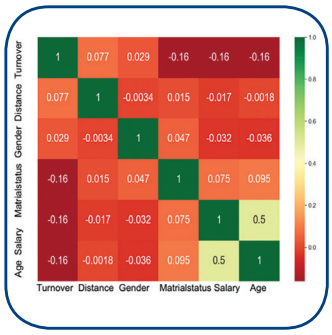
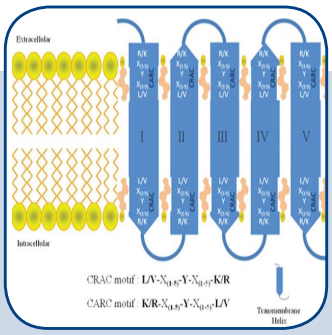
Volume 13, Number 2, 2022 ISSN 1847-6996

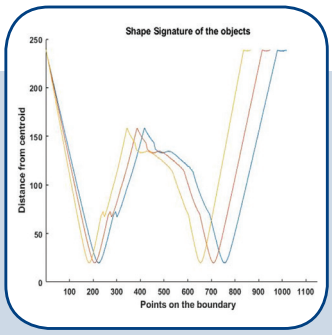
FERIT
FACULTY OF ELECTRICAL ENGINEERING, COMPUTER SCIENCE AND INFORMATION TECHNOLOGY **OSIJEK**

IJECES International Journal of Electrical and Computer Engineering Systems

International Journal of Electrical and Computer Engineering Systems







ISSN: 1847-6996

INTERNATIONAL JOURNAL OF ELECTRICAL AND COMPUTER ENGINEERING SYSTEMS

Published by Faculty of Electrical Engineering, Computer Science and Information Technology Osijek, Josip Juraj Strossmayer University of Osijek, Croatia

Osijek, Croatia | Volume 13, Number 2, 2022 | Pages 87-163

The International Journal of Electrical and Computer Engineering Systems is published with the financial support of the Ministry of Science and Education of the Republic of Croatia

CONTACT

International Journal of Electrical and Computer Engineering Systems (IJECES)

Faculty of Electrical Engineering, Computer Science and Information Technology Osijek,
Josip Juraj Strossmayer University of Osijek, Croatia
Kneza Trpimira 2b, 31000 Osijek, Croatia
Phone: +38531224600, Fax: +38531224605
e-mail: ijeces@ferit.hr

Subscription Information

The annual subscription rate is 50€ for individuals, 25€ for students and 150€ for libraries.
Giro account: 2390001 - 1100016777,
Croatian Postal Bank

Journal is referred in:

- Scopus
- Web of Science Core Collection (Emerging Sources Citation Index - ESCI)
- Google Scholar
- CiteFactor
- Genamics
- Hrčak
- Ulrichweb
- Reaxys
- Embase
- Engineering Village

Bibliographic Information

Commenced in 2010.
ISSN: 1847-6996
e-ISSN: 1847-7003
Published: quarterly
Circulation: 300

IJECES online

https://ijeces.ferit.hr

TABLE OF CONTENTS

A Multi-objective Hybrid Optimization for renewable energy integrated Electrical Power Transmission Expansion Planning

Original Scientific Paper

Shereena Gaffoor
Department of Ship Technology,
Cochin University of Science and Technology
South Kalamassery, Ernakulam, India
shereena.g@gmail.com

Mariamma Chacko
Department of Ship Technology,
Cochin University of Science and Technology
South Kalamassery, Ernakulam, India
mariamma@cusat.ac.in

Abstract *– Due to the large size of conventional electrical power transmission systems and the large number of uncertainties involved, achieving the most favourable Transmission Expansion Planning solution turns out to be almost impossible. The proposed method intends to develop a novel method to solve Transmission Expansion Planning problems in electric power systems incorporating renewable energy sources like wind turbines and Photo Voltaic array using IEEE 24 Reliability Test System. For enhancing the efficiency of search processes and to make its use easier on diverse networks and operations, the hybridization of two renowned meta-heuristic algorithms known as Grey Wolf Optimization (GWO) and Genetic Algorithm (GA) termed as Grey Wolf with Genetic Algorithm (GWGA) is adopted. A novel distance factor based on the best position and current position of the solution in Grey wolf optimization is introduced and proposed for the hybridization technique and gives a quick and promising solution with reduced computational time. The GWO and GA algorithms are combined suitably to achieve the advantages of both algorithms. With this proposed model, the investment cost of the transmission line and the maximum amount of power that can be distributed to the consumer is optimized with an objective of minimum load shedding. Among the state-of-the-art optimization techniques considered, a remarkable performance percentage improvement in the expansion plan in terms of cost reduction and load shedding minimization has been obtained in GWO, but when hybridized with GA, an improvement of 13.42% in cost function and 18.65% in load shedding is achieved for a population size of 60. Hence, the proposed method guarantees to generate the best solution with a faster convergence resulting in reduced computational time.*

Keywords*: genetic algorithm, grey wolf optimization, multi-objective, transmission planning, renewable integration*

1. INTRODUCTION

Transmission systems exists as the foremost element of electric power production [1-2]. It not only offers an association for distribution and generation but also provides consistent and non-discriminative surroundings to consumers and suppliers. The intention of a power transmission network is to convey power from production plants to load centres efficiently, economically, reliably and securely [3-5]. As transmission systems expand, Transmission Expansion Planning (TEP) entails recognizing where to include new circuits to congregate the increased requirement by transferring the power to a new network from the old one. In the last few years, investigations in the area of transmission planning designs have seen an expansion. Several reports and publications regarding novel designs have been available in the technical literature owing to the development of novel optimization algorithms, computer power accessibility and the larger uncertainty level caused by the power sector deregulation.

Transmission system planners exploit mechanical expansion designs to find out a favourable expansion system by reducing the numerical objective functions with respect to the number of constraints. Power system planning is essential to provide reliable, sustain-

able and affordable energy appropriately in developed countries. As a result, power system planning and associated energy issues have captured the attention of the research society [6-8].

From an upcoming scenario of power generation and requirement, the major intention of the TEP crisis solution is to discover the best set of fortifications for transmission that guarantees sufficient distribution of power to users. Therefore, the static TEP crisis solution should indicate how much and where the transmission equipment must congregate the requirements of energy. On the other hand, with the minimum feasible expenses, TEP crisis solutions should meet particular quality stipulations regarding the services. Conventionally, with the intention of guaranteeing the security criterion, TEP issues have been brought to a solution by deterministic models like renowned "N-1" and "N-2" [9-11]. In several cases, the attained arrangement can assist in the increased cost of investment.

A novel evolutionary algorithm [1] to resolve TEP issues in electric power systems was suggested to deploy various operators and a system for selecting the feasibilities of these operators. The arithmetical formulation regards a DC network design together with 'N-1'deterministic criterion and the transmission losses. The implemented technique was applied to a renowned test method and has proven its efficiency. An innovation model [7] was proposed for the planning and modelling of internally connected power systems. The suggested planning model focuses on Carbon Capture and Sequestration technologies, renewable energies and Demand-Side Management as well as CO2 and reserve emission restraints. The innovation of this model relies on an estimation of the above-mentioned proposals which were integrated that could expose feasible synergies and interactions contained by the power system. A Constructive Heuristic Algorithm (CHA) [8] was suggested to discover the most applicable route from candidate expansion routes set for minimizing the exploration space and as a result, enlarge the effectiveness of the PSO process. The suggested methodology was practiced with two real equivalent systems and with the Garver system for the Southeast and South of Brazil in which the effectiveness of the proposed scheme can be confirmed. A preliminary Generation Expansion Planning (GEP) design [9] with obtainable input data from a variety of sources was proposed. Discrete circumstances of feasible climate change effects were described and optimization designs were configured to distinctively design uncertainty. Associations among GEP parameters and climate change were described for the entire situation to regard their effects. A multi-objective shuffled frog leaping scheme [10] was suggested to manage with the bi-level, multi-objective and non-linear nature of the design. The achieved outcomes demonstrate that the implementation of an appropriate approach for TEP could lead to more private investment in the absorption of wind power devoid of a noteworthy transmission investment cost. A proficient and enhanced GA [11] suggested was capable of functioning with various updated operators to guarantee its suitable computation in attaining constructive and best possible coordinated planning crisis solutions. A model [12] that executes optimal TEP powerfully in a Stochastic Optimization context was suggested. The design exploits a customized model of Benders' decomposition that remunerates from numerous developments that were portrayed. It manages with the integration of eventualities using a double structural design for Benders cuts and an improved contingency incorporation scheme. Additionally, it includes the capability to discover the potentially interesting candidate transmission lines mechanically which was particularly exciting in large-scale inconveniences. A static representation for Coordinated Generation and Transmission Expansion Planning (CGTEP) was proposed [13]. On minimizing the investment cost, energy and operation were not provided inside the system, the design intends to alleviate the susceptibility of a power system in opposition to physical intentional attacks in the planning horizon. In addition, the peak load of twelve days in a year was taken as a sample of the months to consider the impacts of load distinctions over a year. The physical intentional attacks and their resulting impacts were measured through the scenario building process. Accordingly, in any given month, the entire scenario was constructed as an attack plan aiming at the transmission system and therefore, they were assigned weights, which were proportional to the subsequent damage imposed on the power system. The importance of the suggested scheme in mitigating the susceptibility of power system was well established by mathematical results. A possible reason for this may be technical security solutions because they rarely include the influence of the human factor on the system security level.

Several research works on hybridized optimization of Transmission Expansion problem have been done recently. A hybrid algorithm that combines Benders decomposition and a Bees algorithm is introduced in [14] and has been tested using the transmission network expansion and energy storage planning model. The hybridized model is designed to produce good solutions quickly while still retaining a guarantee of optimality when run for a sufficiently long time. A non-linear control parameter based on cosine function is presented to replace the original linear parameter of Grey Wolf Optimizer in [15]. Also, the crossover and mutation operation of Genetic Algorithm is introduced into GWO to avoid local minima and premature convergence. A novel hybrid optimization technique based on Differential Evolution and Continuous population based Incremental Learning is proposed in [16] with load Shedding formulation as well as it includes the optimization of shunt compensation. The min conflict local search algorithm hybridized with Grey wolf optimizer for the power scheduling problem is proposed in [17]. The proposed method is compared with twenty state-of-the-art methods. Stochastic transmission expansion planning in the presence of

wind farms considering reliability and N-1 contingency using Grey wolf Optimization technique is proposed in [18]. To check the effectiveness of the proposed method it is compared with other optimization techniques. A new hybrid GA with linear modelling is proposed for TEP Problem in [19]. It is tested on Garver 6 bus system, IEEE 24 bus and South Brazilian test which showed a rapid convergence on the test problem. A hybrid Genetic gradient algorithm is proposed in [20]. The effectiveness and practicability of the proposed model are verified by choosing a Central region of China.

This paper proposes a technique for solving the problems associated with TEP in electrical power systems incorporating renewable energy sources such as wind turbines and Photo Voltaic (PV) array using IEEE 24 Reliability Test System (RTS). For enhancing the exploitation of the transmission systems in various operating conditions and to improve the exploring efficiency, two renowned algorithms namely, Genetic Algorithm (GA) and Grey Wolf Optimization (GWO) are hybridized and adopted.

GWO has proven better capability to not fall in local optima in the initial search [21, 22]. GA can converge to get an overall good solution [23-26]. Hence an algorithm is proposed to have GWO followed by GA by including a novel distance factor for hybridization. Using this proposed GWGA algorithm, the reinforcement line has been optimized and the load has been maximized with reduced investment cost. Among the state-of-the-art optimization methods compared, GWO gives the best cost function and minimum load shedding but when hybridized with GA it gives a superior solution compared to when GWO and GA performed individually.

Next, to the implementation, this scheme is compared with the traditional algorithms such as GA, Particle Swarm Optimization (PSO), Artificial Bee Colony (ABC), Firefly (FF) and GWO, and the results are presented.

The paper is organized as follows: Section 2 illustrates the TEP problem statement and Section 3 explains the proposed GWGA algorithm. Section 4 demonstrates the TEP enhancement by hybrid mechanism, Section 5 provides the results and discussion to examine the performance of the proposed methodology and Section 6 concludes the research work.

2. TEP PROBLEM STATEMENT

Several mathematical models for modelling TEP found in the literature are transportation model, DC model, AC model, hybrid model and disjunctive model [3-4]. These models differ from each other in the level of complexity and accuracy. In this work, DC power flow model is used as it is widely used in Transmission Expansion Planning [27, 28]. However, the methodology presented in this work can be extended to the other models also. The formulation of the classic load flow problem [29, 30] requires considering four variables at each bus i of the power system. These variables are the P_i (Net active power injection), Q_i (Net reactive power injection), V_i (Voltage magnitude) and ∂_i (Voltage angle).

Direct current Load flow (DCLF) gives an estimation of line power flows on AC power systems. DCLF looks only at active power flows and neglects reactive power flows. In DCLF, the non-linear model of the AC system is simplified to a linear form.

The suggested technique for the TEP crisis [31] is dependent on deterministic and static modelling [12] with single-stage production and power requirement. The TEP comprises in formatting such branches of the network which must be made stronger by integrating the required amount of reinforcement. Reinforcements are tools to be mounted on the system and could not be fractionated so that the candidate plan variable z must be an integer value [12].

The actual crisis of the TEP problem is partitioned into two sub-problems [1] which are Operational Problem (OP) and Investment Problem (IP). IP is categorized as a non-linear constrained optimization problem in which the concern is to find an expansion model, which reduces the investment expenditure meeting security measures and upcoming demand.

For a specified expansion plan z, the objective function of the crisis is given by $g(z)$. IP is modelled by Eq. (1).

$$\text{Minimize} \quad g(z) = \sum_{(i,j)\in\mu} C_{ij} n_{ij} \tag{1}$$

$$\text{Subject to} \quad \begin{aligned} &B(z) = 0 \\ &C(z) = 0 \\ &0 \le n_{ij} \le n_{ij\max} \end{aligned}$$

C_{ij} is the unit cost related to the reinforcements to branch connecting bus i to j, μ indicates the group of entire candidate branches of the system network, $n_{ij\max}$ is the maximum number of reinforcements to be integrated to branch connecting bus i to j. Function $B(z)$ indicating load shedding and $C(z)$ indicating overload are described subsequently.

For a specified expansion plan z, the OP is considered as a problem of LP, where the objective is to find the economic transmit of generators which reduces the total load shedding $B(z)$ and is given by Eq. (2).

$$\text{Minimize} \quad B(z) = \sum_{p=1}^{\alpha} r_p \tag{2}$$

$$\text{Subject to} \quad \begin{aligned} &f + r + A\theta = h \\ &\left|g_{ij}\right| \le (n_{ij}^0 + n_{ij\max}) g_{ij\max} \\ &f_{\min} \le f \le f_{\max} \\ &0 \le r \le h \end{aligned}$$

The amount of generation produced by renewable energy sources (Wind turbine and PV array) are represented by r whose constituent r_p indicates the load shed at bus p.

Hence r_p can be represented by Eq. (3).

$$r_p = \begin{Bmatrix} r^w & if\ RE = 1 \\ r^{pv} & if\ RE = 2 \end{Bmatrix} \quad (3)$$

where RE refers to renewable energy sources, r^w indicates energy from wind turbine, r^{pw} indicates energy from PV array.

Moreover, the generating vector is indicated by f and its constituent f_b denotes the production at bus b, the network susceptance matrix is denoted by A, the vector of load demand is given by h where its constituent h_p denotes the load at bus p, the voltage angle vector is denoted by θ and its constituent θ_p indicates the angle at bus p, g_{ij} is the active power flow of the branch connecting bus i to j given by $g_{ij} = (\theta_i - \theta_j)/x_{ij}$ and x_{ij} denotes the reactance. The maximum capacity of each constituent is represented by $g_{ij\max}$ for branch connecting bus i to j; the term $f_{\max}$ and $f_{\min}$ denote the vectors that symbolise the maximum and minimum capacity of generators; the number of elements in the branch connecting bus i to j at the base case is indicated by n^0_{ij} and α is the group of entire network buses. The values of f,r,h,g and x_{ij} are given in p.u and θ values are denoted in radians.

Overload function is represented by $C(z)$. The value of $C(z)$ is obtained by solving DC power flow using the generator vector found by the OP solution. For each branch of the system, the occurrence of overloads is verified. Thus overload function will depend on the reinforcements present in the plan z and should be nil.

The total energy from wind turbine r^w which is generated depending on the texture speed [32] is given by Eq. (4).

$$r^w = \sum_{s=1}^{\omega} \frac{M_0 D v_s^3 L_p}{2 O_L T} e^{\frac{-gH}{O_L T}} \quad (4)$$

ω indicates the number of wind turbines, M_0 is the standard sea-level atmospheric pressure (101325 Pa), D is the swept area in m^2, v indicates velocity of the wind turbine in m/s, L_p is the wind turbine power coefficient defined as the ratio of actual power produced by a wind turbine divided by the total wind turbine power flowing into the turbine blades, O_L indicates specific gas constant for air (287J/(Kg. K)), T is the temperature in Kelvin which can be computed at any height in meter as $T = T_0 - LH$, where T_0 is the temperature at sea level (288K), L is the temperature lapse rate (0.0065°C/m), H is the altitude above sea level in m, g is the gravity constant (9.8m/s2).

The electrical energy generated from the PV array [33] can be found out as given in Eq. (5).

$$r^{PV} = \sum_{t=1}^{q} \eta_m C_m I(t) \quad (5)$$

For the proposed research work, identical PV modules are considered. Here η_m indicates PV module efficiency, C_m indicates area of PV module in m^2, $I(t)$ refers to solar intensity (W/m^2) and q is the number of PV modules.

3. PROPOSED GWGA ALGORITHM

The conventional Grey Wolf Optimization algorithm has certain disadvantages such as low solving precision, slow convergence and bad local searching ability. In order to overcome these disadvantages of GWO, it is hybridized with GA algorithm, as it can provide global optimal solutions. By hybridizing both algorithms, the solution for minimizing the investment cost with the objective function can be obtained.

The process of GWO algorithm [34-35] describes the hierarchy of grey wolves hunting and leadership characteristics. There are four classifications of grey wolves known as α, β, ω and δ that are exploited for executing the leadership hierarchy. Penetrating, encircling and attacking the prey are the three chief practices in hunting that are deployed to progress optimization.

A genetic algorithm [36-37] is a search heuristic that is inspired by Charles Darwin's theory of natural evolution. This algorithm reflects the process of natural selection where the fittest individuals are selected for reproduction in order to produce offspring of the next generation. GAs are the ways of solving problems by mimicking processes that nature uses, i.e., Selection, Crossover, Mutation and accepting to evolve a solution to a problem.

GWGA algorithm is the hybridization of GWO and GA algorithms. Here, the crossover rate r_c is fixed at 0.6, and the bounding factor is given by b as in Eq. (6), where b_{max} and b_{min} indicates the upper bound and lower limits of the solution. Moreover, a distance factor 'd' based on the best position and current position of solution is introduced and can be evaluated as given by Eq. (7), where α_{po} is the best position of the solution and ς_{po} is the current position of the solution.

$$b = \sqrt{mean(b_{\max} - b_{\min})^2} \quad (6)$$

$$d = \sqrt{mean(\alpha_{po} - \varsigma_{po})^2} \quad (7)$$

Moreover, the distance threshold d_{th} can be evaluated by Eq. (8), where $Iter$ denotes current iteration and $Iter_{\max}$ indicates the maximum iteration.

$$d_{th} = r_c \times b \times \frac{Iter}{Iter_{\max}} \quad (8)$$

Also, if the distance d is greater than d_{th} the solution can be updated using GWO. Otherwise, crossover operation has to be performed using the principle of GA algorithm, and the updated solution can be obtained as shown in Eq. (9).

$$Z_2^* = \frac{child1 + child2}{2} \quad (9)$$

Thus, the optimized solution of TEP is obtained as Z^*, which attain the minimized cost and balanced power generation in order to meet the demand of the power system.

The flow chart for the proposed GWGA algorithm is shown in Fig. 1

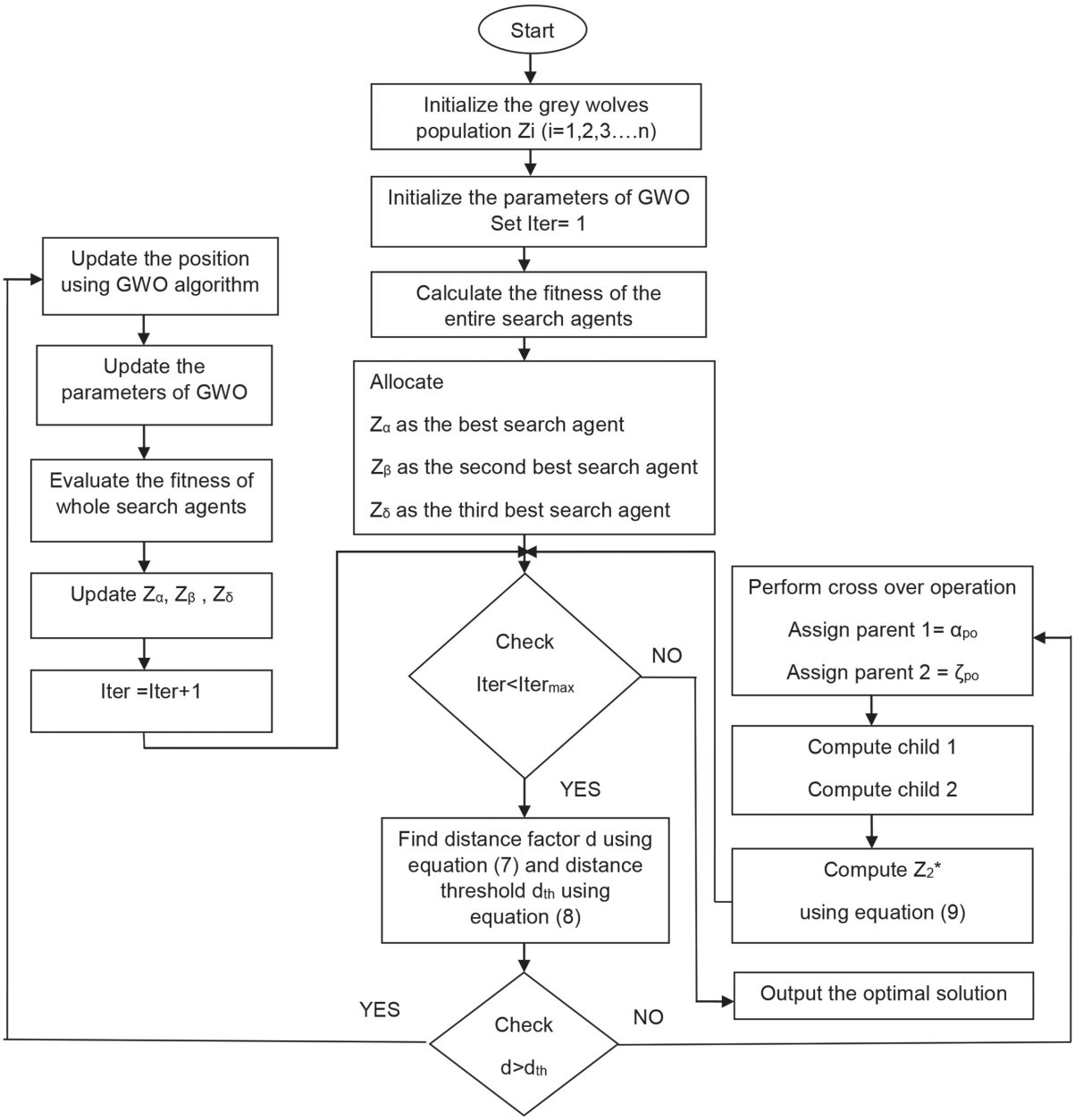


Fig. 1.Flowchart of proposed GWGA Algorithm

4. TEP ENHANCEMENT BY HYBRID OPTIMIZATION

For solving the TEP problem, the number of branches, generators and renewable energy sources is given as solutions for encoding, which has to be optimized. The bounding limit of the number of branches ranges from $n_{ij\min}$ to $n_{ij\max}$. Similarly, the bounding limit of generator bus ranges from $g_{ij\min}$ to $g_{ij\max}$ and the bounding limit of renewable energy sources ranges from 0 to 2, i.e. $RE_i \in\{0,1,2\}$. If RE_i =0, it means that there were no renewable energy sources connected to the bus. If RE_i =1, the wind turbine is connected to the bus and if RE_i =2, PV array is connected to the bus system. Fig. 2 reveals the solution given for encoding process, where N_B indicates the number of branches, N_G denotes the number of generators and N_R is the count of renewable sources connected to the IEEE 24 bus systems. The length of the solution is the summation of N_B, N_G and N_R. Collectively, the solution that is to be optimized is termed as Z.

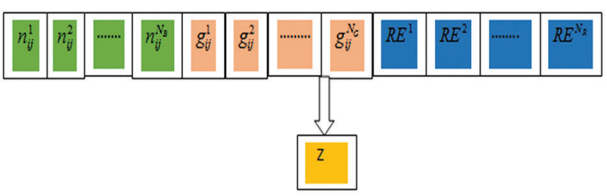


Fig. 2. TEP solution encoding

The IEEE 24 RTS shown in Fig. 3 includes 11 synchronous generators with 17 load points and 38 branches, where each branch can receive a maximum of three reinforcements; the total demand is 2850 MW and the maximum generation capacity is 3405 MW.

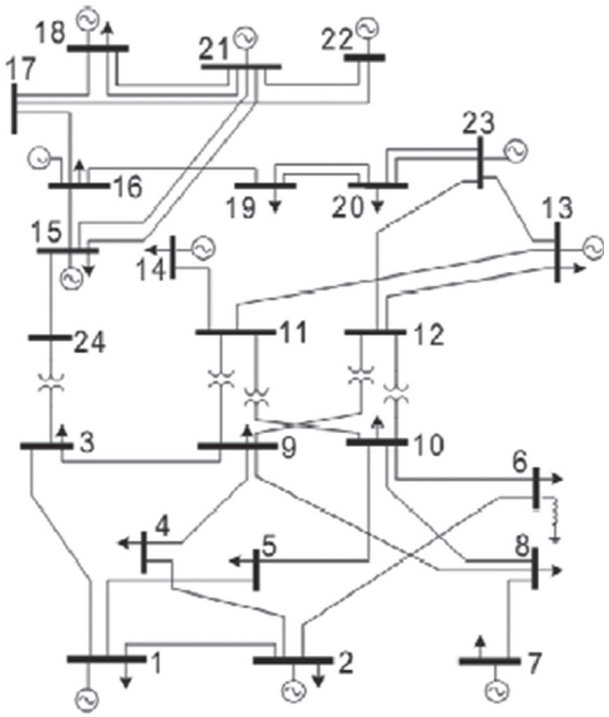


Fig. 3. Single line diagram of IEEE 24 bus Reliability Test System

Table 1 gives the transmission line data indicating the branch numbers. For the optimization process the length of the chromosome (solution size) is taken as 61. Chromosome number 1 to 38 indicate the number of branches, 39 to 47 represents the nodes where generators are connected and 48 to 61 represents the nodes where renewable energy sources are connected.

All wind farms are equipped with three 1.5MW wind turbines. Table 2 shows the manufacturer only specification of a 1.5 MW GE wind turbine [32].

Various design parameters taken for calculating the electrical energy from PV array are: Am = 1.3264 m^2, η_m= 15% and I(t) =1000 W/m^2 [33].

Table 1. Transmission line data indicating branch number

Line No.	From Bus No.	To Bus No.	Line No.	From Bus No.	To Bus No.
1	1	2	20	12	13
2	1	3	21	12	23
3	1	5	22	13	23
4	2	4	23	14	16
5	2	6	24	15	16
6	3	9	25	15	21
7	3	24	26	15	21
8	4	9	27	15	24
9	5	10	28	16	17
10	6	10	29	16	19
11	7	8	30	17	18
12	8	9	31	17	22
13	8	10	32	18	21
14	9	11	33	18	21
15	9	12	34	19	20
16	10	11	35	19	20
17	10	12	36	20	23
18	11	13	37	20	23
19	11	14	38	21	22

Table 2. Specification of a GE 1.5SLE wind turbine

Parameter	Value
Rated Output	1.5MW
Rotor diameter	77m
Cut in wind speed	3.5m/s
Rated wind speed	14m/s
Cut out wind speed	25m/s
Hub height	80m

5. RESULTS AND DISCUSSIONS

The proposed TEP model with renewable energy sources [38-40] such as wind turbine and PV array using GWGA method has been simulated in MATLAB, tested in IEEE 24 bus system and the results obtained are analysed. The performance of the proposed model has been compared with traditional algorithms such as GA [23-26], PSO [41-46], ABC [47-48], FF [49-50] and GWO [22-28] based on cost function and load shedding function. The corresponding outcomes were evaluated to certify the performance of the proposed model.

5.1. COST FUNCTION ANALYSIS

The performance comparison of the cost function of the proposed GWGA model with conventional algorithms is given in Fig. 4. For evaluating the effectiveness of the solution, i.e. to check whether the solution is local optimal or global optimal, the number of iterations is extended to 500 for a population size of 10 and the performance is shown in Fig. 5. The unit cost is randomly generated for the given population size. The proposed method shows the effectiveness in reducing the computational time and hence shows better convergence speed.

