error are shown in Table 4. The mean absolute percentage error was 6.65% when df was 50, was 6.65% when df was 20, and was 6.65% when df was 10. The simulation results show that the changing of df parameter had no effect on the accuracy of wheel wear prediction.

Table 4: Actual wheel wear and predicted using RBFNN, and error; with different values of df

Sample Number	Actual wheel wear (mm)	Wear predicted using RBFNN (mm) df = 50	Error%		Wear predicted using RBFNN (mm) df = 20	Error%	Wear predicted using RBFNN (mm) df = 10	Error%
1	0.73	0.72	1.07		0.72	1.07	0.72	1.07
2	1.56	1.47	5.88		1.47	5.88	1.47	5.88
3	2.21	2.16	2.14		2.16	2.14	2.16	2.14
4	2.40	2.38	1.15		2.38	1.15	2.38	1.15
5	2.36	2.41	1.93		2.41	1.93	2.41	1.93
6	2.35	2.45	4.05		2.45	4.05	2.45	4.05
7	2.32	2.32	0.29		2.32	0.29	2.32	0.29
8	2.29	2.42	5.49		2.42	5.49	2.42	5.49
9	2.54	2.53	0.24		2.53	0.24	2.53	0.24
10	2.31	2.41	3.87		2.41	3.87	2.41	3.87
11	2.48	2.53	2.07		2.53	2.07	2.53	2.07
12	2.34	2.38	1.85		2.38	1.85	2.38	1.85
13	2.26	2.10	7.65		2.10	7.65	2.10	7.65
14	2.25	2.19	2.80		2.19	2.80	2.19	2.80
15	2.63	2.50	5.05		2.50	5.05	2.50	5.05
16	2.28	2.136	6.76		2.13	6.76	2.13	6.76
17	1.55	1.55	0.01		1.55	0.01	1.55	0.01
18	0.77	0.46	67.40		0.46	67.40	0.46	67.40

5. Conclusion

This work investigated the effect of RBFNN parameters on wheel wear. The effect of spread, mn, goal, and df parameters on wheel wear prediction were examined. The mean absolute percentage error (MAPE) was calculated as:

- The MAPE was 10.98% when spread was 30, it was 14.70% when spread was 10, and it was 32.36% when spread was 5 during the change of spread parameter test.
- The MAPE was 9.11% when mn was 500, it was 36.39 % when mn was 300, and it was 49.31% when mn was 20 during the change of mn parameter test.
- The MAPE was 9.01% when goal was 0.01, it was 9.45% when goal was 0.015, and it was 14.55% when goal was 0.1 during the change of goal parameter test.
- The MAPE was 6.65% when df was 50, it was 6.65% when df was 20, and it was 6.65% when df was 10 during the change of df parameter test.


The effects of the RBFNN parameters such as spread, goal, maximum number of neurons, and number of neurons to add between displays on wheel wear prediction was investigated. The simulation results show that the accuracy of wheel wear prediction was influenced by change of spread, mn, and goal; while the change of df parameter has no effect on the wheel wear prediction using RBFNN. Therefore, it can be concluded that the railway wheel wear prediction using neural network is dependent on the correct selection of the neural network parameters.

Finally, the VAMPIRE vehicle dynamic software can assist in using the neural network in railway wheel wear prediction; where several simulations were carried out in this work using VAMPIRE software to produce the data to train, validate, and test the neural network model.

References

- [1] A. Khudhair and N. A. Talib, "Neural Network Analysis For Sliding Wear of 13% Cr Steel Coatings by Electric Arc Spraying," in *Diyala Journal of Engineering Sciences-First Engineering Scientific Conference, College of Engineering–University of Diyal*, 2010, pp. 157-169.
- [2] K. Mehrotra, C. K. Mohan, and S. Ranka, *Elements of artificial neural networks*: MIT press, 1997.
- [3] M. K. Kundu, Ed., *Advanced Computing, Networking and Informatics*. Switzerland, Springer publishing, 2014.
- [4] M. Patel, V. Honavar, and K. Balakrishnan, *Advances in the evolutionary synthesis of intelligent agents*: MIT press, 2001.
- [5] S. Sabnam, D. Kunal, and K. Gitosree, *Emerging Trends in Computing and Communication*: Springer 2014.
- [6] R. Matignon, *Neural network modeling using SAS enterprise miner*: AuthorHouse, 2005.
- [7] G. M. Fathalla, "Analysis and implementation of radial basis function neural network for controlling non-linear dynamic systems," PhD thesis Faculty of Engineering University of Newcastle, UK, 1998.
- [8] K. L. Priddy and P. E. Keller, *Artificial neural networks: an introduction* vol. 68: SPIE Press, 2005.
- [9] S. Fosseng, "Learning Distance Functions in k-Nearest," Norwegian University of Science and Technology, Department of Computer and Information Science, 2013.
- [10] S. S. A. Ali, M. Moinuddin, K. Raza, and S. H. Adil, "An adaptive learning rate for RBFNN using time-domain feedback analysis," *The Scientific World Journal*, vol. 2014, 2014.
- [11] L. D. Fredendall and E. Hill, *Basics of supply chain management*: CRC Press, 2000.
- [12] C. Cătălina-Lucia and G. Hakob, "An Artificial Neural Network for Data Forecasting Purposes," *Informatica Economica*, vol. 19, 2015.
- [13] E. Diaconescu, "The use of NARX neural networks to predict chaotic time series," *WSEAS Transactions on Computer Research*, vol. 3, pp. 182-191, 2008.

- [14] Mathworks, "Design radial basis network - newrb " 2015.
- [15] D. Li and Y. Chen, *Computer and Computing Technologies in Agriculture VII: 7th IFIP WG 5.14 International Conference, CCTA 2013, Beijing, China, September 18-20, 2013, Revised Selected Papers* vol. 419: Springer, 2014.
- [16] M. Iskander, V. Kapila, and M. A. Karim, *Technological Developments in Education and Automation*: Springer Science & Business Media, 2010.
- [17] DeltaRail, *VAMPIRE manual V5.60*, UK, 2012.
- [18] DeltaRail, *VAMPIRE manual V6.30*, UK,



**Feasibility Study of
Photovoltaic Power Plant in
Libya; Location, Technology
and Economics**

2

Feasibility Study of Photovoltaic Power Plant in Libya; Location, Technology and Economics

Ahmed S. Kagilik,
Faculty of Engineering, Sabratha University, Libya
as.kagilik@gmail.com

Khairy Agha
Faculty of Engineering, Tripoli University, Libya
k_r_gha@yahoo.com

Christian Bornhauser
GIZ Gesellschaft für Internationale Zusammenarbeit, Germany,
cb@bornhauser24.de

Abstract

Steeply falling installation costs during the last decade makes the Photovoltaic technology highly competitive particularly, of course, in a sunny region like the Libyan Desert. Photovoltaic power plants are easy to maintain and, once installed, consume practically no additional natural resources. New and innovative jobs will be created and local companies will take part in certain ways and benefit.

The aim of this paper is to demonstrate how efficient and cost-saving solar energy can be exploited in the deserts of Libya. A feasibility analysis to the large-scale grid connected PV project with total capacity of 14 MW is dedicated. The power plant delivers approximately 27 GWh of electrical power in one year. It is shown that the cost for one kWh is in the range of 0.08 USD, which is approximately four times less than the cost for electricity produced

with Diesel-generators in the same region. This leads to savings of Diesel (light oil) of approximately 6,400 tons per year. The value of the saved diesel is approximately 5.7 Million US Dollars (at a world market price of 900 USD per ton). Within its minimum lifetime of 25 years, the PV power plant saves 150,000 tons of Diesel or 133 Millions of US Dollars. Compared to the saved oil the break-even is within 7 years and the Return of Investment (ROI) is in the range of 14 %.

Keywords photovoltaics, electric power, fossil fuel, Return of Investment.

1. Introduction

Renewable energy sources are now seen as an important asset, in particular for the electricity sector, which is still entirely driven by fossil fuels. RE sectors are generally based on mechanical, technically-intensive production technology and are mainly labor-intensive. A domestic market for RE infrastructure acts as driver for manufacturing RE components domestically. Today, photovoltaics are well developed technology with fast growing market with highly competitive to fossils if one considers the world market prices for light oil. The world market for PV technologies and system components is mainly fully developed. Many PV systems have been applied in many countries and implemented successfully in different locations and environments throughout the planet. The compound annual growth rate of PV installations is up to 45 % during the period from 2006 to 2016 with annual installations in the range of at least 40 GWp and the total cumulative installations of about 300 GWp at the end of the year 2016 [1, 2].

Electricity demand in Libya is rising quickly, at an annual rate of more than 6 %. Oil consumption within the country is rising strongly too. Renewable energies for power generation are also expected to

alleviate power supply bottlenecks, since the electricity consumption in Libya rapidly increases, with power supply shortages becoming more and more frequent. Libya is also endowed with an enormous potential for renewable energies (RE) and may enable Libya to diversify from its current strong reliance on fossils towards a more technology-driven model. However, Libya has fully recognized the potential of its excellent wind and solar conditions in combination with empty land in the South of Libya. But despite excellent solar radiation, wind conditions and an abundant availability of undeveloped land, renewable options have never been seriously considered in the past. Today, the new Libyan state is reconsidering its energy policies. Renewable power plants offer an opportunity to reduce the national consumption of domestic fossil fuel resources for electricity generation, thereby saving oil and natural gas for exports that are needed to sustain. By creating a domestic market for RE infrastructure, Libya will realize several advantages in its domestic employment market. The RE share in 2025 is expected to reach approximately 10 % [3].

In this work, a feasibility/pre-Tender - costs and benefits are getting estimated based upon industry proven data and experiences. The first milestone is “Feasibility confirmed” by assessing all relevant influences and judging the successfulness of the PV system at Hun. Numerous stakeholders are involved in the approach. Most important are Libyan governmental authorities, such as the General Electric Company of Libya (GECOL) and the Renewable Energy Authority of Libya (REAol), providing the necessary framework and empowerment to the project. An external drivers and supporters such as Gesellschaft für Internationale Zusammenarbeit GIZ and the Desertec Initiative Dii are providing strategic, structural and technological input for the project. On the other hand, an expert consultants (e.g. Fichtner GmbH, Wuppertal Institut, Dii) have to be involved which bring in additional expertise [4, 5].

2. Site Examination

In parallel the geographic and environmental conditions of Libya are almost ideal for PV, Libya provides vast areas of land with excellent solar (and wind) conditions. As a result many attractive sites can be identified providing excellent opportunities establishing a network of RE power plants [6]. In this work, the proposed site of the PV power plant will be close to the city of Hun (N29°07'07'', E15°56'12), located in the desert area of Libya. The area is flat, arid, highly exposed to the sun and no high mountains or buildings may influence the power system.

Solar irradiance is the single most important parameter for evaluation of PV system performance, so for simulation and evaluation purposes it is important to select the best possible data series. The global horizontal irradiation (GHI) is the most critical resource for solar PV plant. For optimum design of a PV plant, it is important to know the distribution on intensity and wavelength. Solar maps or irradiation tables are essential tools in planning and dimensioning of solar energy installations. In this study we have collected most of the available data sources for the solar climate of the project site in Hoon. However, we have to realize that all the data used in these studies were based on the satellite data. The general meteorological data at Hun site being used in this work for the yield simulations is being sourced via EMPower/NASA data, as shown in figure 1, [7]. One can see clearly that the desert conditions at Hun are very advantageous for PV power generation. If one compares the quality of the Hun location to other locations in Libya, such as al Kufrah, Sebha, Darnah and Tripoli. We can see that Hun is well positioned. Hun's average daily solar radiation (horizontal) is among the best in a set of excellent locations.

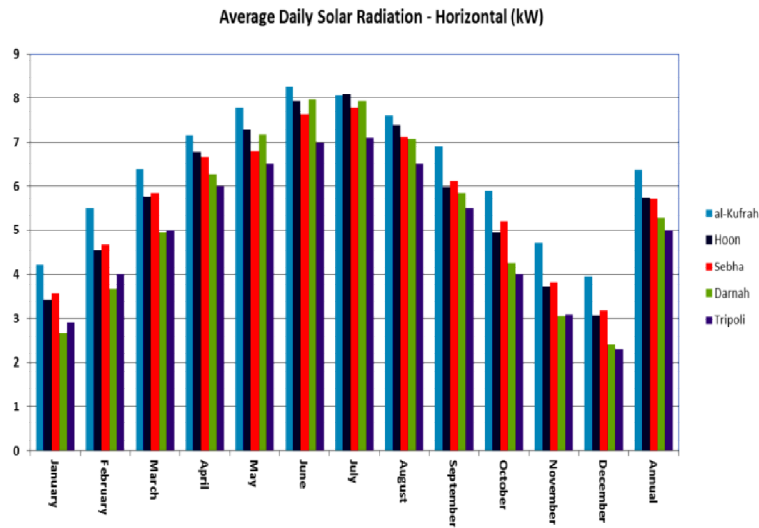


Figure 1: Hun solar radiation compared with other locations in Libya

The 14 MW PV plant shall be connected to the grid close to Hun. A substation is already in use at the construction site. It is obvious that PV power plant without additional power storage delivers electricity solely during daylight hours. Comparing the grid load in Libya with the daylight hours one can observe, that the maximum load is later in the day and sometimes (during summer) even after sunset. This means that PV cannot displace residual power generation, as shown in figure 2.

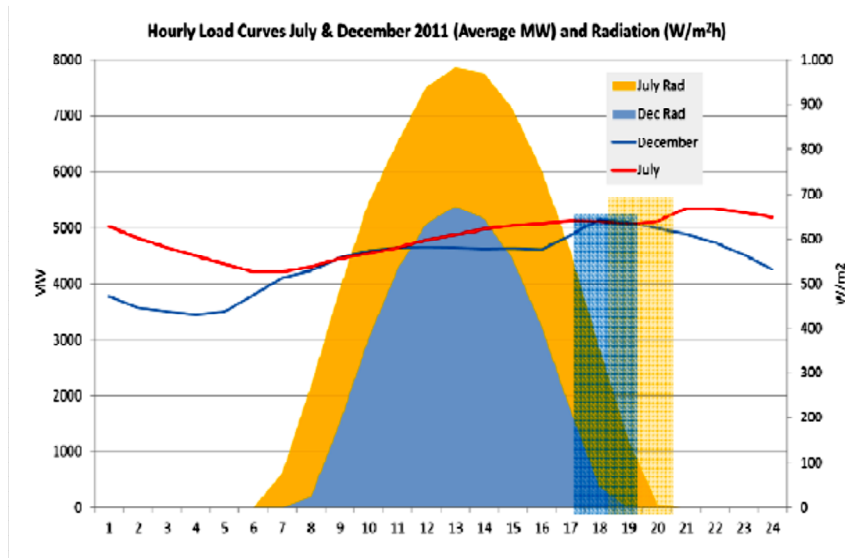


Figure 2: Exemplarily grid load and solar radiation in July and December

3. Yield Assessment Check and Cumulated power output during 25 years:

Solar photovoltaic PV modules are divided into two main categories: crystalline (mono- and poly-silicon) and thin film. Mono-silicon PV modules are the most efficient but also the most expensive. Poly-silicon PV modules are slightly less efficient and less expensive. Thin film PV modules are cheaper and consume much less silicon. In addition, thin film PV modules are less sensitive against high temperatures and more sensitive at ambient/diffuse light. This might lead to an overall higher output of electricity, but there are shortcomings as well. Most of the photovoltaic modules currently in the field are either mono-silicon or poly-silicon and have a proven track record for long lifetimes. Thin film PV modules are gaining market share; however, there are only little experiences for operating times of 10 years or more yet. In this work, technical design of the

solar PV system and an exemplary assessment for the Hun location based upon their CdTe (thin film) technology and the meteorological site conditions in Hun has made by Calyxo company in Germany. The results are determined by a mathematical model calculation. The actual yields of the photovoltaic system can deviate from these values due to fluctuations in the weather, the efficiency of modules and inverters, and other factors. Table 1 shows the main Hoon project design characteristics and yield assessment check results for 1MW AC variant ground mounted rack PV system.

Table 1: Main Hun project design characteristics and results

Project Block:	1MW (AC) System Variant ground mounted rack system	
Climate Data Record:	Hun (1986-2005)	
Gross/Active PV Surface Area:	92,228.20 / 94,051.09	ft ²
PV Array Irradiation:	21,391.580	kWh
Energy Produced by PV Array (AC):	1,902.986	kWh
Grid Feed-in:	1,902.986	kWh
System Efficiency:	8.9	%
Performance Ratio:	87.0	%
Inverter Efficiency:	98.1	%
PV Array Efficiency:	9.1	%
Specific Annual Yield:	2,131	kWh/kWp

The table above does not represent a full technical design of the solar system. However, these results show the huge potential of solar energy investment in Libya [8].

Most important for the evaluation of the power plant is the electrical output during its expected lifetime of at least 25 years. Using data provided by the Hun mast and meteorological data one can simulate different module technologies with regard to the actual electrical power fed to the grid. Figure 3 shows the calculated output amount of electricity produced for three different module technologies (Mono-silicon, poly-silicon and CdTe thin film) within the first year of operation. The diagram shows a certain advantage of the thin film technology that little higher electricity production of thin film modules, mainly caused by differing sensitivity of crystalline and thin film modules at higher temperatures and ambient light. Nevertheless, there is a considerable error margin slightly higher compared to the silicon modules. The error margin is relatively high and there are some drawbacks of the thin film technology, which outweigh the higher power production. Specifically the cumulated output over 25 years is of critical interest as this is the expected minimum lifetime of the PV solar plant. Differences of the three technologies (Mono-Si, Poly-Si and CdTe) are within the range of 5%, which can be assumed to be within the error margin.

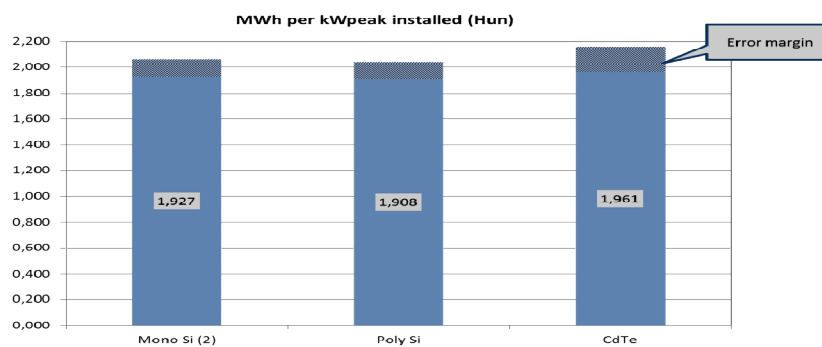


Figure 3: Calculated electricity output within the first year of operation for Poly-, Mono- and CdTe modules considering high temperature degradation

The long-term degradation of PV modules is an important parameter for the calculation of the cumulated power fed to the grid during the expected minimum lifetime (25 years) of the power plant. The expected (calculated) cumulated amount of the supplied electricity for 5, 10 and 25 years is shown in figure 4. The total output of the 14MW power plant in Hun can be expected as approximately 600 GWh over 25 years under the premises that the power station is operated and maintained in an orderly manner. During the minimum of 25 years, operating phase of the PV system, the output has to be measured and compared regularly to the guaranteed value. Degradation is assumed to occur in a linear way over time, which makes yield calculations rational. However, significant parameter to be taken into consideration for long-term operation of a PV power plant is the temporal degradation of efficiency. Silicon cells as well as thin film modules loose efficiency caused by chemical ageing effects. These ageing effects lead to a temporal degradation of efficiency.

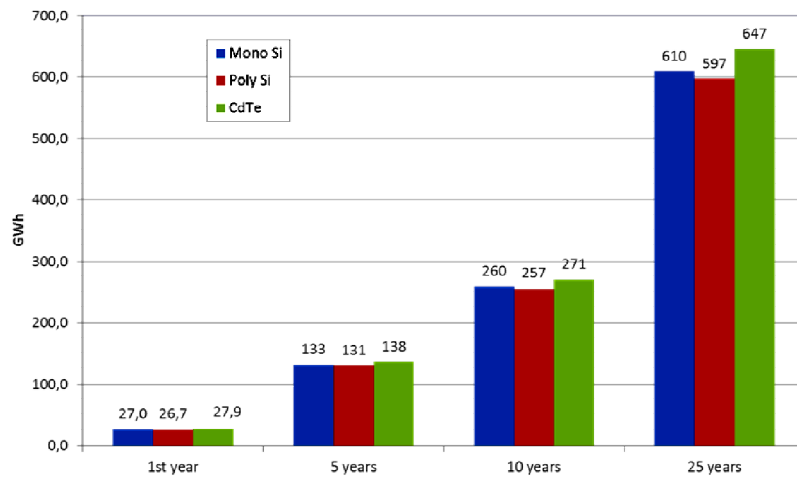


Figure 4: Yield. Calculated electricity output within 25 years of operation for Poly-, Mono- and CdTe modules considering degradation effects (temperature and ageing)

There are only less verified long-term data for thin film. Nevertheless thin film manufacturers (e.g. calyxo, first solar and others) guarantee a degradation of not more than 10% over 10 years and not more than 20% over 25 years for their modules, equivalent to crystalline module manufacturers [8]. Thus, the degradation factors for crystalline and thin film modules can be assumed to be on the same level. The degradation factors need to be specified in the contract with the supplier of modules. Degradation also occurs due to various other reasons and it is possible to diminish some of them by a suitable design, proper operation and consequent maintenance of the PV system during its lifetime. Figure 5 illustrates a relatively high degradation generated by e. g. dust from sand or damage. As a consequence implementation of frequent cleaning devices and trained operation personal is mandatory. There are different methods in place; but mainly cleaning by water gets applied to avoid touching the surface of the modules. In the arid desert area of the PV plant the amount of water needed is an important factor. Thin film technology needs more labor and more water as the surface needed for 14 MW is significantly higher.