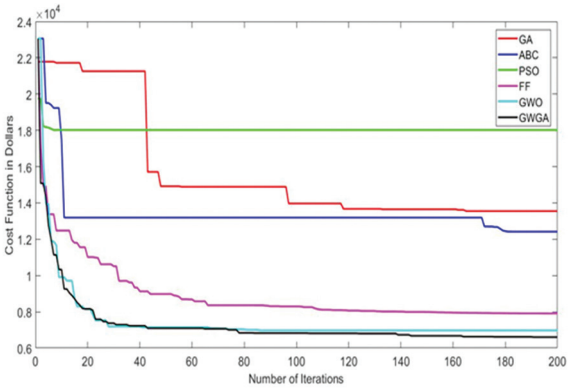


Fig. 4. Cost function analysis of different optimization techniques for iterations 0 to 200

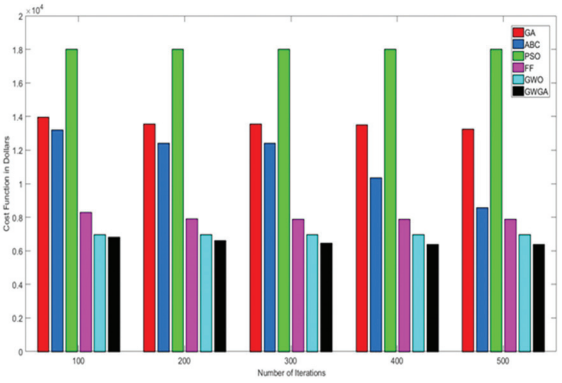


Fig. 5. Cost function analysis of different optimization techniques for iterations 100 to 200

The performance analysis is also done for varying population sizes 10, 30, 60, 80 and 100 for 200 iterations and the comparison result is shown in Table 3 and the performance improvement percentage of the proposed GWGA method compared with conventional algorithms is given in Table 4.

The results show that the proposed GWGA gives better performance than the state-of-the-art optimization methods compared in the proposed work in TEP. For all the population size considered, GWO gives the best performance and the quality solution is obtained for a population size of 100. When hybridized with GA, the best solution is obtained for a population size 60 giving a 15.76% improvement in cost function than the individual approach. The result also shows that the best solution obtained with GWGA is 13.42% better than the best solution obtained for a population size 100 when GWO performed individually. Similarly, with GA approach, the best solution is obtained for a population size of 100. When hybridized with GWO, the best solution is obtained for a population size 60 giving a 61.95% improvement in cost function than the individual approach. The result also shows that the best solution obtained with GWGA is 54.42% better than its best solution when GA has been done individually. Hence the proposed hybridization gives the best solution for less population size thereby computation time is reduced.

The results also indicate the effectiveness of the proposed method to deal with different sizes of the network and hence the proposed method is scalable. Thus, from the simulation analysis, the functionality of the proposed scheme to meet the considered power demand with minimal cost can be attained.

Table 3. Cost function analysis using different optimization techniques with varying population size

Population Size	10	30	60	80	100
	Transmission Cost in Dollars				
GA	15795.9	16486.9	17050.68	14957.14	14235.5
ABC	10229.5	12221.01	13512.73	12673.2	10908.36
PSO	18492.77	17738.34	18137.93	14957.14	14235.5
FF	9385.805	15187.47	12271.21	9566.683	13014.52
GWO	7782.325	7826.439	7701.398	7764.962	7492.607
GWGA	7197.423	7197.894	6487.372	6493.273	7233.44

Table 4. Performance Improvement of GWGA with respect to conventional algorithms (%)

Population Size	10	30	60	80	100
GA	54.43	56.34	61.95	56.59	49.19
ABC	29.64	41.1	51.99	48.76	33.69
PSO	61.08	59.42	64.23	56.59	49.19
FF	23.32	52.61	47.13	32.13	44.42
GWO	7.52	8.03	15.76	16.38	3.46

5.2 LOAD SHEDDING ANALYSIS

The performance comparison of load shedding of the proposed GWGA model with conventional algorithms is given in Fig. 6. The number of iterations is extended to 500 for a population size of 10 giving the convergence speed performance as shown in Fig. 7.

The performance analysis is also done for varying population sizes 10, 30, 60, 80 and 100 for 200 iterations and the comparison result is shown in Table 5 and the percentage performance improvement in load shedding of the proposed GWGA methodology compared with conventional algorithms is shown in Table 6.

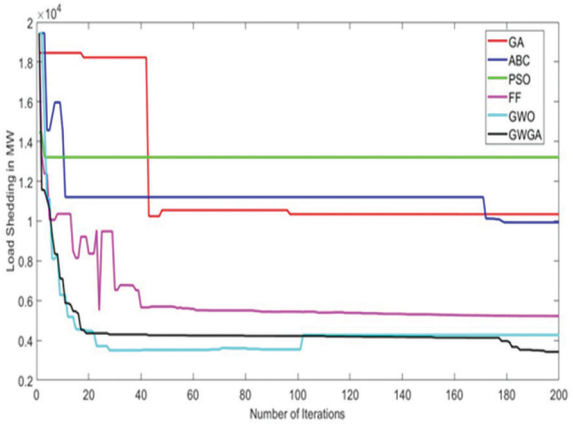


Fig. 7. Load shedding analysis of different optimization techniques for iterations 100 to 500

Table 5. Load shedding analysis using different optimization techniques with varying population size

Population Size	10	30	60	80	100
	Load Shedding in MW				
GA	10345.05	10913.29	11228.2	9279.316	8907.103
ABC	5597.098	7674.407	8855.18	8391.756	6657.085
PSO	13206.99	12081.91	12193.9	9279.316	8907.103
FF	5201.581	9837.438	7143.93	4780.685	7857.024
GWO	4269.435	4339.245	4163.62	4260.662	4215.924
GWGA	4097.052	4097.522	3387	3392.901	4100.289

Table 6. Performance Improvement of GWGA with respect to conventional algorithms (%)

Population Size	10	30	60	80	100
GA	60.4	62.45	69.83	63.44	53.97
ABC	26.8	46.61	61.75	59.57	38.41
PSO	68.98	66.09	72.22	63.44	53.97
FF	21.23	58.35	52.59	29.03	47.81
GWO	4.04	5.57	18.65	20.37	2.74

For all the population size considered, GWO gives the best performance and the quality solution is obtained for a population size of 60. When hybridized with GA, the best solution is obtained for a population size 60 giving an 18.65% improvement in load shedding function than the individual approach. Similarly, with the GA approach, the best solution is obtained for a population size of 100. When hybridized with GWO, the best solution is obtained for a population size 60 giving a 69.83% improvement in load shedding than the individual approach. The result also shows that the best solution obtained with GWGA is 61.97% better than its best solution when GA has been done individually. Hence the proposed hybridization gives the best solution for less population size thereby computation time is reduced.

When compared with ABC, FF and PSO, GWGA gave the best performance for all the population size considered. Hence, the proposed method gives the best solutions of IEEE 24 bus system for load shedding analysis for a population size of 60. The results depicted in Table 4 and Table 6 shows that GWGA converges much faster than GWO and GA algorithms. The computational time of the proposed method is less and gives the best results in the population size in the range of 30 to 60.

Table 7 shows the final expansion plan obtained by the methodology proposed. Table 8 compares the generation capacity obtained before and after TEP with GWGA methodology at different nodes and Table 9 shows the nodes where the renewable energy sources [51-55] are connected for optimal operation of the proposed work.

Table 7. Final Expansion Plan for IEEE 24 bus using the proposed methodology

Line No.	From Bus No.	To Bus No.	Line No.	From Bus No.	To Bus No.
1	3	20	2	12	13
2	2	21	2	12	23
3	2	22	1	13	23
4	2	23	2	14	16
5	3	24	3	15	16
6	1	25	1	15	21
7	2	26	2	15	21
8	3	27	1	15	24
9	2	28	3	16	17
10	2	29	3	16	19
11	3	30	2	17	18
12	2	31	2	17	22
13	2	32	3	18	21
14	1	33	2	18	21
15	3	34	1	19	20
16	3	35	1	19	20
17	1	36	2	20	23
18	3	37	3	20	23
19	1	38	3	21	22

Table 8. Comparison of Generation capacity before and after TEP at different nodes

Generator Bus No.	Generation Capacity before TEP without GWGA (MW)	Generation Capacity after TEP with GWGA (MW)
2	67	140.5605
7	64	148.5388
13	200	554.2946
15	274	24.0616
16	245	71.6325
18	144	351.6817
21	294	315.4005
22	150	300
23	200	200.7639

Table 9. Node numbers of buses where renewable sources are connected

Types of renewable Energy Sources	Node number
Wind turbines	5 8 9 10 11 14 17 19
PV array	6

6. CONCLUSION

This paper has presented a novel scheme to resolve the issues residing in TEP of electric power systems with renewable energy sources such as wind turbines and PV array by optimizing the reinforcement line, transmit of generators and renewable energy sources. It was applied and simulated in IEEE 24 test bus system. In addition, with the intention of improving the effectiveness of transmission in diverse networks and operation conditions, the hybridization of two well-known metaheuristic algorithms said to be GWO and GA named as GWGA is adopted. In order to achieve the advantages of both the methods and obtain a faster convergence at the same time, they are combined suitably and used. Consequently, with this novel hybridized model, the investment cost of the transmission line and the maximum amount of power that can be extended to the consumer is optimized.

The proposed scheme has been compared with the conventional algorithms such as GA, PSO, ABC, FF and GWO and the results are tabulated. The proposed GWGA method shows better optimization results considering expansion cost minimization and minimum load shedding when GWO and GA optimization methods are performed individually. The method is scalable in terms of the size of the network. The GWGA gives the best solution in the population range of 30 to 80 for 200 iterations, which reduces the computation time. The GWGA appears to be a very effective hybrid algorithm for highly complex transmission expansion problems.

In further studies, the proposed method will be applied for AC analysis as well as to a real power system with real data considering uncertainties.

7. REFERENCES:

[1] A. M. L. da Silva, M. R. Freire, L. M. Honorio, "TEP optimization by adaptive multi-operator evolutionary algorithms", Electric Power Systems Research, Vol. 133, 2016, pp. 173-181.

[2] R. Romero, A. Monticelli, "A Zero-One Implicit Enumeration Method for Optimizing Investments in Transmission Expansion Planning", IEEE Transactions on Power Systems, Vol. 9, 1994, pp. 1385-1391.

[3] M. V. F. Pereira, L. M. V. G. Pinto, S. H. F. Cunha, G. C. Oliveira, "A Decomposition approach to automated generation/transmission expansion planning", IEEE Transactions on Power Apparatus and Systems, Vol. 104,1985, pp. 3074-3083.

[4] T. Sum-Im, W. Ongsakul, "A Self-Adaptive Differential Evolution Algorithm for Transmission Network Expansion Planning with System Losses Consideration", Proceedings of the IEEE International Conference on Power and Energy, Kota Kinabula, Malaysia, 2-5 December 2012, pp. 152-156.

[5] C.Tian, X. Lu, L. Chu, T. Dong, D. Li, "Multi-Objective Transmission Network Planning with Consideration of Power Grid Vulnerability and Wind Power Accommodation", Journal of Engineering Science and Technology Review, Vol. 3, 2013, pp. 30-34.

[6] M. J. Rider, A. V. Garcia, R. Romero, "Power system transmission network expansion planning using AC model", IET Generation Transmission and Distribution, Vol. 5, 2007, pp. 731-742.

[7] O. J. Guerra, D. A. Tejada, G. V. Reklaitis, "An optimization framework for the integrated planning of generation and transmission expansion in interconnected power systems", Applied Energy, Vol. 170, 2016, pp. 1-21.

[8] I. M. de Mendonca, I. C. Silver Junior, A. L. M. Marcato, "Static planning of the expansion of electrical energy transmission systems using particle swarm optimization", International Journal of Electrical Power and Energy Systems, Vol. 60, 2014, pp. 234-244.

[9] S. Li, D. W. Coit, F. Felder, "Stochastic optimization for electric power generation expansion planning with discrete climate change scenarios", Electric Power Systems Research, Vol. 140, 2016, pp. 401-412.

[10] M. Jadidoleslam, A. Ebrahimi, Mohammad Amin Latify, "Probabilistic TEP to maximize the integration of wind power", Renewable Energy, Vol. 114, 2017, pp. 866-878.

[11] H. K. Rad, Z. Moravej, "An approach for simultaneous distribution, sub-transmission and transmission networks expansion planning", International Journal of Electrical Power and Energy Systems, Vol. 19, 2017, pp. 166-182.

[12] S. Lumbreras, A. Ramos, F. Banez-Chicharro, "Optimal transmission network expansion planning in real-sized power systems with high renewable penetration", Electric Power Systems Research, Vol. 149, 2017, pp. 76-88.

[13] H. Nemati, M. A. Latify, G. R. Yousefi, "Coordinated generation and TEP for a power system under physical deliberate attacks", International Journal

of Electrical Power and Energy Systems, Vol. 96, 2018, pp. 208-221.

[14] C. A. G. MacRae, M. Ozlen, A. T. Ernst, "The Bee-Benders Hybrid Algorithm with application to Transmission Expansion Planning", Proceedings of the Genetic and Evolutionary Computation Conference Companion, Lille, France, 10-14 July 2021, pp. 1275-1282.

[15] W. Gai, C. Qu, J. Liu, J. Zhang, "An Improved Grey Wolf Algorithm for Global Optimization", Proceedings of the 30th Chinese Control and Decision Conference, Shenyang, China, 9-11 June 2018, pp. 2494-2498.

[16] E. G. Morquecho, S. P. Torres, N. E. Matute, Fabian Astudillo-Salinas, Julio C. Lopez, Wilfredo C. Flores, "AC Dynamic Transmission Expansion Planning using Hybrid Optimization Algorithm", Proceedings of the IEEE PES Innovative Smart Grid Technologies, The Hague, Netherlands, 26-28 October 2020, pp. 499-503.

[17] S. N. Makhadmeh, A. T. Khader, M. A. Al-Betar, S. Nain, A. K. Abasi, Z. A. A. Alyasseri, "A novel hybrid grey wolf optimizer with min conflict algorithm for power scheduling problem in a smart home", Swarm and Evolutionary Computation, Vol. 60, 2021, pp. 2210-2227.

[18] A. A. Ghadimi, M. Amami, M. Bayat, S. Ahmadi, M. R. Mireh, F. Jurado, "Stochastic Transmission Expansion Planning in the presence of wind farms considering reliability and N-1 contigency using Grey Wolf Optimization Technique", Electrical Engineering, 2021, https://doi.org/10.007/s00202-021-01339-w

[19] E. S. Yigit, S. Mutlu, B. Babayigit, "Transmission Expansion Planning based on a hybrid genetic algorithm approach under uncertainty", Turkish Journal of Electrical Engineering and Computer Science, Vol. 27, 2019, pp. 2922-2937.

[20] L. Ma, Y. Wu, S. Lou, S. Lang, M. Liu, Y. Gao, "Coordinated Generation and Transmission Expansion Planning with High Proportion of Renewable Energy", Proceedings of the 4th IEEE Conference on Energy Internet and Energy System Integration, Wuhan, China, 30 October-1 November, 2020, pp. 1019-1024.

[21] A J. Wang, S. Li, "An Improved Grey Wolf Optimizer Based on Differential Evolution and Elimination Mechanism", Scientific Reports, Vol. 9, 2019, pp. 1-24.

[22] Z. Gao, J. Zhao, "An Improved Grey Wolf Optimization Algorithm with Variable Weights", Hindawi Computational Intelligence and Neuroscience, Vol. 2019, 2019, pp. 1-13.

[23] M. Nikolic, J. Jovic, "Implementation of genetic algorithm in map-matching model", Expert Systems with Applications, Vol. 72, 2017, pp. 283-292.

[24] R. A. Gallego, A. Monticelli, R. Romero, "Transmission system expansion planning by an extended genetic algorithm", IEE Proceedings Generation, Transmission and Distribution, Vol. 145, 1998, pp. 329-335.

[25] H. A. Gil, E. L. Silva, "A reliable approach for solving the transmission network expansion planning problem using genetic algorithms", Electric Power Systems Research, Vol. 58, 2001, pp. 45-51.

[26] Luis A. Gallego, Marcos J. Rider, Marina Lavorato, Antonio Paldilha-Feltrin, "An Enhanced Genetic Algorithm to Solve the Static and Multistage Transmission Network Expansion Planning", Journal of Electrical and Computer Engineering, Vol. 2012, 2012, pp. 1-12.

[27] G. Latorre, R. D. Cruz, J. M. Areiza, A. Villegas, "Classification of publications and models on transmission expansion planning", IEEE Transactions on Power Systems, Vol. 18, 2003, pp. 938-946.

[28] I. J. Silva, M. J. Rider, R. Romero, A. V. Garcia, C. A. Murari, "Transmission network expansion planning with security constraints", IEEE Proceedings on Generation, Transmission and Distribution, Vol. 152, 2005, pp. 828-836.

[29] Jizhong Zhu, "Optimal Power Flow", Optimization of Power System Operation, pp. 297-364, A John Wiley and Sons, Inc., Publication, 2009.

[30] Ahmed A. Fathy, Mohamed S. Elbages, Ragab A. El-Sehiemy, Fahmy M. Bendary, "Static transmission expansion planning for realistic networks in Egypt", Electric Power Systems Research, Vol. 151, 2017, pp. 404-418.

[31] M. S. El-bages, W. T. Elsayed, "Social spider algorithm for solving the transmission expansion

planning problem", Electric Power Systems Research, Vol. 143, 2017, pp. 235-243.

[32] H. M. K. Al-Masri, M. Ehsani, "Impact of wind turbine modeling on a renewable energy system", Proceedings of the North American Power Symposium, Denver, CO, USA, 18-20 September, 2016, pp. 1-6.

[33] Sumit Tiwari, Jasleen Bhatti, G. N. Tiwari, I. M. Al-Helal, "Thermal modelling of photovoltaic thermal (PVT) integrated greenhouse system for biogas heating", Solar Energy, Vol. 136, 2016, pp. 639-649.

[34] Seyedali Mirjalili, Seyed Mohammad Mirjalili, Andrews Lewis, "Grey Wolf Optimizer", Advances in Engineering Software, Vol. 69, 2014, pp. 46-61.

[35] Attia A. El-Fergany, Hany M. Hasanien, "Single and Multi-objective Optimal Power Flow Using Grey Wolf Optimizer and Differential Evolution Algorithms", Electric Power Components and Systems, Vol. 43, 2015, pp. 1548-1559.

[36] Theodoros D. Vrionis, Xanthi I. Koutiva, Nicholas A. Vovos, "A Genetic Algorithm-Based Low Voltage Ride-Through Control Strategy for Grid Connected Doubly Fed Induction Wind Generators", IEEE Transactions on Power Systems, Vol. 29, 2014, pp. 1325-1334.

[37] John Mc Call, "Genetic algorithms for modelling and optimization", Journal of Computational and Applied Mathematics, Vol. 184, 2005, pp. 205-222.

[38] Daiki Min, Jong-hyun Ryu, Dong Gu Choi, "A long-term capacity expansion planning model for an electric power system integrating large-size renewable energy technologies", Computers and Operations Research, Vol. 96, 2018, pp. 244-255.

[39] C. Roldan, A. A. Sanchez de la Nieta, R. Garcia-Bertrand, R. Minguez, "Robust dynamic transmission and renewable generation expansion planning: Walking towards sustainable systems", International Journal of Electrical Power and Energy Systems, Vol. 96, 2018, pp. 52-63.

[40] Angela Flores-Quiroz, Rodrigo Palma-Behnke, Golbon Zakeri, Rodrigo Moreno, "A column generation approach for solving generation expansion planning problems with high renewable energy penetration", Electric Power Systems Research, Vol. 136, 2016, pp. 232-241.

[41] Junhao Zhang, Pingi Xia, "An improved PSO algorithm for parameter identification of nonlinear dynamic hysteretic models", Journal of Sound and Vibration, Vol. 389, 2017, pp. 153-167.

[42] Kennedy, R. Eberhart, "Particle swarm optimization", Proceedings of the IEEE International Conference on Neural Networks, Perth, WA, Australia, 27 November-1 December 1995, pp. 1942-1948.

[43] Y. X. Jin, H. Z. Cheng, J. Y. Yan, L. Zhang, "New discrete method for particle swarm optimization and its application in transmission network expansion planning", Electric Power Systems Research, Vol. 77, 2007, pp. 227-233.

[44] Ping Ren, Nan Li, "Optimal expansion planning of high-voltage transmission network using the composite particle swarm optimization", Proceedings of the International Conference on Artificial Intelligence and Computational Intelligence, Sanya, China, 23-24 October 2010, pp. 170-173.

[45] Saeid Jalilzadeh, Ali Kimiyaghalam, Amir Bagheri, Ahmad Ashouri, "Application of IDPSO Approach for TNEP Problem Considering the Loss and Uncertainty in load growth", Proceedings of the International Congress on Ultra Modern Telecommunications and Control Systems and Workshops, Moscow, Russia, 18-20 October 2010, pp. 232-240.

[46] Baris Kocer, "Bollinger bands approach on boosting ABC algorithm and its variants", Applied Soft Computing, Vol. 49, 2016, pp. 292-312.

[47] Kanendra Naidu, Hazlie Mokhlis, Ab Halim Abu Bakar, Vladimir Terzijia, "Performance investigation of ABC algorithm in multi-area power system with multiple interconnected generators", Applied Soft Computing, Vol. 57, 2017, pp. 436-451.

[48] Hui Wang, Wenjun Wang, Xinyu Zhou, Hui Sun, Zhihua Cui, "Firefly algorithm with neighborhood attraction", Information Sciences, Vol. 382, 2017, pp. 374-387.

[49] Abdollah Rastgou, Jamal Moshtagh, "Application of Firefly Algorithm for multi-stage Transmission Expansion Planning with Adequacy-security considerations in Deregulated Environments", Applied Soft Computing, Vol. 41, 2016, pp. 373-389.

[50] J. Qui, J. Zhao, Z.Y. Dong, "Probabilistic transmission expansion planning for increasing wind power penetration", IET Renewable Power Generation, Vol. 11, 2017, pp. 837-845.

[51] S. Majumder, R. M. Shereef, S. A. Khaparde, "Two-stage algorithm for efficient transmission expansion planning with renewable energy resources", IET Renewable Power Generation, Vol. 11, 2017, pp. 320-329.

[52] S. Dehghan, N. Amjady, "Robust Transmission and Energy Storage Expansion Planning in Wind Farm-Integrated Power Systems Considering Transmission Switching", IEEE Transactions on Sustainable Energy, Vol. 7, 2016, 1981, pp. 765-774.

[53] Roy Z. Wu, P. Zeng, X. P. Zhang, "Two-stage stochastic dual dynamic programming for transmission expansion planning with significant renewable generation and N-k criterion", CSEE Journal of Power and Energy Systems, Vol. 2, 2016, pp. 3-10.

[54] I. Fuchs, S. Voller, T. Gjengedal, "Improved method for integrating renewable energy sources in to the power system of Northern Europe: Transmission expansion planning for wind power integration", Proceedings of 10th International Conference on Environment and Electrical Engineering, Rome, Italy, 8-11 May 2011, pp. 1-4.

[55] Jin Lin, Lin Cheng, Yao Chang, Kai Zhang, Guangyi Liu, "Reliability based power systems planning and operation with wind power integration: A review to models, algorithms and applications", Renewable and Sustainable Energy Reviews, Vol. 31, 2014, pp. 921-934.

Low Power Embedded System Sensor Selection for Environmental Condition Monitoring in Supply Chain

Original Scientific Paper

Josip Zidar
J. J. Strossmayer University of Osijek,
Faculty of Electrical Engineering, Computer Science and Information Technology Osijek
Cara Hadrijana 10b, Osijek, Croatia
josip.zidar@ferit.hr

Ivan Aleksi
J. J. Strossmayer University of Osijek,
Faculty of Electrical Engineering, Computer Science and Information Technology Osijek
Cara Hadrijana 10b, Osijek, Croatia
ivan.aleksi@ferit.hr

Tomislav Matić
J. J. Strossmayer University of Osijek,
Faculty of Electrical Engineering, Computer Science and Information Technology Osijek
Cara Hadrijana 10b, Osijek, Croatia
tomislav.matic1@ferit.hr

Filip Sušac
J. J. Strossmayer University of Osijek,
Faculty of Electrical Engineering, Computer Science and Information Technology Osijek
Cara Hadrijana 10b, Osijek, Croatia
filip.susac@ferit.hr

Abstract *– In the modern world different products and goods are available throughout the world thanks to the complex supply chain system. Often products are transported on long journeys with different transportation systems where products can be damaged or spoiled. Smart Sticker is a concept for product environmental condition monitoring that can resolve some problems in the supply chain. Smart Sticker will record product environmental data in the supply chain and enable producer/consumer product monitoring. Because of ultra-low power design, Smart Sticker component selection must satisfy ultra-low power specifications, besides standard accuracy, and real-time implementation. In this paper we give an overview of the necessary measured environmental parameters and the selection of sensors with emphasis on low power design. We provided a model for the calculation of the maximum operating time, which is applied for the two Smart Sticker instances with significantly different energy consumptions. In the worst-case scenario operation time is 198 days which can be increased with a higher capacity battery.*

Keywords*: low power, embedded system, sensor, supply chain, environmental condition monitoring*

1. INTRODUCTION

Due to globalization, the supply chain has become a very complex process. It consists of several stages depending on the goods produced (frozen foods, fresh foods, electronic devices, etc.) and on the consumer that receives the goods. Transportation, warehousing, inventory, packaging, information, and control are the main elements of the supply chain, cf. Fig. 1. Depending on the type of goods and on the journey to the customer, different elements (e.g. goods can be transported multiple times via truck) can repeat multiple times [1], [2].

In the supply chain goods can get damaged because of different hazards. Most common hazards include spillage, shocks, vibrations, temperature change, accidents, handling, etc. Some of the hazards can be prevented with proper packaging, but this is not always true. Moreover, the type of goods influences the type of hazards. For vegetable commodities and products, spillage during transportation and degradation during handling are most common. Fish products are more prone to degradation during icing in the process of packaging, storage, and transportation, while for milk products spillage hazard is most common [3], [4].

In most of the stages of the supply chain information about the environmental conditions of the goods (temperature, humidity, vibration, etc.) is unknown, due to the complexity of the supply chain. This information can help in the reduction of the hazard probability, and therefore the reduction of the number of damaged goods. Additionally, it would be possible to detect the supply chain stage where the damage has occurred. If, for example, frozen food is transported from a producer to a consumer. On its journey, frozen food is stored

in a chiller which sometimes fails to keep the required temperature. Because of this, some of the products become spoiled. In this example information about the environmental conditions would help to pinpoint the reason and possibly the location of the problem.

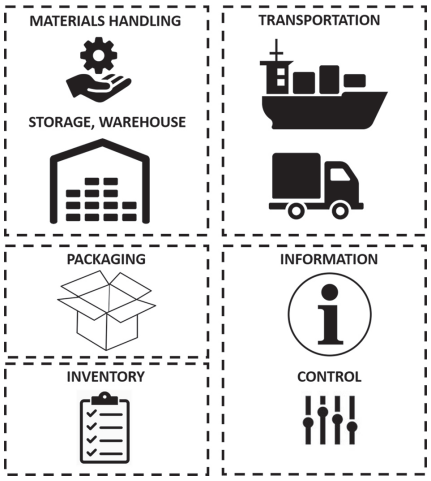


Fig. 1. Main components of the supply chain.

In [1] authors present a Smart Sticker concept that solves the problem of environmental conditions monitoring of goods in the supply chain. It is an ultra-low-power embedded device equipped with sensors for environmental conditions monitoring and RFID (Radio Frequency Identification) technology for data transfer. Device logs sensors data, date, and time which enable the consumer to track the environmental conditions of the received goods.

In this paper, parameters for the monitoring of the environmental conditions of goods in the supply chain are analyzed. The analysis is done based on the Smart Sticker concept and the suggestion from the local Croatian companies. Sensors are suggested to meet the required functionalities, with minimal power consumption. Minimal power consumption is necessary to enable long operation on battery during environmental conditions monitoring of the goods. The long battery life and the small size of the Smart Sticker ensure that it could be placed on a product or a pallet of products for monitoring. For such a device to be able to perform efficiently, it is necessary to analyze the limitations such as energy consumption, dimensions, and price. To measure proposed environmental conditions, state-of-the-art sensors are proposed, which use the latest manufacturing technologies and have the lowest energy consumption. Selected sensors are implemented on a Smart Sticker prototype shown in Fig. 2, and all presented measurements in this paper are done on the developed prototype. Additionally, power consumption is analyzed, and the total battery life of the Smart Sticker prototype is given based on two working scenarios.

The rest of this paper is organized as follows. Section 2 gives the literature overview of the paper topic. Detailed analysis of the monitoring parameters in the supply chain is given in Section 3. Proposed sensors setup and power analysis are outlined in Section 4. Finally, Section 5 concludes the paper by stating the drawn conclusions and providing remarks for future work.