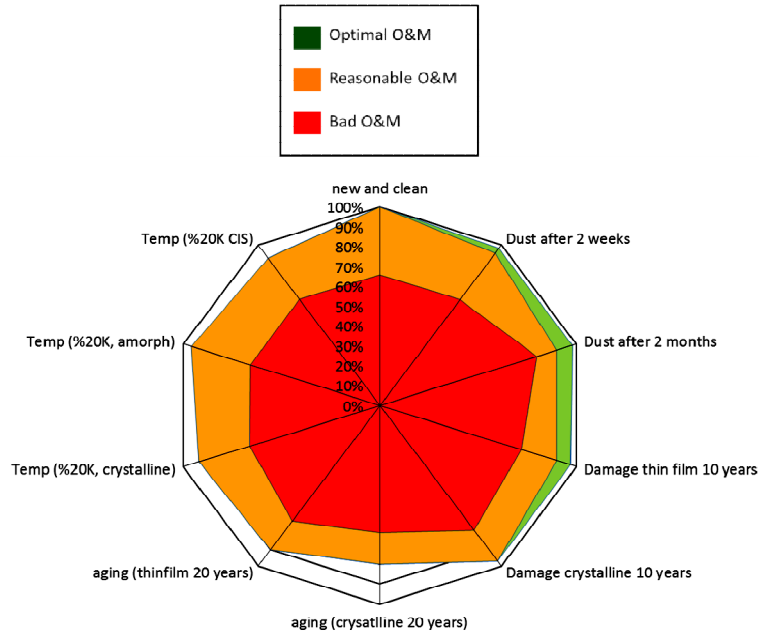


Figure 5: Degradation effects cumulated. The impact of good and bad maintenance

4. Economic Feasibility: Costs, Earnings, ROI and Break Even

Currently there is no open market for electricity in Libya. Both electricity production and distribution are operated and owned by the state. In principle, new power plants have to be benchmarked against the most expensive power production facilities which can be replaced by the new power plant. In the south of Libya (e.g. Hun. Sebha, Ghat) mobile Diesel generators support the security of electricity supply. These generators can be replaced by PV power generation or, at least, the Diesel consumption can be reduced drastically. Therefore, all calculations are based on the savings of Diesel and the actual world market price for Diesel. Thus, we have to calculate the savings

induced by the 14MW PV power plant multiplied by the estimated world market price of 900 USD per ton average which is the lower average world price within the time period 2011-2013. In the next 25 years, it is expected that the world price rather rises or stays steady instead of declining. Table 2 shows how the cost for one MWh can be calculated on basis of the Diesel world market price. These values are taken from the Sebha Diesel generator station. Taking the numbers from the yield estimation, we have the electricity produced by 1 MW installed capacity of 1,927 MWh per year and the production of the 14MW PV power plant in Hun accordingly within a time period of 5, 10, 15, 20 and 25 years is determined. The electricity produced within 25 years period is about 625,890 MWh at 87% performance ratio including all system losses and module degradation.

Table 2: Diesel consumption and USD cost for one MWh (1MWh = 284USD)

Electricity Production Cost – Diesel Generator in Sebha	
Generator Consumption equals to	0.283 liter Diesel per KWh
equals to	3.54 KWh per liter Diesel
equals to	4.21 MWh/ton Diesel
1 ton Diesel =	1,190.5 liter
Diesel WM Price:	900 USD/ton
equals to	0.756 USD/liter
≈	0.214 USD/KWh (Diesel)
Generator rent & OH +	0.070 USD/KWh
Generator Electricity Cost total	0.284 USD/KWh (Diesel)
Generator Electricity Cost total	284 USD/MWh (Diesel)

Assuming that 1 ton of diesel produces approximately 4.21MWh and 1 ton of diesel costs 900USD, the diesel savings from the 14MW PV power plant in tons within a time period of 1, 5, 10, 15, 20 and 25 years are calculated. For example, the diesel saving within 25 years is about 148,573 tons which is equivalent to about 133,715,885 USD.

Fully in operation the PV plant delivers power compared to approximately 6,400 tons of Diesel per year which is equivalent to 200 truck-loads. In 25 years – the guaranteed system lifespan - approximately 150,000 tons of Diesel will be saved. If the world market price is 900USD per ton, this correlates to more than 130 mio USD. To calculate the production cost for the 14 MW PV plant we assume a minimum life time of the plant of 25 years and an investment of 2.70 USD for 1Wpeak. The annual cost of operation and maintenance is set to 1% of the total investment, as shown in table 3. The calculations clearly show that the production of electricity with Diesel generators in the region of Sebha is four times more expensive than production with PV.

Table 3: PV plant production cost

Production Cost Breakdown		
Invest for 1 Wp PV (estm.)	USD	2.7
Total investment 14 MWp	USD	37,800,000
Minimum operation period	Years	25
Annual depreciation	USD	1,512,000
Annual cost of operation	% of invest	1.0 %
Annual cost of operation	USD	378,000
Annual production cost		1,890,000
Production cost 1 (net)	USD/KWh	0.070
Production cost 1 (net)	USD/MWh	70.06

A sensitivity analysis shall demonstrate the variance of the production cost when input parameters vary, as shown in figure 6.

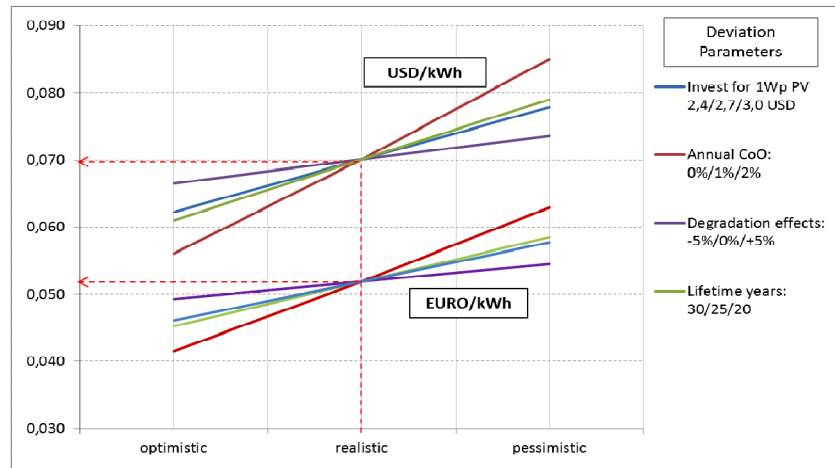


Figure 6: PV plant production cost sensitivity analysis for different input parameters.

Two categories of costs occur in the upfront investment costs and the ongoing maintenance / operating costs. As described above “income” in this project gets defined as avoided tons of diesel multiplied by expected diesel world market price. The point in time “break even” when the cumulated income exceeds the sum of costs is of particular interest for the bankability. Additionally the “pay-back-percentage”, called Return on Investment (ROI), makes transparent the profitability of the PV plant and can be compared to the profitability of other projects in the investment market. Of course the calculation is dynamical due to deviation and other effects [9]. In this project, number of financial elements such as, system procurement and installation, site preparation, project office and logistics (1.5-2.5 % of System) are used for ROI calculation. Ongoing costs and income during operation phase (Maintenance/Operation, Insurance, Diesel saved etc.) are also included. It has found that the ROI for the first year is equal to about 14 % while the break even is about 7 years at degradation of about 0.8% year.

5. Conclusions

The experiences with PV power in Libya are limited, in particular, there are no large-scale PV power plants installed yet. Recently, the promotion and development of renewable energies has recently become an important priority for Libya. Therefore, the electricity from the PV power plant offer an attractive platform for the future to achieve long-term energy independence and can help to stabilize the power supply, while reducing environmental impact and sustaining a strong economy and society for future generations. In this direction large scale grid-connected PV projects will be established as a national plan to improve the diversification of the Libyan energy production portfolio. To verify the feasibility of PV power systems, the first project with 14 MW grid-connected PV system which will be executed in Hun city was presented and a comprehensive study including the plant design and its performance analysis and behavior was performed. The 5, 10 and 25 years analysis performance was investigated and the final yield assessment, cumulated power output and performance ratio considering the degradation effects and system losses for three different module technologies were calculated. Our results revealed that the differences of the three technologies (Mono-Si, Poly-Si and CdTe) are within the range of 5%, which can be assumed to be within the error margin. The 14 MW PV plant production cost, assuming a minimum life time of the plant of 25 years was calculated. The calculations clearly show that the production of electricity with Diesel generators is more expensive than production with PV with the ROI for the first year is equal to about 14 % at degradation of about 0.8% year, demonstrating the benefits and challenges associated with facilitate increasing the penetration levels of PV systems in the electric network in Libya.

Acknowledgement

The work has been supported by the Renewable Energy Authority of Libya (REAol). The authors are thankful to Dipl.-Ing. Rainer Broscheit from Calyxo GmbH who was a part in accomplishing this work and has made technical design and an exemplary assessment of the solar PV project.

References

- [1] REN21, Renewables 2017 Global Status Report, 2017.
- [2] Fraunhofer ISE: Photovoltaics Report (2014).
- [3] Renewable Energy in Libya, Renewable Energy Authority of Libya, annual activity report, 2013.
- [4] GIZ leaflet. Beratungsleistungen für libyschen Energieversorger, 2013, Berlin, Germany.
- [5] Dii GmbH, brochure. The economic Impacts of Desert Power, 2013, Munich, Germany.
- [6] Lahmeyer International GmbH. PIM Solar PV Plant – Shahat – Ghadamis, Libya” 2010.
- [7] NASA data, Utility Toolkit:
<http://empower-ph2.com/EMPowerToolkit/>.
- [8] Rainer Broscheit. Calyxo GmbH, Bitterfeld, Germany
- [9] Doruk Sen, Murat Tunc, Taylan Ozilhan, Investment analysis of a new solar Power Plant, International Journal of Renewable and Sustainable Energy, 2013; 2(6): 229-241.



**Novel Architecture of Self-
organized Wireless Sensor
Network**

3

Novel Architecture of Self- organized Wireless Sensor Network

Hamzza Dawd

Computer science department, University of Bridgeport, Bridgeport,
Connecticut, USA

Hdawd@bridgeport.edu

Rabia Lamami

Electronic System and Programming Center, Tripoli, Libya

Rabia.m.Lamami@gmail.com

Shuyang Li

Computer science department, University of Bridgeport, Bridgeport,
Connecticut, USA

shuyangl@bridgeport.edu

Abstract

Self-organization for Wireless Sensor Network (WSN) is critical issue because of each sensor node's limited energy, limited bandwidth and WSN's scalability. Therefore, how to manage wireless sensor networks effectively is a big challenge task. This paper presents a novel self-organized architecture which is capable to avoid these problems. In this architecture, we suppose that each node does not know its location (Random clustering). Random clustering is practical to implement on some applications which deploy nodes into inaccessible unknown environment. We propose two algorithms to divide sensor nodes into cells. The first algorithm Active-Tree uses tree topology assign different role and node ID to each sensor node. The second algorithm Drawn-Grid divides sensor nodes into cells according to the radio coverage and the roles get from the Active-Tree algorithm. Based on

sensor nodes with different role play different tasks in WSN. The result of numerical simulations will show that our algorithm performs better.

Keywords component; formatting; self-organization; Wireless Sensor Network; self-organized architecture

1. Introduction

A Wireless Sensor Network consists of large number of sensor nodes. The sensor nodes are small, inexpensive, low power, distributive devices and have computation, sensing and wireless communication capabilities. But those sensor nodes are limited in resource i.e. memory, energy and computation power [1]. The sensor nodes mainly use battery. Maybe some use solar battery or limited recharge. For the bandwidth of the sensor nodes, only have a few hundred kbps. WSN has been applied in many military and civilian applications such as machine monitoring, seismic detection, disaster recovery [3]. In some applications, sensor nodes are deployed predetermined like machine monitoring. This kind of applications is application-oriented deployment. And sensor nodes are mainly randomly deployed in most applications such as seismic detection, disaster recovery. It increases the complexity of design for WSN. This paper focuses on the WSN with randomly deployed sensor nodes.

WSN use sensor nodes to sensing the events and transmit it out to base stations. So the feature of the sensor nodes is the feature of the WSN which is limited resources. The challenge for WSN is to complete complex application based on the limited resources especially when the sensor nodes are randomly deployed.

Self-organization is involved to make sensor nodes used efficiently. Generally, Self-organization is the process of autonomous formation of connectivity, addressing and routing structures [1]. As limited resource

of the sensor nodes, and it is practical to distribute these management tasks to different sensor nodes. And there is no need to make every sensor nodes to do management tasks, as sensor nodes only possess limited energy, computation power and limited bandwidth. So the sensor nodes must be clustered to a cell or a group [1, 2, 6, 7], to be an easily manageable network.

Another feature of clustered architecture WSN is that Well-Clustered architecture WSNs consume less energy than self-organization WSNs but only independent nodes [2].

A. Problem Identification

WSNs have following problems: collision, idle-listening, overhearing and disabled sensor nodes. Figure 1 shows the collision. Idle-listening is the most power-intensive action. It could be solved by set sleep and wake model to sensor nodes [8]. Sensor nodes will be set to sleep model when there is no events and it will wake up periodically or wake up by management nodes. When sensor nodes receive the sensing information not belongs to it, overhearing happens. Management nodes could use node ID to solve it [5]. Disabled sensor nodes refer to the sensor nodes with low energy or damaged by environment. Clustered architecture WSN must be capable to avoid these problems to maximize the life time of the WSN [11]. And these problems are mainly solved by the management nodes.

A well-clustered WSN should divide into groups or cells and in a group or cell appropriate amount of sensor nodes are management nodes. As management sensor nodes in this paper are same devices as normal sensor nodes, so resource limited decides that management nodes cannot take much more tasks. Also, if more sensor nodes have chosen to be management nodes, then extra energy will be consumed for the communication between the management nodes. i.e. if 3 sensor nodes take one management task like recording the node ID which

could be done by one sensor node, then when other sensor node needs the information about node ID the 3 sensor nodes need to communication to provide the information, so extra energy is wasted. So in a group or cell the management sensor nodes must be relatively appropriate, not less or more.

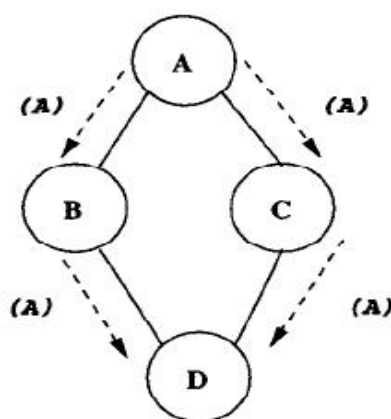


Figure 1: Both sensor nodes B and C desire to send message to sensor node D. Resource (energy, bandwidth) wasted.

As management nodes are same as normal nodes, so management nodes also have the chance of being damaged or low power. How to maintain the architecture of WSN also needs to be considered in our solution.

The rest of this paper is organized as follows. In Section 2, we review several related works, in Section 3 describes our proposed a novel tree self-organized architecture for Wireless Sensor Network is introduced here. Section 4. Finally, conclusions and future works are given in Section 5.

2. Related work

Self-organization has been an important research topic in wireless sensor networks, so there are many researchers who interested in solving its problems.

Duan and Yuan develop a task oriented Clustering algorithm to reduce the granularity of Wireless sensor networks. They assume that each sensor node has a unique ID and holds its one-hop neighbor list. Each sensor node in one sensor network runs a specific task including one or more data to form a task data pattern [8]. The clustering algorithm in this paper is clustering algorithm is designed based on data which related to the task code in each sensor node. It uses Vector Quantization Algorithm to divide the sensor nodes into cells, and choose the sensor node which has the minimum node id to be the cluster head but ignoring the radio coverage which would cause the cluster head may be not able to reach some of sensor nodes in its domain.

Hsu, King, and Banerjee propose that a scheme that is able to find the neighbors on the rim of the irregular radio coverage region to narrow down the choices of the forwarders [9].

Cell ID has been distributed to each node before deployment; it is impractical for some applications which deploy nodes into inaccessible environment [1], however, when the environment is unknown or hostile such as remote harsh fields, disaster areas and toxic urban regions, sensor deployment cannot be performed manually [10].

3. Proposed Solution and Simulation Modle

A. Proposed Solution

In this section, a novel tree self-organized architecture for Wireless Sensor Network is introduced here.

In this architecture, the nodes are clustered into cells. And each cell is based on a tree topology [12] which helps to initiate the cell size automatically. Not like [1] needs cell id predetermined. We assume the each node does not know its location. Figure2 shows the initial state of the nodes.

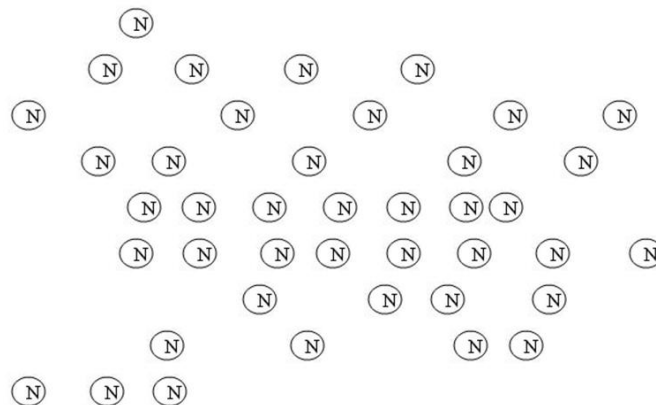


Figure 2: Initial State

The following process is called Active-Tree algorithm, all nodes will be processed to be a child of tree.

Step1. The Base Station which used to gather the information from the nodes will set up to active the closest nodes to be Root nodes. The level-id of the Root nodes is 0.

Step2. Root nodes gather the node id and signal length of the nodes which in their radio coverage and define this set as child-nodes. The level-id of these child-nodes is equal to the level-id of their parents' level-id plus 1.

Step3. Choose the longest child node as a child-root node. The level-id of these child-nodes is equal to the level-id of their parents' level-id plus 1.

Step4. The child-node use step2 to choose the longest node from it as the other child-root node.

Step5. Then the 2 child-root node repeats step2 and step3 and step5. Figure 3 shows the result after this process

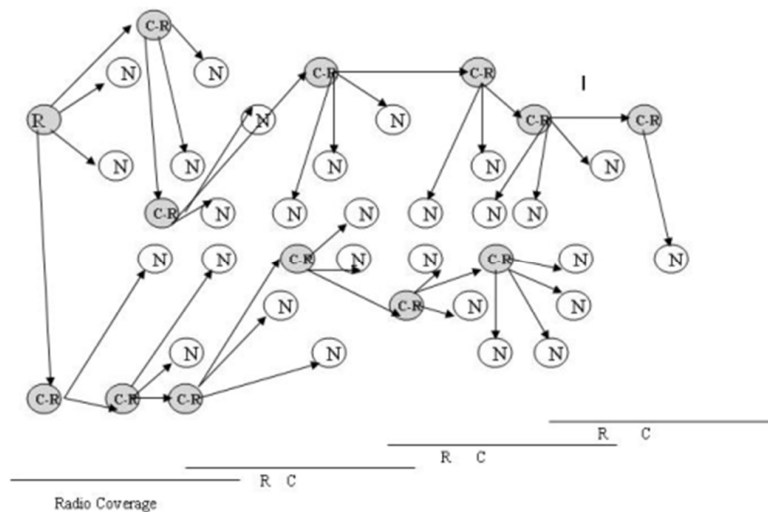


Figure 3: Shows the result after this process.

Active-Tree Algorithm:

Base Stations finds the closest sensor node

Sensor node = Root node;

Level-id = 0;

When still exist sensor node without Level-id

Do {**If** the length between Root node and Sensor node

<=Root node's radio coverage.

Then these sensor nodes Level-id = parent's Level-id + 1;

If the sensor node is the longest from its parent

The sensor node to be the childe-root node}

Till now, we could assume that each node has been classified as Root node, Child-Root node and Normal-Child node. We will use this hierarchy to clustering nodes to be cells. The clustering process is called Draw-Grid algorithm. However, in fact this algorithm is implemented into the step2 of Active-Tree algorithm, but not after the Active-Tree algorithm done. According to the radio coverage, each cell is an irregular grid from the Root node to the Child-Root node which is in its radio coverage but its children nodes are outside of radio coverage. If we get more than one Child-Root node like this, the one which is closest to Root will be the Root for the next cell. The And Root node will assign a cell-id to each node in the grid. The Figure4 shows how the nodes are clustered into cells.

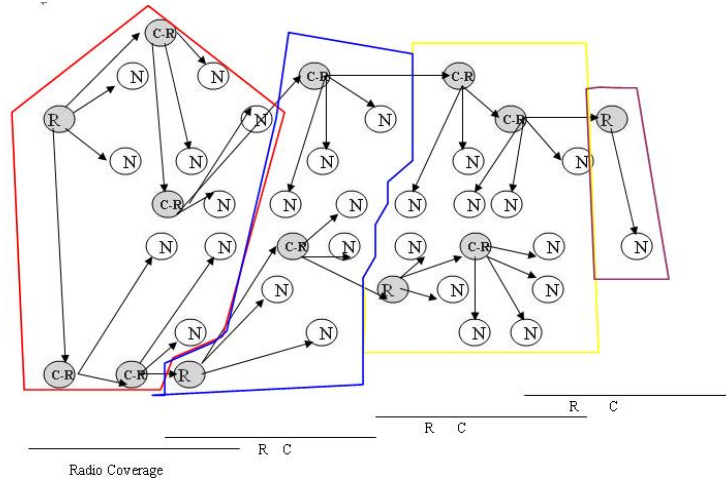


Figure 4: Clustered nodes

Draw-Grid Algorithm:

When still exists sensor node without Cell-id

Do {**If** (the child-root node is in the radio coverage of the cell's Root node) && (the its child-root node is out of the radio coverage)

Then from Root node to this child-root node is a cell. Assign cell Id to each sensor node. }

According to the above algorithm, each cell's Root is in the radio coverage of the 'upper-level' cell's Root. The Root of each cell is the management node for this cell. In each cell, the Root node (R) is on the top of the hierarchy and Child-Root (C-R) node is the second in the hierarchy; the last level is the Normal-Child node (N).

C-R node is responsible to gather messages from N nodes and to manage the N nodes. N nodes are used to process data, detect environment and set sleep and wakeup modes. When N nodes turn into

the sleep mode, then it needed to be wake up by other N nodes with same cell-Id. C-R node will hold its N nodes' node id in its one-hop neighbor list [8]. C-R node will also hold the list of the N nodes' energy status. And N nodes have the responsibility to report their energy status to C-R nodes every certain time. If C-R nodes do not receive energy status from some N nodes, C-R nodes will declare them as died nodes and erase them from the one-hop neighbor list.

Root node is responsible to gather messages from C-R nodes and to manage the C-R nodes. As C-R nodes also need to do management tasks, so C-R nodes consume energy more quickly. Root nodes have the task to re-select C-R nodes. C-R nodes will report their energy status to Root nodes periodically. Once C-R nodes' energy status reaches the threshold, Root nodes will reselect C-R nodes from their N nodes which have the most energy in the energy list. The threshold must be enough for C-R nodes to send their N nodes' node id and one-hop neighbor list to their substitute C-R nodes. However, there is a situation that some C-R nodes go to dead as external environment damage. Root nodes will have a list of their C-R nodes' C-R nodes. So in this situation, Root nodes will directly link the dead C-R nodes' C-R nodes to the substitute C-R nodes. That means, the dead C-R nodes' other N nodes will be discarded.