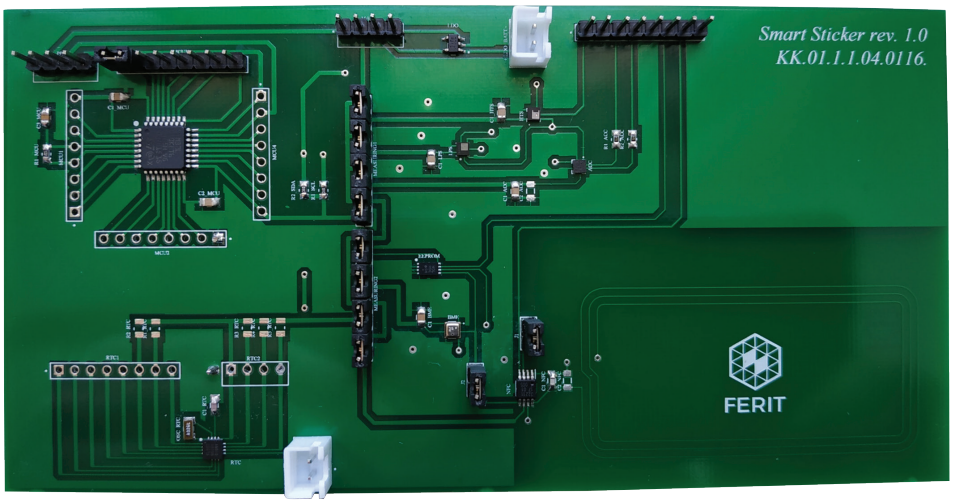


Fig. 2. The Smart Sticker prototype with the microcontroller (top-left), sensors (top-middle), EEPROM (center), NFC (bottom-right), and RTC (bottom-left).

2. RELATED WORK

There have been several attempts in resolving the problem of environmental conditions monitoring in the supply chain. Smart labels are one of the latest trends in emerging Industry 4.0, numerous prototypes of smart labels are presented and developed in the last decade. Nowadays, smart labels go beyond identification purposes, nor sophisticated context-aware labels with embedded modules are currently state of the art [5]. Authors in [6] give a specification for an intelligent product as a part of the Auto-ID project. In the presented concept, goods can store data about themselves, and the ability to communicate with their environment. In the paper, the authors don't give any recommendation about the implementation of the intelligent product or the possible lifespan.

In [7] authors present a system for real-time traceability and cold chain monitoring based on RFID technology. The developed tag is battery-powered and can measure light, temperature, and humidity. The authors used a CPLD device for low-power communication between sensors and a microcontroller. Device dimensions are relatively large and not convenient for easy implementation. Therefore, to achieve the IP65 protection standard, a large box is used. The system is tested on frozen fish products and it showed accurate measurements for temperature. The authors didn't give the maximum battery lifespan of the device and didn't analyze the power consumption under different working conditions or measurement scenarios. In [8] authors discuss traceability systems that will help in increasing the safety and the quality of food products. In the paper, several definitions of traceability are discussed, concentrated on the history, location, and data records of the food products. According to EU regulation (178/2002), traceability is the ability to trace and follow all stages of the supply chain. According to presented research, most papers and systems are focused on traceability until the retail point of the food chain, thereby the consumer part of the food chain is still not covered enough. All trends suggest that food traceability from "farm to fork" is going to become standard soon. In [9] authors present a start-up pilot project for food traceability based on passive RFID tags. Although the presented system detailed capture all stages of the supply chain no environmental conditions are considered at all.

In [10] authors present two typologies of the non-destructive electronic RFID-based tracking system with application in the cheese industry. During the cheese maturation process, only 4 measurements are acquired for 9 months period. Quality information, chemical, and spectrophotometric analysis are conducted with external systems and linked with the RFID identification tag placed on cheese and stored in a web-based application. From the customer's perspective, this system can be used only to validate the maturation process of cheese products. Information about storing, transporting, and other environmental conditions are not considered in this paper. In [11] RFID-based monitoring system for tracking fruit quality is presented. Flexible Tag Microlab (FTM) is communicating with fixed RFID reader AC powered equipped with gas sensors. Although this system reduces the costs of the tags themselves, all information acquired during monitoring is saved on an internal database and cannot be retrieved by the user. Active RIFD tags commonly lack long-life battery performance. To overcome this problem authors in [12] proposed a design for a self-powered RFID tag based on a piezoelectric power supply (PPS). According to their research using PPS integrated inside RFID temperature monitoring system, 0.283 mW of energy per 1 second can be generated during mobile transport of perishable items. For better monitoring and control of environmental conditions, authors in [13] present "Intelligent Container". The system where all transport conditions are measured inside the cargo box, which can help to control active devices such as air conditioners to maintain desired environmental conditions. This approach is useful to decrease food loss in the cold chain but no information about transport itself is available to the customer. With the emergence of printing technologies, new concepts are presented.

In [14] authors present an all-printed smart label with an integrated humidity sensor and power supply. Although the presented prototype can acquire environmental humidity no communication and data storage are integrated. In [15] authors present a different approach, a sensitive time-temperature indicator. A self-healing nanofiber mat is devised that serves as a temperature sensor and a display. The proposed mat changes light transmittances with the temperature where the speed of the change can be regulated based on mat polymer composition and film thickness. The presented device cannot output or save the temperature data and therefore there is no traceability option. Similar devices are also presented in [16], [17]. In [18] authors present a smart RFID label with a printed multisensor platform. Using inkjet-printed technology multisensor platform for humidity, temperature, and ammonia is created. Evaluation of the presented platform confirms the possibility for commercial use of inkjet-printed sensors. Although sensors can achieve commercial standards, presented smart RFID labels with the one-hour interval between the measurements on 150 mAh has an expected lifetime of only 57 days, which can't satisfy the requirement for mass scale usage. Another similar smart label based on flexible PCB (Printed Circuit Board) technology and self-created sensors is presented in [19]. As the presented prototype doesn't include internal storage, the expected battery life is not considered. Although information can be acquired using RFID technology, no information about the history of the environmental conditions is saved at all. Another smart label is presented in [20]. Using printed technology smart sticker for time-temperature history is presented, with a new R2R printed battery proposed label's lifetime is up to 7 days, with one minute between each data logging. As the only temperature is considered, this approach is useful only for highly temperature-sensitive products such as fresh food, dairies, etc. Another RFID-based temperature monitoring system is presented in [21]. The presented system is designed in form of the biocompatible plaster using 130 nm CMOS technology, powered by a rechargeable lithium-ion battery. The system is used for constant body temperature monitoring in a hospital environment but the system itself can be applied to various temperature-sensitive processes. In [22] authors present a fully integrated passive UHF RFID tag for temperature environment monitoring. Using the energy of the incident RF signal from the base station, energy is converted to dc supply voltage and stored in capacitors, so the battery is not included in the design. In the active state, the total current dissipation of the proposed tag is 15.4 μA. The proposed design of the temperature sen-

sor has a measuring error of ±2 °C, which is not sufficient for temperature-sensitive products. Although using RF signal to power tag is an effective approach, in lack of RF signal and no battery all monitoring data will be lost.

Considering all the afore mentioned, in this paper we consider commercially available sensors for low power embedded systems. Based on a created questionnaire, suppliers gave different parameters for environmental condition monitoring in the supply chain. Based on the acquired data, several sensors are presented that fulfill the set characteristics and can be implemented in Smart Sticker design. Final sensor selection is carried out based on the power analysis in different environmental conditions monitoring scenarios. Several Smart Sticker design instances are suggested, and the estimate of maximum time of system operation is given for different measurement sampling times.

3. ENVIRONMENTAL CONDITION PARAMETERS

Firstly, it is necessary to analyze the existing market and monitoring needs and propose a specified solution with selected sensors for environmental condition monitoring. The percentage of sold products in the EU by category is illustrated In Fig. 3. Other products include wood, paper, furniture, computers, electronic and optical products, textiles, wearing apparel, leather, pharmaceutical products, mining. The intended purpose of the Smart Sticker can be found in transport and storage in the food, pharmaceutical, chemical, electrical, and electronic engineering industries.

Damage to goods traveling from manufacturers to stores is a possible occurrence [3]. In the case of damaged electronics, defects can be detected even after the customer buys the product. This includes sensitive electronic devices, laptops, TVs, refrigerators, washing machines, etc. To reduce inconvenience to the buyer, seller, and the manufacturer, an environmental condition monitoring system is required. In this way, it would be possible to detect the supply chain stage where the damage has occurred. Other examples are in the frozen foods industry. After the freezing process, products are stored at -18 °C. Smart Sticker can monitor the storage temperature and it would be possible to detect when the product temperature changed above set threshold value. If the product is spoiled when the consumer buys it, the temperature history of the product can be traced and the supply stage located where the temperature hazard has occurred.

To analyze the need for a Smart Sticker system in practice, a questionnaire was sent to selected local companies. Selected companies have delivery and/or production of goods as the main part of their business. In this way, the company can ensure quality in the production and/or delivery of goods or analyze deficiencies in the supply chain. The questionnaire sent included inquiries about general company information, a general description of the business model, a list of goods to be monitored in the supply chain, mode of transport, method of packaging, storage, and transport conditions, and required monitoring intervals.

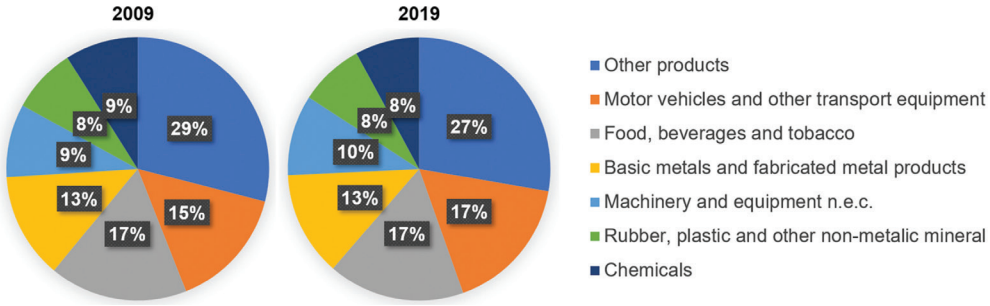


Fig. 3. Percentage of sold products for the years 2009 and 2019 in the EU

Table 1. Summarized questionnaire results from local companies

Product type	Bakery	Meat	Meat	Fruit	Electrical
Product name	Meat pie	Dry sausage	Burger	Apple	Cabinet
Pieces per package	55	7	8	78	1
Package size w/h/d [mm]	291/256/212	1200/800/725	550/378/145	600/400/180	2000/1000/1000
Packages per pallet	84	30	40	48	1
Storage temperature {min,typ,max} [°C]	-25, -18, -17	4, 6, 8	4, 6, 8	0, 1, 4	-10, 25, 50
Storage humidity {min, typ, max} [%]	0, 50, 95	0, 50, 95	0, 50, 95	70, 77, 85	0, 50, 95
Sensitivity to vibrations	No	No	No	Yes	No
Sampling period [hours]	1	1	8	1	1
Storage duration [days]	10	1	1	15	7
Transhipments during transport	3	2	2	0	0
Transport duration [days]	5	1	1	3	1

Table 2. Low-power sensors which are considered in the design of an instance of the Smart Sticker

Sensor ID	Measures
Sensor 1	Temperature and humidity
Sensor 2	Pressure
Sensor 3	Temperature, humidity and pressure
Sensor 4	Impact detection 5 g
Sensor 5	Impact detection 20 g
Sensor 6	Impact detection 60 g

Table 1 shows the main results of the questionnaire answered by local companies. To obtain more accurate data for Smart Sticker usage in practice, the results of questionnaires from companies are analyzed. Companies have expressed the need to monitor the temperature when transporting frozen foods that must be stored within a certain temperature range. Vibration sensitivity has also been a requirement, for companies transporting shock and vibration sensitive goods. According to the stated needs, the list of functional parts has been compiled and shown in Table 2.

The exchange of data between the microcontroller and the peripheral units (sensors, NFC, RTC) is necessary, hence it is important to consider an appropriate low power communication protocol. Most of the peripheral devices support I2C (Inter-Integrated Circuit) and SPI (Serial Peripheral Interface) communication. SPI communication is more beneficial for applications that need a data stream, because of the full-duplex mode. Both I2C and SPI offer excellent integration when implementing communication between the microcontroller unit and peripherals with low data throughput requirements. I2C is more easily implemented for multiple device communication and has a smaller footprint on the PCB (only two wires are necessary) [23], [24]. In this paper, we use peripherals that support I2C communication.

4. SMART STICKER MODEL FOR THE BATTERY LIFE ESTIMATION

The Smart Sticker device that is considered in this work has three modes: Sleep, Measure, and Transfer Mode, cf. Fig 4. It monitors environmental conditions over a long time, for a period of several months or even years, depending on the type of goods that are monitored. Using a battery as a power source the Smart Sticker significantly relies on a concept of low energy consumption. Subsequently, with reduced energy consumption, the space occupied by the battery also reduces. In this way, the Smart Sticker device is feasible in small dimensions and applies to products with small packaging in the product supply chain. Based on the measuring values, Smart Sticker comes in two proposed variants, Type A and Type B. Type A measures environmental conditions, while Type B is used for impact and shock detection. To reduce the energy consumption and increase the maximum time of operation, the sensors are turned on periodically with a relatively long delay between the two measurement samples. The sampling time t_s depends on the application and the dynamics of measured physical property. For example, a temperature sensor can have a much larger t_s (longer delay between samples), while a vibration sensor requires a short t_s due to the higher frequency of the measured signal. In this work, the following sampling time t_s values are considered:

- once a minute,
- once every 10 minutes,
- once every 30 minutes,
- once an hour (60 minutes),
- once every 12 hours (720 minutes),
- once a day (1440 minutes),
- once a week (10080 minutes).

The Smart Sticker has three modes of operation: Measure Mode (MM), Sleep Mode (SM), and Transfer Mode (TM), cf. Fig 4. Each mode of operation has a specific energy consumption regime. During the MM, the microcontroller is communicating with the sensors which are measuring physical properties and are sending data via an I2C communication protocol. The data is then stored in the EEPROM. During the SM, minimal energy is used by the Smart Sticker, only the RTC is consuming power from the battery. Finally, during the TM, the NFC device is providing the power for the MCU to read the data from the EEPROM.

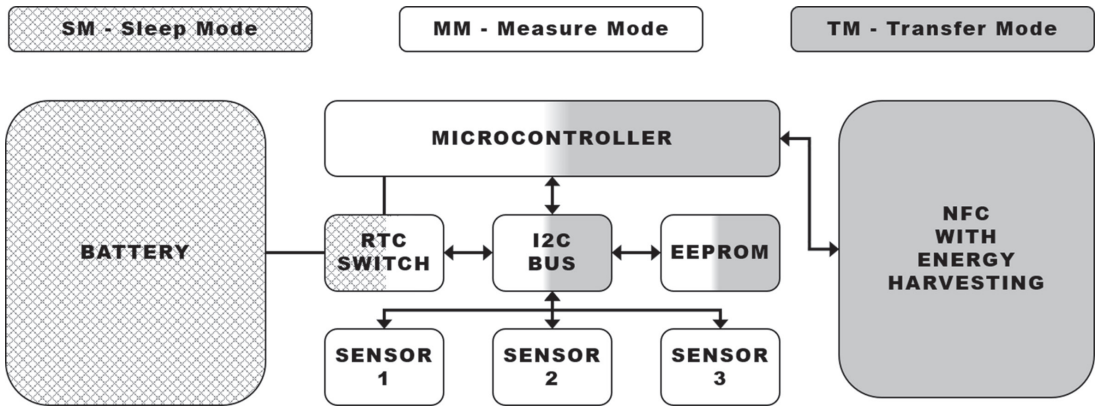


Fig. 4. The Smart Sticker design with three working modes.

Integrated circuits used in the Smart Sticker design are selected from the set of devices with the smallest possible energy consumption on the market. The energy consumption of each component is measured, and it is illustrated in Table 3. Each measured value is obtained from an average of 100 measurements. For the measurements illustrated in Table 3, the Smart Sticker is powered with a 3 V CR2016 battery with E_{BAT}=90 mAh. In this instance, the Smart Sticker has two temperature, humidity, and pressure sensors for redundancy reasons. Redundant sensors are only active in MM, thus having a very low impact on battery life. Additionally, in the case of redundant measurement of the same parameter, the memory requirement remains unchanged. After comparing measured redundant values with each other, only one parameter value is stored based on a voting procedure (e.g. two out of three voting).

Table 3. Current consumption of all components considered for an instance of the Smart Sticker.

	I_MM [µA]	I_SM [µA]	I_TM [µA]
Sensor 1	36.332	0.000	0
Sensor 2	25.280	0.000	0
Sensor 3	38.921	0.000	0
Sensor 4	6.219	6.219	0
Sensor 5	6.219	6.219	0
Sensor 6	6.219	6.219	0
RTC	0.255	0.022	0
I2C	9.181	0.000	0
EEPROM	15.802	0.000	0
MCU	124.605	0.275	0

Further on, a calculation is done to estimate the maximum time of operation t_{OP} of a Smart Sticker design instance. The calculation is done under the ideal room temperature conditions with 25°C. The t_{OP} is estimated by the measurements of the energy consumption of its components, the consumption of the battery self-discharge, and the time when the components are consuming the energy at a certain operation mode MM, SM, and TM, respectively.

The task of the Smart Sticker is to measure environmental conditions and utilize sleep function. With the measured energy consumption results in each mode of operation, the operating time can be calculated using

$$t_{OP} = \frac{E_{BAT}}{I_{OP}}, \tag{1}$$

$$I_{OP} = \frac{I_{MM} \cdot t_{MM} + I_{SM} \cdot t_{SM}}{t_S} + t_S \cdot I_{SD}\,, \tag{2}$$

$$t_S = t_{MM} + t_{SM}, \tag{3}$$

$$I_{SD} = p_{SD} \cdot E_{BAT}\,. \tag{4}$$

In equation (1), E_{BAT} is the battery stored energy [Ah], and I_{OP} is the total current [µA] consumed during the one sampling period t_S. During t_S, one MM and SM cycle is completed. The third operation mode, TM, is not considered in the equation since it is an asynchronous event, while SM and MM are repeating in cycles with period t_S. Additionally, in the TM mode, the Smart Sticker is powered from the RF field of the NFC chip and the onboard antenna. In Equation (2), sleep time t_{SM} and measurement time t_{MM} correspond to the operating modes SM and MM, respectively. The sampling time t_S is the sum of sleep and measure times, cf. eq. (3). Current consumptions in MM and SM are I_{MM} and I_{SM}, respectively, cf. Table 3. The self-discharge current I_{SD} is calculated from the percentage p_{SD} of discharge over the unit of time, which is usually expressed in percentage per month or percentage per year. Since E_{BAT} is usually in [mAh], with hours as the unit of time, the p_{SD} must be converted to the appropriate percentage per unit of time, for CR (Coin Manganese Dioxide Lithium Battery) battery p_{SD} is 1% per year [25].

Table 4. Constant values used for calculations of estimated operation time.

	Value	Description
E_{BAT}	90 mAh	Battery capacity in mA hours
E_{BAT}	5400000 µAmin	Battery capacity in µA minutes
E_{BAT}	10.273 µAyear	Battery capacity in µA years
p_{SD}	1 % / year	Battery self-discharge rate
t_{BATSL}	10 years	Battery shelf life
I_{SD}	0.103 µA / year	Self-discharge current in micro Amps per year
I_{SD}	1.95E-07 µA / min	Self-discharge current in micro Amps per minute
t_{MM}	0.017 min	Time required to perform one measurement cycle

Table 5. Estimated operation time of the Smart Sticker Type A with varying t_S sampling period.

t_S [min]	t_{MM} [min]	t_{SM} [min]	I_{SD} [µA]	I_{OP} [µA]	t_{OP} [year]
1	0.017	0.983	1.95E-07	4.195	2.45
10	0.017	9.983	1.95E-06	0.439	10 (23.39)
30	0.017	29.983	5.86E-06	0.161	10 (63.78)
60	0.017	59.983	1.17E-05	0.092	10 (112.23)
720	0.017	719.983	1.41E-04	0.028	10 (369.63)
1440	0.017	1439.983	2.81E-04	0.025	10 (412.65)
10080	0.017	10079.983	1.97E-03	0.022	10 (458.37)

Calculations using eq. (1) to (4) results with values illustrated in Table 5. Calculations predict the maximum operating time t_OP of the Smart Sticker device depending on the stored energy in the battery and on the energy consumption of used components. As can be seen, the maximum operating time t_OP is 10 years, due to battery shelf life t_BATSL. For a battery provided for a Smart Sticker, after 10 years it is assumed that the battery is no longer functional.

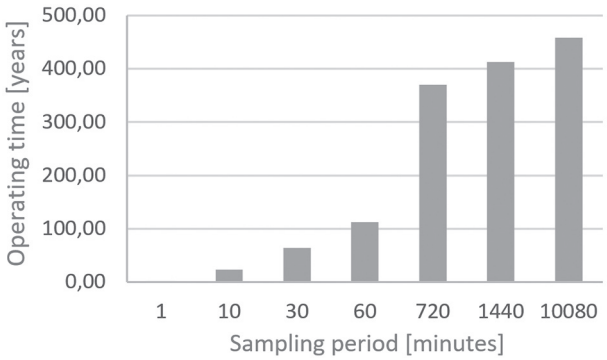


Fig. 5. Estimated operating time in dependence of the sampling period.

The estimated operation time of Smart Sticker Type A is presented in Fig. 5 and Table 5, where Real-Time Clock (RTC) achieves significant energy savings using its alarm function to wake up the microcontroller periodically. RTC is the only working component in SM for Smart Sticker Type A.

As can be seen in Table 5, with more frequent measurements, i.e. with smaller sampling time, energy consumption increases. Increase in the energy consumption results with a shorter operating time, as it is shown in Fig. 5. Sampling period t_S is a parameter that is adjustable by the user of the Smart Sticker, therefore it can vary depending on the application. Frequent measurements are usually preferred to achieve data continuity.

Furthermore, for detecting shocks and impacts during goods transportation and handling, Smart Sticker Type B is presented. Such detection must continuously check for the occurrence of a shock or impact, therefore consuming more energy in comparison to Smart Sticker Type A.

Table 6. Estimated operation time of the Smart Sticker instance types with t_S=60 minutes.

Smart Sticker	E_{BAT} [µAyear]	I_{OP} [µA]	t_{OP} [year]
Type A	10.273	0.092	10.00000
Type B		18.932	0.54268

Impact detection is an event-based real-time system, where it is impossible to predict the number of times or when the shock or impact will occur in practice. When the Smart Sticker Type B registers a shock or an impact with its impact detection sensor, it wakes up from the SS to the MM to store the relevant data into the EEPROM. Subsequently, it goes back to the SM. The sensor for impact detections is continuously working in SM, while the MCU is in standby mode. In TM, impact detection data is transferred to the NFC device which provides the power for the microcontroller to read the stored data from the EEPROM, identical to TM in Smart Sticker Type A.

For estimating the operating time of Smart Sticker Type B, only SM calculations will be performed, because it is impossible to predict an exact number of times when impact or shock will occur in practice and the device will enter MM. The estimated operation time for both Smart Sticker Type A and Type B is illustrated in Table 6. If we assume 1000 impacts per year t_{OP} of Smart Sticker Type B will be 0.54254 years or 198 days which could be increased using a larger capacity battery. Depending on the need for impact or shock detection, specific threshold detection for certain products can be set. More sensitive (fragile) goods can have a lower threshold value, while less sensitive goods can have a higher threshold value. Damage of goods can vary for different products, which use different types of packaging materials and protection during transport.

5. CONCLUSION

The supply chain in the global world has become a very complex process. Goods on their journey from the producer to the customer go thru different supply chain stages, and can get damaged because of different hazards. Information about the environmental conditions of the goods in all stages of the supply chain can help in the reduction of the hazard probability and the detection of hazards.

In this paper we analyze the parameters for the monitoring of the environmental conditions of goods. Local suppliers answered a questionnaire about the environmental conditions of the produced goods. Based on the acquired answers several commercially available sensors are selected with low power operation. All sensors are implemented and tested on the Smart Sticker device.

Smart Sticker model for the battery life estimation is developed based on several parameters. Two Smart Sticker types are developed. Type A for environmental conditions monitoring and Type B for impact detection. Based on real-time measurements of the currents and different sampling times, operation time is estimated. In the worst-case scenario (sampling time 1 min) for Type A, the estimated operation time is 2.45 years, and for Type B (1000 impacts per year) is 198 days.

Future work will include testing the prototypes in an industrial temperature and humidity chamber and on a vibration desk. Additionally, real-time measurements of the currents in different environmental conditions will be carried out. The aforementioned tests will help improve robustness for industrial applications. Finally, the developed prototype will be tested in real supply chain conditions where it can provide food safety and quality assurance.

6. ACKNOWLEDGEMENT

This work was funded by the European Union through the European Regional Development Fund, under project "Smart Sticker for measuring and monitoring storage and transportation conditions of products" KK.01.1.1.04.0116.

7. REFERENCES

[1] F. Sušac, T. Matić, I. Vidović, I. Aleksi, and Ž. Hocenski, "Smart Sticker: Concept for Better Storage and Transportation", Proceedings of the 2020 International Conference on Smart Systems and Technologies, Osijek, Croatia, 14-16 October 2020, pp. 121–126.

[2] A. Rushton, P. Croucher, and P. Baker, "The handbook of logistics & distribution management", 5th Ed., Kogan Page, 2014.

[3] S. A. Alsobhi, K. K. Krishnan, D. Gupta, and A. T. Almaktoom, "Analysis of damage costs in supply chain systems", International Journal of Industrial and Systems Engineering, Vol. 28, No. 1, 2018, p. 70.

[4] J. Gustavsson, C. Cederberg, U. Sonesson, "Global food losses and food waste: extent, causes and prevention", Interpack 2011, Düsseldorf, Germany, 16-17 May 2011.

[5] T. M. Fernandez-Carames, P. Fraga-Lamas, "A Review on Human-Centered IoT-Connected Smart Labels for the Industry 4.0", IEEE Access, Vol. 6, 2018, pp. 25939–25957.

[6] C. Y. Wong, D. McFarlane, A. Ahmad Zaharudin, V. Agarwal, "The intelligent product driven supply chain", Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, Yasmine Hammamet, Tunisia, 6-9 October 2002, pp. 1-6.

[7] E. Abad et al., "RFID smart tag for traceability and cold chain monitoring of foods: Demonstration in an intercontinental fresh fish logistic chain", Journal of Food Engineering, Vol. 93, No. 4, 2009, pp. 394–399, 2009.

[8] M. M. Aung, Y. S. Chang, "Traceability in a food supply chain: Safety and quality perspectives", Food Control, Vol. 39, 2014, pp. 172–184.

[9] I.-H. Hong et al., "An RFID application in the food supply chain: A case study of convenience stores in Taiwan", Journal of Food Engineering., Vol. 106, No. 2, 2011 pp. 119–126.

[10] P. Papetti, C. Costa, F. Antonucci, S. Figorilli, S. Solaini, P. Menesatti, "A RFID web-based infotracing system for the artisanal Italian cheese quality traceability", Food Control, Vol. 27, No. 1, 2012, pp. 234–241.

[11] A. Vergara et al., "An RFID reader with onboard sensing capability for monitoring fruit quality", Senssors and Actuators B: Chemical, Vol. 127, No. 1, 2007, pp. 143–149.

[12] H. Chu, G. Wu, J. Chen, F. Fei, J. D. Mai, W. J. Li, "Design and simulation of self-powered radio frequency identification (RFID) tags for mobile temperature monitoring", Science China Technological Sciences, Vol. 56, No. 1, 2013, pp. 1–7.

[13] R. Jedermann, T. Poetsch, W. Lang, "Smart Sensors for the Intelligent Container", Proceedings of the Smart SysTech 2014: European Conference on Smart Objects, Systems and Technologies, Dortmund, Germany, 1-2 July 2014, pp. 1–2.

[14] N. Pereira, V. Correia, N. Peřinka, C. M. Costa, and S. Lanceros-Méndez, "All-Printed Smart Label with Integrated Humidity Sensors and Power Supply", Advanced Engineering Materials, Vol. 23, No. 3, 2021.

[15] S. Choi et al., "A Self-Healing Nanofiber-Based Self-Responsive Time-Temperature Indicator for Securing a Cold-Supply Chain", Advanced Materials, Vol. 32, No. 11, 2020.

[16] A. Pavelková, "Time temperature indicators as devices intelligent packaging", Acta Universitatis Agriculturae et Silviculture Mendelianae Brunensis, Vol. 61, No. 1, 2013, pp. 245–251.

[17] H. Vaikousi, C. G. Biliaderis, K. P. Koutsoumanis, "Applicability of a microbial Time Temperature Indicator (TTI) for monitoring spoilage of modified atmosphere packed minced meat", International Journal of Food Microbiology, Vol. 133, No. 3, 2009, pp. 272–278.

[18] A. V. Quintero et al., "Smart RFID label with a printed multisensor platform for environmental monitoring", Flexible and Printed Electronics, Vol. 1, No. 2, 2016.

[19] E. Smits et al., "Development of printed RFID sensor tags for smart food packaging", Proceedings of the IMCS, Nuremberg, Germany, 20-23 May 2012, pp. 403–406.

[20] B. B. Maskey et al., "A Smart Food Label Utilizing Roll-to-Roll Gravure Printed NFC Antenna and Thermistor to Replace Existing "Use-By" Date System", IEEE Sensors Journal, Vol. 20, No. 4, 2020, pp. 2106–2116.

[21] C. Kollegger, C. Steffan, P. Greiner, M. Wiessflecker, G. Holweg, and B. Deutschmann, "Intelligent plaster for accurate body temperature monitoring and investigations regarding EMI using near-field magnetic scan", Elektrotechnik und Informationstechnik, Vol. 133, No. 1, 2016, pp. 25–31.

[22] Hongwei Shen, Lilan Li, Yumei Zhou, "Fully integrated passive UHF RFID tag with temperature sensor for environment monitoring", Proceedings of the 7th International Conference on ASIC, Guilin, China, 22-25 October 2007, pp. 360–363.

[23] A. Gloria, F. Cercas, N. Souto, "Comparison of communication protocols for low cost Internet of Things devices", Proceedings of the South Eastern European Design Automation, Computer Engineering, Computer Networks and Social Media Conference, Kastoria, Westen Macedonia, 23-25 September 2017, pp. 1–6.

[24] A. K. Oudjida, M. L. Berrandjia, R. Tiar, A. Liacha, K. Tahraoui, "FPGA implementation of I^2C & SPI protocols: A comparative study", Proceedings of the 16th IEEE International Conference on Electronics, Circuits and Systems, Yasmine Hammamet, Tunisia, 13-16 December 2009, pp. 507–510.

[25] muRATA, "Coin Manganese Dioxide Lithium Battery Technical note", https://www.mouser.com/pdfDocs/tcn-cr-001-200722-2.pdf (accessed: 2021)

Design and implementation of shape-based feature extraction engine for vision systems using Zynq SoC

Original Scientific Paper

Navya Mohan
Department of Electronics, Cochin University of Science and Technology,
Kochi, Kerala, India
navyamohan@cusat.ac.in

James Kurian
Department of Electronics, Cochin University of Science and Technology,
Kochi, Kerala, India
james@cusat.ac.in

Abstract *– With the great impact of vision and Artificial Intelligence (AI) technology in the fields of quality control, robotic assembly and robot navigation, the hardware implementation of object detection and classification algorithms on embedded platforms has got ever-increasing attention these days. The real-time performance with optimum resource utilization of the implementation and its reliability as well as the robustness of the underlying algorithm is the overarching challenges in this field. In this work, an approach employing a fast and accurate vision-based shape-detection algorithm has been proposed and its implementation in heterogeneous System on Chip (SoC) is discussed. The proposed system determines centroid distance and its Fourier Transform for the object feature vector extraction and is realized in the Zybo Z7 development board. The ARM processor is responsible for communication with the external systems as well as for writing data to the Block RAM (BRAM), the control signals for efficient execution of the memory operations are designed and implemented using Finite State Machine (FSM) in the Programmable Logic (PL) fabric. Shape feature vector determination has been accelerated using custom modules developed in Verilog, taking full advantage of the possible parallelization and pipeline stages. Meanwhile, industry-standard Advanced Extendable Interface (AXI) buses are adopted for encapsulating standardized IP cores and building high-speed data exchange bridges between units within Zynq-7000. The developed system processes images of size 32 × 64 in real-time and can generate feature descriptors at a clock rate of 62MHz. Moreover, the method yields a shape feature vector that is computationally light, scalable and rotation invariant. The hardware design is validated using MATLAB for comparative studies.*

Keywords*: hardware-software codesign, hardware accelerators, feature extraction, vision systems, embedded systems*

1. INTRODUCTION

Machine Vision is one of the most frontiers and revolutionary technology in computer science as a plethora of industrial activities has been potentially benefitted. It ensures consistent and continuous excellence in automating monotonous chores- visual inspection in manufacturing [1-3], localization and navigation for robotic guidance [4], real-time measuring and sorting in factory floors and production lines [5-7]. Vision systems bring operational benefits by reducing human involvement in a manufacturing process and excels in quantitative analysis of structured scenes because of their speed, accuracy and repeatability.

Traditionally, visual inspection and quality control were accomplished by trained experts though it produces inconsistent results. Vision systems may effectively replace human inspection in such demanding cases, so automation has become inevitable to improve precision and reliability. Machine vision systems reckon on cameras to acquire images so that computer hardware and software can process, analyse and measure the various characteristics for decision making. Typically the initial step in such applications is to localize the object or feature of interest within the 2D images. In the past few years, various research activities have been carried out to propose intelligent systems for real-time

analysis of characteristic image features like colour, texture and regions. A machine vision assisted sorting has been achieved with colour image processing [8]. Classification based on texture analysis has been investigated in [9]. The paper also discusses the advantages and disadvantages of texture image descriptors and covers the discrimination performance, computational complexity and resistance to challenges such as noise, rotation etc. Automated visual-based defect detection approaches applicable to various materials such as metals, ceramics and textiles are discussed in [11]. A survey of textural defect detection based on statistical, structural and other approaches are considered. Several efforts have been made in areas of image processing to implement algorithms in hardware. Employing dedicated hardware structures accelerates these vision modules. The work [12] discusses the description of a simple fast shape detection algorithm and its implementation in hardware structures like FPGA. The detection algorithm is based on the concepts of Hu´s moments that are invariant to similarity transformations. [13- 19] describes the realization of standard edge descriptors that are computationally intensive.

As the demand for computationally intensive algorithms for real-time industrial applications is on a steep rise, a trend has been witnessed towards utilizing Field Programmable Gate Arrays (FPGAs) or Graphics Processing Units (GPUs) for implementing the vision algorithms instead of using CPUs, which are inherently sequential processing devices. Compared with general-purpose CPUs, realizing applications in FPGAs can act as hardware accelerators that can offload the computational burden from the CPU and also help in developing a prototype system before ASIC implementation. Although modern FPGAs have excellent hardware capabilities, high development difficulty is the main disadvantage. Image processing systems realized in FPGAs involve the pipeline data processing approach, where the pixel stream passes through different computing elements. In addition, while handling significant quantities of data, resource constraints often jeopardise the real-time performance of the system. To address these challenges, modern FPGA chips usually embed microprocessor cores to achieve the convenience and flexibility of software.

This paper presents a shape-based feature extraction system and adopts the Zynq SoC platform launched by Xilinx where it integrates the software programmability of an ARM-based processor with the hardware programmability of an FPGA, enabling hardware acceleration on a single device. Through this combination, Zynq has the advantages of having ARM and FPGA. The ARM core eases interfacing the peripherals while FPGA performs parallel processing as well as dynamic reconfiguration. The work discusses the hardware realization of the algorithm on Zynq architecture [20] so that the unit acts as a standalone shape feature extractor. Feature extraction algorithms generally employed include steps like pre-processing, feature detection and feature descriptor building. In the current work, Centroid Distance calculation and Fast Fourier Transform (FFT) are employed for feature building and the output obtained is in the form of feature descriptors. The proposed method effectively determines the shape Fourier descriptors by tackling the problems related to the hardware implementation, such as the requirement of nonlinear operations. These descriptors which uniquely characterises the object based on its shape is later utilized by classification algorithms. Although hardware realization of different features is reported, a computationally efficient algorithm is proposed, which employs custom modules developed using Verilog along with the Xilinx IP cores. However, hardware realization of the algorithm using a reconfigurable platform is a complex task that needs validation of the proposed methods in a simulation. The results of the proposed work are also validated by implementing the algorithm in MATLAB. The main contributions of the paper can be summarised as follows:

1) A vision system on Zynq heterogeneous platform is built, which covers shape-based feature detection, and process images of size 32×64 and generates feature descriptors/vectors at a clock rate of 62MHz
2) Shape feature using FFT offers translation, rotation and scale-invariant attributes and ensures less dedicated hardware resource utilization, thus making the robotic vision module compact.

Robotic systems with an embedded camera can make decisions based on the observations of the inputs around them. The proposed vision-enabled shape feature extraction technique integrated with robotic systems can be widely used in industrial sectors for automation by intelligent sensing of production lines. Hardware realization in FPGA ensures real-time processing of the captured images with sufficient speed and accuracy that satisfy ever-increasing production and quality requirements, consequently aiding in the development of totally automated processes. In addition, such systems can also be trained to work in extreme environments without involving human intervention.

The rest of the paper is organized as follows. Section 2 explains the overall structure of the algorithm used in the proposed shape detection. Section 3 explains the details of the proposed hardware implementation. Section 4 and 5 report the results obtained and conclude by discussing the future scope.

2. SYSTEM OVERVIEW

The proposed vision engine for shape feature determination includes the following: acceleration module for pixel stream processing and analysis, storage modules and associated interfaces. The ZYNQ-based image processing system establishes a connection with the host computer through the UART communication pro-

tocol, which enables data communication between the platform and the host computer. The feature extraction system is the core of the developed system. The working process for the entire system is described as follows: First, the robotic system is allocated the workbench where the industrial camera captures the scene with the aid of lens and associated light source. The camera image is processed and loaded into the Block RAM. Then, object boundaries, as well as the features, are extracted. Finally, these features are used for their classification. The overall architecture of the work is illustrated in Fig. 1.

2.1 ALGORITHMS FOR EFFICIENT COMPUTATION OF FEATURE DESCRIPTORS

Fourier descriptors are derived by applying Fourier Transform on a 1D Shape Signature. A shape signature z(t) is a 1D function representing 2D areas or boundaries [21]. A shape signature usually captures the perceptual feature of the shape. In the following, we assume that the shape boundary coordinates *(x(t), y(t)), t= 0,1,2,....., N-1* has been extracted in the preprocessing stage where t usually represents the arc length.

The position function of an object from its 2D view is derived from the boundary coordinates using the equation

$$z(t) = |x(t) - x_c| + i|y(t) - y_c| \quad (1)$$

where (x_c, y_c) is the centroid of the shape, which is the mean of the boundary coordinates. For a sequence (x_1, y_1), (x_2, y_2), . . . , (x_N, y_N) of uniform contour points in order, where N refers to the sampling points on the contour, the centroid point of these contour points has to be calculated first. The centroid coordinates corresponding to the object in a binary image is the arithmetic mean of all boundary coordinates as described in the equation (2) and (3) where x_i and y_i are the x and y coordinates along the boundary of the object and N is the total number of boundary coordinates.

$$x_c = \frac{1}{N}\sum_{i=1}^{N} x_i \quad (2)$$

$$\text{and} \quad y_c = \frac{1}{N}\sum_{i=1}^{N} y_i \quad (3)$$

$z(t)$ represents the shape boundary which is also translation invariant. Rotation causes $z(t)$ circular shift and scaling of shape only introduces linear changes in $z(t)$. Employing position function (complex coordinates) as shape signature involves little computation and is calculated using Centroid Contour Distance (CCD).

The CCD function is the distance of the boundary points to the centroid (x_c, y_c) of the shape and $r(t)$ is translation invariant.

$$r(t) = ((x_t - x_c)^2 + (y_t - y_c)^2)^{1/2} \quad (4)$$

Rotation introduces circular shift while the scaling of shape only changes r(t) linearly.

The CCD feature is a distance sequence that describes the contour feature of the object by using the vector composed of the distance from the shape centroid to the contour point. The coordinates of the traversing point along the contour can be represented as a function whose period is defined by the perimeter of the shape boundary. This period function is represented by the Fourier series expansion.

This 1D signature which is derived from 2D shape boundary coordinates is subjected to discrete Fourier Transform of M points (centroid-distance points) and is given by the equation

$$F(\mathrm{u}) = \frac{1}{M}\sum_{\mathrm{x}=0}^{\mathrm{M}-1} \mathrm{f}(\mathrm{x})\mathrm{e}^{-\frac{\mathrm{j}2\pi\mathrm{ux}}{\mathrm{M}}}, \quad (5)$$

For u=0,1,...,M-1

The results are a set of Fourier coefficients, which represents the shape using the feature vectors and these normalized Fourier coefficients are used as shape descriptors.

3. DESIGN OF PROPOSED VISION MODULE

3.1 ZYNQ HARDWARE PLATFORM

The proposed image processing system is realized with Xilinx fully programmable SoC chip ZYNQ-7000 7z010clg400-1 and is integrated with a high-performance dual-core ARM Cortex-A9 processor. It provides a wealth of peripheral interfaces and general expansion pins, and can generate the required processing and computing performance for high-end embedded system applications such as industrial control, machine vision, image and video processing and automotive driving assistance. The corresponding hardware block design for the module developed is shown in Fig. 2.

3.2 VISION MODULE FOR THE SHAPE FEATURE EXTRACTION

Image processor block forms the core of the architecture where real-time implementation of feature extraction algorithm of objects in 2D images is carried out. The proposed block determines the centroid as well as computes the Fourier descriptors. To compute the feature, a pipeline process is carried out for each frame and the digital machine for the feature extractor is shown in Fig. 3. In the first stage, the conversion from gray to binary is carried out and stored in Block RAM (referred to as BRAM1). In the second stage, contour coordinates (row, column indexes) of the region of interest are calculated and stored in another Block RAM (referred to as BRAM2). In the third stage, the indices are averaged to determine the x and y coordinates of the centroid (x_c, y_c). The centroid contour distance is determined using equation (4) in the fourth stage. Finally, the FFT of the centroid distance is determined which yields the shape feature descriptors. Starting from the next frame, the feature is obtained in each subsequent frame. BRAM1 is for storing the images which

are captured using the cameras from the processor side. BRAM2 is for storing the indexes of the boundary pixels of the object. The feature vector can be stored in a BRAM that is subsequently accessed for classification.

Following are the main steps involved in the implementation of the image processor block:

1) Initialize row and column counters for reading each row of the binary 2D image stored in BRAM1.
2) Identify pixels (equal to 1) from each row that corresponds to the boundary; the position of row and column index is stored in BRAM2.
3) Average the row and column index to determine the centroid coordinates.
4) CCD distance is calculated by finding the square root of distances between each index position and centroid coordinates, which is the 1D shape signature of the object.
5) Determine the Fourier descriptors by applying FFT on the shape signature

The finite state machine is used to implement BRAM reading and writing, while square root and FFT operations are performed using the Xilinx IP Cores. The efficient data transfer between the BRAM IP core and ZYNQ hard core is very important as the image pixels are stored in the BRAM initially and it is configured in BRAM controller mode. The data is passed through AXI interconnect from the processor/ Processing System (PS) and read out again using custom controller module for further processing developed using Verilog HDL and the state diagram for exchange of data between BRAM is shown in Fig. 4.

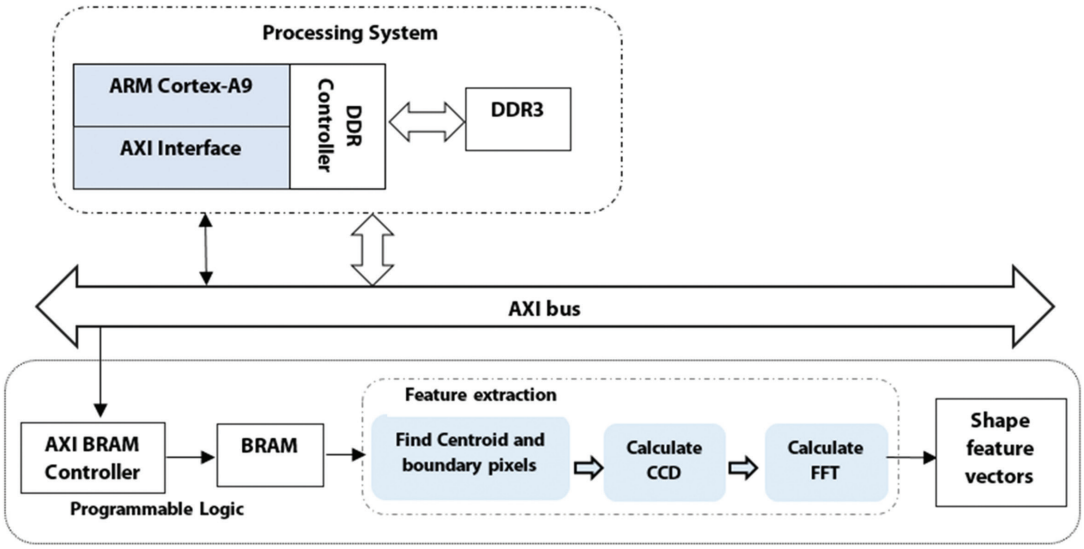


Fig. 1. The overall architecture of the proposed feature extraction module

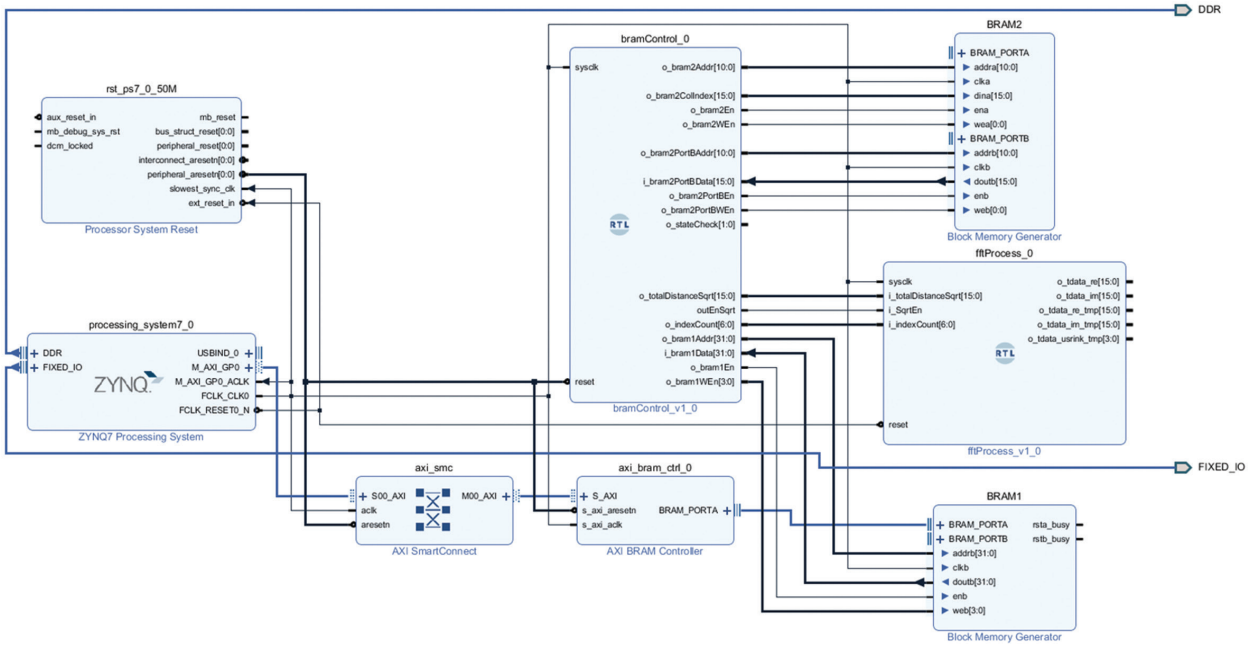


Fig. 2. Hardware Block Design of the proposed shape feature extraction module in Vivado

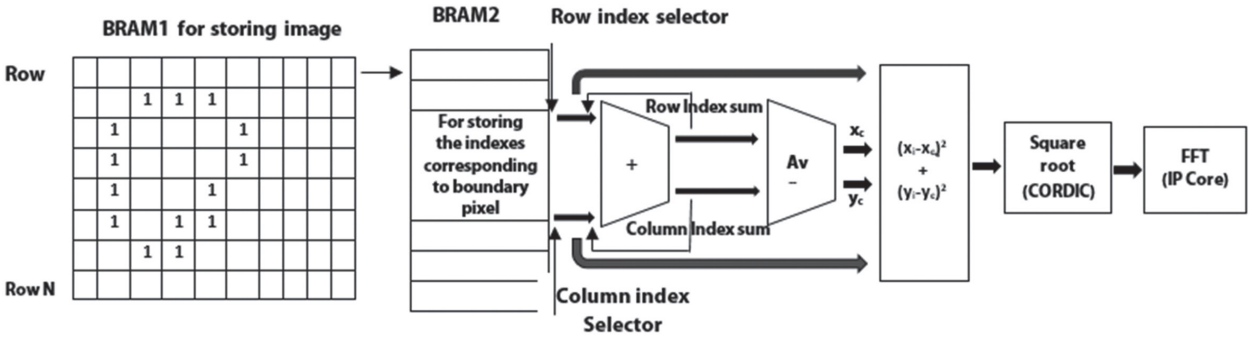


Fig. 3. Digital Machine for CCD –FD feature extractor

The image pixels are written to BRAM from the PS side and the access to the contents of the BRAM1 (configured in BRAM controller mode) is done using the Finite State Machine technique. The memory size of BRAM is initially calculated based on the standard image size from the camera interface. Four states are used to manipulate the pixels and the output of the state machine gives the position of the boundary pixels in terms of row and column index. In the **IDLE** state, all the signal initialisation for the data fetches is done. This state resets the row and column position counters for identifying the pixels on the object shape contour. The address is initialized to read the image data from the storage element. In the **RD_BRAM1** state, enable is made high while the read signal is tied low for its reading. The pixel value is checked if it is in object periphery, in such cases, its coordinates like row and column have to be stored which is done in **WR_BRAM2** state. Another BRAM2 (in standalone mode) is configured for storing the row and column index for each contour pixel. The process is repeated until the entire image frame is processed when it enters the **DONE** state. Likewise, another state machine is maintained to read the contents from the standalone BRAM for further processing.

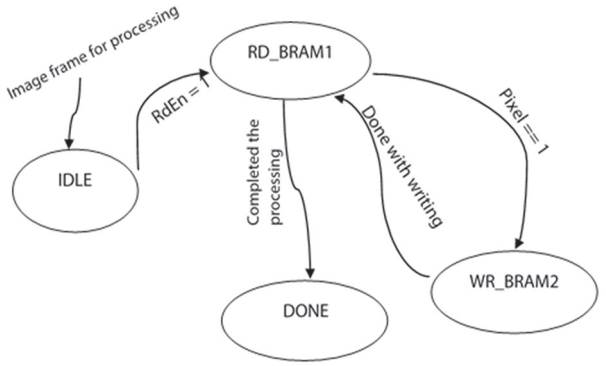


Fig. 4. State diagram showing the pixel manipulation of images

3.2.1 Centroid Calculator

It is necessary to identify the object's boundary pixels to obtain the (x_c, y_c) values for equations (2) and (3). It is determined by averaging up the row_index_sum and column_index_sum separately each time the clock changes.

3.2.2 CORDIC IP Core Configuration

Characterization of IP cores is important while concerning the timing for various configurations. This information is critical in the design flow and it affects all stages, from IP core selection to the synthesis of interfaces to calculation of latency and throughput. CORDIC IP Core is for determining the square root operation while calculating CCD distance. For the CORDIC IP core to be used as the square root calculator block parameters like pipelining mode, data format, input width and round mode has to be properly configured.

Not all configurable parameters will be available for all functions supported by the Xilinx CORDIC IP core. At the same time, some parameters cannot be configured when a particular parameter is selected. For instance, when configured for square root operation with 32-bit unsigned integer input, the output width is automatically set to 17 in the core and the architectural configuration is set to parallel. The square root of the distance between the boundary and centroid coordinates has to be computed for which CORDIC (6.0) IP core is employed and is optimized for FPGA fabrics. The data format chosen is unsigned integer, truncate rounding mode has opted.

3.2.3 FFT IP Core Mode Selection

The Xilinx FFT core (9.1) supports four architectures -Pipelined, Radix-4, Radix-2 and Radix-2 Lite and for the proposed image processing work, a pipeline architecture is adopted. The FFT core is configured to a transform length with the target of 50Millions of Floating-point operations per second (MPS) throughput. Input ports are connected to input source block through appropriate blocks keeping the data and signal integrity issues in mind. Out ports of the FFT processor are terminated, and only the output data such as real, imaginary and index values are captured. The outputs (real and imaginary components) of the FFT IP core are captured from the port m_axis_data (TDATA) by using the configuration settings. The real part of the FFT IP core is obtained by concatenating the bits {m_axis_data_tdata[20],m_tdata[14:0]}. The imaginary part is identified using the bits configuration {m_axis_data_tdata[44], m_axis_data_tdata[38:24]}. The index of the FFT Core is identified by m_axis_data_tuser[3:0].

The designed hardware architecture was functionally correct, simulation was performed with the image (rows) used for feature extraction and was evaluated to improve the simulation time of the design. Simulation of the algorithm is also done in MATLAB.Then the testbench results of the algorithm stages are compared to their counterpart implementation to ensure that the implemented hardware architecture has correct functionality.

The verilog implementation considers the image as a continuous stream of pixel values. The simulation and MATLAB results for the CCD step are almost identical even if the rounding errors associated with square root operations are considered. While performing FFT operations, suitable rearrangement of the output bits and conversion is performed to make the results comparable with software simulation. By performing proper bit alignment to get the real and imaginary components of Fourier descriptors, the implemented hardware architecture for the algorithm can be proved functioning correctly.

4. SIMULATION AND ANALYSIS RESULTS OF THE PROPOSED VISION MODULE

To evaluate the performance of the proposed shape feature extraction algorithm, it is tested on the KTH dataset [22] of hand tools and was collected using the Yumi pedestral robot platform under vision setup. Data set consists of handtools of category- hammer, plier and screwdriver that are taken under different illumination and background as in Fig. 5

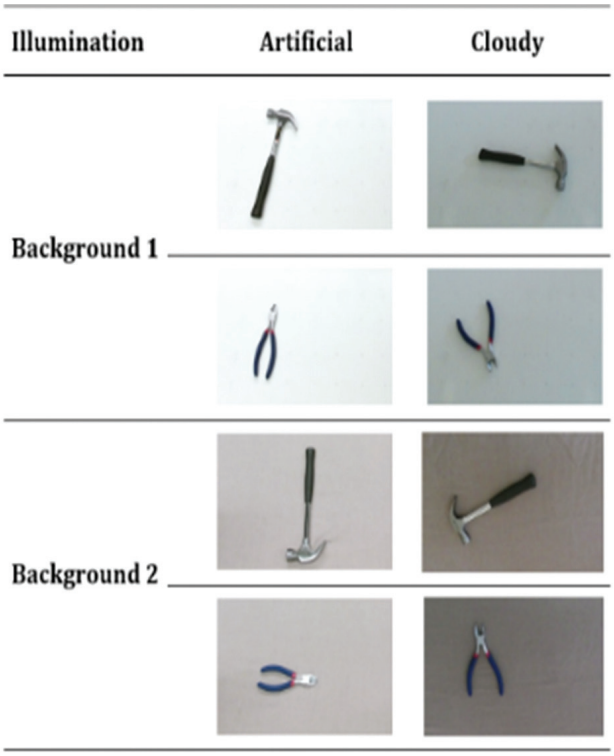


Fig. 5. Sample images of the KTH database captured under the Vision Setup

The shape feature vectors determined are tested using eight different objects: two hammer variants, three distinct pliers and three different screwdrivers. The evaluation used a total of 640 images divided into 8 different classes with 40 images per class. Each of the datasets is captured under artificial and cloudy background. The feature vectors so calculated for the hand tools of the KTH database under different illumination and background setting are translation, rotation and scale-invariant.

To test the robustness of the proposed scheme, we determine the Fourier coefficients of rotated and translated postures of a sample object. The accuracy plot is drawn to prove the feasibility and effectiveness of the proposed algorithm, as shown in Fig. 6

The accuracy was around 86% with 10 feature points (descriptors) and was chosen as the optimum number of points for the classification. Fig. 7a and 7b show the processed binary images and the extracted shape signatures respectively. Taking the Fourier Transform of each of the shifted signatures produced the same normalized Fourier coefficients and can be shown that our implementation is invariant to translation and rotation as shown in Fig. 8.

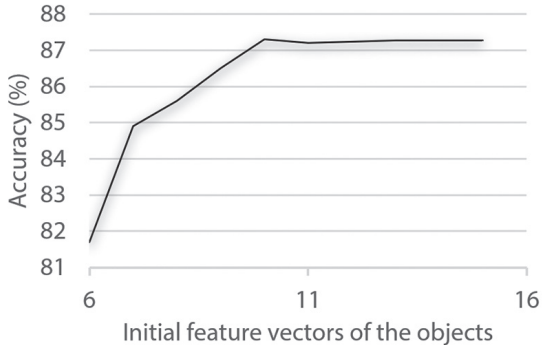
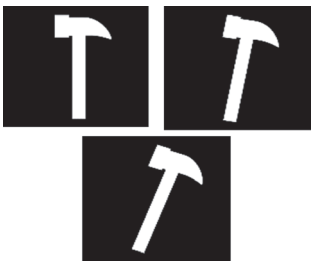


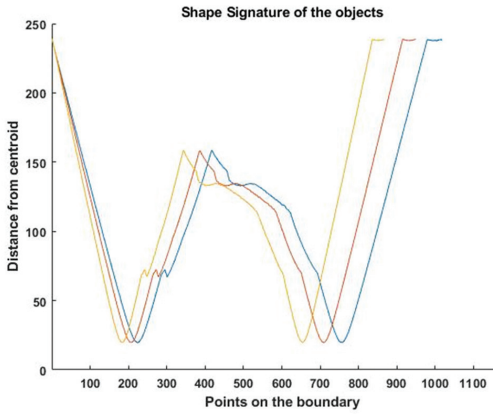
Fig. 6. Accuracy plot using SVM classifier

The hardware implementation of the proposed architecture is synthesized for the Zybo board using Xilinx Vivado Design Suite 2019.2. Simulation is done to verify the functionality of the shape-based feature extraction algorithm. The implementation was also done in MATLAB and both the results were evaluated with the same frames. The results obtained from the simulator are validated by converting the fixed-point output to integer format and are almost identical.