B. Simulation Model

In our simulation model, we define a certain schema according to our algorithm with particular characteristics:

- The task of the certain application is pre-embedded into the sensor nodes.
- Radio coverage is equal to the sensing node. And radio coverage is constant.

- The communication protocol could be CAN protocol or dedicated one [2].
- In our proposed architecture, C-R node gather data from its child nodes no matter it is also a C-R node or Root node from another cell. So we can say, it is one-hop communication, collisions are largely reduced. So we assume the response time is constant.
- Our proposed architecture can be used on many applications, especially for remote harsh fields, disaster areas and toxic urban regions which sensor deployment cannot be performed manually. Like earthquake and disaster detection. We define one earthquake or one diction as an “event”.
- We assume the transmitting, activating, receiving and managing energy consumption ratio is 10:1:4.5:3.
- We assume the transmitting time from one node to the other nodes with closest level-id is 0.6sec. We ignore the receiving process time.
- We assume the damage probability is 10 among 100 sensor nodes.
- We assume sleep and wake up mode used periodically. When N nodes stay in wake-up mode with no activity in 5secs, then turn to sleep mode until be woke up by other N nodes with same cell-Id for event happened. We the wake-up response time as 0.5 sec.
- We assume the total energy of each sensor node is 20.

Our architecture, each cell is a tree topology. For all sensor nodes, energy consumption are depends on the transmitting consumption, receiving consumption. For C-R nodes and Root nodes, energy consumption are also depends on management tasks. So frequency of event happens and management task frequency decides the whole energy consumption of the WSN.

We assume E_u represents utilized energy. E represents the whole energy of each sensor node. NC represents the number of the cells which we get after running our algorithms. In each cell, the response time to an event is:

$$TR = (\text{wake-up time}) 0.5\text{sec} \times \text{number of sensor nodes in the cell}$$

$$E_u = \sum C_{\text{active}} \times ER + \sum C_{\text{active}} \times ET + P_{\text{damage}} \times EM + \sum C_{\text{active}} \times EA$$

C_{active} represents the total number of the active sensor nodes in each cell. ER represents the energy used in receiving, here we assume the $ER = 4.5$. ET represents the energy used in receiving, here we assume the $ET = 10$. EM represents the energy used in receiving, here we assume the $EM = 3$. EM is related to the probability of damage, also related to low energy pick up process which should be calculated in the whole process. P_{damage} demonstrates the probability of damage.

Comparing with static architecture; static architecture has less flexibility. Especially, when damaging occurs, radio coverage constraints the backup nodes candidates. In our paper, the fixed size architecture also has Gateway same as R-node in our paper and they do the same management tasks. And Gateway of each cell could communicate. And each also has 1 Manager which is same as our C-R

node and they do the same management tasks. We will use the same schema and parameters for fixed size architecture.

But as the cell size is fixed, so every node's back up must be picked from the same cell. The position of gateway and manager is determined randomly. Also as gateway and manager may be not a one-hop communication, the routing of the gateway is also determined randomly.

Below is the assumption for our architecture:

Cell_Id	Role	Node_Id	Level_Id	Active or not	When be damaged
1	R	1	1	N/A	N/A
1	C-R	3	2	N/A	N/A
1	N	4	2	N/A	N/A
1	N	5	2	N/A	N/A
2	R	2	2	N/A	N/A
2	C-R	6	3	N/A	N/A
2	N	8	3	N/A	N/A
2	N	10	4	N/A	N/A
2	N	11	4	A	N/A
3	R	7	3	N/A	N/A
3	C-R	9	4	N/A	After 1sec
3	N	12	5	N/A	N/A
3	N	13	6	N/A	N/A
3	N	14	6	N/A	N/A
3	N	15	5	A	N/A

The architecture will be shown as below:

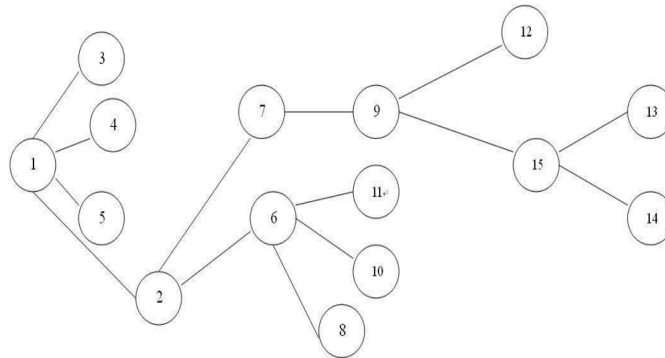


Figure 5: Scenario 1 Proposed architecture

Below is the assumption for our architecture:

Cell_Id	Role	Node_Id	Active or not	When be damaged
1	G	1	N/A	N/A
1	M	3	N/A	N/A
1	N	4	N/A	N/A
1	N	5	N/A	N/A
2	G	2	N/A	N/A
2	M	6	N/A	N/A
2	N	8	N/A	N/A
2	N	10	N/A	N/A
2	N	11	A	N/A
3	G	7	N/A	N/A
3	M	9	N/A	After 1sec
3	N	12	N/A	N/A
3	N	13	N/A	N/A
3	N	14	N/A	N/A
3	N	15	A	N/A

The architecture will be shown as below:

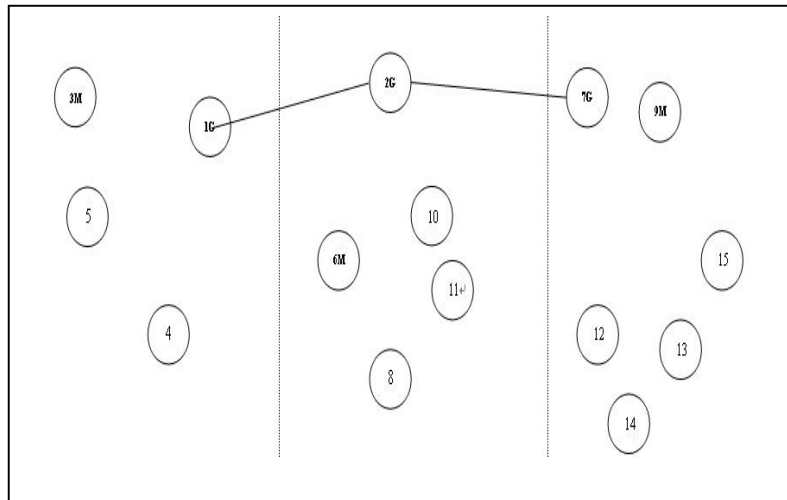


Figure 6: Scenario 2 (Static architecture)

4. Simulation Result

According to our assumption and the table shows above. In the first scenario which is played by our architecture, the sensor node 11 and 15 are activated by the event. And sensor node 1 is the closest one to the base station. Firstly, 15 and 11 will activate the other normal sensor nodes in the same cell, like 15 will activate 12, 13, and 14 from the sleep node. And 11 will activate 10 and 8, as 6 is the C-R node which is responsible to gather the data from 10, 11 and 8. This action will take $TR = 3 \times 0.5\text{sec} + 2 \times 0.5\text{sec} = 2.5\text{sec}$. $E_u = 4 \times E_R + 3 \times E_R = 31.5$. Sensor nodes 8, 10, 11, 13, 14 and 15 use energy 4.5 to receive the data respectively. This action takes the same time and energy for both scenarios. After 1 second from the beginning, the sensor node 9 is damaged. 13, 14 have been active.

In scenario1, 9 damaged, sensor node 7 will choose 12 to be the C-R node. As 12 has more energy at that time. Then 12 will start to gather data from 13,14,15. 12 needs to take $3 \times ER = 13.5$ to receive data from 13,14 and 15. Then 12 sends data to 7, 7 takes $ET = 10$ to receive the data. Next 0.5 second, 7 detects 12 is in low energy, at that time 7 has energy 6. Sending data to 2 has higher priority, so 7 sends the data out firstly. After that, 7 only has energy 3. Then 2 detects 7 in low energy. In the cell which R node is 2, 11 uses 1 second active the 8 and 10. 6 gathers data from 8 and 10. And 6 has enough energy to send data to the 2. So when 2 detects 7 in low energy, 2 still has energy 15.5 to do management task. So 2 takes energy 3 to choose 10 which left energy 5 to replace 7. Then 10 picks up 13 which left 5.5 as the C-R node. After that, 2 sends the data to the Root node 1 which will send the data directly to the Base station.

The Figure 7 below will show the energy consumption with time.

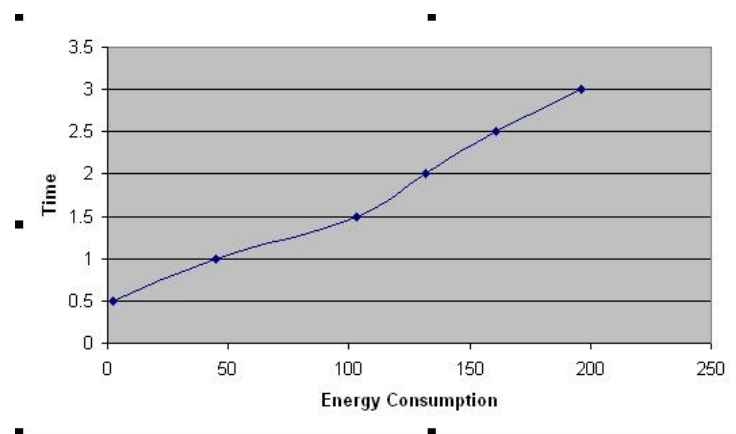


Figure7: Energy Consumption for scenario1

In scenario2 has the same energy consumption with scenario 1 before 3 seconds. After 2 sends data to 1, 2 needs to pick up Gateway to cell3 to replace 7. 13,14,15 are the candidates. 13 has more energy

comparint with 14 and 15. But 13 in our assumption is out of the radio coverage of 2 and 10. So there generates 2-hop communication. 2 will choose 10 and then 10 chooses 14 and 14 chooses 13. As we set a low total energy to each sensor node, so in this situation, we will lose cell3 to investigate its region. We will put this lose to the energy consumption as this energy is unable to be used.

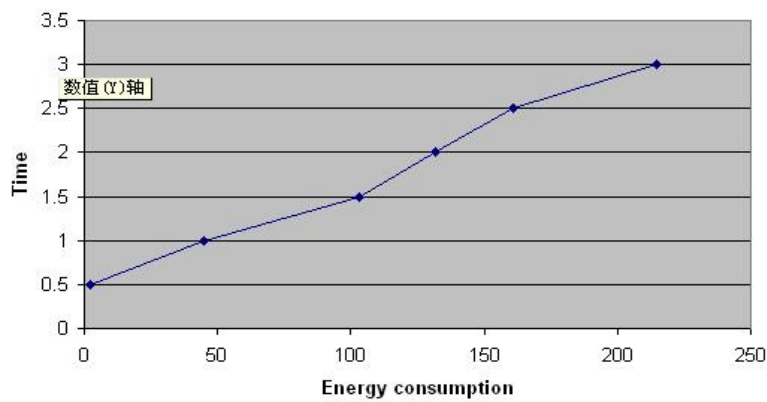


Figure 8: Energy Consumption for Scenario 2

From the simulation results, as the flexibility of our proposed architecture, we could archive less energy consumption.


5. Conclusion

This paper proposes a novel architecture for WSN. Although there are many researches for this aspect, our architecture is suitable for more applications especially for the applications apply for harsh environment situation. Our architecture is a dynamic architecture and changes with time changes. Simulation results shows that our dynamic architecture with appropriate management tasks distribution could archive a better energy saving in given time than static architecture. Future work will be focused on investigate the MAC (Medium Access Control) protocols to optimize our solution.

References

- [1] M.Asim, H.Mokhtar, M.Merabti, “Cellular self-organization architecture for wireless sensor networks” PGNet 2008, Liverpool UK, June 2008.
- [2] Mario Limon Mendoza,Victor Hugo Zarate Silva, “Evaluation of Self-organization architecture for Special WSN using geometrical arrays of nodes”,2009 First International Conference on Computation Intelligence, Communication Systems and Networks, 2009
- [3] F. L. LEWIS, “Wireless Sensor Networks” The University of Texas at Arlington, <http://arri.uta.edu/acs>.
- [4] Wendi Rabiner Heinzelman, Joanna Kulik,Hari Balakrishnan “Adaptive Protocols for Information Dissemination in Wireless Sensor Networks” Mobicom ‘99 Seattle Washington USA ,ACM ,1999
- [5] Partial Fulfillment, “ALGORITHMS FOR SELF-ORGANIZING WIRELESS SENSOR NETWORKS,” ElMoustapha Ould-Ahmed-Vall, Georgia Institute of Technology,May 2007
- [6] Hongke zhang, stephan olariu, jiannong cao and david B. johnson Mobile ad-hoc and sensor networks. Beijing, China, in December 2007.
- [7] John Wiley & Sons Ltd, “Self organization in sensor and actor networks”, The Atrium, Southern Gate, Chichester , West Sussex, England.
- [8] Shangwei Duan and Xiaobu Yuan, “Exploring Hierarchy Architecture for Wireless Sensor Networks Management”,School of Computer Science, University of Windsor, Canada 2006,1-4244-0340-5/06 02006 IEEE
- [9] Li-Chun Hsu, Chung-Ta King, Amit Banerjee “On Broadcasting in Wireless Sensor Networks with Irregular and Dynamic Radio Coverage”, Department of Computer Science,

- National Tsing Hua University, 2007 International Conference on Parallel Processing ,2007
- [10] Pradnya Gajbhiye, Anjali Mahajan” A Survey of Architecture and Node deployment in Wireless Sensor Network”, Post Graduate Department of Computer Science and Engineering G.H.Raisoni College of Engineering, Nagpur, M.S., INDIAr, Applications of Digital Information and Web Technologies, 2008. ICADIWT 2008. First International Conference, pp426-430, IEEE,2008
- [11] G“unter Reise and Gerald Matz “Clustered Wirless Senso Networks for Robust Distributed Field Recognition Based on Hybrid Shift-Invariant Spaces”, Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology 978-1-4244-3696-5/09/ pp.67-70 IEEE 2009
- [12] Yosi Ben-Asher, Moran Feldman and Sharoni Feldman, “Ad-Hoc Routing Using Virtual Coordinates Based on Rooted Trees” Computer Science Department, Haifa University, Proceedings of the IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC’06),IEEE,2006



**Drift Region Voltage Drop
in SiC VDIMOS - The
Influence of Anisotropy**

4

Drift Region Voltage Drop in SiC VDIMOS - The Influence of Anisotropy

Abedalhakem Alkowash
University of Sabratha, Faculty of Engineering, Sabratha

Abstract

Silicon Carbide (SiC) is an important indirect wide band gap semiconductor with outstanding electronic properties. This work focuses on an investigations of silicon carbide (SiC) based vertical Double Implanted Metal Oxide Semiconductor Field Effect Transistor (DIMOSFET). Silicon Carbide (4H SiC as well as 6H SiC) is known to be highly anisotropic material. Among others, the transport parameters like low field mobility and saturation velocity are considerably different in c direction compared to a and b directions. The aim of this paper is to investigate the influence of variation of mentioned parameters, as well as the variation of parameters describing specific model, on "drift region voltage drop" in vertical DIMOS structure.

Keywords: DIMOSFET, SiC, Transport parameters, SiC anisotropy, drift region voltage drop.

1. Introduction

In the last fifteen years, silicon-carbide (SiC) has emerged as one of the most promising semiconductor materials for design and fabrication of different devices widely used in microelectronics and nanoelectronics [1]. It turned out to be extremely successful in overcoming the limitations and shortcomings of Si-based devices. Due to its superior thermal and electric properties and satisfactory transport performances if compared to silicon, silicon-carbide has become a promising candidate for use in high-temperature and high-power, as well as in switching devices. One of the best developed structures is a vertical double implanted metal oxide semiconductor (field effect) transistor. It can carry large currents during "on" state, while its vertical section (drift region) is capable of sustaining large blocking voltages in the "of" state [2]. In spite of many papers reporting on the investigation of such structures, their level of development can still be regarded as modest and insufficient. Therefore, the aim of this paper is to include the anisotropic character and temperature dependence of transport parameters such as low field mobility and saturation velocity, as well as the behavior of parameter β describing specific features of drift-diffusion transport model in silicon and similar materials [3]-[13]. The focus of the paper is on vertical section (drift region), and its investigation has been carried out analytically as long as possible. The horizontal channel itself has been described and modeled in many conventional MOS-structures, hence no special attention has been paid to it in this paper.

2. Theoretical Basis of "Drift" Region Model

The conventional vertical SiC DIMOSFET investigated in this paper is shown in figure 1. The domain of interest is a vertical "drift" region and it is assumed to be divided into three sections A, B, C (figure 1.). The total current in each of these three regions should be the same. It is vertical and equal to drain current. The previous analyses of "drift" region have led to the development of its widely accepted model

suggesting following expressions for voltage drop in each of its sections [2]:

$$V_B = \frac{I_D \cdot W_A}{e \cdot W \cdot N_D \cdot \mu_n \cdot L_d - \frac{I_D}{E_C}} \quad (1a)$$

$$V_B = \frac{I_D}{2 \cdot W \cdot e \cdot N_D \cdot \mu_n \cdot ct g \alpha} \cdot \ln \frac{e \cdot W \cdot N_D \cdot \mu_n \cdot (L_d + L_P) - \frac{I_D}{E_C}}{e \cdot W \cdot N_D \cdot \mu_n \cdot L_d - \frac{I_D}{E_C}} \quad (1b)$$

$$V_C = \frac{I_D \cdot \left(W_T - W_j - W_d - \frac{L_P}{2} \cdot t \alpha \right)}{W \cdot e \cdot N_D \cdot \mu_n \cdot (L_d + L_P) - \frac{I_D}{E_C}} \quad (1c)$$

where W_A is the accumulation region depth, W_T is the total thickness of epilayer, L_d is the accumulation region length, L_P is the p-body length, N_D is the concentration of the ionized donors, $W_A = W_j + W_d$, μ_n is the low field mobility, $E_C = v_s / \mu_n$. Other geometric parameters are labeled in figure 1.

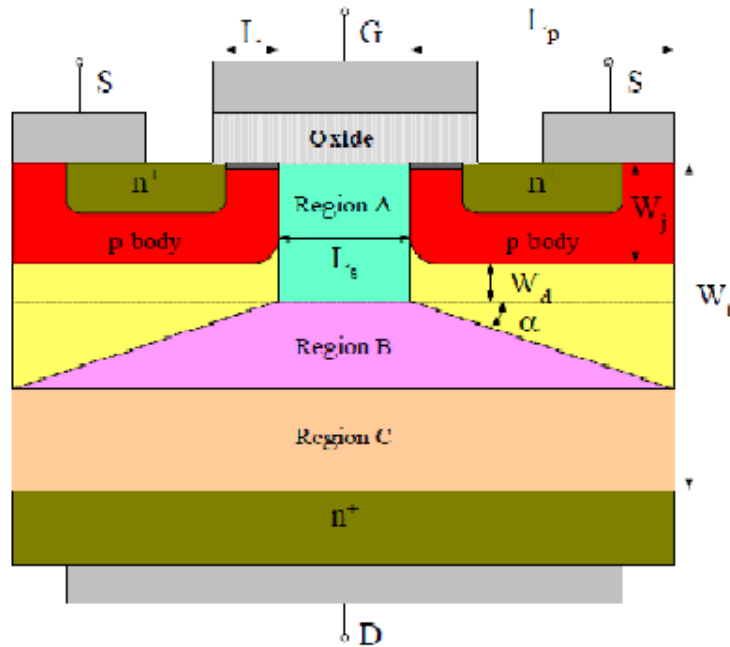


Figure 1: Cross section of the conventional vertical SiC DIMOSFET

The drift region itself is assumed to be n-doped by concentration N_D and completely ionized. The above expressions (1) have been obtained by means of drift-diffusion model and the most prominent formula [2].

$$v = \frac{\mu_n \cdot E_x}{1 + \frac{\mu_n}{v_s} \cdot E_x} = \frac{\mu_n \cdot \frac{dV}{dx}}{1 + \frac{\mu_n}{v_s} \cdot \frac{dV}{dx}}; \quad E_c = \frac{v_s}{\mu_n} \quad (2)$$

Describing carriers' transport in silicon unipolar devices. This model is realized from equation (2) with the limited number of parameters to be fitted to the experimental data. It also provides the analytical treatment of its expressions far enough, as well as its straightforward implementation in more complex circuit simulation. But this model also suffers from serious shortcomings to be dealt with in this paper. First, silicon-carbide (both 4H-SiC and 6H-SiC) is known to be a highly anisotropic material, i. e. its transport properties, and hence corresponding coefficients, are different in \underline{c} -direction compared to \underline{a} and \underline{b} directions. These coefficients also depend on the operating temperature, what begins to play an important role having on mind that silicon-carbide devices are supposed to be used in the high temperature range. The model can also be improved by introducing the parameter β responsible for transport characteristic $v(E)$ fine tuning. The parameter itself is also anisotropic and shows the considerable temperature dependence [3], [4].