The operations were performed on a 32×64 image and the proposed hardware acceleration module computes the feature vectors within 16.13ns(62 MHz). In this context, it can be recalled that the software implementation of the technique requires 26.3 ms to extract the features. The simulator results of the feature extractor are shown in Fig. 9. The output lines u_ila_0_o_tdata_re_tmp[15:0] and u_ila_0_o_tdata_im_tmp[15:0] are the user defined signals from the ILA (Integrated Logic Analyser) corresponding to real and imaginary components of the FFT IP Core.



(a)



(b)

Fig. 7. Translation and rotation invariance of objects (a) Hand tools for extracting feature vectors (b) The shape signatures of the images with the starting points shifted.

These signals correspond to o_tdata_re_tmp[15:0] and o_tdata_im_tmp[15:0] (of the block fftProcess_0 of the hardware block design). u_ila_0_o_tdata_usink_tmp[3:0] in the simulated waveform provides the index of the Fourier coefficients. The hardware utilization after the place & route of the implemented algorithm is given in Table 1. The FPGA essentially consists of hardware resources such as memory, slice registers, slice LUTs, LUT flip flop pairs and DSP blocks. A full post-synthesis and post-routing test-bench simulation are performed to measure the frame processing time. Here, it is shown that 21.08% of LUTs and 25% of available BRAM is used and this utilisation includes image processing, centroid determination and feature extractor.

Table 1. Resource Utilisation of the proposed shape feature extraction architecture

Resource	Utilisation	Available	Utilization (%)
LUT	3710	17600	21.08
LUTRAM	605	6000	10.08
FF	5025	35200	14.28
BRAM	15	60	25.0
DSP	6	80	7.50
BUFG	2	32	6.25

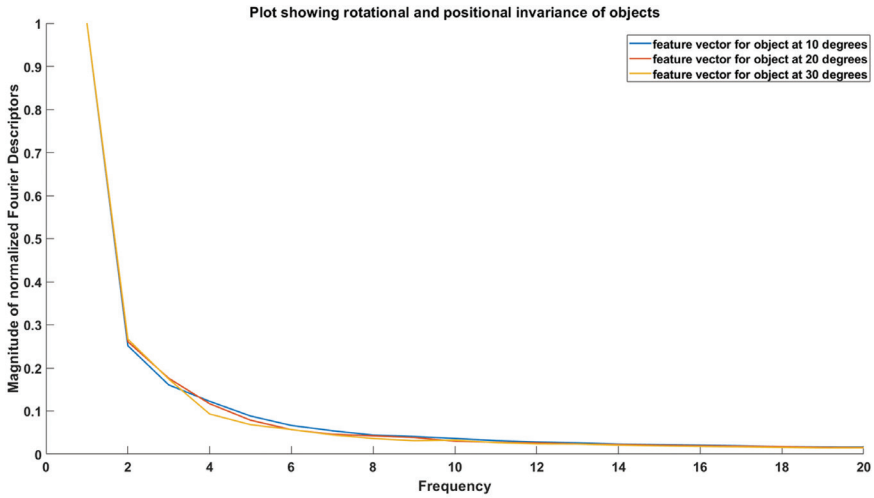


Fig. 8. The initial normalized Fourier coefficients of the shifted shape signatures

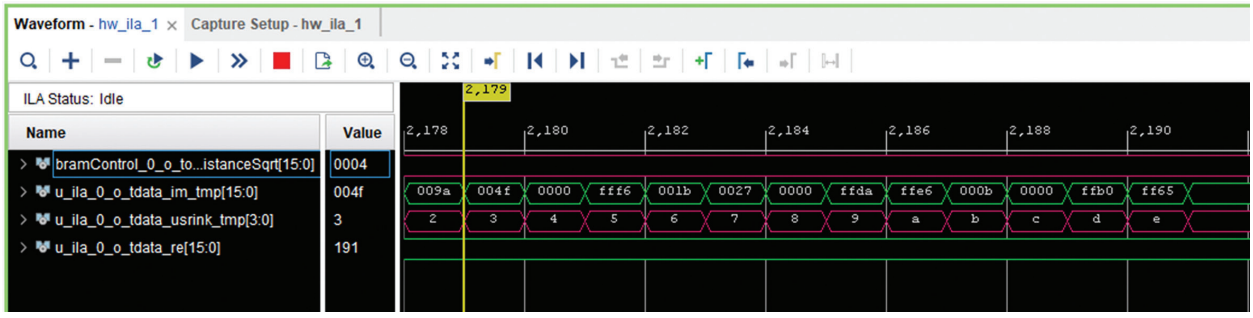


Fig. 9. Simulated results of the proposed Shape Fourier descriptor

5. CONCLUSION

This paper has presented a real-time shape feature extraction algorithm for embedded platforms. This algorithm combines the centroid distance and Fast Fourier Transform for determining the feature descriptors. These are accelerated by using Field Programmable Gate Arrays, which satisfies the real-time performance of the algorithm in embedded platforms. Algorithms for the shape feature extraction have been implemented and tested by hardware modules developed using Verilog language.

Hardware implementation of shape feature vector has been demonstrated using a sample image where processing was done at a speed of 16.13ns as compared to software implementation. The interaction using standard AXI4 interfaces allows different modules in the system to be exchanged or configured easily. Direct Memory Access (DMA) will be investigated for carrying out image transfer for future study.

6. REFERENCES

[1] M. Lu, C. L. Chen, "Detection and Classification of Bearing Surface Defects Based on Machine Vision", Applied Sciences, Vol.11, No.4, 2021.

[2] Y.J Chen, J. C. Tsai, Y. C. Hsu, "A real-time surface inspection system for precision steel balls based on machine vision", Measurement Science and Technology, Vol. 27, No. 7, 2016, pp. 74010-74019.

[3] Z.Ren, F. Fang, N. Yan, Y. Wu, "State of the Art in Defect Detection Based on Machine Vision", International Journal of Precision Engineering and Manufacturing-Green Technology, 2021.

[4] L. Purez, I. Rodriguez, N. Rodriguez, R. Usamentiaga, D. F. Garcia, "Robot Guidance Using Machine Vision Techniques in Industrial Environments: A Comparative Review", Sensors, Vol.16, No.3.2016

[5] G. Maier, F. Florian, M. Wagner, C. Pieper,R. Gruna, B. Noack, H.K. Emden, T. Langle, U. D. Hanebeck, S. Wirtz, V. Scherer, J. Beyerer, "Real-time multitarget tracking for sensor-based sorting", Journal of Real-Time Image Processing, Vol. 16, No. 6, 2019, pp. 2261-2272.

[6] M Schluter, C Niebuhr, J Lehr, J Kruger, "Vision-based Identification Service for Remanufacturing Sorting", Proceedings of the 15th Global Conference on Sustainable Manufacturing, Haifa, Israel, 25-27 September 2017, 2018, pp.384-391.

[7] K. Xia, Z. Weng, "Workpieces sorting system based on industrial robot of machine vision", Proceedings of the 3rd International Conference on Systems and Informatics, Shanghai, China, 19-21 November 2016, pp. 422-426.

[8] A. A. Eissa, A. A. Khalik, "Understanding color image processing by machine vision for biological materials", Structure and Function of Food Engineering, InTech Publishers, 2012.

[9] Q. Luo, X. Fang, L. Liu, C. Yang, Y. Sun, "Automated Visual Defect Detection for Flat Steel Surface: A Survey", IEEE Transactions on Instrumentation and Measurement, Vol. 69, No. 3, 2020 pp. 626-644.

[10] L. Armi, S. F.-Ershad, "Texture image analysis and texture classification methods - A review", International Online Journal of Image Processing and Pattern Recognition, Vol. 2, No.1, 2019, pp. 1-29.

[11] T. Czimmermann, G. Ciuti, M. Milazzo, M. Chiurazzi, S. Roccella, C. M. Oddo, P. Dario "Visual-Based Defect Detection and Classification Approaches for Industrial Applications—A Survey", Sensors, Vol. 20, No.6, 2020.

[12] M. I. Al Ali, K. M. Mhaidat, I. A. Aljarrah, "Implementing image processing algorithms in FPGA hardware", Proceedings of the IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies, Amman, Jordan, 3-5 December 2013, pp. 1-5.

[13] N. Nausheen, A. Seal, P. Khanna, S. Halder, "A FPGA based implementation of Sobel edge detection", Microprocessors and Microsystems, Vol.56, 2018, pp. 84-91.

[14] R. Maini, H. Aggarwal, "Study and Comparison of Various Image Edge Detection Techniques", International Journal of Image Processing, Vol. 3, No. 1, 2009, pp. 1-11.

[15] A. B. Amara, E. Pissaloux, M. Atri, "Sobel edge detection system design and integration on an FPGA based HD video streaming architecture", Proceedings of the 11th International Design & Test Symposium, Hammamet, Tunisia, 18-20 December 2016, pp. 160-164.

[16] Y. Zheng, "The Design of Sobel Edge Extraction System on FPGA", Proceedings of the 7th International Conference on Information Science and Technology, Washington, USA, 16-19 April 2017.

[17] G. B. Reddy, K. Anusudha, "Implementation of image edge detection on FPGA using XSG", Proceedings of the International Conference on Circuit, Power and Computing Technologies, Nagercoil India, 18-19 March 2016, pp. 1-5.

[18] S. Taslimi, R. Faraji, A. Aghasi, H. R. Naji, "Adaptive Edge Detection Technique Implemented on FPGA", Iranian Journal of Science and Technology – Transactions of Electrical Engineering, Vol. 44 No.4, 2020, pp. 1571-1582.

[19] A. G. Mahalle, A. M. Shah, "An Efficient Design for Canny Edge Detection Algorithm Using Xilinx System Generator", Proceedings of the 3rd International Conference on Research in Intelligent and Computing in Engineering, Universidad Don Bosco, El Salvador, 22-24 August 2018, pp. 1-4.

[20] Zybo Z7 Board Reference Manual https://reference.digilentinc.com/programmable-logic/zybo-z7/reference-manual (accessed: 2021)

[21] D. S. Zhang, G. Lu, "Study and evaluation of different Fourier methods for image retrieval", Image and Vision Computing, Vol. 23, No.1, 2005, pp. 33-49.

[22] M. Mancini, H. Karaoguz, E. Ricci, P. Jensfelt, B. Caputo, "Kitting in the Wild through Online Domain Adaptation", Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Madrid, Spain, 1-5 October 2018, pp. 1103-1109.

Multispectral Image Classification Based on the Bat Algorithm

Original Scientific Paper

Almas Ahmed Khaleel
University of Mosul,
College of Basic Education, Department of Mathematics
Mosul, Iraq
diamond-12@uomosul.edu.iq

Joanna Hussein Al-Khalidy
Ninevah University,
College of Electronic Engineering, Department of Computer and Informatics Engineering
Mosul, Iraq
joanna.abdulhakeem@uoninevah.edu.iq

Abstract *– There are many traditional classification algorithms used to classify multispectral images, especially those used in remote sensing. But the challenges of using these algorithms for multispectral image classification are that they are slow to implement and have poor classification accuracy. With the development of technologies that mimic nature, many researchers have resorted to using intelligent algorithms instead of traditional algorithms because of their great importance, especially when dealing with large amounts of data. The bat algorithm (BA) is one of the most important of these algorithms. This study aims to verify the possibility of using the BA to classify the multispectral images captured by the Landsat-5 TM satellite image of the study area. The study area represents the Mosul area located in the Nineveh Governorate in northwestern Iraq. The purpose is not only to study the ability of the BA to classify multispectral images but also to obtain a land cover map of this region. The BA showed efficiency in the classification results compared to Maximum Likelihood (ML), where the overall accuracy of classification when using the BA reached (82.136%), while MLreached (79.64%).*

Keywords*: Bat algorithm; Multispectral image; Classification; Land cover classification*

1. INTRODUCTION

In recent years, algorithms based on swarm intelligence (SI) have grown rapidly and significantly. They are being used extensively in many mathematical and engineering applications as well as in digital image processing. The bat algorithm (BA) is one of the most widely used techniques for solving clustering problems especially in remote sensing images where traditional classification algorithms do not provide high accuracy [1].

The idea of the BA is taken from nature. When bats fly at night in search of food, they emit echoes to guide them and determine their correct path, thus avoiding any obstacles they might encounter in their path. During its flight, the bat uses a technology that is somewhat similar to a sonar. It listens to the echoes of the sound to guide it on its way. A bat's echolocation is determined by emitting high-frequency sound pulses through its mouth or nose and listening for the echo [2].

Many researchers have studied the bat-inspired algorithm and below are short summaries of their works:

In [3], the authors presented a bat-inspired algorithm (BA), which relied on the behaviour of echo sites, to select multiple thresholds for different levels using the principle of maximum entropy (MaxEnt). The experimental results obtained in this study indicated that the proposed BA algorithm could find multiple thresholds similar to the optimal ones that were determined using a comprehensive search approach. Compared with Cuckoo search(CS), the computational times show that the BA algorithm is superior to the CS algorithm.

In [4] the authors provide a comprehensive review of the new BA and its variants. In this review, the following three main features are summarized by analyzing the main features and updating the equations:

1- Frequency control: This feature may be similar to the one used for particle swarm optimization and harmony search.

2- Auto Zoom: The BA can automatically zoom in on the area where promising solutions have been found. For this reason, he concluded that BA has a fast rate of convergence.

3- Parameter control: BA uses parameter control, which enables it to change parameter values while iterations continue.

In [5] the authors used the BA to solve the multi-level image threshold problem. From the results they obtained, they concluded that the BA outperformed all other tested algorithms.

In [6], the authors used the BA to improve the contrast of the grayscale fingerprint image. The purpose of the BA was to map gray level distributions of contrast enhancement endings.

In [7], the authors discussed in their paper a variety of research articles, most of which dealt with the possibility of modifying the BA and its available applications. Despite the characteristics and capabilities of this algorithm, the researchers argue, there are still some problems that require additional research, such as parameter tuning and parameter control in this algorithm. However, BA remains an efficient algorithm.

In [8], the authors used simulation experiments on six test functions to verify and improve the Binary Bat algorithm (BBA). According to the simulation results, it was found that the convergence velocity of the algorithm is relatively sensitive to the parameter setting of the algorithm used.

In the field of remote sensing, many methods are proposed to extract the important features from the image, the most important of which are the following:

In [9], the authors presented in their paper a method for extracting features of remote sensing images based on BA and natural chromatic aberration. Initially, the contrast of remote sensing images was improved by the BA. Then, the colour feature was extracted using normal chromatic aberration, and then the features were encoded in a binary system.

In [10], the authors used BA to solve the transport network design problem so as to get the best objective function value solution. BA outperformed all intelligent techniques such as Genetic Algorithms (GA) and, Ant Colony Optimization (ACO).

In [11], the authors explored BAT Algorithm (E-BA) with a 2-D histogram is based on multi-level image thresholding for effective picture thresholding which is achieved by maximising the Renyi entropy. E-BA was successfully tested on standard test pictures to demonstrate the algorithm's performance. The E- BA's obtained results were compared to Renyi entropy algorithms such as BAT and Particle Swarm Optimization (PSO). According to these comparisons, the algorithm (E-BA) has the highest fitness value among all the algorithms.

In [12], the authors presented a robust digital image watermarking scheme based on Static Wavelet Transform (SWT) using the Bat Optimization BA and the powerful Speedup (SURF) feature. The results showed that it is possible to achieve a targeted balance between watermark visibility and durability.

In [13] the authors presented a BA analysis based on the preliminary mathematical analysis and statistical comparisons of the first-hit time performance measurement distributions obtained on a test set of five carefully selected objective functions. The results show that BA is not an original contribution and does not generally outperform the PSO algorithm when making a fair comparison. Finally, the results indicate that the best BA version is a simple combination between PSO and simulated annealing.

In [14], the authors propose in their paper several Objective Bat algorithms (MaOBAT) that solve multi-objective optimization problems through effective approximation of Pareto approximation (PA) with high diversity and good convergence. A pilot study shows that MaOBAT works efficiently and has significant advantages over MaOPSO, SMPSO, and NSGA III.

Some previous studies in the field of multispectral and classification based on image processing were reviewed: Perumal and Bhaskaran compared the performance of different classifiers and found that the Mahalanobis classifier is superior even to the advanced classifiers. In terms of successful image classification, the simple and accurate workbook demonstrates the importance of considering the relationship between the data set and the classifier. Also, the researchers stated that additional studies are needed to improve the use of classifiers to increase the applicability of such methods [15].

In [16], Some researchers revealed the loss of vegetation cover for Nineveh Governorate by using remote sensing images based on the algorithm of an ant colony According to the results, the agricultural area and floodplain decreased from about 31% in 1987 to 11.2% in 2009, while the origin of the early sandy plate and desert region increased from 42.7% to 49%. Sand dunes also appeared in 2009 in about 26.47% of the overall study area.

In [17], the authors proposed a bat-based clustering algorithm. They used multispectral satellite images to help solve crop type classification problems. The performance of the proposed method is compared to that of three other techniques: K-mean clustering, GA, and PSO. Since BA successfully converges to optimal centres of mass, it was concluded that BA can be successfully applied to address crop type classification problems.

In [18], the authors presented a method for detecting buildings in densely populated urban areas by integrating data from the first pulsed laser scanner. The purpose of this method is to achieve a land cover classification.

In [19], the authors address the problem of multispectral satellite image classification of forest plants.

Features were investigated based on wavelet transform and a classification method involving consideration of the significance of each feature.

In [20], the author investigated the neural network matrix classifier and its application in multispectral image classification. By the results obtained, the effectiveness of this classifier has been demonstrated compared to standard classifiers.

In [21], the authors optimized the K-mean algorithm according to the problems found in visual tagging when performing supervised classification. The weighting coefficient and K-mean clustering method were combined to characterize the training samples, which reduced the subjective influence of visual interpretation on the classification results.

In [22], the authors proposed an automated approach to detecting water bodies based on the Dempster-Shafer theory of combining supervised learning with a specific property of water in the spectral range in a completely unsupervised context, thus providing valuable information for further land cover classification.

In [23], the authors presented a method of describing and classifying multispectral images that allows recognition of classes of different organisms was presented. The algorithm has been successfully applied to forest plants.

Using multispectral sensors, In [24], the authors proposed a new method for recording captured images and introduced a new fusion scheme that combines images differently for high and low frequencies.

In [25], the authors proposed a multispectral image classification based on an object-based active learning approach which is a novel method for object-based sampling.

In [26], the authors responded to the difficulty of detecting a change in multispectral images by proposing a methodology (GBF-CD) based on the merging of graphic data.

For our current study, Mosul city is the study area. It is located in Nineveh. Mosul is about 465 km to the north of Baghdad, as shown in Fig. 1 [24].

Fig. 1. Study Area

This study proposes a supervised classification method for multispectral image data based on the BA. The aim of this study is to use the BA for classifying multispectral satellite image datasets and compare their overall accuracy with the conventional image classification method.

The study organization is as follows: Section 1 presented an introduction; Section 2 presents the bat algorithm concept; Section 3 provides an image classification based on the BA. Section 4 discusses the proposed work and section 5 lists the results of the study.

2. BAT ALGORITHM (BA)

The bat algorithm (BA) is developed by Xin-She Yang in 2010 and it is established according to the following rules [3]:

Bats have an echolocation behaviour similar to that of microbats as they measure the distance and the difference between prey and background block. To locate a prey, bats fly indiscriminately in space vi at position x_i, with a frequency f_i and loudness A_o. They may change the wavelength (frequency) of the pulses and the rate at which they release them based on the proximity of their prey. The loudness should range from a high positive value A_o to a low constant value A_{min}. As the number of samples grows, this rule can be changed. The following will be used to determine the new locations xi(t) and velocities vi(t) [17]:

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (1)$$

$$v_i(t+1) = v_i(t) + (x_i(t) - p(t)).f_i \quad (2)$$

$$f_i = f_{min} + (f_{max} - f_{min}).\beta \quad (3)$$

When it is a regular random number the range will be [0, 1], $f_{min} = 0$, $f_{max} = 1$ and p(t) represents the current global best solution (location). When the bats' positions are updated, a random number is created. If the generated random number is more than the pulse emission rate r_i, a new solution is generated using a local random walk around the existing global best solution.

$$x_i(t+1) = \vec{p}(t) + \varepsilon\overline{A}(t) \quad (4)$$

When the random number is ε, which is between [-1, 1], the population average loudness is $\overline{A}(t)$.

Furthermore, by regulating the loudness A_i (t+1) and the pulse rate r_i(t+1).

$$A_i(t+1) = \alpha A_i(t) \quad (5)$$

$$r_i(t+1) = r_i(0)[1 - \exp.(-\gamma t)] \quad (6)$$

The constants are α and γ while the initial values of loudness and pulse rate are α>0,γ>0. A_i (0)and r_i (0), respectively. The steps of the standard BA are described as follows: The first step: Initialize the position, velocity, and parameters for each bat, then use Equation to produce the frequency at random (3).

The second step: Eq. (1) and Eq. (2) should be used to update each bat position and velocity.

The third step: Create a random number for each bat (0 < rand1 < 1). If rand1 $r_i(t)$, update the temp position and use Eq. (4) to get the fitness value for the corresponding bat.

The fourth step: Create a random number for each bat (0 rand2 1). If rand2 $A_i(t)$ and $f(x_i(t))$ $f(p(t))$, update $A_i(t)$ and $r_i(t)$ with Eq. (5) and Eq. (6), respectively.

The fifth step: Select each individual according to their fitness levels and save It.

The sixth step: If the requirements are met, the algorithm is complete; otherwise, proceed to the second step[11].

3. Multispectral Image Classification using the BA

Based on the above descriptions, the main steps of the BA consist of a single loop with some probabilistic switching during the iteration. Thus, the BA procedure is summarized in the following flowchart, as shown in Fig. 2.

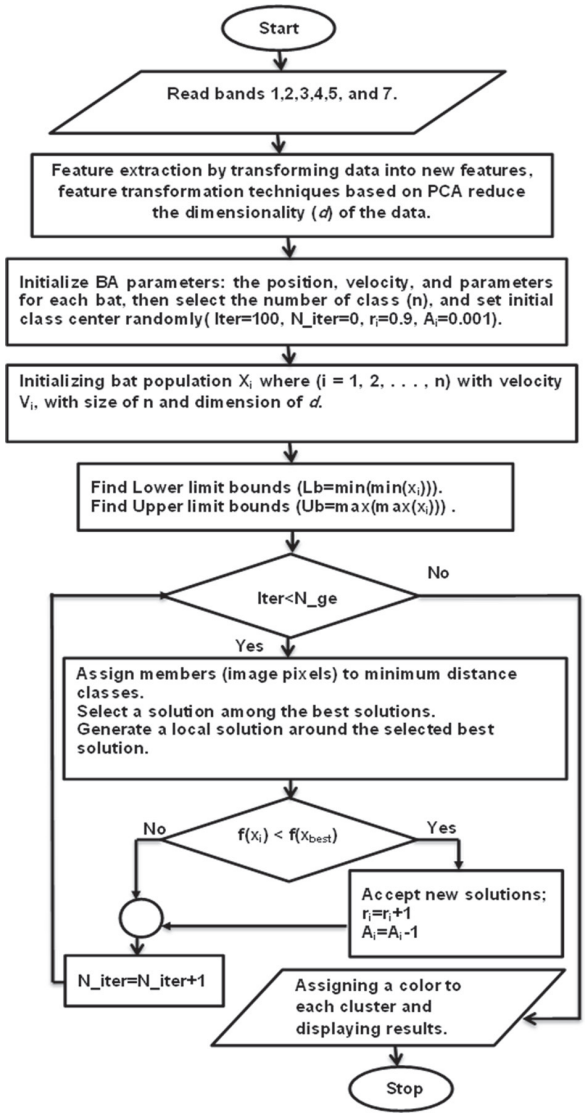


Fig. 2. BA flowchart

4. PROPOSED METHODOLOGY

Due to the large amount of data contained in multispectral images, dealing with the amounts of information contained in them is necessary. Dealing with it, though, is complicated. Unlike interpreting aerial imagery, these multi-spectral images do not allow for immediate object detection of the type of feature. As a result, a remote control is required. To help the user understand the features in the image, sensor data should be classified first, then processed by using various data optimization techniques. Depending on the classification algorithm used, this classification is a difficult task that needs careful validation of training samples. The classification algorithms in this study come in a variety of shapes and sizes. The simulation algorithm used is the BA.

Classification algorithms are gaining popularity in the field of remote sensing and land and land cover classification. Because of the complexity of extracting information from images and satellite data, the task of classifying images is a daunting task. But with the huge advances in swarm intelligence simulation in nature and artificial intelligence algorithms, it has helped solve classification problems. For this reason, the current study aims to verify the possibility of using the BA in land cover classification.

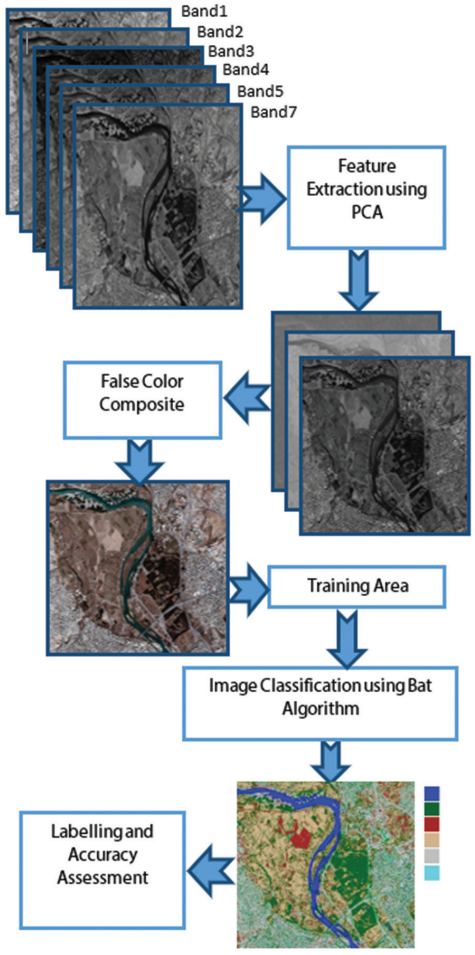


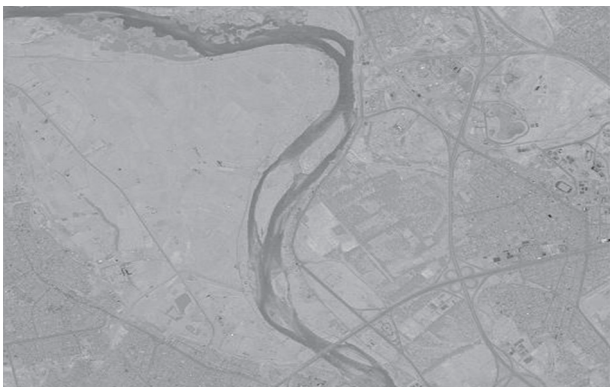
Fig. 3. Steps of multispectral image classification

Figure 3 illustrates the basic steps that were applied in order to classify the multispectral images of the Mosul area in north western Iraq.

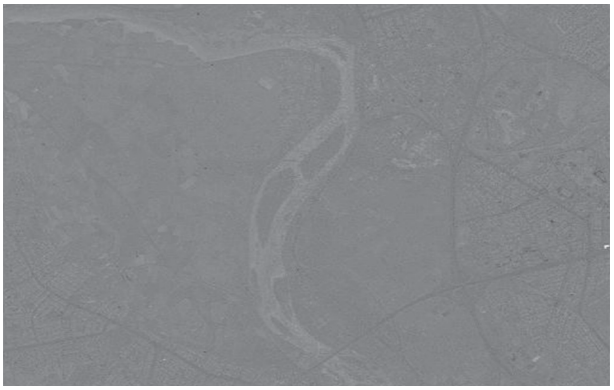
To control the amount of red, green, and blue in a colour image, three LANDSAT spectral bands were used to produce three-band composite images. The spectral range of vision of the human eye is roughly represented by the False-color composite images with band groups 7, 4, and 2; therefore these images appear close to what would be expected in a normal colour photograph. This phase is essential for determining the training area based on field information. A synthetic representation of a multi-spectral image is a false-colour image. The range numbers used for red, green, and blue in a given order are often used to identify the specific ranges used in tri-band combinations. So, an image that uses Band 7 will be assigned to red, Band 4 to green, and Band 2 to blue (7, 4, 2). As shown in Fig. 4, the short wave complex of ranges 7, 4 and 2 represents the Tigris River in bluish-green, and the cultivated areas with different degrees of brown. The roads appear as straight, light lines, and so is the case with the rest of the other classes of land cover. The image classification process includes some basic steps such as optimization of multi-spectral images, feature extraction, use of the principal component analysis (PCA) algorithm to extract features, and reduction in the dimensions included in the classification. Instead of entering six images, three components are used (PCA1, PCA2, PCA3), as shown in Fig. 5.

Fig. 4. False-color composite images with band groups (7, 4, 2)

(a)



(b)



(c)

Fig. 5. Principal Component Analysis Results: (a) PCA1, (b) PCA2 and (c) PCA2

5. RESULTS AND DISCUSSIONS

The results of this study show the importance of relying on remote sensing data and Landsat satellite images with the sensor (TM) to detect discrimination and identification of land cover types in the study area, and the importance of using the BA programmed using Matlab to classify multispectral images.

To train the classifier, the training areas were selected from false images based on the field information and the GIS system available about the study area,. Finally, the BA is used to classify multispectral images by drawing on field information and generating colour land cover classification images. Fig. 6 illustrates the classification results.

The BA was used as a supervised classification that takes advantage of the collected training samples. Choosing spectral signatures for different classes of training areas is very important for training the supervised classifier. In this study, training areas were selected based on the false-colour composite in addition to the available field and geographic information about the study area. Since biased selection in supervised classification negatively affects classification accuracy, the accuracy of selection of training regions is important and that is why ERDAS IMAGINE was used to select training regions. Depending on the nature of the study area, the number of classes required to implement the BA was divided into six categories: Water, Agriculture, Barren, Grasslands, Roads/Birds, and Urban

Areas. Fig. 6 and Fig. 7 show the result of the classification using the BA and ML after giving a label for each class and obtaining images of the land cover of the study area.

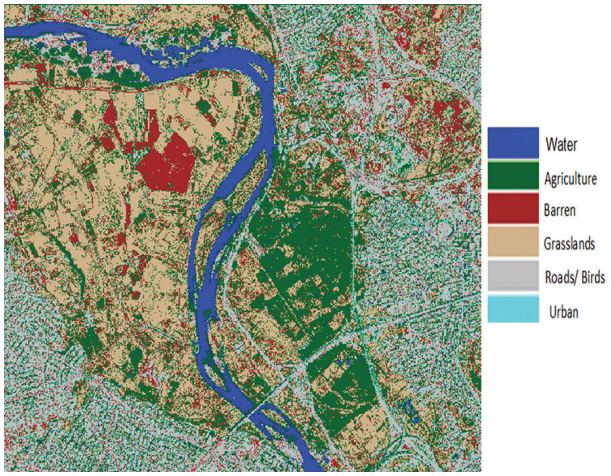


Fig. 6. BA Classification Result

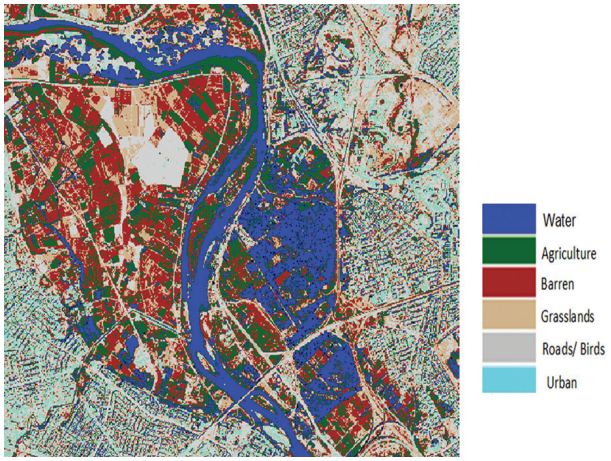


Fig. 7. ML Algorithm Classification Result

To create a land cover map, it is necessary to use a huge amount of data about the study area with the use of one of the classification algorithms. In this study, multispectral images were used, and they were classified using the BA. But the classification will not be useful to use unless accuracy is evaluated. As a result, the objectives of accuracy assessment are to evaluate the accuracy of the classification and document its usefulness so that it can be properly understood by others.

In most cases, accuracy assessment consists of two steps:

1. Collection of reference data. This data enables the identification of land cover classes at specific locations using a methodology independent of the data used for remote sensing classification. Reference data sources include high-resolution, remote-sensed spatial data, such as field or geographic information system information, and field survey measurements with GPS-recorded location.

$$AC=CR/SR, \quad (7)$$

where AC denotes the correct rate. CR represents the number of correctly classified pixel samples. SR represents the total number of pixels in each training class.

Omission error: pixels that belong to actual classes but aren't classified into them (e.g., 15 pixels which should have been classified as Grasslands, were classified as Agriculture).

After calculating the accuracy and error in the performance of the classifier for each class and calculating the final classification accuracy, the results were very acceptable as shown in Table 1 and Table 2. When calculating the classification accuracy, we noticed that the correctly marked classes were the first, second and fourth classes, while the worst marked were the Fifth and sixth classes. But the overall accuracy of the BA classifier was better than the maximum likelihood algorithm, scoring 82.136% and 79.64%, respectively.

Table 1. Accuracy Assessment (BA)

Class No.	Class Name	Accuracy	Omission Error
Class1	Water	96.08	3.92
Class2	Agriculture	87.61	12.39
Class3	Barren	84.4	15.6
Class4	Grasslands	87.51	12.49
Class5	Roads/ Birds	71.9	28.1
Class6	Urban	65.32	34.68
Overall Accuracy= 82.136%		**Overall Error=17.863%**	

Table 2. Accuracy Assessment (ML)

Class No.	Class Name	Accuracy	Omission Error
Class1	Water	90.16	5.71
Class2	Agriculture	84.83	10. 9
Class3	Barren	83.5	18.99
Class4	Grasslands	88.33	10.32
Class5	Roads/ Birds	70.8	29.23
Class6	Urban	60.25	40.68
Overall Accuracy= 79.64%		**Overall Error= 24.805%**	

6. CONCLUSION

In this study, the multispectral image was used to classify the land cover with acceptable accuracy. We used an unconventional classification method based on the BA to produce the land cover map. This study shows that relying on multispectral images is better than using single images; however, a method to extract the characteristics without losing information is needed. PCA was successfully used to extract the three compounds and enter them into the classification algorithm based on the BA and produce a land cover map for the study area. It has been concluded that the use of algorithms that mimic nature such as the BA is useful for classifying multispectral images. The BA proved successful and accurate in classifying multispectral images and producing a ground cover image of the study

area. In the future, we will apply the BA to classify hyperspectral images. As well as using other types of different land cover features.

7. ACKNOWLEDGEMENTS

The authors are very grateful to University of Mosul / College of Basic Education, and Ninevah University/ College of Electronic Engineering for their provided facilities, which helped to improve the quality of this work.

8. REFERENCES:

[1] X.-S. Yang, "Swarm intelligence-based algorithms: A critical analysis", Evolutionary Intelligence. Vol. 7, 2014, pp. 17-28.

[2] S. Nandy, P. P. Sarkar, "Bat algorithm-based automatic clustering method and its application in image processing", Elsevier Ltd, 2016.

[3] A. Alihodzic, M. Tuba, "Bat Algorithm (BA) for Image Thresholding", Recent Researches in Telecommunications, Informatics, Electronics and Signal Processing, 2013, pp. 364-369.

[4] X.-S. Yang, "Bat algorithm: Literature review and applications", International Journal of Bio-Inspired Computation, Vol. 5, No. 3, 2013, pp. 141-149.

[5] A. Alihodzic, M. Tuba, "Improved bat algorithm applied to multilevel image thresholding", The Scientific World Journal, Vol. 2014, 2014.

[6] A. Bouaziz, A. Draa, S. Chikhi, "Bat Algorithm for Fingerprint Image Enhancement", Proceedings of the 12th International Symposium on Programming and Systems, Algiers, Algeria, 28-30 April 2015.

[7] S. L. Yadav, M. Phogat, "A Review on Bat Algorithm," International Journal of Computer Science and Engineering, Vol. 5, No. 7, 2017, pp. 39-43.

[8] X. X. Ma, J. S. Wang, "Optimized Parameter Settings of Binary Bat Algorithm for Solving Function Optimization Problems", Journal of Electrical and Computer Engineering, Vol. 2018, 2018.

[9] Y. Cao, Y. Xun, Y. Han, J. Chen, S. Wang, Z. Zhang, N. Du, H. Meng, "Feature Extraction of Remote Sensing Images Based on Bat Algorithm and Normalized Chromatic Aberration", IFAC-Papers On Line, Vol. 52, No. 24, 2019, pp. 318-323.

[10] S. Srivastava, S. K. Sahana, "Application of Bat Algorithm for Transport Network Design Problem", Applied Computational Intelligence and Soft Computing, Vol. 2019, 2019.

[11] V. Manohar, G. Laxminarayana, T. S. Savithri, "Image compression using explored bat algorithm by Renyi 2-d histogram based on multilevel thresholding", Evolutionary Intelligence, Vol. 19, No. 3, 2019, pp.1-11.

[12] A. Pourhadi, H. Mahdavi-Nasab, "A robust digital image watermarking scheme based on bat algorithm optimization and SURF detector in SWT domain", Multimedia Tools and Applications, Vol. 79, No. 29, 2020, pp. 1-25.

[13] I. Gagnon, A. April, A. Abran, "A critical analysis of the bat algorithm", Engineering Reports, July, 2020.

[14] U. Perwaiz, I. Younas, A. A. Anwar, "Many-objective BAT algorithm", PLoS One, Vol. 15, No. 6, 2020, pp. 1-20.

[15] K. Perumal, R. Bhaskaran, "Supervised Classification Performance of Multispectral Images", Journal of Computing, Vol. 2, No. 2, 2010, pp. 124-129.

[16] Y. M. Abbosh, T. A. Khaleel, J. H. Al-Khalidy, "Vegetation Loss Detection of Nineveh Province Using Remote Sensing Images Based on Ant Colony Algorithm", Al-Rafidain Engineering Journal, Vol. 21, No.6, 2013, pp. 6-7.

[17] J. Senthilnath, S. Kulkarni, J. A. Benediktsson, X. S. Yang, "A Novel Approach for Multispectral Satellite Image Classification Based on the Bat Algorithm," IEEE Geoscience and Remote Sensing Letters, Vol. 13, No. 4, 2016, pp. 599-603.

[18] F. Rottensteiner, J. Trinder, S. Clode, K. Kubik, "Using the Dempster-Shafer method for the fusion of LIDAR data and multi-spectral images for building detection", Information Fusion, Vol. 6, 2005, pp. 283-300.

[19] A. I. Nazmutdinova, V. N. Milich, "Dependence of the results of classification of multispectral images of forest vegetation on wavelet-transform parameters", Optoelectronics, Instrumentation and Data Processing, Vol. 52, 2016, pp. 231-237.

[20] X. Fu, "Multispectral image classification based on neural network ensembles", Proceedings of the 8th

International Conference on Advanced Computational Intelligence, Chiang Mai, Thailand, 14-16 February 2016, pp. 275-277.

[21] X. Jing, S. Y. Chen, L. L. Fan, "Semi-supervised classification of multi-spectral images based on density", Proceedings of the 9th International Conference on Digital Image Processing, Vol. 10420, Hong Kong, China, 2017, pp.1-6.

[22] A. M. Li Na, R. Estival, "An automatic water detection approach based on Dempster-Shafer theory for multi-spectral images", Proceedings of the 20th International Conference on Information Fusion, Xi'an, China, 10-13 July 2017, pp. 1-8.

[23] A. I. Nazmutdinova, A. G. Itskov, V. N. Milich, "Description of the process of presentation and recognition of forest vegetation objects on multi-spectral space images", Pattern Recognition and Image Analysis, Vol. 27, 2017, pp. 105-109,.

[24] N. Ofir, S. Silberstein, D. Rozenbaum, Y. Keller, S.D. Bar, "Registration and fusion of multi-spectral images using a novel edge descriptor", Proceedings of the 25th IEEE International Conference on Image Processing, Athens, Greece, 7-10 October 2018, pp. 1857-1861.

[25] T. Su, S. Zhang, T. Liu, "Multi-spectral image classification based on an object-based active learning approach", Remote Sensing, Vol. 12, No. 3, 2020, pp. 1-37.

[26] D. A. J. Sierra, H. D. B. Restrepo, H. D. V. Cardonay, J. Chanussot, "Graph-based fusion for change detection in multi-spectral images", arXiv:2004.00786, 2020.

Prediction of Plasma Membrane Cholesterol from 7-Transmembrane Receptor Using Hybrid Machine Learning Algorithm

Original Scientific Paper

Rudra Kalyan Nayak
School of CSE
VIT Bhopal University, Sehore, Madhya Pradesh, India
rudrakalyannayak@gmail.com

Ramamani Tripathy
Department of Master of Computer Application,
United School of Business Management,
Bhubaneswar, Odisha, India
ramatripathy1978@gmail.com

Hitesh Mohapatra
Department of Computer Science and Engineering,
Koneru Lakshmaiah Education Foundation,
Vaddeswaram, AP, India.
hiteshmahapatra@gmail.com

Amiya Kumar Rath
Department of Computer Science and Engineering,
Veer Surendra Sai University of Technology, Burla,
Sambalpur, Odisha, India
amiyaamiya@rediffmail.com

Debahuti Mishra
Department of Computer Science and Engineering,
Siksha 'O' Anusandhan (Deemed to be) University,
Bhubaneswar, Odisha, India
mishradebahuti@gmail.com

Abstract *– The researches have been made on G-protein coupled receptors (GPCRs) over the long-ago decades. GPCR is also named as 7-transmembrane (7TM) receptor. According to biological prospective GPCRs consist of large protein family with respective subfamilies and are mediated by different physiological phenomena like taste, smell, vision etc. The main functionality of these 7TM receptors is signal transduction among various cells. In human genome, cell membrane plays significant role. All cells are made up of trillion of cells and have dissimilar functionality. Cell membrane composed of different components. GPCRs are reported to be modulated by membrane cholesterol by interacting with cholesterol recognition amino acid consensus (**L/V**-X $_{(1-5)}$-**Y**-X $_{(1-5)}$-**R/K**) (CRAC) or reverse orientation of CRAC (**R/K**-X $_{(1-5)}$-**Y**-X $_{(1-5)}$-L/V) (CARC) motifs present in the TM helices. Among all, cholesterol is one who is regulated by membrane proteins. Here we took GPCR as membrane proteins and this protein modulates membrane cholesterol. According to cell biology, GPCR regulates a wide diversity of vital cellular processes and are targeted by a huge fraction of approved drugs. In this paper we have concentrated our investigation on membrane protein with membrane cholesterol. A hybrid algorithm consisting of spectral clustering and support vector machine is proposed for prediction of membrane cholesterol with GPCR. Spectral clustering uses graph nodes for calculating the cluster points and also it considers other concept such as similarity matrix, low-dimensional space for projecting the data points and upon this parameter at last construct the cluster centre. Supervised learning method is used for solving regression and classification problems. From the analysis we found that our result shows better prediction accuracy in terms of time complexity when compared with two existing models such as fuzzy c-means (FCM) and rough set with FCM model.*

Keywords*: GPCR, TM, Membrane cholesterol, FCM, Rough Set, Spectral clustering, SVM*

1. INTRODUCTION

In mammalian cells, so many important components are included with their diversified functionality. In recent decades, all researches have been going on cell biology. Because huge amount of unsolved issues are still there and varieties of challenges were emerged day by day. In this manuscript our focal point of research is on membrane cholesterol with plasma membrane protein. Plasma membrane is also known as biological membrane or cell membrane which surrounds every living cell to separate the internal stimuli from the outside stimuli and it is made up of bilayer, membrane proteins and carbohydrates in addition with phospholipids shown in figure 1 [1-7]. Basically, plasma membrane is semi-permeable in nature.

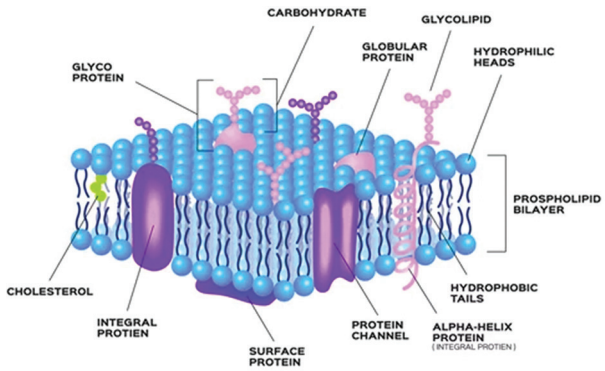


Fig. 1. Schematic representation of cell membrane [5]

By nature cholesterol is amphipathic and it has both hydrophilic and hydrophobic regions. Foremost work of cholesterol in plasma membrane is that it impacts the fluidness and facilitates to produce an effectual dispersal barrier. In the plasma membrane all gaps amongst phospholipids are filled up by cholesterol and also it forbids water soluble molecules from diffusive all around the plasma membrane which is shown in figure 2. A vital purpose of cholesterol is hormone production, Vitamin D production and bile production. Due to much deposit of lipoproteins (LDL) in cell wall, heart disease and other forms of cardiovascular diseases are basically shown in case of human body [8-13].

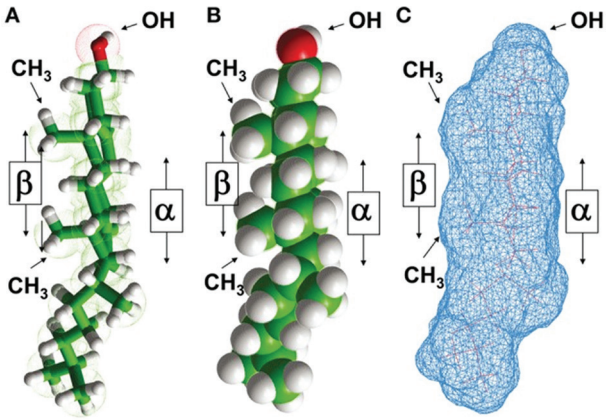


Fig. 2. Cholesterol Structural properties [13]

GPCRs are also called as 7 transmembrane (7-TM) receptors. They are treated as the most important diverse protein families in mammalian genomes.

GPCR is a bigger super family among all cell membrane proteins. It includes above 820 genes with their sub family and symbolized main targets in the development of novel drug candidates in all clinical areas. This family mostly known as larger receptor protein family and are involved in transmitting signals from a diversity of stimuli exterior part to its inside part of cells [13-18]. These families take part in a vital task in physiology by facilitating interaction among cell through recognition of dissimilar ligands, together with nucleosides, bioactive peptides, lipids and amines. Membrane Cholesterol is another imperative component of cellular membrane and has been reported to have a modulatory role in the function of a number of GPCRs. Due to novel functionality of GPCR protein with membrane lipids; it has come out as an exciting domain of research. Cholesterol is a waxy like substances and it is hydrophobic in nature. All cellular cholesterols are distributed unlikely inside the membrane from N-C terminus and to identify cholesterol binding sites of all motifs among seven helices named as helix 1 to helix 7 which is shown if figure 3 below [19-22].

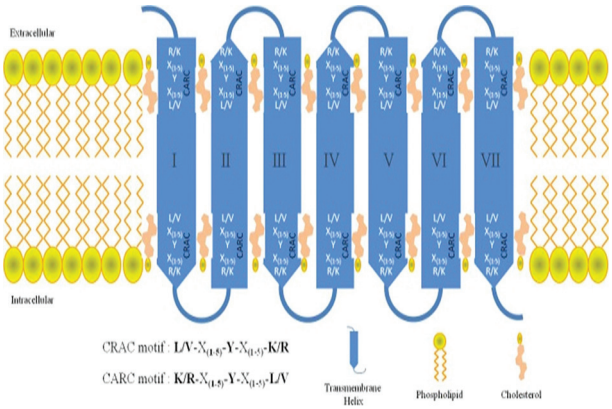


Fig. 3. The 7 helices with N-terminus and C-terminus [1]

Most of the researchers focused their work only on plasma membrane receptor like GPCR. Because GPCR is the largest super family among all the receptors in cell biology and has much functionality such as cell signaling, drug targeting etc. We know that it is an emerging area of research so many researchers have implemented varieties of algorithms upon it like support vector machine, naive Bayes, neural network, fuzzy c-means [21-30] etc. Therefore we have concentrated our experiment on GPCR along with membrane cholesterol which is an innovative idea. Here we proposed a hybrid spectral clustering based FCM model to predict the valid motif sequence(s) from different helices and also compared our proposed model with other two traditional models to showcase the efficacy of proposed model. The said model discovered improved result like more number of targeted promising motif types which could be used for drug design for patients. The rest part of the paper is ordered as follows. In part 2 data set of cholesterol with GPCR proteins and proposed model is discussed. Part 3 explains methodologies with experimental work and finally in part 4 we conclude our work with valid results.

2. DATASET AND PROPOSED MODEL DESCRIPTION

2.1 DATASET DESCRIPTION

From uniprot database [31] we have collected the total helical sequences of each protein. Length of each helix may vary according with their gene ID. Total 820 known proteins with their amino acid sequences reside in database. All helices have individual trans-

membrane region which is the combination of different amino acids. All database genes contain 7 helices that means from helix 1 to helix 7. Another dataset is membrane cholesterol motif sequence. We prepared a cholesterol dictionary on basis of two algorithms that is CRAC for forward orientation and CARC for backward orientation. Table 1 denotes the possible cholesterol motif using forward and backward sequences with the presence of amino acid. Figure 4 shows one helical file snapshot which was retrieved from database.

Table 1. All probable motif type that is mixture of cholesterol plus X which signifies the arrangement of amino acid that can be different from (one to twenty).

Type of Motif	FORWARD MOTIF FORMULA	BACKWARD MOTIF FORMULA
11 to 15	L/V1Y1R/K,…., L/V1Y5R/K	K/R1Y/F1L/V, ….., K/R1Y/F5L/V
21 to 25	L/V2Y1R/K,……, L/V2Y5R/K	K/R2Y/F1L/V, ……, K/R2Y/F5L/V
31 to 35	L/V3Y1R/K, …, L/V3Y5R/K	K/R3Y/F1L/V, ……., K/R3Y/F5L/V
41 to 45	L/V4Y1R/K, ……, L/V4Y5R/K	K/R4Y/F1L/V, ……, K/R4Y/F5L/V
51 to 55	L/V5Y1R/K, ……, L/V5Y5R/K	K/R5Y/F1L/V, ..., K/R5Y/F5L/V

```
>sp|P28222|50-75 SISLPWKVLLVMLLALITLATTLSNAFVIATVYRTRKLHT
>sp|P28566|23-47 PKTITEKMLICMTLVVITTLTTLLNLAVIMAIGTTKKLH
>sp|P28221|39-64 RTLQALKISLAVVLSVITLATVLSNAFVLTTILLTRKLHT
>sp|P28223|76-99 LHLQEKNWSALLTAVVIILTIAGNILVIMAVSLEKKLQ
>sp|P28335|53-78 FKFPDGVQNWPALSIVIIIIMTIGGNILVIMAVSMEKKLH
>sp|Q13639|20-40 GFGSVEKVVLLTFLSTVILMAILGNLLVMVAVCWD
>sp|P41595|57-79 QGNKLHWAALLILMVIIPTIGGNTLVILAVSLEKKLQ
>sp|P34969|84-104 GRVEKVVIGSILTLITLLTIAGNCLVVISVCFVKK
>sp|P30542|11-33 SISAFQAAYIGIEVLIALVSVPGNVLVIWAVKVNQAL
>sp|P29274|8-32 MPIMGSSVYITVELAIAVLAILGNVLVCWAVWLNSNLQN
>sp|P33765|15-37 LSLANVTYITMEIFIGLCAIVGNVLVICVVKLNPSLQ
>sp|Q01718|24-49 NNSDCPRVVLPEEIFFTISIVGVLENLIVLLAVFKNKNLQ
>sp|P30556|28-52 AGRHNYIFVMIPTLYSIIFVVGIFGNSLVVIVIYFYMKL
>sp|P50052|46-71 PSDKHLDAIPILYYIIFVIGFLVNIVVVTLFCCQKGPKKV
>sp|Q16581|24-46 PWNEPPVILSMVILSLTFLLGLPGNGLVLWVAGLKMQ
>sp|Q9P296|39-61 DPLRVAPLPLYAAIFLVGVPGNAMVAWVAGKVARRRV
>sp|P21730|38-60 NTLRVPDILALVIFAVVFLVGVLGNALVVWVTAFEAK
>sp|P30988|172-191 VLYYLAIVGHSLSIFTLVISLGIFVFFRKLTTIF
>sp|Q16602|147-166 NLFYLTIIGHGLSIASLLISLGIFFYFKSLSCQR
>sp|P41180|613-635 LSWTEPFGIALTLFAVLGIFLTAFVLGVFIKFRNTPI
>sp|P32238|42-67 PSKEWQPAVQILLYSLIFLLSVLGNTLVITVLIRNKRMRT
>sp|P41597|43-70 DVKQIGAQLLPPLYSLVFIFGFVGNMLVVLILINCKKLKCLT
>sp|P51677|35-62 DTRALMAQFVPPLYSLVFTVGLLGNVVVVMILIKYRRLRIMT
>sp|P51679|40-67 GIKAFGELFLPPLYSLVFVFGLLGNSVVVLVLFKYKRLRSMT
>sp|P46092|53-68 SVSLTVAALGLAGNGLVLATHLAARRAARS
>sp|P51681|31-58 NVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKRLKSMT
>sp|P51684|48-74 VRQFSRLFVPIAYSLICVFGLLGNILVVITFAFYKKARSMT
>sp|P51685|36-63 LIQTNGKLLLAVFYCLLFVFSLLGNSLVILVLVVCKKLRSIT
>sp|O00421|44-64 SAQLVPSLCSAVFVIGVLDNLLVVLILVKYKGLKR
>sp|Q99788|42-64 EARVTRIFLVVVYSIVCFLGILGNGLVIIIATFKMKK
>sp|P21554|117-142 VLNPSQQLAIAVLSLTLGTFTVLENLLVLCVILHSRSLRC
>sp|P34972|34-59 ILSGPQKTAVAVLCTLLGLLSALENVAVLYLILSSHQLRR
>sp|Q13324|109-139 LDDKQRKYDLHYRIALVVNYLGHCVSVAALVAAFLLFLALRSIRC
>sp|P34998|112-142 ILNEEKKSKVHYHVAVIINYLGHCISLVALLVAFVLFLRLRPGCT
```

Fig. 4. Snapshot of one helical protein

2.2 PROPOSED MODEL

In eukaryotic membrane, cellular cholesterol plays a vital role and is modulated by GPCR which is renowned as cell signalling among intracellular with extracellular leaflets. So many receptors and transporters are available in mammalian cell. But our research focuses on superfamily GPCR receptor with membrane cholesterol. GPCR regulate a wide diversity of vital cellular processes and are targeted by a huge fraction of approved drugs. The workflow is explained in Figure 5 and step wise elaboration is mentioned below.

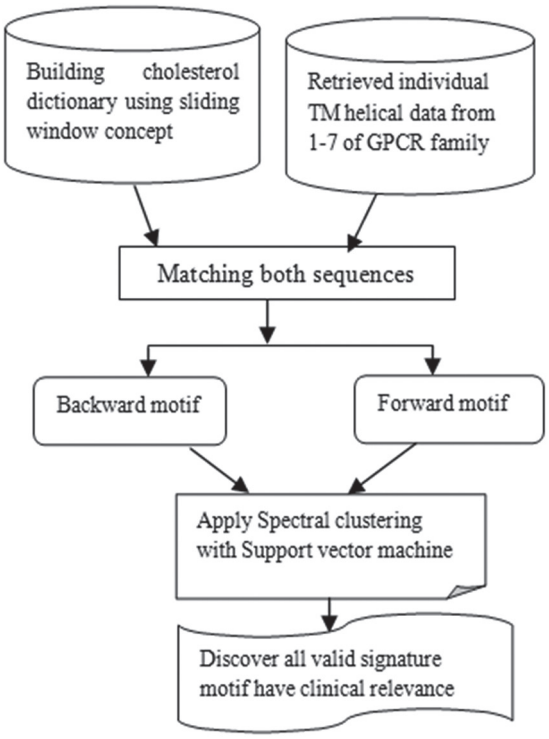


Fig. 5. Proposed model of cholesterol with GPCR family

Step-1: Firstly, we have collected GPCR proteins with their respective helices from uniprot database. And also have constructed cholesterol dictionary with the help of CRAC and CARC algorithm for both forward and backward direction.

Step 2: Then we took matching position of both dataset to find the backward and forward motif using Rabin-Karp string matching algorithm [1].

Step 3: In the next step we have applied hybrid spectral clustering with support vector machine method for both CRAC and CARC algorithm.

Step 4: Finally, we found our targeted valid motif which have clinical relevance.

3. METHODOLOGY WITH EXPERIMENTAL DISCUSSION

3.1 SPECTRAL CLUSTERING

In various fields like bioinformatics, image processing, networking, data mining etc. clustering approaches have been widely used for solving the numerous problems. In every aspect clusters are formed according with their similarity of data objects. As we know that clustering is an unsupervised machine learning-based algorithm that comprises a group of data points into clusters so that the objects belong to the same group. In our paper we used spectral clustering algorithm on forward and backward motif of membrane cholesterol to distinguish the motif sequences with the help of cluster. Spectral clustering uses graph nodes for calculating the cluster points and also it considers other con-

cepts such as similarity matrix, low-dimensional space for project the data points and upon this parameter, at last constructs the cluster centre [32-34].