The modern literature concerning drift-diffusion model suggests replacing the expression (2) by a more detailed one [3]:

$$v = \frac{\mu_n \cdot E_x}{\left[1 + \left(\frac{\mu_n}{v_s} \cdot E_x\right)^\beta\right]^{1/\beta}} = \frac{\mu_n \cdot \frac{dV}{dx}}{\left[1 + \left(\frac{\mu_n}{v_s} \cdot \frac{dV}{dx}\right)^\beta\right]^{1/\beta}} \quad (3)$$

This formula is also consistent with the main features of hydrodynamic model, which is based on the Boltzmann transport equation and its moments leading to balance equations. Therefore, the relation (3) can be regarded as a product of serious theoretical consideration rather than the consequence of simple fitting to the experimental data. Once having this formula adopted, its implementation in the "drift region" model causes no unavoidable difficulties in its accumulation layer A:

$$I_D = e \cdot W \cdot N_D \cdot L_d \cdot v$$

$$= \frac{e \cdot N_D \cdot \mu_n \cdot W \cdot L_d \cdot \frac{dV}{dx}}{\left[1 + \left(\frac{\mu_n}{v_s} \cdot \frac{dV}{dx}\right)^\beta\right]^{1/\beta}} \quad (4a)$$

Straight-forward leading to expressions:

$$\frac{dV}{dx} = \frac{I_D}{\left[(e \cdot W \cdot N_D \cdot L_d \cdot \mu_n)^\beta - \left(\frac{\mu_n \cdot I_D}{v_s}\right)^\beta\right]^{1/\beta}} \quad (4b)$$

and:

$$V_A = \frac{I_D \cdot W_A}{\left[(e \cdot W \cdot N_D \cdot L_d \cdot \mu_n)^\beta - \left(\frac{\mu_n \cdot I_D}{v_s} \right)^\beta \right]^{1/\beta}} \quad c)$$

The calculation of the voltage drop over varying cross-section layer B causes some more effort:

$$I_D = \frac{e \cdot N_D \cdot \mu_n \cdot W \cdot [L_d + (x - V_A) \cdot ctg\alpha] \cdot \frac{dV}{dx}}{\left[1 + \left(\frac{\mu_n}{v_s} \cdot \frac{dV}{dx} \right)^\beta \right]^{1/\beta}} \quad (5a)$$

and hence:

$$- \int_{W_A}^{W_A + \frac{L_p}{2} \cdot ctg\alpha} \frac{I_D \cdot dx}{\left\{ [e \cdot W \cdot N_D \cdot \mu_n \cdot (I_D + 2 \cdot (x - W_A) \cdot ctg\alpha)]^\beta - \left(\frac{\mu_n \cdot I_D}{v_s} \right)^\beta \right\}^{1/\beta}} \quad V_B \quad (5b)$$

The above expression becomes simpler after introducing a new variable:

$$e \cdot W \cdot N_D \cdot \mu_n \cdot [L_d + 2 \cdot (x - W_A) \cdot ctg\alpha] = z \quad (5c)$$

whit a result:

$$V_B = \frac{I_D \cdot \operatorname{tg} \alpha}{2 \cdot e \cdot W \cdot N_D \cdot \mu_n} \int_{e \cdot W \cdot N_D \cdot \mu_n \cdot L_d}^{e \cdot W \cdot N_D \cdot \mu_n (L_d + L_p)} \frac{dz}{\left[z^\beta - \left(\frac{\mu_n \cdot I_D}{v_s} \right)^\beta \right]^{1/\beta}} \quad (5d)$$

The analytical evaluation of V_B would be possible if we used approximate values for coefficient β . But, the procedure is tedious and hence not worth performing. It is much more convenient to calculate V_B by means of any of adequate software tools (MATHLAB, MATHEMATICA...). The expression for V_B can be reorganized as follows:

$$V_B = \frac{I_D \cdot \operatorname{tg} \alpha}{2 \cdot e \cdot W \cdot N_D \cdot \mu_n} \cdot \{f[e \cdot W \cdot N_D \cdot \mu_n (L_d + L_p)] - f(e \cdot W \cdot N_D \cdot \mu_n \cdot L_d)\} \quad (6a)$$

With the abbreviation:

$$f(z) = \int \frac{dz}{\left[z^\beta - \left(\frac{\mu_n \cdot I_D}{v_s} \right)^\beta \right]^{1/\beta}} \quad (6b)$$

Similar to the accumulation layer A, the voltage drop over the region C can easily be calculated:

$$V_C = \frac{I_D \cdot \left(W_T - W_A - \frac{L_p}{2} \cdot tg\alpha \right)}{\left\{ [e \cdot W \cdot N_D \cdot \mu_n (L_d + L_p)]^\beta - \left(\frac{\mu_n \cdot I_D}{v_s} \right)^\beta \right\}^{1/\beta}}$$

drop over the whole drift region turns out to be:

$$V_{drift} = V_A + V_B + V_C \quad (8)$$

and is calculated by means of relations (4c), (5d) and (7). In each of these expressions, low field electron mobility μ_n , saturation velocity v_s and the curvature coefficient β in overall formula (3), play an important role. As already mentioned these coefficients show considerable level of anisotropy, i. e. their values for c - direction and for a and b directions are remarkably different. Same parameters also depend on temperature. Low field mobility has been taken from state - of - the art measurements and lined by the following expressions [4], [5]:

$$\mu_{n\perp c} = 40 + \frac{350 \cdot \left(\frac{T}{300} \right)^{-2.4} - 40}{1 + \left(\frac{T}{300} \right)^{-0.76} \cdot \left(\frac{N_D}{2 \cdot 10^{17}} \right)^{0.76}} \frac{cm^2}{Vs} \quad (9a)$$

$$\mu_{n\parallel c} = 48 + \frac{1140 \cdot \left(\frac{T}{300} \right)^{-2.4} - 48}{1 + \left(\frac{T}{300} \right)^{-0.76} \cdot \left(\frac{N_D}{2 \cdot 10^{17}} \right)^{0.76}} \frac{cm^2}{Vs} \quad (9b)$$

The high field mobility parameters ν_s , β are extracted from the full band Monte-Carlo simulations and can be comprised by following formulae:

$$\nu_{s\perp c} = \frac{2.77 \cdot 10^7}{1 + 0.23 \cdot e^{\left(\frac{T}{600}\right)}} \frac{cm}{s} \quad (10a)$$

$$\nu_{\parallel c} = \frac{2.55 \cdot 10^7}{1 + 0.30 \cdot e^{\left(\frac{T}{600}\right)}} \frac{cm}{s} \quad (10a)$$

$$\beta_{\perp c} = 0.60 + 1 \cdot 10^{-3} \cdot T \quad (11a)$$

$$\beta_{\parallel c} = 0.01 + 1 \cdot 10^{-3} \cdot T \quad (11b)$$

The expressions (9), (10) and (11) have been inserted into main relations of the developed model ((4c), (5d) and (7)) and so far the drift region voltage drops in various cases have been calculated.

3. Numerical Results and Discussion

The model described above is accompanied with structure parameters used in this simulation (and usually met in the relevant references [2]):

$$W = 400\mu m \quad L_d = 20\mu m$$

$$W_A = 32\mu m \quad L_p = 50\mu m$$

$$W_T = 70\mu m$$

$$N_D = 4 \times 10^{21} m^{-3}$$

The model itself regards some general remarks:

- a) for a specific value of drift current I_D , the drift region voltage drop V_{DRIFT} is still proportional to μ_n^{-1} ;
- b) drift region voltage drop V_{DRIFT} should decrease if saturation velocity v_s increased;
- c) current parameter β models the shape of the $I_D(V_{DRIFT})$ (V_{DRIFT}) characteristic and thus makes a considerable influence on specific values.

The results of the calculation performed according to the proposed model are given in figures 2 and 3. The figure 2 exposes the effect of anisotropy for different values of temperature with the main consequences:

- for each value of temperature, the "normal" orientation ($\perp c$) provides larger values of quasi-saturation drain current I_{Dqs} compared to parallel orientation ($\parallel c$);
- for each value of temperature, the characteristic $I_D(V_{DRIFT})$ has a bigger slope for small values of drain current in the case of "parallel" orientation ($\parallel c$) than in the case of "normal" orientation ($\perp c$);
- for higher values of drain current, the situation turns out to be quite different (due to the anisotropy of the curvature coefficient β); the slope of the characteristic suddenly becomes smaller in the case of "parallel" orientation ($\parallel c$) compared to the case of "normal" orientation ($\perp c$), thus resulting in the smaller values of quasi-saturation current for "parallel" orientation ($\parallel c$) compared to the latter one ($\perp c$).

The figure 3 shows $I_D(V_{drift})$ characteristic calculated for different values of temperature for each of the investigated orientations. In both cases, for the same value of drain current, the drift region voltage drop remarkably increases with the increase of temperature, mostly due to the μ_n^{-1} dependence described above.

So far only the characteristic drain current versus drift region voltage drop has been considered. Naturally, drain current has its upper cut-off caused by the channel saturation appearance. For gate voltage values usually met in such structures the section of the characteristic concerning small values of drain current (and drift region voltage drop consequently) is expected to play an important role. The drain current values comparable to quasi-saturation one (inevitably accompanied for greater gate voltages) appears rather as an exception.

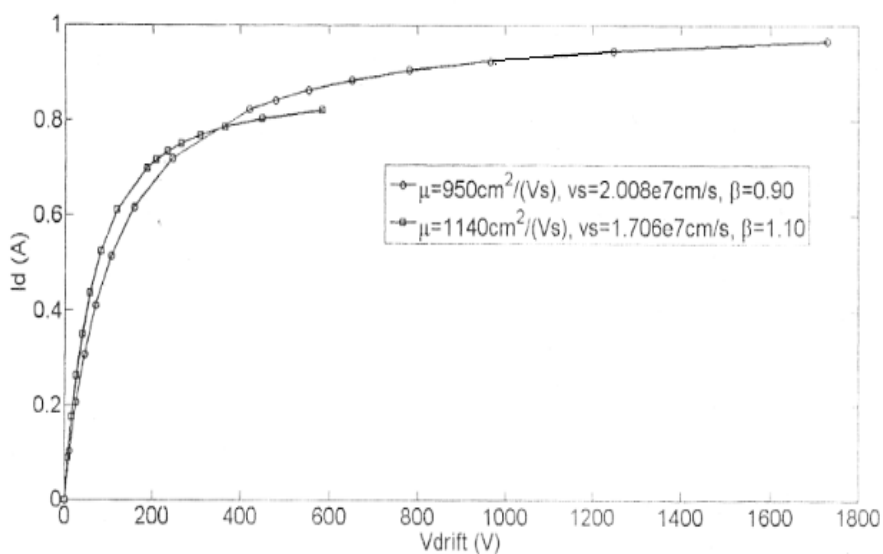


Figure 2a: Drain current versus drift region voltage drop- the influence on anisotropy (temperature=300K)

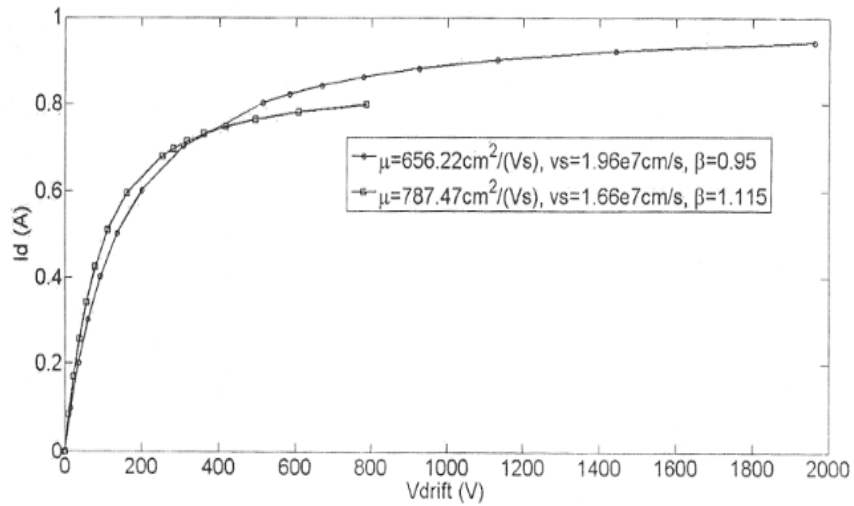


Figure 2b: Drain current versus drift region voltage drop- the influence on anisotropy (temperature=350K)

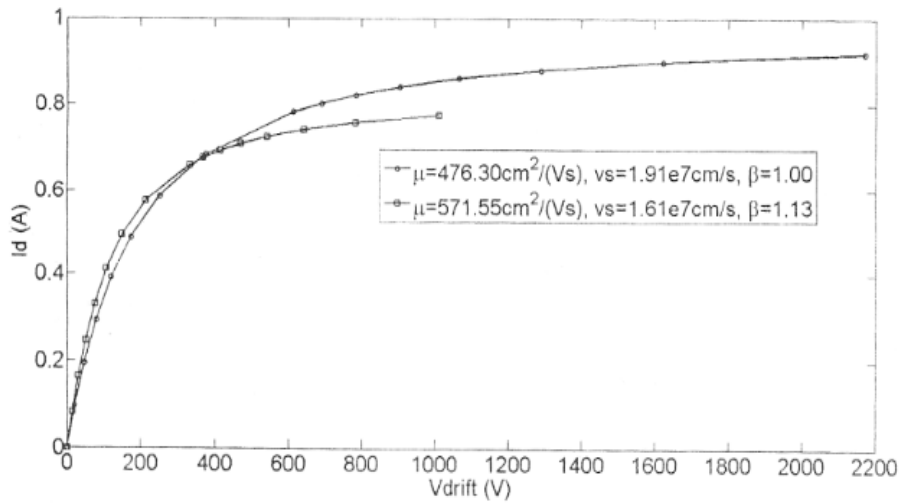


Figure 2c: Drain current versus drift region voltage drop- the influence on anisotropy (temperature=400K)

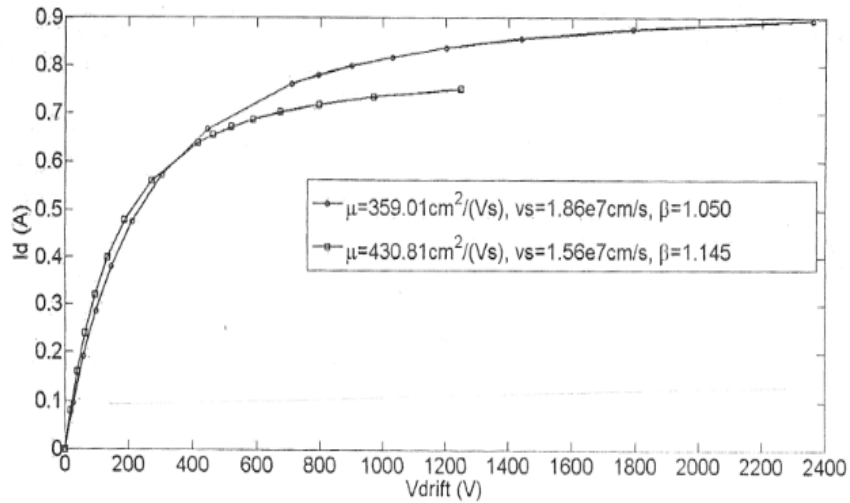


Figure 2d: Drain current versus drift region voltage drop- the influence on anisotropy (temperature=450K)

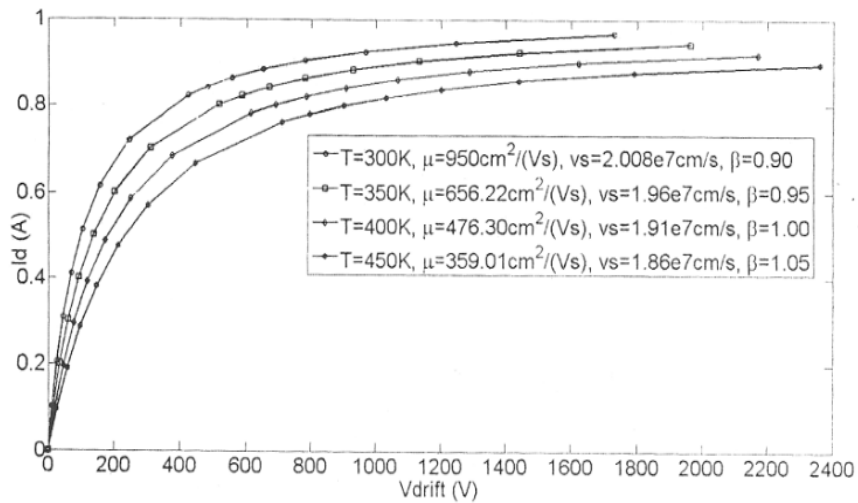


Figure 3a .Drain current versus drift region voltage drop for different values of temperature ("normal" orientation \perp c))

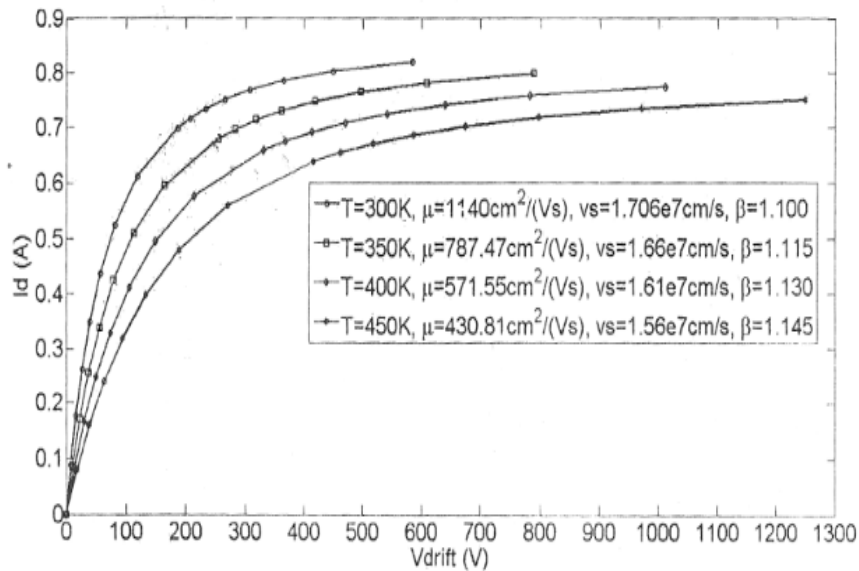


Figure 3b. Drain current versus drift region voltage drop for different values of temperature ("parallel" orientation (c))


4. Conclusions

This paper has given its contribution to the improvement of the existing theory of VDIMOS by introducing more accurate transport model including the fine tuning curvature coefficient. By means of such developed model, the drift region voltage drop has successfully been calculated. The coefficient, as well as saturation velocity and low electric field mobility possess a considerable level of anisotropy, which mainly affected the steepness of calculated current-voltage characteristics. It has also been proven that the change of working temperature strongly affected the investigated characteristics. This way obtained expressions form an adequate basis for constructing final model of VDIMOS, i.e. unification with the model describing channel.

References

- [1] MD Hasanuzzaman, Sred K. Islam, Leon M. Tolbert, Burak Ozpineci: "Design, Modelling , Testing and Spice Parameter Extracting of Transistor in 4H-Silicon Carbide", *Frontiers in Electronics*, pp. 733.-746.,2005.
- [2] A. Elasser and T.P. Chow, "Silicon Carbide Benefits and Advantages for Power Electronics Circuits and Systems," *Proc. IEEE*, 2002: 90(6): 969-86.
- [3] Zau J., D kotchetkov and AA Balankin "Thermal conductivity of GaN films: Effects of impurities and dislocations" *Journal of Applied Physics* 92, (2002) p.2534
- [4] A. Power and L. Rowland: "SiC Materials- Progress, Status and Potential Roadblocks, *Processing IEEE*, pp.942.-955.,90(6).2002.
- [5] J.R. Brews, A charge sheet model of the MOSFET, *Solid State Electron.* 21 (1978) 345–355.
- [6] M . Roschke, F. Schwierz, Electron mobility model for 4H, 6H and 3C SiC, *IEEE Trans. Electron Dev.* 48 (2001) 1442–1447.
- [7] Hasanuzzaman M, Islam SK, Tolbert LM, Ozpineci B. Model simulation and verification of a vertical double implanted (DIMOS) transistor in 4H-SiC, In: *Proc the IASTED International Conference on Power System Palm Spring 2003* p. 313–6.
- [8] Nakashima and H.Harima, "Characterization of structural and electrical properties in SiC by Raman spectroscopy", in *Inst. Phys. Conf. Series*, p. 269,(1969).
- [9] T. Kinoshita, K. Muto, M. Schadt, G. Pensl, and K. Takeda,"Calculation of the anisotropy of the Hall mobility in N-type 4H- and 6H-SiC", *Mat. Sci. Forum*, vols.264-268,p.295,1998.
- [10] M. Schadt, G. Pensl. R . P. Devaty, W, J , Choyke, R. Stein. And D. Stephany, "Anisortropy of the electron Hall in 4H,6H,

- Choyke, and 15R silicon carbide", *Appl. Phys. Lett.*, vol. 65, p. 3120.1994.
- [11] Krishna Shenai, Effect of P-Base sheet and contact Resistances on Static current-voltage Characteristics of Scaled low-voltage Vertical Power DMOSFETs, *IEEE Electron Devices Letters*, no. 6, pp. 270-272,1991.
- [12] Hasanuzzaman M, Islam S K, Tolbert L M and Ozpineci B 2002 Analytical modelling of vertical double implanted power MOSFET (DIMOS) in 4H-SiC *Proc. of Connecticut Symp. on Microelectronics and Optoelectronics Connecticut Marc.*
- [13] Hasanuzzaman M D, Islam S K, Tolbert L M and Ozpineci B 2005 Design, modeling, testing and spice parameter extraction of DIMOS transistor in 4H-SiC *Front. Electron.* 733–46.



**Assessment of different Growth
Techniques of Strained Germanium
Heterostructures for Electronic and
spintronic Devices**

5

Assessment of different Growth Techniques of Strained Germanium Heterostructures for Electronic and spintronic Devices

A.H. A. Hassan
Department of Physics, Faculty of science,
University of Tripoli.
amnahassan53@yahoo.co.uk

A. Diyaf
Department of Physics, Faculty of science,
University of Tripoli.

U. Elfurawi
Department of Physics, Faculty of science,
University of Tripoli.

A. E. Abubkr
Department of Physics, Faculty of science,
University of Tripoli.

Abstract

This paper, emphasis different growth techniques of two-dimensional hole gas of strained germanium (sGe) heterostructure, molecular beam epitaxy (MBE) and chemical vapor deposition (CVD). sGe heterostructure has become an important material as a replacement material to Silicon in P-type devices because of its higher hole mobility and lower effective mass. Researchers study this material in terms of electrical and spintronic devices according to technology demands for devices with higher efficiency and low power consumption. High hole mobility up to $1 \times 10^6 \text{cm}^2/\text{Vs}$ at temperature of 1.5 K has been reported for normal structure declaring high quality samples with low density dislocation and low interface roughness.

These samples were grown using Reduced Pressure Chemical Vapour Deposition (RP-CVD) indicating high purity system and the affordability of this technique in terms of electronic and spintronic devices.

Keywords Strained Germanium, MBE, CVD, High purity system.

1. Introduction

Although the first transistor was made of germanium (Ge), silicon (Si) is the main component in device fabrication mainly due to the good quality of the interface of Si with its natural oxide (SiO_2), and it is abundant compared to Ge therefore cheaper. On the other hand, Ge does not possess a good interface with its natural dielectric, GeO_2 . Nevertheless, Ge is again considered as a significant candidate for future Metal-Oxide-Semiconductor Field Effect Transistor (MOSFETs), as it has a higher bulk mobility being more than 4 times that of Si for holes and 2 times for electrons as seen in Table 1.

Gordon Moore in 1965[1] has observed that the number of transistors on an integrated circuit doubles approximately every couple of years, an observation that become known as Moore's Law. This remarkable feature has been achieved through a phenomenon known as scaling a systematic reduction of the MOSFET dimensions. Which, in turn has been attained by advancements in device fabrication techniques and, the advancements in lithographic techniques that helps in decreasing dimensions to be determined [1].