Algorithm: Spectral Clustering

Input:

Data set Y={y_1,…,y_n}, Initailize σ *scaling* start on:

Step 1: All data are preprocessing with the help of scaling method

Step 2: Make an Affinity matrix $A_{jk} = \exp\left(-\frac{||d_j - d_k||^2}{2\sigma^2}\right)$

Step 3: Put up a Laplacian matrix $L^{norm} := I - P^{-\frac{1}{2}} A P^{-\frac{1}{2}}$

Step 4: Work out the k largest Eigen vectors $x_1,..,x_k$ of L

Step 5: matrix is X=$[u_1,..,u_k] \in R^{(m \times k)}$

Step 6: Outline a matrix W from X as $W_{ij} = \frac{x_{ij}}{(\sum_j x_{ij}^2)^{1/2}}$

Step 7: Cluster every one W by k-means

Step 8: Allocate the X_i to cluster *j*iffW_i is assign to cluster *j*

3.2 SUPPORT VECTOR MACHINE (SVM)

In recent days, classification approach has been treated as one of the powerful tool for dissimilar applications like protein structure prediction, text categorization, face recognition, fingerprint recognition, speech recognition, data classification, micro-array gene expression, etc. In this paper we have applied a novel approach spectral with SVM algorithm on our dataset. Basically, this supervised learning method is used for solving regression and classification problems. SVM theory is always characterizing the decision boundaries using the decision planes concepts [35-37].

A training set that includes label pairs (w_i, v_i), i=1,...,n every w here $w_i \in R^n$ and $v \in \{profit, loss\}^i$, SVM algorithm wants outcome with the help of optimization problem that is stated below.

$$\min_{x,y,\xi} \frac{1}{2} U^M u + C \sum_{i=1}^{n} \xi_i \quad (1)$$

$$\text{Subject to: } v_i(z^M \Phi(w) + y \geq 1 - \xi_i, \xi_i \geq 0. \quad (2)$$

In equation (3) decision function is denoted as

$$P = pfn(\sum_{i=1}^{n} x_i \alpha_i B(x_i, x) + \rho) \quad (3)$$

With help of the kernel function ϕ, training vector w_i is mapped with their dimensional space. According to SVM concept all data points are classified using hyper planes even if it is impracticable for receiving linear solution in case of two dimensional spaces. For this reason we are using kernel function $k(u_i u_j) \equiv \Phi(u_i)^N \Phi(u_j)$ for multidimensional data. Utilizing divergent kernels function, this algorithm is trained which is shown in below equations (4), (5) and (6).

(i) *Linear kernel:*

$$(u_i, u_j) = u_i^N u_j \quad (4)$$

(ii) Polynomial kernel:

$$f(u_i, u_j) = (\gamma u_i^N u_j + r)^e, \gamma > 0, \quad (5)$$

and

(iii) Radial Basis kernel

(RBF): $(u_i, u_j) = exp(-\gamma \parallel u_i - u_j \parallel^2), \gamma \quad (6)$

The entire kernel arguments such as C,γ,r, and e are initialized by utilizing the dataset. All kernel parameters areaffected based upon the size of training data [35-40].

3.3 EXPERIMENTAL PART ELABORATION

The work flow of our manuscript is finished using windows 10 operating system plus Intel i5 processor with hard disk of 8 GB for finishing the experimental part and here Python 3.7 is used for coding purpose. To compute the overall performance here we took helical data of GPCR receptor with dictionary of cholesterol. The work flow of our model is elaborated step wise manner.

Step 1: In very beginning step, first of all extracted the individual helix protein data of GPCR receptor from uniprot database like, protein Id, helix name (h1-h7) and length of the protein whichever is the mixture of different amino acid. Next dataset cholesterol is also computed based on CRAC/CARC approach.

Step 2: After collecting both dataset sequences we used sliding window concept on both to find out the separate motif sequences of backward and forward region. Window size is as W = {*w5, w6, w7, w8, w9, w10, w11, w12, w13*}. The formula for cholesterol dictionary is CRAC and CARC.

Step 3: After completion of step 2 work, we move to next step where we applied our proposed algorithm spectral clustering and SVM for prediction of membrane cholesterol with membrane receptor GPCR. Our proposed algorithm is well suited for both the datasets. The foremost objective of this paper is to find out valid signature motif from prediction.

Here in below, Table 4 shows resultant prediction of cholesterol from GPCR receptor for forward motif. In this table we have taken two columns which have represented as Id number proteins with their motif type, motif sequences with helix name. The entire predicted motifs are found using CRAC (L/V X_{1-5} Y X_{1-5} K/R) formula. Here we can explain one result. With protein id P30550 which is included in motif type 55 (*L/V XXXXX Y XXXXXK/R*) which means first position amino acid *L/V* is present and last position amino acid *K/R* is present. Expect 1st and last position another position is their where another amino acid Y is present. Therefore, our motif sequence is like LSIISVYYYFIAK. In this case first position is followed by L, after then 5 different amino

acid are present such as SIISV and then middle position is Y and after that another dissimilar amino acid are there like YYFIA. Finally, in last position K is present. And all this sequence is present in helix 5 region of GPCR proteins. In this way all forward motif sequences are predicted using the proposed algorithm

Table 4. Resultant prediction of cholesterol from GPCR receptor for forward motif

Identification number of protein with Motif Type	Sequence of motif and helix name	Identification number of protein with Motif Type	Sequence of motif and helix name
P30550 (55)	LSIISVYYYFIAK (helix 5)	P11229(44)	VMCTLYWRIYR (helix 5)
Q13585 (55)	LIVGFCYVRIWTK (helix 5)	P08912(44)	VMTILYCRIYR (helix 5)
P28336 (55)	LAIISIYYYHIAK (helix 5)	Q14833(44)	VTCTVYAIKTR (helix 5)
P49683(55)	LVILLSYVRVSVK (helix 5)	P16473(42)	VIVCCCYVK (helix 5)
P32745(55)	LVICLCYLLIVVK (helix 5)	Q9UBY5(52)	VVNPIIYSYK (helix 7)
P32248(54)	LAMSFCYLVIIR (helix 5)	Q8NH63(52)	VLNPIVYSVK (helix 7)
P25025(54)	LIMLFCYGFTLR (helix 5)	Q14833(52)	VSLGMLYMPK (helix 7)
O43193(54)	LCLSILYGLIGR (helix 5)	O15303(52)	VSLGMLYVPK (helix 7)
P41146(54)	LVISVCYSLMIR (helix 5)	Q14831(52)	VALGMLYMPK (helix 7)
Q96G91(54)	LLTLAAYGALGR (helix 5)	O00222(52)	VSLGMLYMPK (helix 7)
P51582 (54)	LVTLVCYGLMAR (helix 5)	P29275(52)	VVNPIVYAYR (helix 7)
Q9UKP6 (54)	LLIGLLYARLAR (helix 5)	P41968(52)	VIDPLIYAFR (helix 7)
P08908(54)	LLMLVLYGRIFR (helix 5)	P33032(52)	VMDPLIYAFR (helix 7)
Q9NPB9 (54)	LIMGVCYFITAR (helix 5)	Q96R84(52)	VMNPLIYSLR (helix 7)
O43603(44)	VLGLTYARTLR (helix 5)	P46092(52)	LNPVLYAFLGLR (helix 7)
P32239(44)	VMAVAYGLISR (helix 5)	P41231(52)	LDPVLYFLAGQR (helix 7)
P41146(44)	VISVCYSLMIR (helix 5)	P30411(45)	LNPLVYVIVGKR (helix 7)

Table 5. Resultant prediction of cholesterol from GPCR receptor for backward motif

Identification number of protein with Motif Type	Sequence of motif and helix name	Identification number of protein with Motif Type	Sequence of motif and helix name
P49238(55)	KSVTDIYLLNLAL (helix 2)	P03999 (42)	RQPLNYILV (helix 2)
P41143(55)	KTATNIYIFNLAL (helix 2)	Q86VZ1 (42)	RHHWVFGVL (helix 3)
P41145(55)	KTATNIYIFNLAL (helix 2)	Q9BZJ6 (42)	RVSAMFFWL (helix 3)
P41146(55)	KTATNIYIFNLAL (helix 2)	P21453 (42)	REGSMFVAL (helix 3)
P55085(55)	KHPAVIYMANLAL (helix 2)	P21452 (42)	RAFCYFQNL (helix 3)
Q99500(55)	KFHNRMYFFIGNL (helix 2)	Q9NYW4 (42)	RYLSIFWVL (helix 3)
O00421(55)	KRVENIYLLNLAV (helix 2)	P32248 (35)	KMSFFSGMLLL (helix 3)
Q9Y271(55)	KSAFQVYMINLAV (helix 2)	P25024 (35)	KEVNFYSGILL (helix 3)
Q969V1(55)	KTVPDIYICNLAV (helix 2)	P47900 (35)	KLQRFIFHVNL (helix 3)
Q9GZQ4(55)	KTPTNYYLFSLAV (helix 2)	P41231 (35)	KLVRFLFYTNL (helix 3)
P31391(55)	KTATNIYLLNLAV (helix 2)	P51582 (35)	KFVRFLFYWNL (helix 3)
P50052(55)	KKVSSIYIFNLAV (helix 2)	Q9NYW0 (35)	KIANFSNYIFL (helix 3)
O43193(55)	RTTTNLYLGSMAV (helix 2)	P46094 (52)	RTVKLIFAIV (helix 6)
Q9HB89(55)	RTPTNYYLFSLAV (helix 2)	P25024 (52)	RAMRVIFAVV (helix 6)
P23945(54)	KLTVPRFLMCNL (helix 2)	P25025 (52)	RAMRVIFAVV (helix 6)
P22888(54)	KLTVPRFLMCNL (helix 2)	P30559 (52)	RTVKMTFIIV (helix 6)
P16473(54)	KLNVPRFLMCNL (helix 2)	P37288 (52)	RTVKMTFVIV (helix 6)
Q6W5P4(54)	KKSRMTFFVTQL (helix 2)	P47901 (52)	RTVKMTFVIV (helix 6)
P30559(54)	KHSRLFFFMKHL (helix 2)	P49019 (32)	RIHIFWLL (helix 6)
Q9NYW0(54)	KLSTIGFILTGL (helix 2)	P49683 (32)	RRRTFCLL (helix 6)
P21453(54)	RPMYYFIGNLAL (helix 2)	O14514 (32)	RSALFQIL (helix 6)
P43220(54)	RALSVFIKDAAL (helix 2)	Q9UP38 (32)	RIGVFSVL (helix 6)
P34969 (42)	RQPSNYLIV (helix 2)	Q9ULW2v (32)	RIGLFSVL (helix 6)
P29371(42)	RTVTNYFLV (helix 2)	Q14332 (32)	RIGVFSVL (helix 6)
O43613(42)	RTVTNYFIV (helix 2)	Q9NPG1 (32)	RIGVFSIL (helix 6)
O43614(42)	RTVTNYFIV (helix 2)	Q13467 (32)	RIGIFTLL (helix 6)
O95977(42)	RRWVYYCLV (helix 2)	O60353 (32)	RIGVFSGL (helix 6)
P21452(42)	RTVTNYFIV(helix 2)	O75084 (32)	RIGVFSVL (helix 6)
P25103(42)	RTVTNYFLV (helix 2)	Q9H461 (32)	RLGLFTVL (helix 6)
P35368(42)	RTPTNYFIV (helix 2)		

Table 5 shows resultant prediction of cholesterol from GPCR receptor for backward motif. All the motif sequences for forward motif was mentioned in table 4

Like table 4, all the motif sequences of table 5 have been predicted for backward direction. Here the CARC (K/RX1-5 YX1-5 L/V) formula has been implemented for prediction. Here we took one predicted motif sequence

RTVTNYFIV from table 5 for description. This sequence has protein id P21452 and it is included under helix2. First position of sequence motif is either K or R and middle part is constant I that is Y and the last position it contains either amino acid L or V. From this analysis we found backward motifs target the membrane proteins more in comparison to forward motifs. Most of the target sites are under higher motif 55, 52, 45, 42, 35 etc. and helix are 5, 2, 7, 3, 6.

3.4 COMPARATIVE ANALYSIS

In Table 6, we have compared FCM model, rough set with FCM model and our proposed model to showcase the efficacy of the proposed model in terms of discovering suitable types of motifs. Target site of helix by FCM model was found to be h2, h5 and h7 having motif type 11, 12, 21, 54, 34. Further, target site of helix by rough set with FCM model was found to be h2, h3, h5 and h7 having motif type 21, 51, 44, 54, 25, 53. Our proposed model discovers the target site of helix as h2, h3, h5 and h7 and extra helix in h6. Also their motif types are 55, 52, 45, 42, 35 which is higher than the existing models.

Table 6. Motif type comparison by different methods

Methods	Helix Name	Motif Type (Forward/ Backward)
FCM [1]	h2,h5,h7	11,12,21,54,34
Rough Set with FCM [28]	H2,h3,h5,h7	21,51,44,54,25
Spectral with SVM (Proposed)	H2,h3, h5,h7,h6	55,52,45,42,35

4. CONCLUSION

In this paper, we concentrated our work on prediction of uncovering membrane cholesterol from human GPCR super family. Frequently, such receptors are the most significant protein super family in biological membrane and play a substantial role in the transduction of signal across cell membranes. It also signifies as an essential drug target in all clinical fields. About 820 human proteins are included in this family. Membrane cholesterol has a modulatory role in the function of some GPCRs.

In our manuscript we have discussed about cellular receptor with membrane cholesterol. According to biological perspective GPCRs consist of large protein family in mammalian cells. The main functionality of these 7TM receptors is signal transduction among unlike cells. In human genome, cell membrane plays significant role. Cell membrane composed of different components. GPCRs are reported to be modulated by membrane cholesterol by interacting with these CRAC or CARC motifs present in the TM helices. Among all, cholesterol is one who is regulated by membrane proteins. From experiment, we found both forward and backward motif from different helices. Among all, we observed targeted sites are under higher motif 55, 52, 45, 42, 35 etc. and helix are 5, 2, 7, 3, 6.

Here, our experimental analyses conclude that prediction of membrane cholesterol with GPCR receptor using spectral and SVM performs well. Backward motif sequences target the protein sites greater than forward motif that means CARC algorithm has higher valid signature motifs which has clinical relevance.

5. REFERENCES

[1] R. Tripathy, M. Debahuti, V. B. Konkimalla, "A novel fuzzy C-meansapproach for uncovering cholesterol consensus motif from human G-protein coupled receptors (GPCR)", Karbala International Journal of Modern Science, Vol. 1, No. 4, 2015, pp. 212-224.

[2] R. D. DiMarchi, P. Mitra, "Gas regulates Glucagon-Like Peptide 1 Receptor-mediated cyclic AMP generation at Rab5 endosomal compartment", Molecular Metabolism, Vol. 6, No. 10, 2017.

[3] R. Tripathy, M. Debahuti, V. B. Konkimalla, "A computational approach for mining cholesterol and their potential target against GPCR seven helices based on spectral clustering and fuzzy c-means algorithms", Journal of Intelligent & Fuzzy Systems, Vol. 35, No. 1, 2018, pp. 305-314.

[4] O. Mouritsen, M. Zuckermann. "What's so special about cholesterol?", Lipids, Vol. 39, No.11, 2004, pp. 1101-1113.

[5] G. Karp, "The structure and function of the plasma membrane", Cell and Molecular Biology: Concepts and Experiments, 3rd edition, John Wiley and Sons, 2002, pp. 122-82.

[6] Y. Lange, B. V. Ramos, "Analysis of the distribution of cholesterol in the intact cell", Journal of Biological Chemistry, Vol. 258, No. 24, 1983, pp. 15130-15134.

[7] K. G. Burger, G. Gimpl, F. Fahrenholz, "Regulation of receptor function by cholesterol", Cellular and Molecular Life Sciences CMLS, Vol. 57, No. 11, 2000, pp. 1577-1592.

[8] D. Lingwood, K. Simons, "Lipid rafts as a membrane-organizing principle", Science, Vol. 327, No. 5961, 2010, pp. 46-50.

[9] Md, Jafurulla, S. Tiwari, A. Chattopadhya,. "Identification of cholesterol recognition amino acid consensus (CRAC) motif in G-protein coupled recep-

tors", Biochemical and biophysical research communications, Vol. 404, No. 1, 2011, pp. 569-573.

[10] H. Nakashima, Y. Kuroda, "Differences in dinucleotide frequencies of thermophilic genes encoding water soluble and membrane proteins", Journal of Zhejiang University SCIENCE B, Vol. 12, No.6, 2011, p. 419.

[11] T. Schöneberg et al. "Learning from the past: evolution of GPCR functions", Trends in pharmacological sciences, Vol. 28, No. 3, 2007, pp. 117-121.

[12] R. Fredriksson, H. B. Schiöth, "The repertoire of G-protein–coupled receptors in fully sequenced genomes", Molecular pharmacology, Vol. 67, No. 5, 2005, pp. 1414-1425.

[13] C. J. Baier, J. Fantini, F. J. Barrantes. "Disclosure of cholesterol recognition motifs in transmembrane domains of the human nicotinic acetylcholine receptor", Scientific reports, Vol. 1, 2011, p. 69.

[14] C. Ellis, "The state of GPCR research in 2004", Nature Reviews Drug Discovery, Vol. 3, No. 7, 2004, pp. 577-626.

[15] H. Li, V. Papadopoulos, "Peripheral-type benzodiazepine receptor function in cholesterol transport. Identification of a putative cholesterol recognition/interaction amino acid sequence and consensus pattern", Endocrinology, Vol. 139, No. 12, 1998, pp. 4991-4997.

[16] S. Schlyer, R. Horuk, "I want a new drug: G-protein-coupled receptors in drug development", Drug discovery today, Vol. 11, No. 11-12, 2006, pp. 481-493.

[17] T. J. Pucadyil, A. Chattopadhyay, "Role of cholesterol in the function and organization of G-protein coupled receptors", Progress in lipid research, Vol. 45, No. 4, 2006, pp. 295-333.

[18] X. Sun, G. R. Whittaker. "Role for influenza virus envelope cholesterol in virus entry and infection", Journal of virology, Vol. 77, No. 23, 2003, pp. 12543-12551.

[19] R. M. Epand et al. "Cholesterol interaction with proteins that partition into membrane domains: an overview", Cholesterol Binding and Cholesterol Transport Proteins, Springer, Dordrecht, 2010, pp. 253-278.

[20] A. K. Hamouda et al. "Cholesterol interacts with transmembrane α-helices M1, M3, and M4 of the Torpedo nicotinic acetylcholine receptor: photolabeling studies using [3H] azicholesterol", Biochemistry, Vol. 45, No. 3, 2006, pp. 976-986.

[21] C. Banchhor, N. Srinivasu, "FCNB: Fuzzy correlative naïve bayes classifier with MapReduce framework for big data classification", Journal of Intelligent Systems, Vol. 29, No. 1, 2018, pp. 994-1006.

[22] R. Tripathy, R. K. Nayak, P. Das, D. Mishra, "Cellular cholesterol prediction of mammalian ATP-binding cassette (ABC) proteins based on Fuzzy C-Means with support vector machine algorithms", Journal of Intelligent & Fuzzy Systems, 2020. (Preprint).

[23] C. Subbalakshmi, G. Ramakrishna, S. Rao, "Evaluation of data mining strategies using fuzzy clustering in dynamic environment", Proceedings of the 3rd International Conference on Advanced Computing, Networking and Informatics, New Delhi, 2016, pp. 529-536.

[24] S. C. Gopi, K. Kiran, M. Veerabrahmam, Y. Ayyappa, "Fuzzy Based Classification of X-Ray Images with Convolution Neural Network", International Journal of Emerging Trends in Engineering Research, 2020, pp.4433-4436.

[25] S. P. Potharaju, M. Sreedevi, "A novel subset feature selection framework for increasing the classification performance of SONAR targets", Paper presented at the Procedia Computer Science, Vol. 125, 2018, pp. 902-909.

[26] S. Razia, M. R. Narasingarao, "A neuro computing framework for thyroid disease diagnosis using machine learning techniques", Journal of Theoretical and Applied Information Technology, Vol. 95, No. 9, 2018, pp. 1996-2005.

[27] A. Ayushree; S. Kumar, "Comparative Analysis Of Aodv And Dsdv Using Machine Learning Approach In Manet", Journal Of Engineering Science And Technology, Vol. 12, No. 12, 2017, pp. 3315-3328.

[28] R. K. Nayak, R. Tripathy, V. Saravanan, P. Das, D. K. Anguraj, "A Novel Strategy for Prediction of Cellular Cholesterol Signature Motif from G Protein-Coupled Receptors based on Rough Set and FCM Algorithm", Proceedings of the Fourth Interna-

tional Conference on Computing Methodologies and Communication, Erode, India, 11-13 March 2020, pp. 285-289.

[29] R. K. Nayak, R. Tripathy, D. Mishra, D., V. K. Burugari, P. Selvaraj, A. Sethy, A., B. Jena, "Indian Stock Market Prediction Based on Rough Set and Support Vector Machine Approach", Intelligent and Cloud Computing, 2020, pp. 345-355.

[30] R. Tripathy, R. K. Nayak, V. Saravanan, D. Mishra, G. Parasa, K. Das, P. Das, "Spectral Clustering Based Fuzzy C-Means Algorithm for Prediction of Membrane Cholesterol from ATP-Binding Cassette Transporters" Intelligent and Cloud Computing, 2019, pp. 439-448.

[31] UniProt Consortium, "The universal protein resource (UniProt)", Nucleic acids research, Vol. 36, 2007, pp. D190-D195.

[32] E. Arias-Castro, Ery, C. Guangliang G. Lerman, "Spectral clustering based on local linear approximations", Electronic Journal of Statistics, Vol. 5, 2011, pp. 1537-1587.

[33] H. Zare et al. "Data reduction for spectral clustering to analyze high throughput flow cytometrydata", BMC bioinformatics, Vol. 11, No. 1, 2010, p. 403.

[34] R. K. Nayak, D. Mishra, K. Shaw, S. Mishra, "Rough set based attribute clustering for sample classification of gene expression data", Procedia engineering, Vol. 38, 2012, pp. 1788-1792.

[35] R. K. Nayak, D. Mishra, A. K. Rath, "A Naïve SVM-KNN based stock market trend reversal analysis for Indian benchmark indices", Applied Soft Computing, Vol. 35, 2015, pp. 670-680.

[36] C. Cortes, "WSupport-vector network", Machine learning, Vol. 20, 1995, pp. 1-25.

[37] X. Zhang, X. Zheng, "Comparison of text sentiment analysis based on machine learning", Proceedings of the 15th International Symposium on Parallel and Distributed Computing, Fuzhou, China, 8-10 July 2016.

[38] V. P. Upadhyay et al. "Forecasting stock market movements using various kernel functions in support vector machine", Proceedings of the International Conference on Advances in Information Communication Technology & Computing, Bikaner, India, 12-13 August 2016.

[39] R. K. Nayak, D. Mishra, A. K. Rath, "An optimized SVM-k-NN currency exchange forecasting model for Indian currency market", Neural Computing and Applications, Vol. 31, No. 7, 2019, pp. 2995-3021.

[40] N. Chandra, "Support Vector Machine Classifier for Predicting Drug Binding to P-glycoprotein", Journal of Proteomics & Bioinformatics, Vol. 2, No. 4, 2009.

Early Prediction of Employee Turnover Using Machine Learning Algorithms

Original Scientific Paper

Markus Atef
Faculty of Management Sciences,
October University for Modern Sciences and Arts (MSA), Giza, Egypt
Business Information Systems Department,
Faculty of Commerce and Business Administration,
Helwan University, Cairo, Egypt
markusdaoud96@outlook.com

Doaa S. Elzanfaly
Information Systems Department, Faculty of Computers and Artificial Intelligence,
Helwan University, Cairo, Egypt
Faculty of Informatics and Computer Science, British University in Egypt, Cairo, Egypt
Doaa.saad@fci.helwan.edu.eg

Shimaa Ouf
Business Information Systems Department, Faculty of Commerce and Business Administration,
Helwan University, Cairo, Egypt
shimaaouf@yahoo.com

Abstract *– Employee turnover is a serious challenge for organizations and companies. Thus, the prediction of employee turnover is a vital issue in all organizations and companies. The present work proposes prediction models for predicting the turnover intentions of workers during the recruitment process. The proposed models are based on k-nearest neighbors (KNN) and random forests (RF) machine learning algorithms. The models use the dataset of employee turnover created by IBM. The used dataset includes the most essential features, which are considered during the recruitment process of the employee and may lead to turnover. These features are salary, age, distance from home, marital status, and gender. The KNN-based model exhibited better performance in terms of accuracy, precision, F-score, specificity (SP), and false-positive rate (FPR) in comparison to the RF-based model. The models predict the average probability percentage of turnover intentions of the workers. Therefore, the models can be used to aid the human resource managers to make precautionary decisions; whether the candidate employee is likely to stay or leave the job, depending on the given relevant information about the candidate employee.*

Keywords*: Prediction Models, Employee Turnover, Machine Learning Algorithms*

1. INTRODUCTION

Employee turnover can be defined as the rate of employees who quit an organization and are substituted by new employees. A high employee turnover rate represents a potentially fatal problem for organizations due to the high costs of separation, vacancy, recruitment, training, and replacement. Moreover, an organization with a high turnover rate eventually becomes understaffed, consequently non-productive and its growth stagnates [1]. Thus, the prediction of employee turnover is a vital procedure for any sustainable organization, where acquiring early information regarding employee turnover status helps organizations to take precautions to such a status.

Artificial intelligence (AI) can be defined as the utilization of the machine instead of human brains to accomplish a required goal or carry out a certain task. Recently, there is an increasing tendency to apply AI in human resource (HR) management [2-3] because using the computational and processing powers of the machine is faster and more accurate than the human brain [4-6]. Nowadays, the AI field is an important aspect and a rapidly growing trend of the technology-driven economy. AI starts to run deeply on the organizational level, affecting some of its structures in some countries. The field of HR has steadily shown interest in the AI technology through baby-steps across the world. An AI-driven HR management will put back the routine jobs and complicated tasks of the human resource personnel;

thus, conserving substantial amounts of time, money, and manpower [7-9].

The literature provides a set of recommendations on how the AI tools and practices could be applied for specific HR management tasks like using machine learning techniques in employee selection [10] and recruitment by information extraction techniques [11, 12].

One of the main branches of artificial intelligence (AI) is machine learning; a scientific development of computer machines, where the machine can learn and adapt based on provided data and experience, without necessarily following programmed instructions. The learning process starts with data analysis of recurring patterns in a dataset. Then, the observed patterns are used to extrapolate a decision or predict an outcome. Various machine learning algorithms can be utilized for turnover prediction such as neural networks, apriori, KNN, extreme gradient boosting, RF, decision tree, logistic regression, support vector machines, etc. [13-17].

Accordingly, the aim of the present research is to construct data-driven prediction models based on machine learning algorithms to predict the likelihood of a candidate quitting his/her job in the future. This would help the organization managers with taking the necessary precautions to diminish the turnover rate. The importance of the present research from a practical standpoint is that as mentioned before, employee turnover is a huge problem, which causes quite a lot of drawbacks. The present work could be a step in diminishing the drawbacks of employee turnover by predicting whether employees will leave the organizations or not before recruitment to help the managers make decisions to diminish the turnover rate. The importance of the present research from an academic point of view is that to the best of our knowledge, there are some studies in the open literature, dealing with the turnover prediction by using machine learning algorithms. But, no academic research attempts have been conducted to tackle the issue at hand, from the same angle, by predicting employee turnover during the recruitment process. The present work's methodology of data-based predictions uniquely stands out amongst several open-literature studies addressing turnover concerns via machine learning. To achieve this goal, the dataset created by IBM was used in the present work. The data was carefully studied and selected to include only the essential features, which should be considered during the recruitment process of the employee and may lead to turnover. The selected features are salary, age, distance from home, marital status, and gender. Models have been built based on KNN and RF algorithms to predict the probability percentage of turnover intention of candidates before recruitment.

2. RELATED WORK

A comparative study involving accuracy and memory utilization of selected algorithms for predicting employee turnover was conducted by Rohit Punnoose et al. [18]. The authors collected the data from the information system used by the human resource department of a retailer with global operations and data from the Bureau of Labor Statistics. Several classification algorithms were applied, namely, extreme gradient boosting (XGBoost), logistic regression, Naïve Bayesian, RF, linear support vector machine, linear discriminant analysis and KNN. The authors have found that XGBoost exhibits the best performance regarding the accuracy and memory utilization.

Jain et al. [19] have carried out research to predict turnover rate using XGBoost. They have found that age, gender, marital status, years at the company, job satisfaction, and distance from home have the most significant effects on turnover among all attributes in the dataset.

Numerous algorithms; namely, logistic regression, gradient boosting classifier, support vector machine, and RF were applied to the IBM dataset prepared by IBM data scientists [20, 21]. After applying the RF classifier, fifteen features were found to be more significant in deciding whether employees quit their jobs or not. Out of the fifteen features, overtime and monthly income exhibit the highest influence on employees to leave their jobs or not. XGBoost was found to have the highest performance (with an AUC of 0.84596) among all the applied algorithms.