The continued performance enhancements in Complimentary Metal-Oxide-Semiconductor (CMOS) circuits cannot be achieved by scaling alone. Indeed, a major concern is the cost of new fabrication plants. However, of even greater importance could be that MOSFET dimensions are rapidly approaching a regime where the key device features consist of just a few hundred atoms. In this regime quantum

effects such as tunneling become significantly problematic, resulting in higher leakage currents and power consumption. There are many materials with intrinsically superior electrical properties compared to silicon, which have found roles in the market applications such as optical devices and the high frequency components of mobile communication devices. One of the most promising materials is Germanium (Ge).

Strained Si technology has been extensively used to improve the performance of advanced Si integrated circuit recently. However, due to the strong demand of high drive current to increase circuit speed, Si has reached its limit especially in p-type MOSFET. For that reason, sGe becomes a promising alternative to Si CMOS devices. Introducing strain to Ge leads to the effective mass reduction and mobility enhancement. The strain also splits the heavy_light hole bands further reducing the (100) in plane effective mass and the interband scattering as well [2]. High dielectric constant materials should be applied between the gate and the substrate such as SiO₂ and HfO₂ in terms of minimising the power consumption of MOSFETs that accounted for gate current leakage [1]. Germanium has emerged as a good alternative to silicon in terms of p-type MOSFETs because it comes from the same IV group as Si, but with a smaller bandgap (0.67eV) and significantly higher hole mobility as shown in Table (1).

High hole mobility devices required a good hole confinement in Ge layer, since the latter grown directly on Si is just thermally stable below 1nm thickness, which is not suitable for device application. Introducing SiGe layer virtual substrate (VS) on Si helps to increase the thickness of Ge layer and improves the hole confinement. This paper aims to introduce some of a significant research area in modulation doped strained Ge heterostructures in terms of improving the quality of the structure by employing different growth techniques, and investigating their effect on structural characterization such as

channel thickness, strain and scattering mechanism of the carriers in Ge quantum well.

Table 1: Si, Ge and GaAs selected property comparison at 300K

	Si	Ge	GaAs
Mobility(cm²/Vs)	$\mu_e=1500$ $\mu_h=450$	$\mu_e=3900$ $\mu_h=1900$	$\mu_e=8500$ $\mu_h=340$
Lattice constant (Å°)	5.431	5.646	5.653
Energy Gap (eV)	1.12	0.67	1.42
Intrinsic carrier concentration (cm³)	1.0×10^{10}	2.0×10^{13}	2.1×10^6

2. Discussion

The key growth technique for Heterostructures including strained Germanium quantum well is Molecular Beam Epitaxy (MBE) and Chemical Vapor Deposition (CVD). Regarding to the requirement of high mobility devices that raise their speed and performance, scientists attempt to acquire high quality material with few defects. Different efforts have been made to investigate the most effective growth technique to give low threading dislocation density, low surface roughness and the highest mobility. There are two main ways of epitaxial growth:

The main characteristic of the MBE growth technique [3, 4] is that it has independent control of growth parameters, low controllable growth temperatures enabling minimisation of solid state out diffusion and

auto doping, also the low growth rates and shutter control permits thin, highly uniform layers to be grown. Whilst solid source MBE (SS-MBE) has demonstrated high wafer uniformity and ultra-sharp doping profiles, the method does have some limitations. A major limitation is the lack of in-line calibration is particularly missed during the growth of thick (several microns) epilayers and arises from the depletion of the solid-source material. Typically, for the growth of SiGe epilayers, the solid atomic sources are evaporated via electron beam impingement. However, as the sources are consumed material flux calibrations are required on a rather frequent basis which is time consuming. SS-MBE is therefore quite a slow growth process and is mainly reserved for research rather than industry.

CVD process offers high growth rates, simultaneous growth of wafers, and it is more stable and reproducible for thick structures, a gas containing the material wished to be deposited, known as a "precursor", is passed over a heated substrate [5]. Under the correct conditions the gas will chemically react with the substrate and deposit epilayers. The standard chemical gases used for the SiGe system are silane and germane, which are passed down a furnace tube using a carrier gas, typically hydrogen. In addition, the growth of thick layers is possible since the gas sources are effectively infinite and the gas ratios can be constantly maintained and monitored throughout the growth via the use of mass flow control units. It is therefore the technique favoured by industry, and the major different between these two Technique are summarized in Table 2.

Table 2: The major different between CVD and MBE growth technique

CVD	MBE
Pressure around 0.5-760 torr	Requires UHV with Pressure below 10^{-8} torr
High growth rate	Low growth rate (1 atomic layer or less at a time) = high uniform layers
Useful for experiments and mass production	Useful for research lab experiments

Growth technique is the main responsible parameters to get an optimal characterisation in the samples, that affects the purity of the channel, the defects like surface and interface roughness, threading dislocation density (TDD) and the accuracy of the layers thicknesses. These parameters reflect on the electrical characterisation such as carrier mobility and sheet density, as well as, the optical properties. The transport properties of the two dimensional hole gas in sGe quantum well (QW) have been the focus of intensive research for many years because of its potential for device applications. Both low temperature (LT) [5, 6-10] and room temperature (RT) [11-19] measurements have been performed in order to obtain a better understanding of the fundamental physics behind these transport properties.

To obtain an improved electrical performance for sGe, a high purity Ge QW needs to be grown on a Ge-rich buffer layer with a low defect density. Hole mobilities at RT of up to $3000 \text{ cm}^2 / \text{V s}$ [10, 12, 15, 17, 19] have been reported for structures grown via Low Energy Plasma

Enhanced Chemical Vapour Deposition (LEPE-CVD) [19, 20] and SS-MBE [13, 17]. However, growth by the Reduced Pressure Chemical Vapour Deposition (RP-CVD) utilizing a low growth temperature methodology has enabled a pure sGe QW (Si concentration less than 0.01 at.%) [11] to be grown, leading to the highest reported hole mobility to date for sGe at 10 K ($1.1 \times 10^6 \text{ cm}^2 / \text{V s}$ at a sheet density of $3 \times 10^{11} \text{ cm}^{-2}$ [6], while at the room temperature (RT) mobility extracted using a simulation methods helps to analysis the data and defined the mobility and sheet density of the channel excluding the carriers in the parallel channels. Mobility of $(3.9 \pm 0.4) \times 10^3 \text{ cm}^2 / \text{V s}$ was determined by maximum entropy-mobility spectrum analysis (ME-MSA) for a sheet density (p_s) $9.8 \times 10^{10} \text{ cm}^{-2}$, and by using another method for mobility simulation (Bryan's Algorithm Mobility Spectrum (BAMS)) mobility of $(3.9 \pm 0.2) \times 10^3 \text{ cm}^2 / \text{V s}$ for a sheet density (p_s) $5.9 \times 10^{10} \text{ cm}^{-2}$ has extracted which confirmed the high RT mobility for this structure [21]. More higher mobility at RT of $4230 \text{ cm}^2 / \text{V s}$ for a sheet density (p_s) $1 \times 10^{11} \text{ cm}^{-2}$ has determined using the ME-MSA technique by Myronove et al [22]. These results were the highest mobility reported for mobility at room temperature as seen in Figure 1.

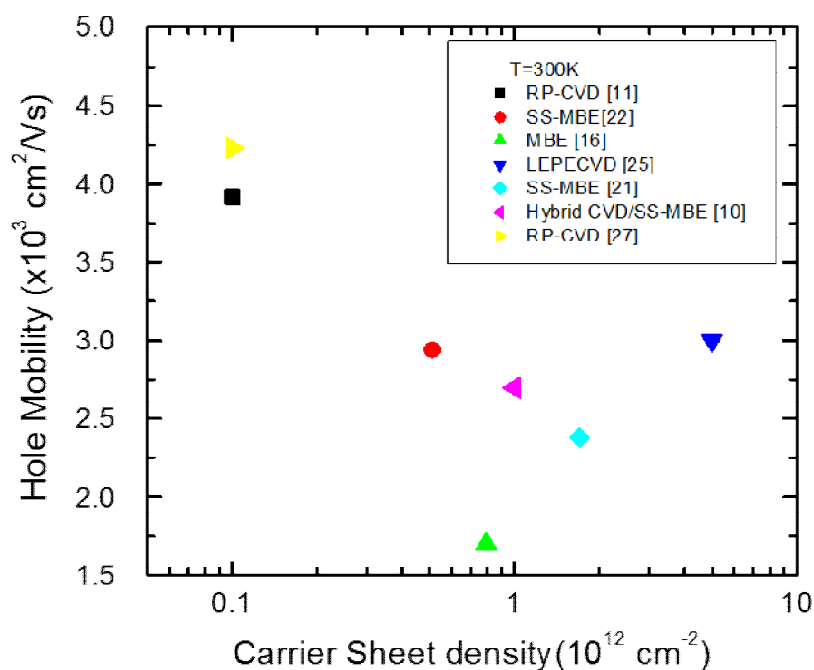


Figure 1: Room temperature hole mobility of strained germanium devices grown by different growth technique.

(LEPE-CVD) [5] which is considered as a good choice for a high growth rate 10 nm/s. Strained germanium modulation doped structure result in a high hole mobility of $1.2 \times 10^5 \text{ cm}^2/\text{Vs}$ at temperature of 2 K with a sheet density of $8.5 \times 10^{11} \text{ cm}^{-2}$, which was thought to be due to the high quality structure, with low interface roughness compared to other techniques[5]. By optimizing this method (LEPE-CVD) for appropriately chosen plasma densities and substrate temperatures, abrupt interfaces can be achieved on both sides of the Ge channels. Additional hydrogen is supplied to the reactive gases, even for channel widths above the critical thickness for dislocation formation and it ended with high mobility $9 \times 10^4 \text{ cm}^2/\text{Vs}$ at sheet

density $6 \times 10^{11} \text{ cm}^{-2}$ at 4.2 K [38]. It has been seen that LT processing below 600°C is essential for high performance devices to avoid Si-Ge interdiffusion at high temperature at the Ge/SiGe interface. Annealing at temperatures above 600°C causes a reduction in mobility and an increase in the sheet density of the structure under study. Moreover, Ge layers beyond the critical thickness easily relax when the annealing temperature is increased above 500°C .

Surface roughness is an important parameter to get high mobility samples and for the inverted structure it has more effect on the top of the channel, for this reason increasing the thickness will drop its effect which separate the roughness from the carriers that located near the bottom of the channel [23]. Introducing two steps LT buffers account for no significant different between the mobility of normal and inverted structure, which shows mobility at low temperature of $1.4 \times 10^4 \text{ cm}^2/\text{Vs}$ and $1.3 \times 10^4 \text{ cm}^2/\text{Vs}$ respectively [4] owing to the smooth surface compared with Ge structure on graded buffer layers. in terms of enhance mobility (higher than $3000 \text{ cm}^2/\text{Vs}$ at room temperature) and reduce surface roughness for samples fabricated by LEPECVD, they suggest more optimisation for the growth temperature [18]. Another kind of CVD growth mechanism, used by Myronov et al [21], is (RP-CVD) with a reverse linearly graded buffer, which improves the root mean square (rms) surface roughness to about 1.5 nm with TDD of $2 \times 10^6 \text{ cm}^{-2}$. This is comparable to that achieved by LEPE-CVD with a thinner virtual substrate (VS) ($3 \mu\text{m}$ to reach 80% Ge), in comparison to the VS grown by LEPE-CVD (above $10 \mu\text{m}$ to reach 80% Ge).

Symmetric doping modify the Ge QW valance band energy from triangular like to rectangular like that increase mobility of room and low temperature by 3.3 and 1.7 times its previous values, respectively as reported by Myronov et al [15]. However, in another study symmetric doping produces two subband energy in quantum well that

increase the interband scattering which drastically reduce the hole mobility [8]. The mobility of two dimensional hole gas (2DHG) for sGe channel has been studied heavily by researchers [6, 9, 11, 13, 18, 24-27]. However, the significant enhancement in mobility of sGe channel has reported by Dobbie et al [6] using RP-CVD, and in the same time the highest mobility for two dimensional electron gas in sSi was revealed to be $(1.6 - 2) \times 10^6 \text{ cm}^2/\text{Vs}$ [28] as shown in Figure 2.

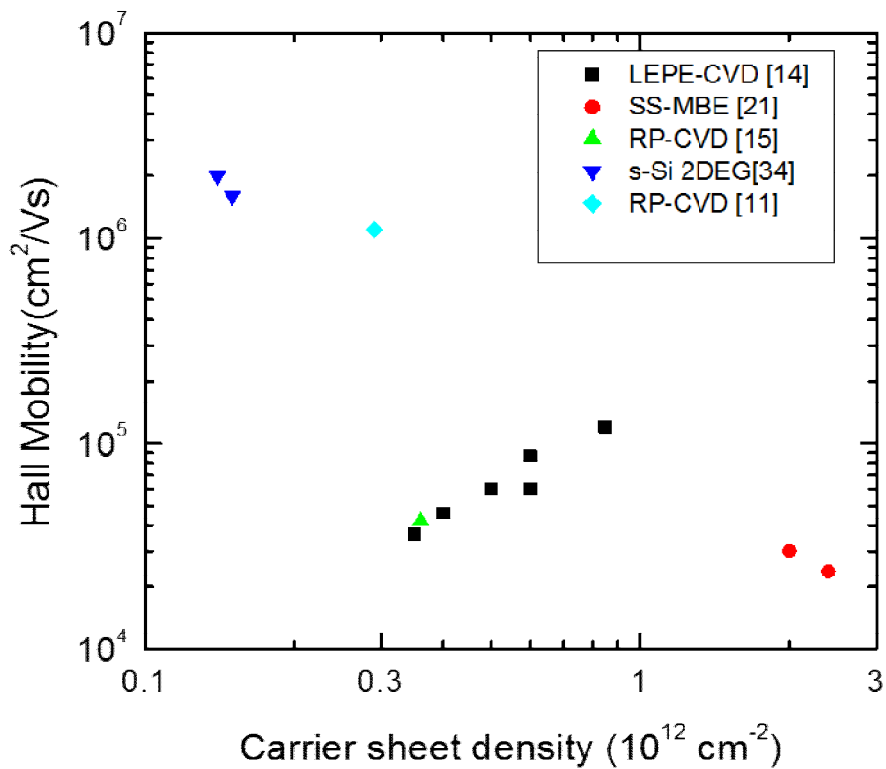


Figure 2: Low temperature hole mobility of strained germanium devices grown by different growth technique.

These high mobilities open a new field of characterisation for 2DHG in sGe such as fractional quantum Hall effect and spin splitting effect properties. The investigation of Rashba spin splitting in 2DHG of sGe materials has opened the door of using these materials in spintronic

devices [30-33] . The zero field spin splitting raises from bulk or structural inversion asymmetry (BIA and SIA respectively), BIA is excluded in Ge because of its crystal symmetry and only SIA has been observed in sGe samples (Figure 3) in low magnetic field of magnetoresistance [29]. The fractional quantum Hall effect has also been investigated in high magnetic field and showed optimistic result for devices with high mobility [34], The different growth techniques show also significant mobility anisotropy in different growth orientations that explains the mobility different in the same sample with different orientation [35].

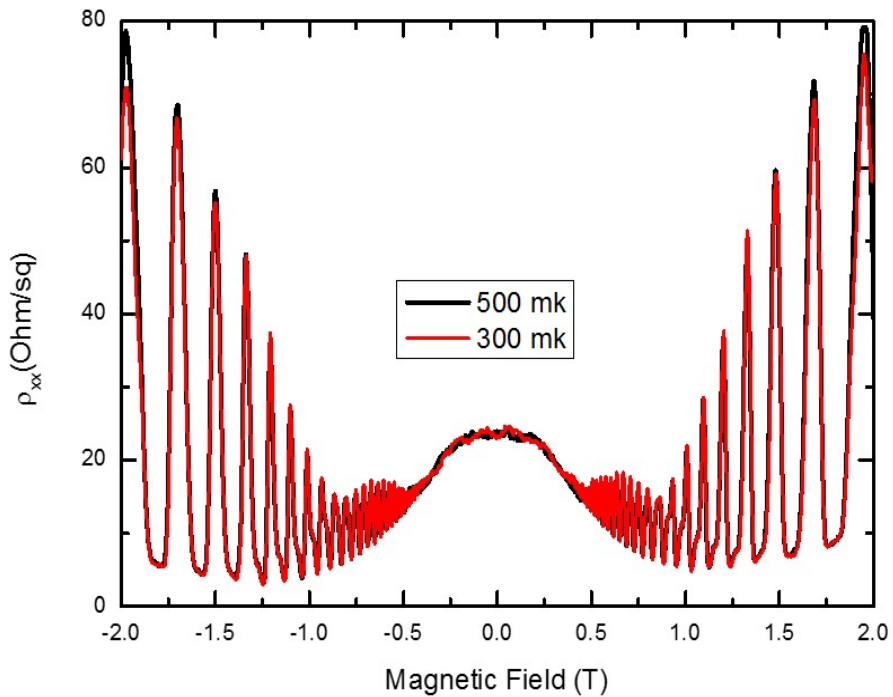


Fig 3. magnetoresistance as a function of magnetic field for sGe channel sample shows Rashba spin splitting for both temperatures 300 mK and 500 mK.

Li et al [36] attempts to get high quality samples using MBE, which investigated a low temperature technique (LT-MBE) for $\text{Si}_{0.7}\text{Ge}_{0.3}$ on Si (001) and found that using this process gives a lower (TDD) in the order of 10^5 cm^{-2} for a smaller overall layer thickness and smoother surface. This development in the quality of the structure was an excellent motivation for Ueno et al [4, 37] who applied a low temperature buffer ($\text{Si}_{0.3}\text{Ge}_{0.7}$) for p-type sGe channel to examine the effect of this method on the properties of a Ge channel. Two paths were followed to grow the sample; one and two step LT techniques and employing SS-MBE. They found that the two step LT buffer produced a single period of surface roughness with a 10 nm amplitude whereas the one step LT technique produced many periods. The one step process leads to a higher drop of mobility, whilst the two step LT buffer generated a high mobility at room temperature ($1700 \text{ cm}^2/\text{Vs}$) because of the low TDD ($1 \times 10^5 \text{ cm}^{-2}$), small surface roughness (1-3 nm), and almost total relaxation (>95%).

Sawano et al found out that sheet density enhancement with temperature depends on the growth technique of SiGe buffer carrier [7]. Structures grown by LT buffer have a higher increase in sheet density than the structure grown with graded SiGe buffer at higher temperature, in other words, the effect of conduction layer is higher for LT SiGe buffer. Also they found that GS-MBE is better for fabricating SiGe layer, because SS-MBE generates an acceptor-like point defect which is a source of parallel conduction, furthermore, these defects are a significant reason for current leakage in MOS devices.

The effect of Ge channel thickness on SiGe (VS) has been studied by Irisawa et al [13] where they reported that the optimized channel thickness is 7.5 nm for high mobility strained germanium channel on $\text{Si}_{0.3}\text{Ge}_{0.7}$. Reducing the channel thickness less than that will drop the mobility dramatically, owing to the boost of the effect of surface

roughness, whereas increasing channel thickness, reduces the mobility because of strain relaxation[23]. Nevertheless, strain relaxation on graded SiGe buffer could be controlled using LT technique with chemical mechanical polishing (CMP). This structure experiences a good improvement in mobility by reducing the effect of interface roughness. Therefore, more reduction in channel thickness does not influence the mobility [38,40,41].

To enhance the performance of the structure in terms of reduced doping, Ge segregation, and reduced surface roughness, researchers have also attempted to apply a combination of two growth methods such as LEPE-CVD and SS-MBE techniques [39]. This resulted in a large increase in room temperature mobility, $2700 \text{ cm}^2/\text{Vs}$ at sheet density $1 \times 10^{12} \text{ cm}^{-2}$, as well as a significant improvement in the conductance that is important for device applications.

3. Conclusion

In summary, different growth techniques have been discussed for 2DHG sGe heterostructure and important different in the structures properties are reviewed. It has been shown that RP-CVD technique produces a high purity sample with low threading dislocations and low interface roughness, which result in high mobility reached of $1 \times 10^6 \text{ cm}^2/\text{Vs}$ at sheet density of $3 \times 10^{11} \text{ cm}^{-2}$ at low temperature which states clean and pure system of growth with very low contamination allows for new feature that make sGe as an important material in the field of electronic and spintronic devices.

References

- [1] G. E. Moore, "Cramming more components onto integrated circuits," *Electronics* vol. 38, no. 8, p. 4, 1965.
- [2] K. Sawano, "strain dependence of hole effective mass and scattering mechanism in strained germanium channel structure" *applied Physics Letters*, vol. 95, p. 122109, 2009.
- [3] J. C. Bean, "Silicon molecular beam epitaxy: 1984–1986," *Journal of Crystal Growth*, vol. 81, no. (1-4), pp. 411-420, 1987.
- [4] T. Ueno, T. Irisawa, and Y. Shiraki, "p-type Ge channel modulation doped heterostructures with very high room-temperature mobilities. ," *Physica E: Lowdimensional Systems and Nanostructures*, vol. 7, no. (3-4), pp. 790-794, 2000.
- [5] D. C. G. Isellaa, B. Rossner , T. Hackbarthc , H.-J. Herzogc ,U. Konig, H. von Kanel, "Low-energy plasma-enhanced chemical vapor deposition for strained Si and Ge heterostructures and devices," *Solid-State Electronics*, vol. 48, pp. 1317-1323, 2004.
- [6] A. Dobbie *et al.*, "Ultra high hole mobility exceeding one million in a strained germanium quantum well" *Applied Physics Letters*, vol. 101, no. 17, p. 172108, 2012.
- [7] K. Sawano, Y. Abe, H. Satoh, K. Nakagawa, and Y. Shiraki, "Mobility enhancement in strained ge heterostructures by planarization of SiGe buffer layers grown on Si substrates," (in English), *Japanese Journal of Applied Physics Part 2-Letters & Express Letters*, Article vol. 44, no. 42-45, pp. L1320-L1322, 2005.
- [8] B. Rossner, H. von Kanel, D. Chrastina, G. Isella, and B. Batlogg, "2-D hole gas with two-subband occupation in a strained Ge channel: Scattering mechanisms," (in English), *Thin Solid Films*, Proceedings Paper vol. 508, no. 1-2, pp. 351-354, Jun 2006.
- [9] B. Rossner, D. Chrastina, G. Isella, and H. von Kanel, "Scattering mechanisms in high-mobility strained Ge channels," (in English), *Applied Physics Letters*, Article vol. 84, no. 16, pp. 3058-3060, Apr 2004.

- [10] M. Myronov, A. Dobbie, V. A. Shah, X. C. Liu, V. H. Nguyen, and D. R. Leadley, "High Quality Strained Ge Epilayers on a $\text{Si}_{0.2}\text{Ge}_{0.8}/\text{Ge}/\text{Si}(100)$ Global Strain-Tuning Platform," (in English), *Electrochemical and Solid State Letters*, Article vol. 13, no. 11, pp. H388-H390, 2010.
- [11] S. Madhavi, V. Venkataraman, and Y. H. Xie., "High room-temperature hole mobility in $\text{Ge}_{0.7}\text{Si}_{0.3}/\text{Ge}/\text{Ge}_{0.7}\text{Si}_{0.3}$ modulation-doped heterostructures," *Applied Physics Letters*, vol. 89, p. 2497, 2001.
- [12] T. Tanaka, Y. Hoshi, K. Sawano, N. Usami, Y. Shiraki, and K. M. Itoh., "Upper limit of two-dimensional hole gas mobility in strained Ge/SiGe heterostructures," *Applied Physics Letters*, vol. 100, p. 222102, 2012.
- [13] T. Irisawa, H. Miura, T. Ueno, and Y. Shiraki, "Channel width dependence of mobility in Ge channel modulation-doped structures," (in English), *Japanese Journal of Applied Physics Part 1-Regular Papers Short Notes & Review Papers*, Proceedings Paper vol. 40, no. 4B, pp. 2694-2696, Apr 2001.
- [14] T. Irisawa, S. Tokumitsu, T. Hattori, K. Nakagawa, S. Koh, and Y. Shiraki, "Ultrahigh room-temperature hole Hall and effective mobility in $\text{Si}_{0.3}\text{Ge}_{0.7}$ $\text{Ge}/\text{Si}_{0.3}\text{Ge}_{0.7}$ heterostructures," *Applied Physics Letters*, vol. 81, no. 5, 2002.
- [15] M. Myronov, K. Sawano, K. M. Itoh, and Y. Shiraki, "Observation of pronounced effect of compressive strain on room-temperature transport properties of two-dimensional hole gas in a strained Ge quantum well," (in English), *Applied Physics Express*, Article vol. 1, no. 5, p. 3, May 2008.
- [16] M. Myronov, Y. Shiraki, T. Mouri, and K. M. Itoh, "Enhancement of room-temperature hole conductivity in narrow and strained Ge quantum well by double-side modulation doping," (in English), *Applied Physics Letters*, Article vol. 90, no. 19, May 2007.
- [17] M. Myronov *et al.*, "Extremely high room-temperature two-dimensional hole gas mobility in $\text{Ge}/\text{Si}_{0.33}\text{Ge}_{0.67}/\text{Si}(001)$ p-type modulation-doped heterostructures," (in English), *Applied Physics Letters*, Article vol. 80, no. 17, pp. 3117-3119, Apr 2002.

- [18] H. v. Kanel, "Very high hole mobilities in modulation-doped Ge quantum wells grown by low-energy plasma enhanced chemical vapor deposition," *APPLIED PHYSICS LETTERS*, vol. 80, no. 16, p. 2922, 2002.
- [19] H. von Kanel, D. Chrastina, B. Rossner, G. Isella, J. P. Hague, and M. Bollani, "High mobility SiGe hetero structures fabricated by low-energy plasma-enhanced chemical vapor deposition," (in English), *Microelectronic Engineering*, Proceedings Paper vol. 76, no. 1-4, pp. 279-284, Oct 2004.
- [20] H. von Kanel, M. Kummer, G. Isella, E. Muller, and T. Hackbarth, "Very high hole mobilities in modulation-doped Ge quantum wells grown by low-energy plasma enhanced chemical vapor deposition," (in English), *Applied Physics Letters*, Article vol. 80, no. 16, pp. 2922-2924, Apr 2002.
- [21] O. A. Mironov *et al.*, "Ultra high hole mobilities in a pure strained Ge quantum well," *Thin Solid Films*, vol. 557, pp. 329-333, 2014/04/30/ 2014.
- [22] M. Maksym *et al.*, "An extremely high room temperature mobility of two-dimensional holes in a strained Ge quantum well heterostructure grown by reduced pressure chemical vapor deposition," *Japanese Journal of Applied Physics*, vol. 53, no. 4S, p. 04EH02, 2014.
- [23] Y. H. Xie, D. Monroe, E. A. Fitzgerald, P. J. Silverman, F. A. Thiel, and G. P. Watson, "Very high mobility two-dimensional hole gas in Si/GexSi1-x/Ge structures grown by molecular beam epitaxy," *Applied Physics Letters*, vol. 63, no. 16, pp. 2263-2264, 1993.
- [24] M. Myronov *et al.*, "Temperature dependence of transport properties of high mobility holes in Ge quantum wells," (in English), *Journal of Applied Physics*, Article vol. 97, no. 8, Apr 2005.
- [25] M. Myronov, K. Sawano, and Y. Shiraki., "Enhancement of hole mobility and carrier density in Ge quantum well of SiGe heterostructure via implementation of double-side modulation doping" *Applied physics letters*, vol. 88, 2006.
- [26] D. Chrastina *et al.*, "High quality SiGe electronic material grown by low energy plasma enhanced chemical vapour

- deposition," (in English), *Thin Solid Films*, Article vol. 459, no. 1-2, pp. 37-40, Jul 2004.
- [27] T. Irisawa *et al.*, "Hole density dependence of effective mass, mobility and transport time in strained Ge channel modulation-doped heterostructures," (in English), *Applied Physics Letters*, Article vol. 82, no. 9, pp. 1425-1427, Mar 2003.
- [28] T.-M. L. S.H. Huang, S.-C. Lu, C.-H. Lee, C. W. Liu, and D. C. Tsui, "Mobility enhancement of strained Si by optimized SiGe/Si/SiGe structures," *APPLIED PHYSICS LETTERS*, vol. 101, no. 042111, 2012.
- [29] A. H. A. Hassan, R. J. H. Morris, O. A. Mironov, S. Gabani, A. Dobbie, and D. R. Leadley, "An origin behind Rashba spin splitting within inverted doped sGe heterostructures," *Applied Physics Letters*, vol. 110, no. 4, p. 042405, 2017/01/23 2017.
- [30] R. Moriya *et al.*, "Cubic Rashba Spin-Orbit Interaction of a Two-Dimensional Hole Gas in a Strained- Ge/SiGe Quantum Well," *Physical Review Letters*, vol. 113, no. 8, p. 086601, 08/21/ 2014.
- [31] C. Morrison and M. Myronov, "Strained germanium for applications in spintronics," *physica status solidi (a)*, vol. 213, no. 11, pp. 2809-2819, 2016.
- [32] C. Morrison, P. Wiśniewski, S. D. Rhead, J. Foronda, D. R. Leadley, and M. Myronov, "Observation of Rashba zero-field spin splitting in a strained germanium 2D hole gas," *Applied Physics Letters*, vol. 105, no. 18, p. 182401, 2014.
- [33] J. Foronda, C. Morrison, J. E. Halpin, S. D. Rhead, and M. Myronov, "Weak antilocalization of high mobility holes in a strained Germanium quantum well heterostructure," *Journal of Physics: Condensed Matter*, vol. 27, no. 2, p. 022201, 2015.
- [34] O. A. Mironov, N. d'Ambrumenil, A. Dobbie, D. R. Leadley, A. V. Suslov, and E. Green, "Fractional Quantum Hall States in a Ge Quantum Well," *Physical Review Letters*, vol. 116, no. 17, p. 176802, 04/27/ 2016.
- [35] A. H. A. Hassan *et al.*, "Anisotropy in the hole mobility measured along the [110] and $[1\bar{1}0]$ orientations in a strained Ge quantum well," *Applied Physics Letters*, vol. 104, no. 13, p. 132108, 2014.

- [36] C. S. P. J. H. Li, Y. Wu, D. Y. Dai, J. M. Zhou "Relaxed Si_{0.7}Ge_{0.3} layers grown on low-temperature Si buffers with low threading dislocation density," *Appl. Phys. Lett*, vol. 71, p. 3132, 1997.
- [37] T. Irisawa, T. Ueno, H. Miura, and Y. Shiraki, "Thermal stability of Ge channel modulation doped structures," (in English), *Journal of Crystal Growth*, Proceedings Paper vol. 227, pp. 796-800, Jul 2001.
- [38] E. M. I. M. Bollania, S. Signorettic , C. Beelic , G. Isellad , M. Kummerc , H. von Kanel, "Compressively strained Ge channels on relaxed SiGe buffer layers," *Materials Science and Engineering* vol. B101, pp. 102-105, 2003.
- [39] M. Myronov, X. C. Liu, A. Dobbie, and D. R. Leadley, "Control of epilayer thickness during epitaxial growth of high Ge content strained Ge/SiGe multilayers by RP-CVD," (in English), *Journal of Crystal Growth*, Proceedings Paper vol. 318, no. 1, pp. 337-340, Mar 2011.
- [40] R. J. H. Morris *et al.*, "High conductance Ge p-channel heterostructures realized by hybrid epitaxial growth," (in English), *Semiconductor Science and Technology*, Article vol. 19, no. 10, pp. L106-L109, Oct 2004.
- [41] H. S. K Sawano, K. Nakagwa, Y. Shiraki, "mobility enhancement in strained- Ge modulation -doped structure by planarization of SiGe buffer layers," *Physica E: Lowdimensional Systems and Nanostructures*, vol. 32, pp. 520-523, 2006.



**Prediction of Wheel and Rail
Roughness Parameters Using
Artificial Neural Network**

6

Prediction of Wheel and Rail Roughness Parameters Using Artificial Neural Network

A. Shebani
College of Computer Technology – Zawya

C. Pislaru
Institute of Railway Research, University of Huddersfield, Huddersfield, UK

Abstract

Wheel and rail reprofiling costs millions of dollars around the world. The wheel/rail roughness is one of the important parameter which be able to effect on wheel/rail wear. The use of an artificial neural network to predict the wheel/rail roughness parameters can help to improve the design of the wheel/rail profiles. Wheel/rail roughness can define as the shorter frequency of real wheel/rail surfaces relative to the troughs. There are several roughness parameters, but the arithmetical mean roughness (Ra) is the common parameter, it is indicating the average of the absolute value along the sampling length. In this paper, both rail and wheel roughness were measured experimentally using Alicona profilometer and replica material, then, the arithmetical mean height of the wheel and rail was predicted using artificial neural networks. The results showed that the neural network predicted the wheel and rail roughness parameter efficiently.

Keywords: Wheel/rail roughness, artificial neural network, Matlab, roughness parameters, Alicona profilometer, and replica material.

1. Introduction to Railway System

1.1 Railway System

The railway train running along a track is one of the most complicated dynamical systems in engineering. Bogie is one of the most important part on the railway system. It consists of the following parts: wheelset, axle box, wheels, suspension, elastic elements, and damping. The wheelset comprises two wheels rigidly connected by a common axle. The wheelset is supported on bearings mounted on the axle journals. The wheelset provides the necessary distance between the vehicle and the track, the guidance that determines the motion within the rail gauge, and the means of transmitting traction and braking forces to the rails to accelerate and decelerate the vehicle. The design of the wheelset depends on the type of the vehicle, the type of braking system used, and the construction of the wheel centre and the position of bearings on the axle. The axle box is the device that allows the wheelset to rotate by providing the bearing housing and also the mountings for the primary suspension to attach the wheelset to the bogie or vehicle frame. The axle box transmits longitudinal, lateral, and vertical forces from the wheelset on to the other bogie element. The wheels and axles are the most critical parts of the railway rolling stock. Mechanical failure or exceedance of design dimensions can cause derailment. Solid wheels have three major elements: the tyre, the disc, and the hub, and mainly differ in the shape of the disc. The suspension is the set of elastic elements, dampers and associated components which connect wheelsets to the car body. If the bogie has a rigid frame, the suspension usually consists of two stages: primary suspension connecting the wheelsets to the bogie frame and secondary suspension between the bogie frame and the bolster or car body [1]. The elastic elements (springs) are components which return to their original dimensions when forces causing them to deflect are removed. The Damping is usually provided in railway vehicle suspension by the use of viscous or friction damping devices [1].

1.2 Railway Track

The typical shape and construction profiles of a ballasted track are illustrated in Figure (1). The rail track sleeper is used to transmit the wheel load to the ballast medium. In addition, it has functions such as maintaining track alignment and gauge, restraining longitudinal and lateral rail movements, and providing strength and stability to track structure. The rail joints are used to join rails depending on the required position of the rails [2].

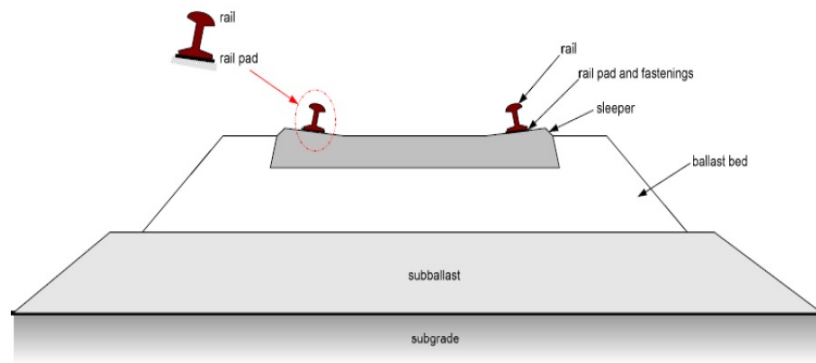


Figure 1: Rail track components and their arrangements [2]

1.3 Wheel Set

The wheel set is placed attached to the railway bogie. Bogie is a structure underneath a train to which axles and hence wheels are attached through bearings. Bogies are classified according to their configurations in terms of the numbers of axles, the design and structure of the suspension systems [2]. Generally, the railway wheel set has 6 degrees of freedom broadly classified as translational and rotational degrees of freedom parts such as in Figure (2). The translational degrees of freedom comprise three components that is translation along: X-axis, Y-axis, and Z-axis.

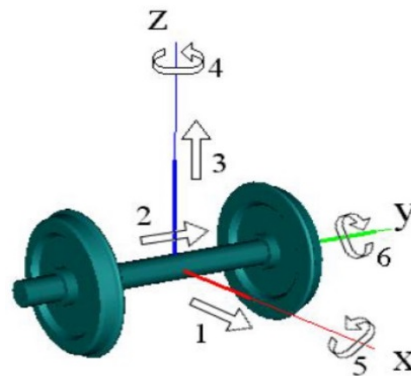


Figure 2: Wheel set degrees of freedom [2]

1.4 Rail

Railway lines are made of straight sections and curves. Train driving on the curves essentially differs from that one on straight sections. On the curves railway gauge is widened (when curve radius is less than 350 m). Rails are longitudinal steel members that are placed on spaced sleepers to guide the rolling stock. Support of traffic load and guidance of vehicles are the two main tasks of the rails. For both tasks the correct contact geometry between wheel and rail is essential. In addition to that rails are used to accommodate and transfer the wheel/axle loads into the supporting sleepers. The most commonly used profile is flat-bottom rail and is divided into three parts such as in Figure (3) [2]:

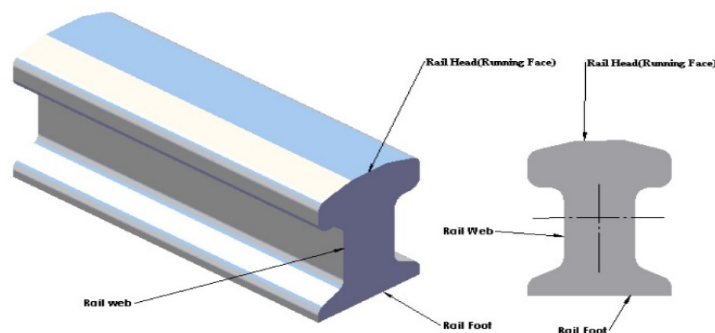


Figure 3: Flat bottom rail parts [2]

1.5 Rail-Wheel Interaction

The dynamic behavior of railway vehicle is greatly affected by the rail-wheel dynamic interactions. This interaction (wheel/rail) mainly depends on wheel/rail contact geometry. The changes in contacting geometry of rail/wheel depends on different parameters like the variation of wheel and rail profile, track gauge, rail inclinations, railhead surface irregularities, and flexibility of rail support. The main parameters influencing the wheel rail contact geometry are the profiles of wheels and rails, rail inclination and track gauge [2]. Figure (4) shows the general wheel rail interactions from the front and side view respectively.

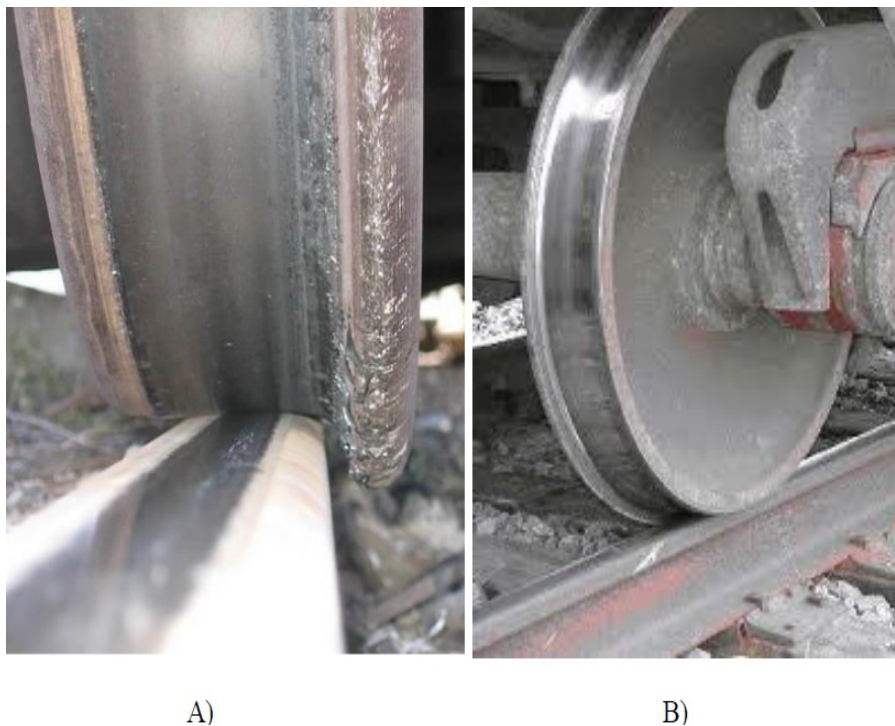


Figure 4: Front views (A), and side view (B) of wheel/rail contact interface [2]

1.6 Wheel/Rail Wear

The profile change of rails on curves makes a large contribution to track maintenance cost. The profile change on wheels can also be significant, especially on a curved track. Damage mechanisms such as wear and plastic deformation are the main contributors to profile change. Another growing problem for many railways is rolling contact fatigue. Wear is the loss or displacement of material from a contacting surface. Material loss may be in the form of debris. Material displacement may occur by transfer of material from one surface to another by adhesion or by local plastic deformation. There are many different wear mechanisms that can occur between contacting bodies, each of them producing different wear rates. Wheel/rail wear is shown in Figure (5).

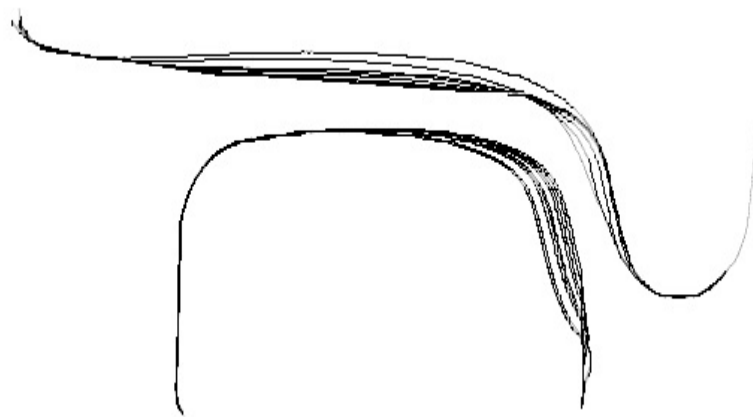


Figure 5 Form change of wheel and rail from the Stockholm test case [1].

1.7 Roughness Parameters:

Roughness average (R_a), it called centre line average value or arithmetic average. Among Height Parameters, the roughness average (Ra) is the most widely used because it is a simple parameter to obtain when compared to others. The roughness average is described as follows [3]:

$$R_a = \frac{1}{L} \int_0^L |Z(x)| \quad (1)$$

Where $Z(x)$ is the function that describes the surface profile analyzed in terms of height (Z) and position (x) of the sample over the evaluation length “ L ”, such as in Figure (6).

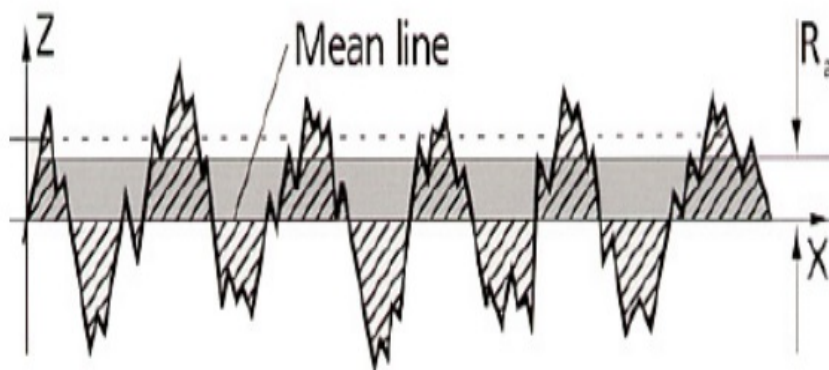


Figure 6: Profile of a surface (Z). It represents the average roughness Ra [3]

2. Artificial Neural Networks (ANN)

2.1 Introduction to Neural Networks

Artificial neural networks (ANN) are, as their name indicates, computational networks which attempt to simulate, in a gross manner, the networks of nerve cell (neurons) of biological (human or animal) central nervous system [4]. An ANN consists of interconnected processing units. The general model of a processing unit consists of a summing part receives N input values, weights each value, and computes a weights sum. The weighted sum is called the activation value. The output part produces a signal from activation value [5]. The use of neural networks offers the following useful properties and capabilities Nonlinearity, input-output mapping, adaptively, evidential response, contextual information, fault tolerance, uniformity of analysis and design, very large scale integrated implement ability, and neurobiological analogy [6].

The activation function is illustrated in Figure (7).

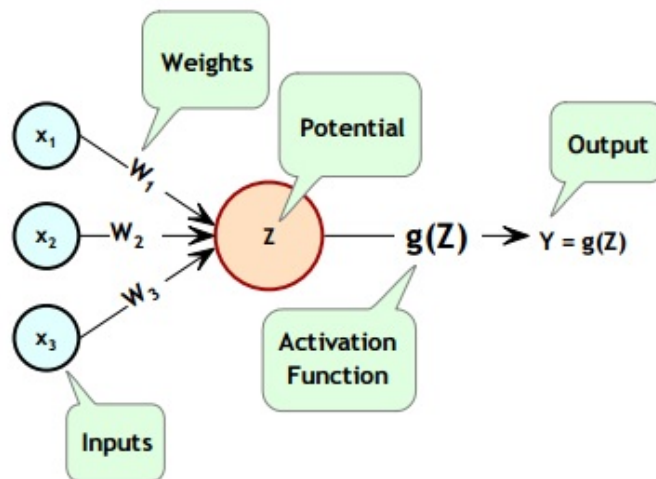


Figure 7: A Neural Net Perceptron [7]

As shown in Figure 7, perceptron consists of the following five components [7]:

Inputs: X_1 , X_2 , and X_3 ; weights: W_1 , W_2 , and W_3 ; potential:

$Y(n) = \sum_{i=1}^3 W_i X_i$; activation function: $g(Z)$; and the output:

$$Y = g(Z).$$

Although theoretically any differential function can be used as an activation function, the sigmoid function is the most commonly used. Figure (8) shows the sigmoid activation function.

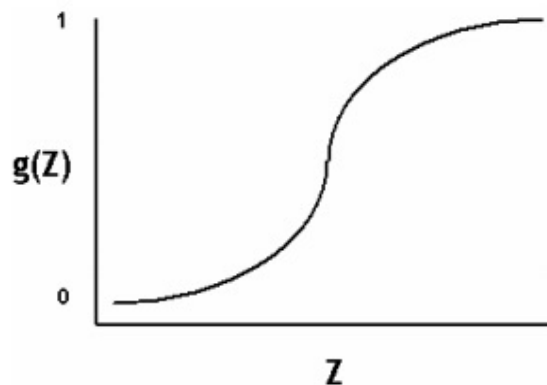


Figure 8: A Sigmoid Activation Function [7]

In practice, the most common sigmoid activation function is the logistic function that maps the potential into the range 0 to 1. The sigmoid function (also known the logistic function) is defined in the general form [7]:

$$f(x) = \frac{1}{1+e^{-z}} \quad (2)$$

Since $0 < g(Z) < 1$, the logistic function is very popular for use in networks that output probabilities.

2.2 The Multilayer Perceptron Neural Network

Multilayer perceptron network is an important class of neural networks. The network consists of a set of sensory units that constitute the input layer and one or more hidden layer of computation modes. The input signal passes through the network in the forward direction. The network of this type is called multilayer perceptron (MLP). The Multilayer perceptron are used with supervised learning and have led to the successful backpropagation algorithm. The disadvantage of the single layer perceptron is that it cannot be extended to multilayered version. In multilayer networks there exists a nonlinear activation function. The widely used non-linear activation function is logistic sigmoid function. The MLP network also has various layers of hidden neurons [8]. Figure (9) illustrates the architecture of the multilayer perceptron [9].

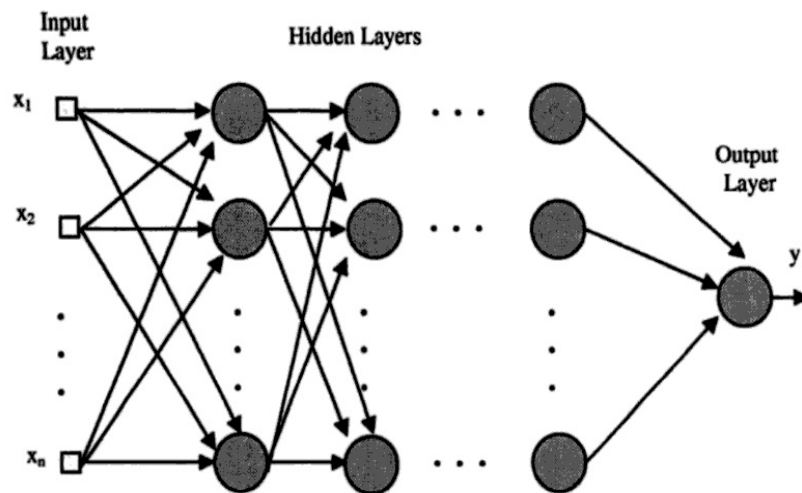


Figure 9: Organization in layers of the multilayer perceptron [9]

1.3 Feedforward neural network

Feedforward neural network is consists of a layered structure with information following from the inputs, at the bottom of the diagram, to the outputs at the top such as in Figure (10) [10].

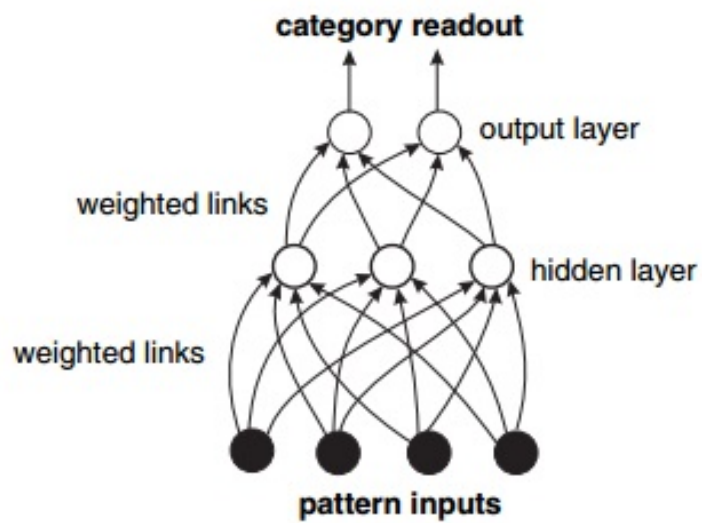


Figure 10: Simple feedforward neural network [10]

3. Prediction of the of wheel/rail roughness parameters using artificial neural network

The twin disc rig, replica material, and Alicona profilometer shown in Figure (11) were used for wheel and rail arithmetic average roughness (R_a) measurement such as shown in Table (1), and then, the neural network was used for wheel/rail arithmetic average roughness prediction.

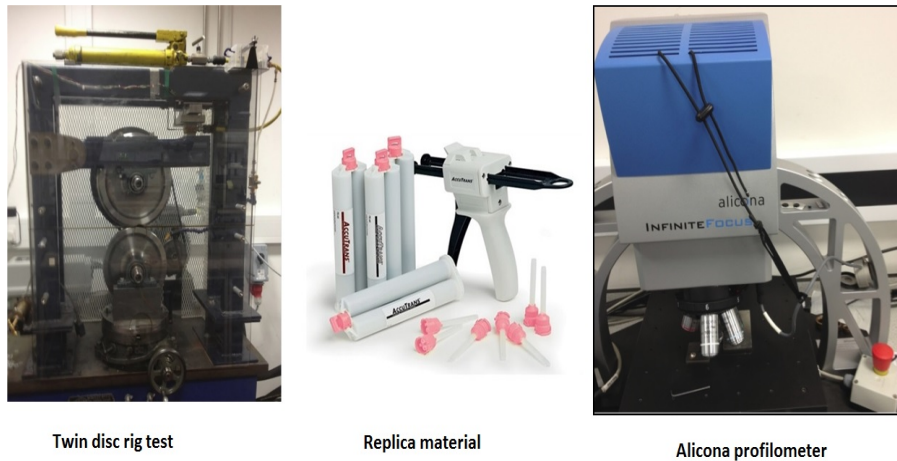


Figure 11: The twin disc rig, replica material and Alicona profilometer

Matlab toolbox used to design and training the neural network, and then, it was used to predict the arithmetic average roughness (R_a). The inputs of the neural network were load, speed, yaw angle, test time, and wheel/rail profile; while the output of the neural network is the arithmetic average roughness (R_a). The load was changed from 1000N to 4000N in steps of 100N, the speed was 660rpm, the yaw angle was 0.4degree, and the test time was 10min. The wheel/rail profiles were measured using Alicona profilometer.

Figure (12) shows the neural network toolbox used to predict the arithmetic average roughness.

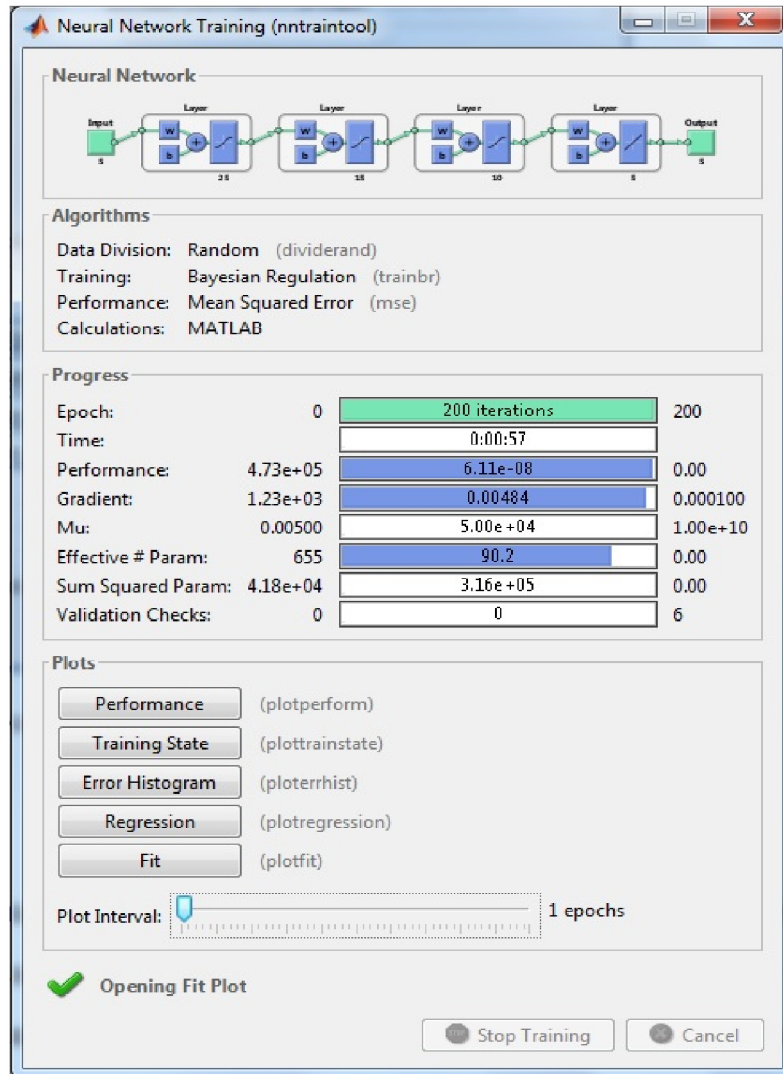


Figure 12: MATLAB training window

Table 1 shows the wheel/rail arithmetic average roughness measured using Alicona profilometer, and predicted using neural network.

Table 1 Arithmetic average roughness (wheel/rail) measured using Alicona profilometer, and predicted using neural network

Sample	R _a measured for wheel (μm)	R _a predicted using NN for wheel (μm)	Error %	R _a measured for rail (μm)	R _a predicted using NN for rail (μm)	Error %
1	1.49	1.46	2.05	1.37	1.36	0.72
2	1.69	1.67	1.19	1.59	1.56	1.88
3	1.88	1.87	0.53	1.76	1.74	1.13
4	1.94	1.91	1.57	1.87	1.88	0.53
5	1.98	1.96	1.02	1.93	1.96	1.55
6	2.03	2.02	0.49	1.96	1.95	0.51
7	2.11	2.12	0.47	1.98	1.97	0.50
8	2.12	2.10	0.95	2.11	2.13	0.94
9	2.13	2.15	0.93	2.13	2.10	1.40
10	2.18	2.17	0.46	2.17	2.15	0.92
11	2.22	2.21	0.45	2.20	2.23	1.36
12	2.23	2.24	0.44	2.22	2.21	0.45
13	2.25	2.27	0.88	2.24	2.22	0.89
14	2.36	2.38	0.84	2.32	2.30	0.86
15	2.41	2.40	0.41	2.38	2.39	0.42

The MAPE for wheel roughness was equal to 0.84%; therefore, the accuracy of NN model was 99.16%.

The MAPE for rail roughness was equal to 0.94%; therefore, the accuracy of NN model was 99.06%.

Figure (13) shows the arithmetic average roughness actual, and predicted using neural network (for wheel surface).

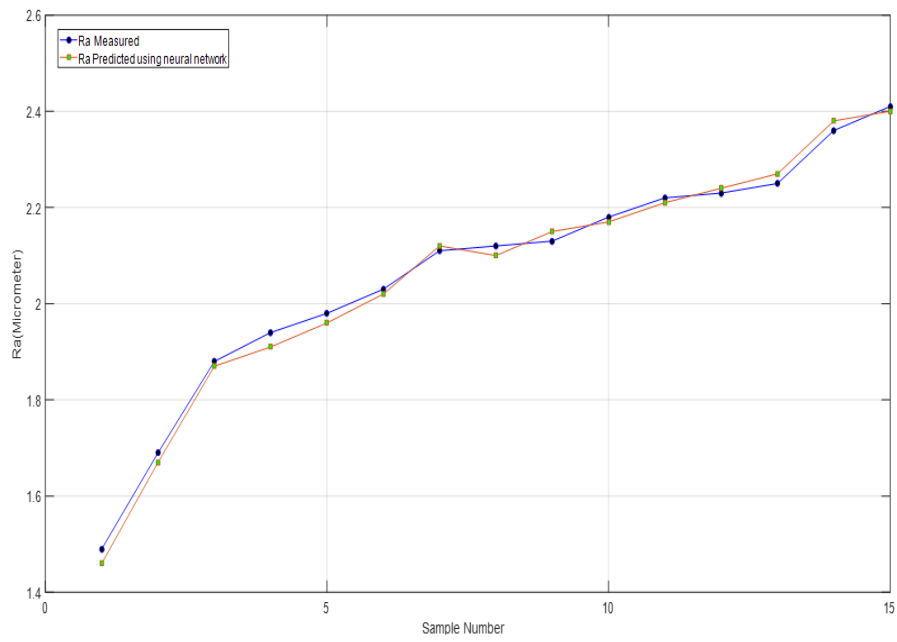


Figure 13: Arithmetic average roughness measured, and predicted using neural network (for wheel surface)

Figure (14) shows the arithmetic average roughness measured, and predicted using neural network (for rail surface).

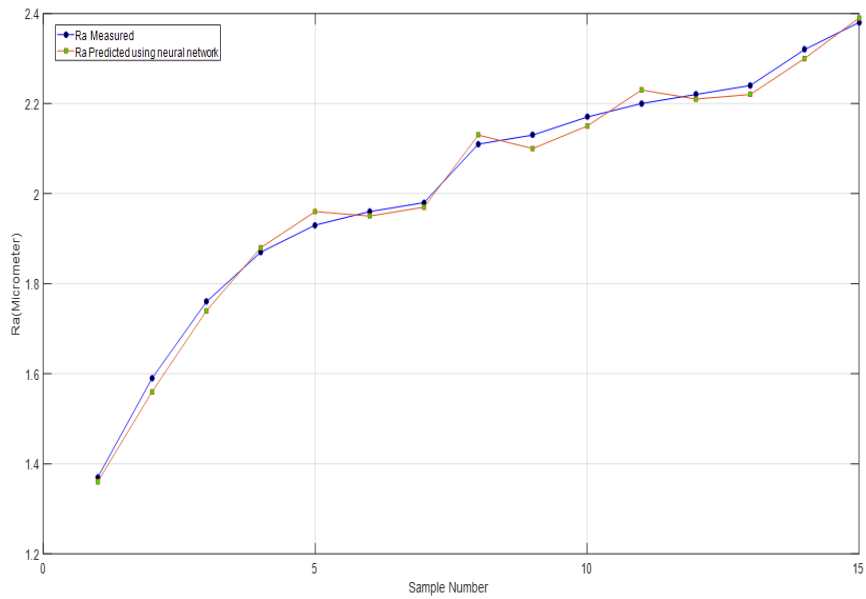


Figure 14: Arithmetic average roughness measured, and predicted using neural network (for rail surface)

As an example, Figure (15) shows the MATLAB performance plot, it shows a good performance for the neural network during roughness parameter prediction.

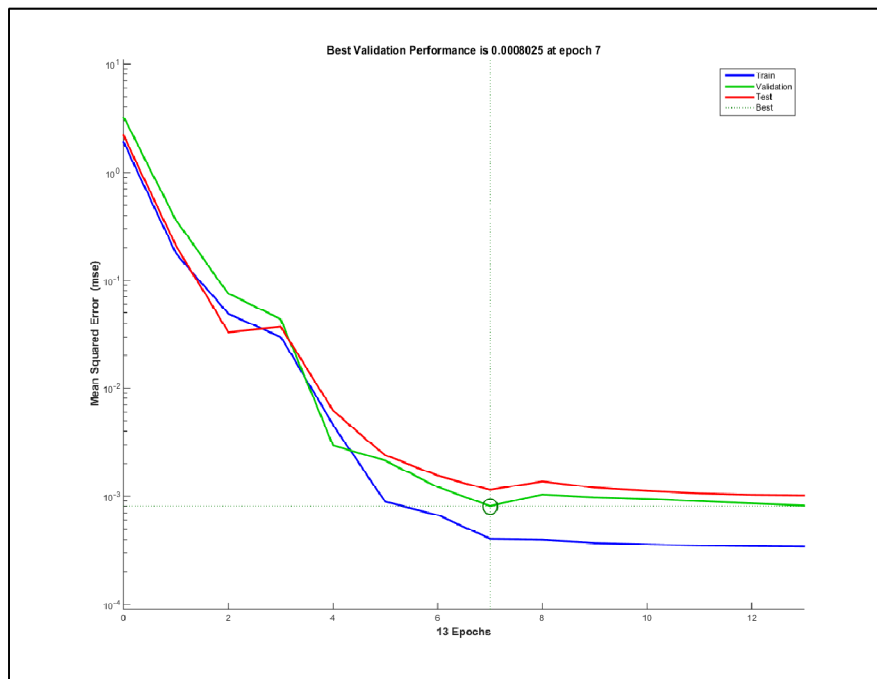


Figure 15: MATLAB performance plot for wheel roughness

4. Discussion and Conclusion

The percentage error for wheel and rail roughness parameter prediction was calculated, and the results show good prediction of wheel and rail roughness parameter in term of percentage of error, where the wheel and rail roughness parameter predicted using the NN was close to wheel and rail roughness parameter measured.


The MAPE for wheel roughness was equal to 0.84%, then, the accuracy of NN model was 99.16%; while the MAPE for rail roughness was equal to 0.94%, then, the accuracy of NN model was 99.06%. Therefore, the accuracy of the artificial neural network model was between 99.06% and 99.16%; (for unseen data).

The wheel and rail roughness can have established using the replica material and profilometer methods presented in this paper and compared with the results from the neural network techniques. The major finding in this work is that the Alicona profilometer can be used for wheel and rail roughness measurements. The replica material and Alicona profilometer are effective tools for the wheel and rail roughness measurements. An advantage of using the replica method is that it is a permanent record of wheel and rail roughness.

The neural network is an effective tool for the wheel and rail roughness prediction. This work can be used to promote the use of predictive maintenance strategies by railway operators. It can for example help in understanding remaining life of wheels or rails and in planning of maintenance interventions.

References

- [1] S. Iwnicki, *Handbook of railway vehicle dynamics*: CRC press, 2006.
- [2] A. Negash, "Analysis of Wheel/Rail Contact Geometry and Applied Load Conditions on the Rail Head Surface," Addis Ababa University, 2012.
- [3] V. Bellitto, *ATOMIC FORCE MICROSCOPY IMAGING, MEASURING AND MANIPULATING SURFACES AT THE ATOMIC SCALE*. Croatia, 2012.
- [4] D. Graupe, *Principles of artificial neural networks* vol. 6: World Scientific, 2007.
- [5] B. Yegnanarayana, *Artificial neural networks*: PHI Learning Pvt. Ltd., 2009.
- [6] S. Hayken, *Neural networks A comprehensive foundation* USA: Macmillan College, 1994.
- [7] E. R. Jones, *An Introduction to Neural Networks* USA: Visual Numerics, 2004.
- [8] S. Sivanandam and S. Deepa, *Introduction to neural networks using Matlab 6.0*: Tata McGraw-Hill Education, 2006.
- [9] N. Nedjah and L. de Macedo Mourelle, *Evolvable machines: theory & practice* vol. 161: Springer Science & Business Media, 2005.
- [10] C. R. Tosh and G. D. Ruxton, *Modelling perception with artificial neural networks*: Cambridge University Press, 2010.



**Performance Comparison of
(AODV and DSDV) Routing
Protocols in Mobile Ad Hoc
Networks.**

7

Performance Comparison of (AODV and DSDV) Routing Protocols in Mobile Ad Hoc Networks.

Mohamed Manita
Zawia College of computer technology
Maneta81@yahoo.com

Abstract

MANET is a collection of wireless nodes that can dynamically form a network to exchange information without using any pre-existing fixed network infrastructure. In an Ad-Hoc network nodes cooperate to maintain network connectivity and perform various functions including routing. This paper focuses on two flat routing protocols the reactive Ad- Hoc on Demand Distance Vector Routing protocol (AODV), and the proactive Destination-Sequenced Distance-Vector Routing (DSDV). Where a comparison between these two protocols was done using the well-known Network Simulator 2 (NS2).

Keywords: Ad-Hoc Network, mobile nodes, Network Simulator 2(NS2)

1. Introduction

The rapid growing of using mobile devices (e.g. laptops, mobile phones, personal digital assistants [PDAs]) in the recent years has made a need of some system or network that collect these omnipresent devices together, therefore people can get access the electronic platforms and searching for the information they need wherever and whenever they might be [1].The nature of these omnipresent devices

makes Mobile Ad-Hoc networks (MANETs) the most appropriate solution for such as this situation.

Mobile Ad-Hoc networks (MANETs) are independent systems which contain a collection of mobile devices that communicate each other over wireless connection. They are self-organized and self-controlled infrastructure-less networks. In this type of network each device is equipped with a wireless vector and receiving system which allows it to communicate with other devices over wireless channels. All devices can behaviour as routers to guarantee data packets to be arrived to their final destination. Which means that, ad-hoc networks allow for multi-hop transmission of data packets between devices out of the transmission range of each other .These networks can be created or deployed in any place and at any time because they are decentralized and they not need any existing network infrastructure. These networks have been using by different community users such as military, researchers, business, students, and emergency services [2].

Ad-hoc network routing protocols can be divided into proactive, reactive and hybrid routing [3]. A proactive routing protocol is also called "table driven" routing protocol. Using a proactive routing protocol, nodes in a mobile ad hoc network continuously evaluate routes to all reachable nodes and attempt to maintain consistent, up-to-date routing information, the Destination-Sequenced Distance-Vector (DSDV) protocol is an example of proactive routing protocols. However, the property of reactive routing protocols only requests a route when it is needed, and does not require mobile nodes to maintain routes to unreachable destinations [4]. The Ad hoc on-demand Distance Vector routing (AODV) is an another example for reactive routing protocols for mobile ad hoc networks.

These protocols have varying qualities for different wireless routing aspects. It is due to this reason that choice of a correct routing protocol is critical. In this paper a simulation model of a simple Ad-Hoc

network is done using two flat routing protocols the reactive Ad- Hoc on Demand Distance Vector Routing protocol (AODV), and the proactive Destination-Sequenced Distance-Vector Routing (DSDV) these two protocols has been chosen to be experimented which one performs better during the run of the simulation.

This paper addresses the question. ‘Which routing protocol provides a better performance in Mobile Ad hoc Networks?’ This question addresses the overall performance of the two routing protocols investigated in this research.

The Network simulator 2 (NS2) is used to do the simulation and the result is focused on the throughput , the routing overhead and the packet loss with same number of nodes and mobility level for both AODV and DSDV routing protocols.

2. Literature Review

Balram Swamia and Ravindar Singh [11] states that DSDV is a table driven routing and this protocols manage the route information in tables and that route information is broadcast to other neighbours through this method DSDV reduces the route detection time periods. DSDV is less energy consumption in mobile communications. OWL is on demand routing protocol which uses the DFS instead of RREQ. The main feature of OWL is less energy and time ingesting in route discovery it includes few nodes in route detection procedure and remaining nodes are able to accepts other route request by this reduce the interruption and recover the delivery ratio.

A.A. Chavana , Prof. D. S. Kurule and Prof. P. U. Dere [12] defined that routing protocols DSDV and AODV are tested in deferent terms PDR, overhead and end to end delay. AODV is superior to DSDV. These protocols compare in different terms routing overhead, PDR and throughput. DSDV is a table driven routing and this protocols manage the route information in tables and that route information is

broadcast to other neighbours through this method DSDV minimize the path discovery time periods. When some node needs to drive any data to its destination, it firstly checks the table to decide if it has route to final node, if sure then it transfer the data packet to next hop node. If not sure then it begins a route detection process that is path discovery and path maintains process. AODV performance is affected by black hole attack.

V. Rajeshkumar, P.Sivakumar [13], presents a performance comparison and study of reactive and proactive protocols AODV, DSR and DSDV based on metrics such as throughput, control overhead ,packet delivery ratio and average end-to-end delay by using the NS-2 simulator. They conclude that AODV performance is the best considering its ability to maintain connection by periodic exchange of data's. As far as Throughput is concerned, AODV and DSR perform better than the DSDV even when the network has a large number of nodes. Overall, their simulation work shows that AODV performs better in a network with a larger number of nodes whereas DSR performs better when the number of nodes is slight. Average End-to-End Delay is the least for DSDV and does not change if the number of nodes is increased. Thus, they find that AODV is a viable choice for MANETs.

3. Ad Hoc Network Routing Protocols Classification

There are many measures to be followed when designing and classifying routing protocols for a network. For example, what routing information is exchanged between hosts, when and how the routing information is exchanged, when and how routes are computed, etc. mainly, for ad hoc networks, routing protocols may be classified according to the routing strategy, which is, proactive (also known as table-driven) and reactive (known as on-demand), or they can be classified according to the network structure underlying routing protocols as it shown in figure (1) [5].

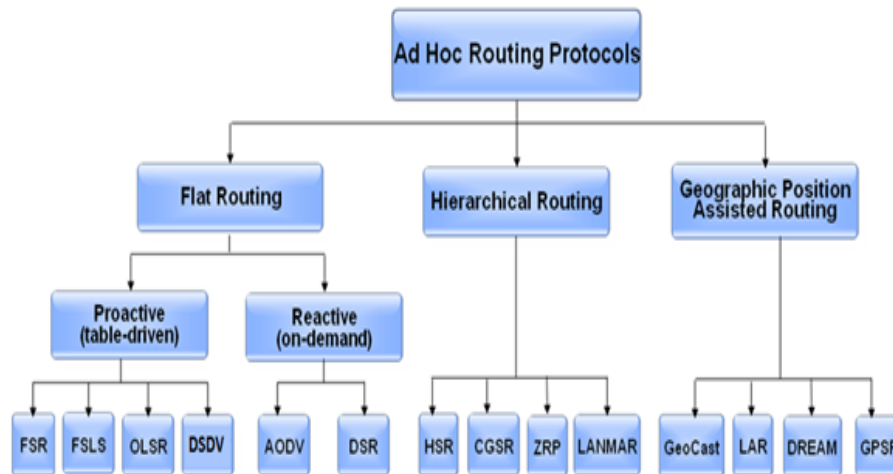


Figure 1: Classification of ad hoc routing protocols

3.1 Flat Routing

Flat routing approaches follow a flat addressing where all nodes in a network are at the same level and have the same routing functionality. Flat routing schemes generally can be divided into two classes proactive and reactive. Proactive protocols purpose to find and maintain consistent, up-to date routes between all source destination pairs regardless of the need or the use of suchlike routes here routing strategies are either link state or distance vector. But in reactive protocols routes are created only when a source device request them and data forwarding is accomplished according to source routing or hop by- hop routing [6].

3.2 Hierarchical Routing

Hierarchical routing is another possible routing approach for mobile ad hoc networks, in contrast of flat routing where every device knows about every other device in the topology, in this routing protocol, each device knows only about those devices in its level, and for all other destinations outside its level it forwards the packets to the border

router of its level. As following the per-mentioned classification four types of this routing protocol will be explained, these are, Hierarchical State Routing (HSR), Clusterhead-Gateway Switch Routing (CGSR), The Zone Routing Protocol (ZRP), and Landmark Ad Hoc Routing Protocol (LANMAR).

3.3 Geographic Position Assisted Routing

Geographic Position Assisted Routing protocols are another routing approach that should be indicated as they can be implemented in ad-hoc networks, and as they use different strategies to work. These protocols unlike other protocols use geographical addresses instead of link-specific addresses such as IP-addresses to find path from source to destination.

4. Comparison of AODV and DSDV

When a node using an on-demand protocol desires a route to a new destination, it will have to wait until such a route can be discovered. This feature, although useful for datagram traffic, incurs substantial signaling traffic and power consumption. Since both bandwidth and battery power are scarce resources in mobile computers, this becomes a serious limitation.

4.1 The Ad hoc On-Demand Distance Vector (AODV)

is a reactive routing protocol [7]. That means it tries to find a route when it is needed. AODV is intended for use by mobile devices in an ad-hoc network, it enables dynamic, self-starting, multi-hop routing between participating mobile devices wishing to establish and maintain an ad- hoc network. AODV allows mobile devices to find routes quickly for new destinations, and does not require devices to maintain routes to destinations that are not in active communication.

Routing Tables: For each routing table the following information have to be contained:

- Destination
- Next hop
- Number of hops
- Destination sequence number
- Active neighboring devices for this route
- Expiration time for this route table entry
- Expiration time, (is also as lifetime), is reset each time the route has been used. The new expiration time is the sum of the current time and a parameter called active route timeout.

Ad Hoc On-Demand Distance Vector Routing (AODV): AODV [8] is essentially a combination of both DSR and DSDV. It borrows the basic on-demand mechanism of Route Discovery and Route Maintenance from DSR, plus the use of hop-by-hop routing, sequence numbers, and periodic beacons from DSDV. It uses destination sequence numbers to ensure loop freedom at all times and by avoiding the Bellman-Ford "count-to -infinity" problem offers quick convergence when the ad hoc network topology changes. In this research paper we attempted to present an overview of two main categories of mobile ad-hoc routing protocols and performance comparison of both the protocols based on Random way point model and the simulation of two routing protocols focusing on their differences in their dynamic behaviours that can lead to performance differences.

4.2 Destination Destination-Sequenced Distance Vector (DSDV)

DSDV is a Routing Algorithm which is based on the classical Bellman-Ford routing algorithm. DSDV can prevent the routing loops. These routing protocols work to discover routes before they are needed that is why these types of protocols called Proactive. Each

node maintains a routing table in which there are possible destinations and possible number of hops for every destination is stored. It also store time between first and best announcement of a path. It also slows down updates if it seems to be unstable. This decision is based on the stored time values the nodes with this routing protocol only know the next hop to destination so these protocols are Table Driven. All nodes with DSDV run identical algorithms so these protocols are Flat protocols. Physical location of the nodes is unknown so DSDV is a non-location based protocol.

One of the proactive or table-driven mobile ad-hoc networks protocols is Destination-Sequenced Distance Vector (DSDV) [9] which is based on the idea of the classical Bellman-Ford routing algorithm with some enhancements.

In Destination-Sequenced Distance Vector protocol each and every node maintains a routing table that stores all potential destinations with number of hops required to access the destination and the sequence number that is allocated by the destination node. So a routing table in DSDV includes:

- Potential destinations
- Number of hops
- Sequence number

Table (1) Sample DSDV Routing Table

Destination Address	Metric	Address of Next hop	Sequence no.
Host 1	2	Host 2	Seq 1
Host 2	1	Host 2	Seq 2

5. Simulation Modelling

5.1 Network Simulator 2

The well-known Network Simulation (NS2) version ns-allinone2.28 has been used to do this simulation. NS2 is an event driven simulation tool that can simulate wired and wireless networks effectively as well as it can evaluate their functions and protocols (e.g., routing algorithms, TCP, UDP) [10].

5.2 Simulation Modelling in General

As shown in table (2) for this study a simulation of a virtual environment of 200 * 200 m for 600 sec of simulation time is used. The channel data rate and transmission range set to 2 Mbps and 100 m, respectively. Each run of the simulator accepts as input a scenario file that describes the exact motion of each node and the exact time at which each change in motion or packet origination is to occur. A total of 50 different scenario files with varying network size, movement patterns and traffic loads were generated and then the two routing protocols were run against each of these scenario files.

Table 2: Simulation parameters

Simulation Parameter	Value
Simulator	NS2
Network Range	200m × 200m
Transmission Range	100m
Mobile Nodes	20
Traffic Generator	Constant Bit Rate
Band Width	2 Mbps
Packet Size	512 Bytes
Packet Rate	10 Packet Per Second
Simulation Time	600s

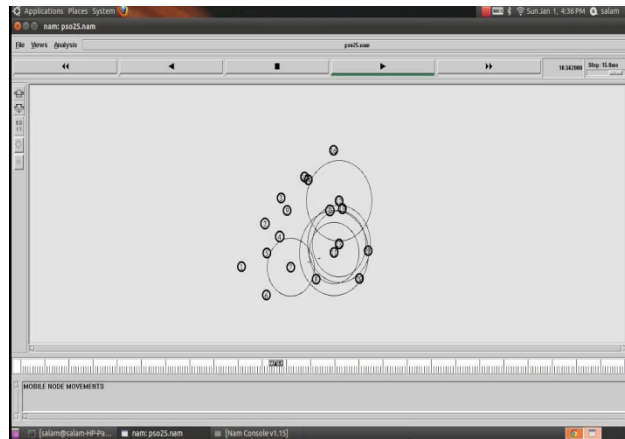


Figure 2: the execution of the simulation

5.3 Mobility Pattern

The nodes in the **simulation** move according to the ‘random way point’ model. At the start of the simulation, each node waits for a pause time, then randomly selects and moves towards a destination with a speed lying from zero and some maximum speed. When reaching this destination it pauses again and repeats the above procedure until the end of the simulation.

6. Results and Discussion

6.1 Throughput

As it can be seen by the table (3) and the graph (4) the performance of the AODV protocol is better for nearly 9% than DSDV for the same number of nodes and same level of mobility, where the throughput of the protocol AODV did not decreased as much as the peer protocol did for the same parameters .which is due to the difference in working mechanism of both of them.

Table 3: Comparison between DSDV and AODV Throughput

Throughput		Mobility (m/s)
DSDV	AODV	
1.52	1.44820	1.000
1.5175	1.43230	2.000
1.4825	1.45770	4.000
1.362	1.34520	10.000
1.1124	1.27560	13.000
1.0263	1.26000	15.000
0.9911	1.25500	18.000
0.9776	1.15500	25.000
0.8994	1.25011	35.000

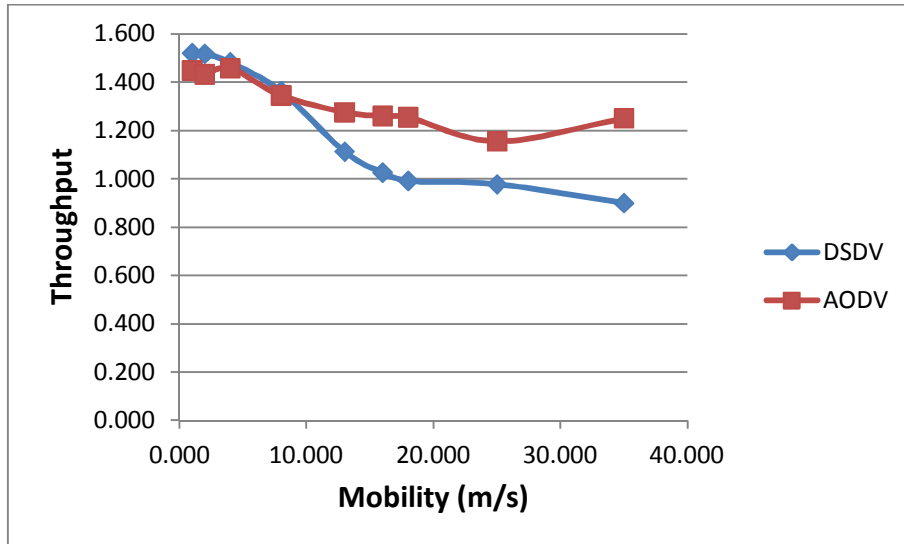


Figure 4: Comparison between DSDV and AODV Throughput

6.2 Packet Loss

It can be easily concluded from figure (5) that as the network traffic and node movement are increased the packet loss is sufficiently increased in DSDV. This is because the packet has to travel more to reach the destination and network is also getting congested.

However, the increase level of mobility results in higher packet loss in AODV. The reason why is simple because the packet have to traverse through hops and the last used link may be not valid because of the change in topology.

Table 4: packet loss in DSDV and AODV

Mobility (m/s)	DSDV	AODV
1	0	0
2	0.0613	0.0669
4	0.0864	0.0799
10	0.088	0.0996
13	0.178	0.1577
15	0.1896	0.1976
18	0.1946	0.513
25	0.196	0.588
30	0.1886	0.598

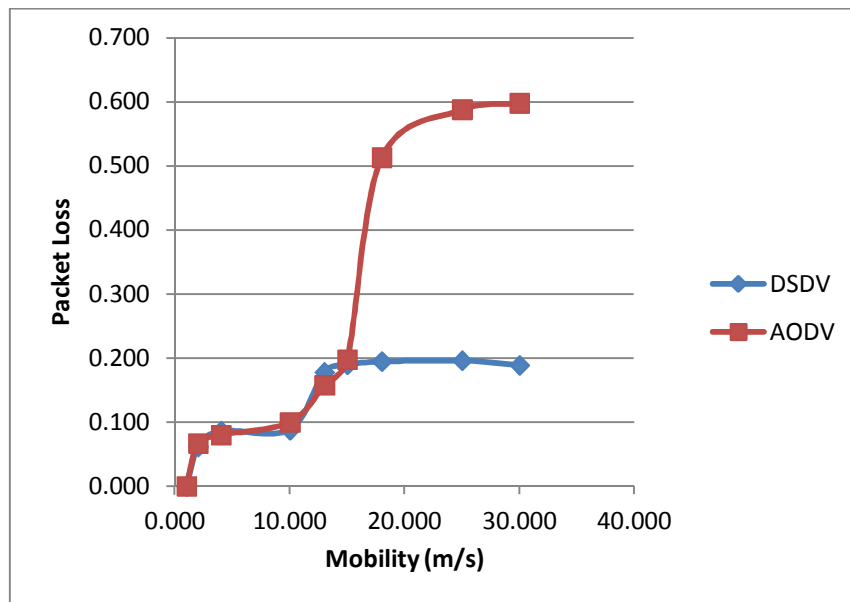


Figure 5: packet loss in DSDV and AODV

6.3 Routing Overhead

As shown in figure (6) DSDV protocol the effect of the increasing the level of mobility can be clearly seen. The routing overhead is increasing drastically by the increase in the level of mobility in the nodes. This is because the DSDV protocol is proactive so it is must to know the route before required but the change in topology will increase the routing overheads.

We can conclude by looking carefully at (figure 6) that the routing overhead in DSDV is significantly higher than the routing overhead in AODV which is nearly 950%.

Table 5: the Routing Overhead in DSDV and AODV

Mobility (m/s)	DSDV	AODV
0	0	0
5	9.55	0.1653
10	10.549	1.2011
15	13.25	1.3996
20	15.78	1.499
25	16.85	1.5985
30	17.9	1.5775
35	18.77	1.8877
40	19.121	2.0175

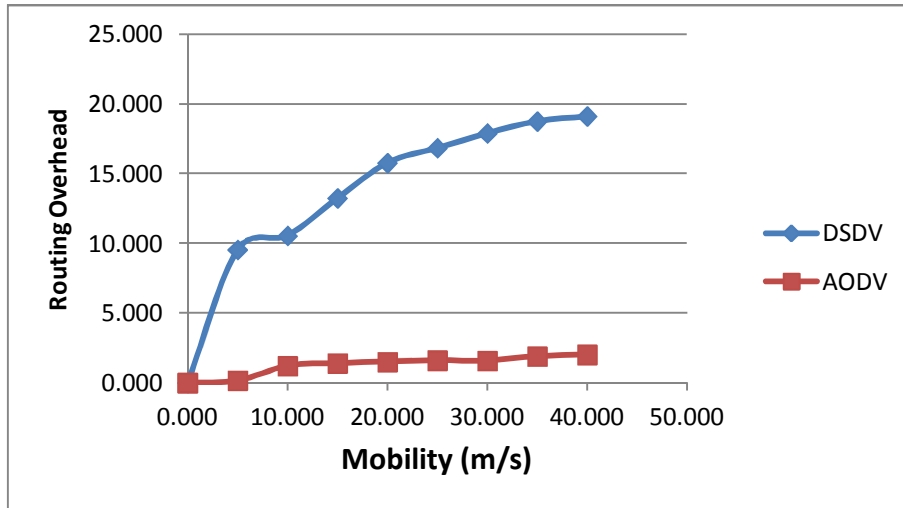


Figure 6: Routing Overhead in DSDV and AODV

7. Conclusion

MANET is a collection of mobile nodes, dynamically establishing short-lived networks where no prefixed infrastructure is needed. In this research we had a look on some routing protocols of this type of networks, which are differentiated in the working mechanism and how and when they build the routing table. And because of the time constraints we have focused on two different grouped protocols which are the reactive AODV and the proactive DSDV routing protocols, these routing protocols are proposed for ad-hoc mobile networks. In DSDV routing protocol, mobile nodes periodically broadcast their routing information to the neighbours. Each node requires to maintain their routing table. AODV protocol finds routes by using the route request packet and route is discovered when needed. The comparison of these protocols is done with random movement of the nodes which is changed over time. The parameters throughput, routing overhead, and packet loss have shown that AODV performs better than DSDV in throughput and routing overhead. However, figures have shown that packet loss in AODV is higher than in DSDV.

8. **Future Work**

As my future studies and to observe the protocols more objectively, it is worth trying different applications with different traffic types in.

A comparison of two routing protocols, AODV and DSDV has been carried out. Which can be proposed to compare all other routing protocols considering the same simulation parameters so that an exhaustive comparison of various routing protocols can be made.

By studying and analysing the building block of Routing Protocols it make possibilities to create better routing protocol by new change in these routing protocols.

References

- [1] M. Ilyas, "The Handbook of Ad Hoc Wireless Networks", CRC Press, 2003.
- [2] S. Sarkar, T. G. Basavaraju, C. Puttamadappa, "Ad hoc Mobile Wireless Networks Principles, Protocols and Applications", Auerbach Publications (Taylor & Francis Group), 2008.
- [3] DevashishRastogi and SachinGanu and Yanyong Zhang and Wade Trappe and Charles Graff, "A Comparative Study of AODV and OLSR on the ORBIT Testbed ". IEEE Military Communications Conference, 2007.
- [4] Pirzada, A.A., Portmann, M., Indulska, J., "Performance Comparison of Multi-Path AODV and DSR protocols in Hybrid Mesh Networks", 14th IEEE International Conference on Networks, vol. 02, Pages: 1-6, 2006.
- [5] X. Hong, Kaixin Xu, and Mario Gerla, "Scalable Routing Protocols for Mobile Ad Hoc Networks" IEEE, pp.11-21, 2002.
- [6] V. Garousi, "Simulating Network traffic in Multi-hop Wireless Ad Hoc Networks based on DSDV (Destination Sequenced Distance Vector) protocol using NS (Network Simulator) Package", University of Waterloo, Fall, 2001.
- [7] C. Perkins, E. Royer, "Ad-hoc on-demand distance vector routing" in Mobile Computing Systems and Applications, Proceedings WMCSA '99, Second IEEE Workshop, pp. 90-100, 1999.
- [8] D. Johnson, D. Maltz, and Yih-Chun Hu, The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks (DSR), Internet- draft, 2004.
- [9] Perkins and P. Bhagwat, "Highly Dynamic Destination-Sequenced Distance-Vector Routing (DSDV) for Mobile Computers", Comp. Comm. Rev., Oct. 1994.

- [10] T. Issariyakul, E. Hossain, “Introduction to Network Simulator NS2”, Springer, 2008.
- [11] Balram Swami and Ravindar Singh, “Simulation Based Comparison between OWL and DSDV”, Published in : International Conference on Emerging Trends in Engineering, Science and Technology, pp. 1575 – 1580,2015.
- [12] A. A. Chavan , Prof. D. S. Kurule and Prof. P. U. Dere, Performance Analysis of AODV and DSDV Routing Protocol in MANET and Modifications in AODV against Black Hole Attack, Published in : 7th International Conference on Communication, Computing and Virtualization, pp. 835 – 844, 2016.
- [13] V. Rajeshkumar, P.Sivakumar , Comparative Study of AODV, DSDV and DSR Routing Protocols in MANET Using Network Simulator-2 : International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 12, December 2013.



Refereed Journal

Journal of Electronic Systems and Programming

Issue No:2 June 2019