Zhao et al. [13] have evaluated the performance of ten supervised machine learning algorithms; namely, RF, gradient boosting trees, XGBoost, support vector machines, decision tree, neural networks, linear discriminant analysis, Naïve Bayesian, logistic regression, support vector machines and KNN on numerous HR datasets. The authors have found that XGBoost is the most reliable algorithm among all the applied algorithms.

Zhang et al. [22] have attempted to find out the most important factors that lead to employee turnover. The authors have found an essential correlation between department and work. Also, they have found that the gender of employees significantly affects turnover. A logistic regression algorithm was applied for predicting the turnover with an accuracy of 87.2%.

Sisodia et al. [23] have carried out an investigation to find out the reasons causing the employee turnover by building models using machine learning algorithms to forecast employee turnover. They used the dataset on Kaggle with ten features. They have found that the main reasons causing high employee turnover rates are time spent with the company, workload, and promotion. The used machine learning algorithms for building the models were decision tree, support vector machine, Naïve Bayesian, KNN, and RF. The accuracy, precision, F-score, recall, specificity, and FPR of the models were compared. In terms of accuracy, F-score, and precision, RF performed better and in terms of recall, the decision tree was better.

3. RESEARCH METHODOLOGY

3.1 RESEARCH FRAMEWORK

The turnover prediction framework is presented in Fig.1. The used methodology comprises several phases; namely, data collection, data cleaning, data selection, data preprocessing, benchmarking the algorithms, and evaluating the predicted outcome.

3.2 DATASET

The dataset was obtained from the Kaggle website (IBM, 2020). The IBM dataset comprises 1470 records with 34 features (6 categorical and 27 numeric), such as monthly salary, experience, distance from home, skills, nature of work, position etc.

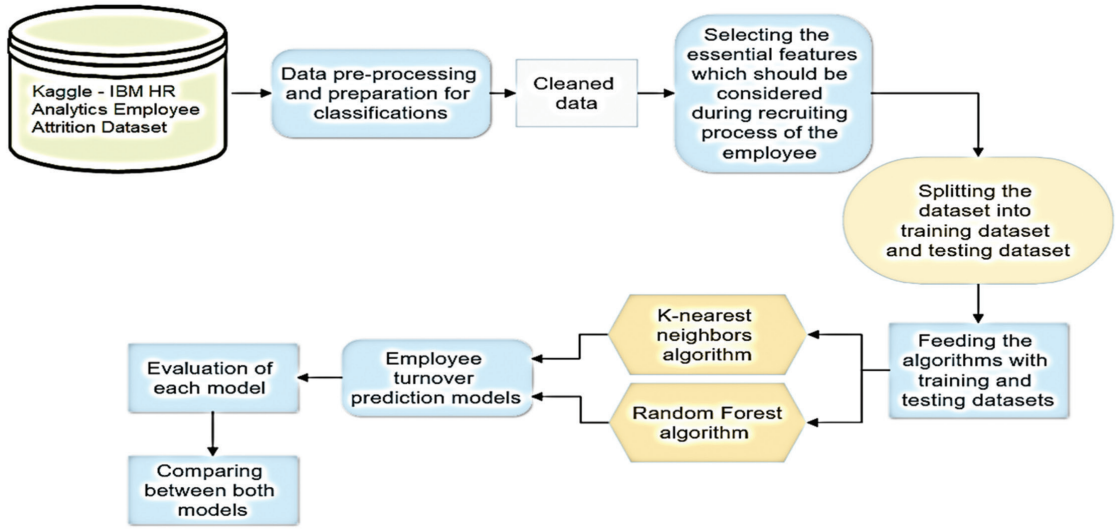


Fig. 1. Turnover prediction framework

3.3 DATA PREPROCESSING

Machine Learning algorithms can typically process only numerical input. Hence, the qualitative variables (gender and marital status) were encoded into quantitative variables (One-Hot Encoding) to input an acceptable format for the machine.

Based on the scientific literature [24-26] and logical reasoning, the features in the employee dataset that are essential for the prediction of turnover before recruitment were cherry-picked. This means that the features were selected based on the information provided by the candidates for the jobs and before they are put to work. For example, features such as job involvement, job satisfaction, job level, overtime, relationship satisfaction, and so on cannot be considered during the recruitment process because the candidates for the jobs have not yet been put to work. To fulfill the aim of the present work, all other features were excluded and solely considered salary, distance from home, marital status, age, and gender, which should be considered during the recruiting process of the employee and may lead to turnover in organizations.

3.4 APPLIED MACHINE LEARNING ALGORITHMS

The two machine learning algorithms applied in the present work are RF and KNN algorithms. The two algorithms were chosen because most of the open literature vouch for their reliability and accuracy [13,15, 27-28]. For example, Badillo et al. [27] have reported that RF and KNN algorithms demonstrate the best combination of performance and interpretability. Zhao et al. [13] have proved that KNN is accurate and reliable with a small number of features. The novelty of the present work lies within the hybridization technique used to tailor the parameters to better represent the model features in a realistic manner. Moreover, the trial-and-error process in the KNN algorithm was carried out on the cross-K validation rather than the algorithm itself to find the optimal K value for the given dataset as opposed to finding the optimal K for a particular trial and error process. In the case of the RF algorithm, the Random Search algorithm was utilized to narrow down the range of each parameter, then the Grid Search algorithm was applied on the relevant set of parameters only, introducing more novelty to the work.

4. RESULTS AND DISCUSSIONS

4.1 DESCRIPTIVE ANALYSIS

The descriptive analyses were performed on the IBM employee turnover dataset, including the selected features that significantly affect the turnover; namely, age, distance from home, marital status, gender, and monthly income. The HR employee turnover dataset was loaded into MySQL workbench, which is a database manage-

ment system, and MS Excel and then preprocessed. The gender was converted to 0 and 1 for male and female, respectively. The marital status was converted to 0, 1, and 2 for single, married, and divorced, respectively. The distance unit was converted from a mile to km. Then, the relationships between the employee turnover and the selected features were generated by using MySQL and MS excel. The distributions of the target variable (employee turnover) related to the selected features; namely, age, distance from home, marital status, gender, and monthly income within the dataset are presented in Figs. 2-6. In the IBM dataset, the total number of employees is 1470, from which 237 employees (16%) left the job. Fig. 2 presents the relationship between age and percentage of employee turnover. It can be observed that as the age increases the total turnover percentage or the turnover percentage in the cluster almost linearly decreases. The highest turnover percentage lies within the cluster of 18-24 years (43.7%) meanwhile, the lowest turnover percentage lies within the cluster of 43-48 years (9.1%). Concerning the percentage of total turnover, the highest turnover percentage lies within the cluster of 31-36 years (29.1%) and the lowest turnover percentage lies within the cluster of 55-60 years (4.6%). These findings indicate that younger employees have a more propensity to leave the job.

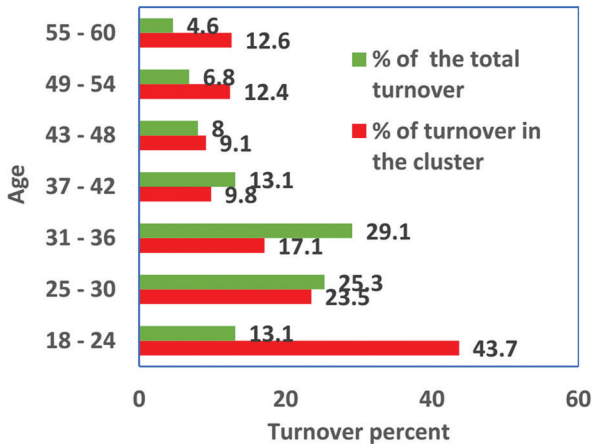


Fig. 2. Age versus percentage of the total turnover and percentage of turnover in the cluster

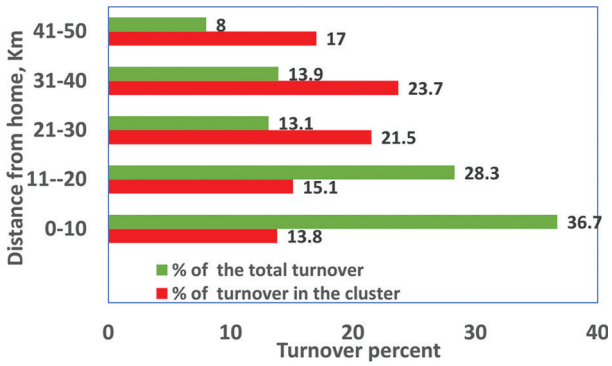


Fig. 3. Influence of distance from home on the percentage of the total turnover and percentage of turnover in the cluster

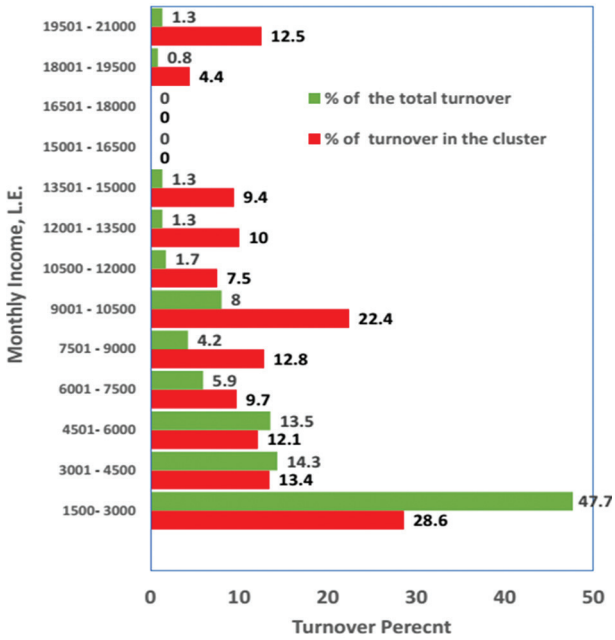


Fig. 4. Influence of monthly income on percentage of the total turnover and percentage of turnover in the cluster

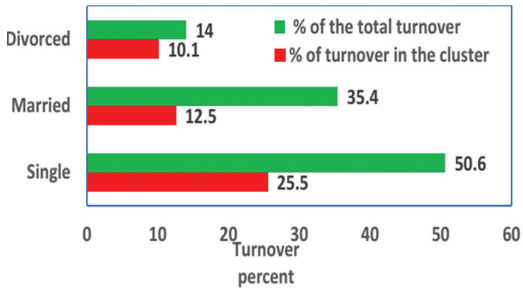


Fig. 5. Martial status versus percentage of the total turnover and percentage of turnover in the cluster

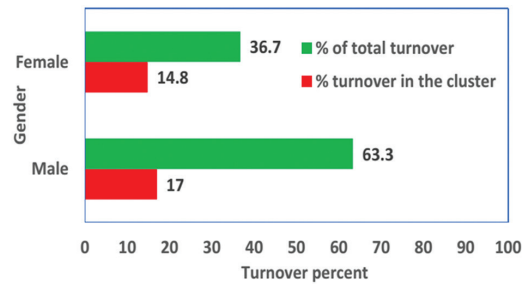


Fig. 6. Gender versus percentage of the total turnover and percentage of turnover in the cluster

Fig. 3 reveals the influence of distance from home on the percentage of employee turnover. The highest turnover percentage lies within the clusters of 21 – 30 and 31 – 40 km. Unexpectedly, the highest percentage (36.7%) of the total turnover lies within the cluster of 0-10 km and the lowest turnover percentage lies within the cluster of 0-10 km. The turnover percentage of employees as a function of salary is presented in Fig.4. The percentage of employee turnover is inversely proportional to the salary with the highest value of the lowest

salary cluster (1500-3000 L.E). The employees with salaries between 15000 and 18000 L.E. do not likely show a tendency to leave the job. Fig.5 shows the distribution of turnover percentage related to marital status. Single employees have a higher desire to leave, but divorced employees are more likely to stay in the job as indicated by the highest turnover percentage occurred with the single employees, whereas the lowest turnover percentage with the divorced employees. Fig. 6 reveals the relation between turnover and gender. It is worth noting that female employees show a higher tendency to stay in the job. However, male employees have a strong propensity to leave as they represent 63.3% of the total employee turnover in the dataset.

4.2 FEATURE IMPORTANCE

After analyzing the dataset, the RF algorithm was applied to find the feature importance score, Fig.7.

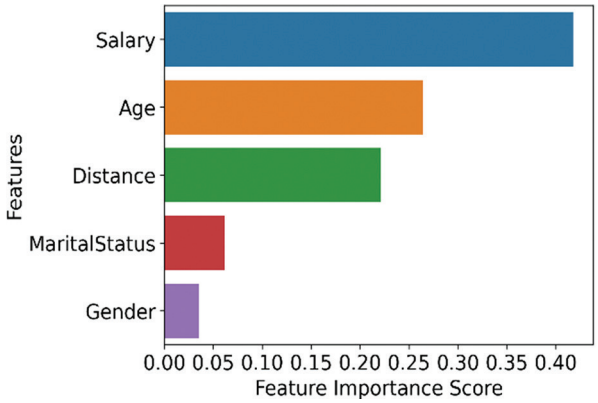


Fig. 7. Feature importance

From Fig. 7, it can be observed that salary exhibits the highest feature affecting the employee turnover and gender has the lowest effect.

4.3 HEATMAP

The heatmap uses a warm-to-cool color spectrum to show the correlations between variables. For obtaining the heatmap, pyplot, pandas, and seaborn libraries were imported and employed. The heat map, demonstrating the correlations between features and target variable (turnover), is presented in Fig. 8. A negative correlation indicates that the alteration of a feature or the target variable is inversely influenced by the other. However, a positive correlation indicates that the variation of a feature or the target variable is directly influenced by the other. The heatmap displays that there is no correlation higher than 0.5, which is between age and salary. The positive correlation between age and salary indicates that as the age increases, the salary increases. The correlation between distance from home and turnover is also positive, indicating that as the distance from home increases, the turnover increases. Moreover, the heatmap reveals a negative correlation between age and turnover, as well as salary and turnover, signifying that as age and salary decrease, the turnover increases (employees are more likely to leave the job).

Fig. 8. Correlation heatmap

4.4 BUILDING OF THE MODEL

4.4.1 KNN-based model

The idea of KNN is to determine the K data points in the training data that are closest to the new instance and categorizes this new instance by a majority vote of its K neighbors. The KNN algorithm is usually denoted as K- Nearest Neighbor classification because it takes more than one neighbor into account. The dataset present in the CSV data was loaded into a DataFrame, which was split into the feature's matrix and the target values vector. Then, the categorical features were one- hot-encoded (OHE) by mapping each categorical feature to a vector basis in the n-space, where n is the cardinality of the feature set. OHE is essential for the machine to interpret the categorical data.

Sklearn implemented the model selection class, the K- Nearest neighbor classifier class, and the score matrix class. The model selection class was used for splitting the dataset into two sets; the training set and the testing set.

Moreover, the Sklearn library defined cross K-validation and grid search classes, which were used for K-folding and hyperparameter tuning. The K -Nearest neighbor classifier class carried out the KNN algorithm by using the training and testing sets. The score metrics class generated the algorithm's classification report and calculated the confusion matrix with the accuracy score per model training instance. The data was randomly split into two sets, X and Y, where X represented the feature matrix and Y was the target variable (turnover); the size of the training set was α, and the testing set was 1- α, where α was kept at 0.84. It

should be mentioned that the target variable (turnover) is a binary representation of No (84%) and Yes (16%) outcomes. Thus, the dataset was split and kept at 84% and 16% for training and test datasets, respectively. Random splitting percentages were avoided because it can change the percentages of the classes present in the training and test datasets from that in the original to conform to the original distribution. Each set was further split into two sets where X was split into X-train, X-test, and Y was split into Y-train, Y-test. Before running the algorithm, the sets must be normalized to minimize the outlier effect on the dataset. In other words, feature scaling was employed to eliminate biases towards the outlier feature values. The feature scaler used for standardization on this data was the z-score method. It should be here mentioned that data normalization is one of the crucial issues in ML because unnormalized data allows one feature to entirely dominate the other features. The most common techniques for normalization are Min-max and Z-score. Min-max normalization ensures that all features will have an equal influence on that data despite the outliers. The Z-score technique handles outliers but does not generate normalized data with the same scale. Thus, in the present work, Z-score normalization was selected to be applied to the dataset.

The grid search algorithm was used for obtaining the optimal K value by performing the cross K-validation algorithm on the original standardized dataset. Finally, the model based on the KNN algorithm ran on the training and testing sets with the chosen K value, generating the classification report, confusion matrix, and accuracy score. The model based on the KNN algorithm ran several times to obtain the mean accuracy and graphs, which were generated based on the classification report, correspondingly.

Cross-validation is considered the most used technique to evade overfitting, diminish the bias of sampling data and guarantee model error randomness; thus, in the present work, ten-fold cross-validation was applied. Cross-validation randomly divided the dataset into ten-fold subsets, where each subset was used as a training dataset once and as a testing dataset nine times. This process was repeated iteratively ten times to ensure that each subdivision is used as the training set once. In each repetition, a different part was selected as the training set. Finally, the average prediction error was obtained by measuring the mean accuracy of the ten iterations performed on each validation set.

The Cross-validation result performed above was used for the grid search algorithm's parameters to determine the optimal K value; this process is known as hyperparameter tuning. The grid search algorithm exhaustively searched for the hyperparameters on all the datasets using cross-validation as a metric. Unlike the mean error method used to obtain the K value as conducted above, grid search guarantees a distinct K value relative to the cv metric. For instance, a cv value of 5 consistently provided an optimal K value of 13, meanwhile, a cv value of 10 provided an optimal K value of 12. As expected, the cross-validation results for each subset revealed low variance in its accuracy score, compared with the other subsets' accuracy scores, signifying a low bias of sampling data; hence, the effect of overfitting was diminished. It should be mentioned that the algorithm runs iteratively using the features (salary, age, distance from home, gender, and marital status) of the candidate employee before recruitment to predict the turnover intentions of the candidate by obtaining the average probability percentage of whether the candidate employee is likely to stay or leave the job, depending on the given features of the candidate employee. It is crucial to run the algorithm several times because a single run is susceptible to false positives or false negatives, which should be mitigated. An independent optimal K value was calculated for each run on the randomly split datasets. However, in the case of the full dataset being used as a training set for the prediction during the recruitment process, the optimal K is calculated only once.

4.4.2 RF algorithm-based model

The basic methodology of RF algorithm is to merge numerous algorithms, which is known as ensemble learning, to resolve a complex problem and enhance the model's performance. RF is based on tree algorithms, where it generates a set of decision trees from a subset of training data and collectively takes the prediction of all the trees instead of one decision tree. Then, a vote is conducted from each tree to conclude a prediction based on the most voted for the outcome. The final output is the prediction, which has the most votes. RF utilizes so-called "Bagging", which means that the consecutive forthcoming trees do not depend on the preceding trees. The accuracy and performance of a model based on RF could be enhanced by increasing the number of trees. This helps subside the effect of data overfitting. RF is a robust algorithm because each node is split depending on some prediction variables and probability functions, where the best subset of trees is split. Like the model based on the KNN algorithm, the CSV data was loaded into a DataFrame, which was split into the feature's matrix and the target values vector. Then, One-Hot-Encoding (OHE) was applied to the categorical features, where each feature was mapped to a vector basis in the n-space. In order to avoid feature bias, where a feature value prominently influences the data, feature scaling (Z-score normalization) was applied to the dataset.

The model selection class, the RF classifier class, and the score metrics class are implemented by Sklearn. The dataset was split into the training and testing sets by the model selection class. The dataset was split into the X-set with a size of α% and the Y-set with a size of 100% - α%, representing the feature matrix and the target variable (turnover), respectively.

The grid search algorithm is typically used to obtain the optimal hyperparameters, which yield the best performance for a model. Alternatively, a random search algorithm would find parameters yielding accurate results on average at risk of occasional non-optimal parameters. Hence, the random search algorithm has a high variance. Improving the random search algorithm results would require several iterations at least proportional to the set's cardinality, which could be inefficient on a set of a large size. On the other hand, the grid search requires a certain range of parameters to exhaustively search for the parameters yielding the most accurate results. Hypothetically, the range of parameters provided to the grid search algorithm could be impractically large. Taking advantage of both algorithms, multiple iterations of the random search algorithm on the dataset were used to prune the parameter ranges later provided to the grid search algorithm. Thus, eliminating the random search algorithm's high variance while maintaining efficiency for the grid search algorithm.

The cross-validation class and grid search class for K-folding and hyperparameter tuning were then applied to the dataset. The K-folding further divided each X- set and Y-set into K training and K testing sets, respectively. Then, the RF classifier was fit using the training and testing sets. Finally, the classification report, the confusion matrix, and the accuracy score were generated per model training instance using the score metrics class. Unlike the model based on the KNN algorithm, the grid search for hyperparameters of the RF algorithm is less methodical since the RF parameters should be tailored for a given dataset to obtain optimal results. Thus, a process of trial and error was necessary to narrow down the grid search range for each parameter. It is important to maintain a balance between the algorithm's efficiency and precision by keeping each parameter range concise. It follows that running the grid search for each RF iteration is inefficient since it increases the algorithm's runtime exponentially. In other words, the grid search of the RF algorithm was applied only once. In order to achieve an accurate prediction without sample bias, the RF algorithm should be applied multiple times with random data splits, where the prediction is the mean prediction probability of all iterations.

4.5 EVALUATION OF THE ALGORITHMS

For evaluating the performance of the present trained models, including accuracy, recall, precision, the area under the curve (AUC) of the receiver operating characteristics curve (ROC) and F-score, 'metrics' module from 'Sklearn' was used. Accuracies of the algorithms by ten-fold cross-validation are given in Table 1 and results of the KNN and RF-based models are presented in Table 2. The AUC-ROC curves are presented in Fig. 9. The AUC under the ROC is used to compare the accuracies of the models, where a higher AUC indicates better performance of the model.

Table 1. Prediction accuracies of KNN and RF-based models by 10-fold cross-validation

Accuracy	KNN	RF
Fold1	0.823	0.728
Fold2	0.857	0.782
Fold3	0.81	0.803
Fold4	0.844	0.776
Fold5	0.816	0.789
Fold6	0.823	0.789
Fold7	0.837	0.776
Fold8	0.837	0.789
Fold9	0.837	0.796
Fold10	0.864	0.81
Average	0.835	0.784

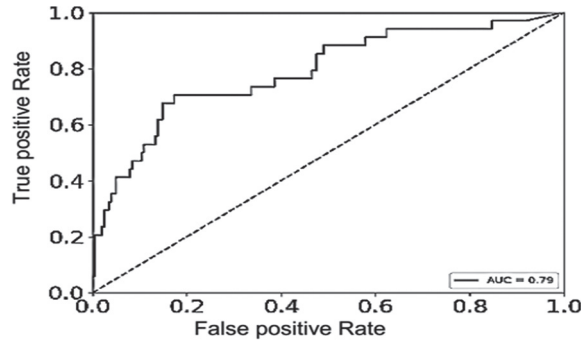
Table 2. Results of KNN (a) and RF(b) based models

Average Confusion matrix values: [[193 5] [33 5]]
Average accuracy score : 0.84
Average precision score: [No :0.853, Yes: 0.481]
Average recall score: [No: 0.973, Yes: 0.126]
Average F1 score: [No: 0.909, Yes: 0.196]
Average support score: [No: 198, Yes: 38]
(a)
Average Confusion matrix values: [[179 19] [29 9]]
Average accuracy score: 0.8
Average precision score: [No: 0.861, Yes: 0.333]
Average recall score: [No: 0.903, Yes: 0.249]
Average F1 score: [No:0.882, Yes: 0.281]
Average support score: [No: 198, Yes: 38]
(b)

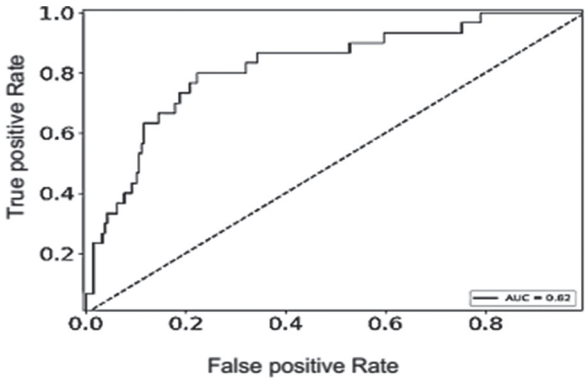
The present study shows that the KNN-based model exhibits better prediction performance in terms of precision, accuracy, F-score, FPR, and SP in comparison to the RF-based model, Table 2. Table 3 presents a comparison between KNN- and RF-based models developed in the present work and those reported in the literature using the same dataset. It should be mentioned that the number of selected features (5) is low, which could have a negative effect on the performance of the model. Despite that, the developed models in the present work show comparable performance in terms of accuracy or better performance in terms of AUC and F-score when compared with those reported in the literature using a much higher number of features. This can be attributed to the fact that the methodology present in the current research is unprecedented due to the hybridization technique used to tailor the parameters to better represent the model features in

a realistic manner. Typically, the KNN algorithm's optimal parameters are obtained through the iteration yielding the least margin of error from multiple trial and error processes. In this paper's KNN algorithm, the trial-and-error process was carried out on the cross-K validation rather than the algorithm itself. This technique consistently finds the optimal K value for a given dataset as opposed to finding the optimal K for a trial and error process Similarly, the RF algorithm conceptually requires a wider span of trial-and-error processes for each parameter. Unlike the typical RF method, this research paper employs the Random Search algorithm to narrow down the range of each parameter, then applies the Grid Search algorithm on the relevant set of parameters only. The conclusion inferred via this methodology is supported by the empirical observations from various experiments on the parameters, especially for the RF, given the specific set of features fulfilling the objective of this research.

Finally, the constructed model based on the KNN algorithm developed in the present work to predict the turnover probability of employees before recruitment could be a useful decision support tool to help the HR managers of the organizations during the recruitment process. This is because the constructed model based on the KNN algorithm reveals high performance to be implemented for a real-life business.



(a)



(b)

Fig. 9 AUC-ROC of KNN algorithm (a) and RF algorithm (b) based models

Table 3. Comparison between the KNN and the RF-based model in the present work and those reported in the literature

Algorithm	Accuracy	AUC	Precision	Recall	F-score	FPR	SP	Ref.
KNN	0.84	0.79	0.47	0.12	0.187	0.025	0.974	present work
RF	0.80	0.82	0.333	0.249	0.281	0.095	0.904	present work
KNN	0.867	-	0.38	0.23	-	-	-	[29]
RF	0.879	-	0.45	0.22	-	-	-	[29]
KNN	0.852	-	0.551	0.09	0.15	-	0.994	[30]
RF	0.861	-	0.658	0.132	0.194	-	0.991	[30]
KNN	0.832	0.52	0.070	0.384	0.119	-	-	[14]
RF	0.85	0.58	0.183	0.619	0.282	-	-	[14]

5. Conclusions

There is a positive correlation between distance from home and employee turnover, signifying that as the distance from home increases, the employee turnover increases. However, there is a negative correlation between age and employee turnover, indicating that as age decreases, employees are more likely to leave the job. The percentage of employee turnover is inversely proportional to the salary. Single employees show a higher desire to leave, but divorced employees are more likely to stay in the job, where the highest turnover percentage occurs in the single employees and the lowest turnover percentage occurs in the divorced employees. Female employees have more tendency to stay in the job. However, male employees have a strong propensity to leave. Salary exhibits the highest feature affecting the turnover of employees and gender has the lowest effect.

The KNN-based model exhibits better prediction performance in terms of accuracy, precision, F-score, FPR, and SP in comparison to the RF-based model.

A data-driven prediction model of the turnover probability of employees before recruitment is constructed to predict the probability percentage of the likelihood of an employee quitting or staying. The constructed model could be a useful decision support tool to help the HR managers during the recruitment process.

6. References

[1] J. Berengueres, G. Duran, D. Castro, "Happiness, an Inside Job? Turnover Prediction Using Employee Likeability, Engagement and Relative Happiness", Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Sydney, Australia, 31 July - 3 August 2017, pp. 509–516.

[2] S. Pandey, P. Khaskel, "Application of AI in Human Resource Management and Gen Y"s Reaction", International Journal of Recent Technology and Engineering, Vol. 8, 2019. pp. 10325-10331.

[3] G. Ginu, M. Thomas. "Integration of Artificial Intelligence in Human Resource", International Journal of Innovative Technology and Exploring Engineering, Vol. 9, 2019, pp. 5069-5073.

[4] S. Kar, S. Srihari, "Impact of Digital HR Practices on Strategic Human Capital Management and Organizational Performance in IT Sector", International Journal of Advance Research in Computer Science and Management Studies, Vol. 6, No. 12, 2018, pp. 17-25.

[5] V. Yawalkar, "A Study of Artificial Intelligence and Its Role in Human Resource Management", International Journal of Research and Analytical Reviews, Vol. 6, 2019, pp. 20-24.

[6] P. Merlin, R. Jayam, "Artificial Intelligence in Human Resource Management", International Journal of Pure and Applied Mathematics, Vol.119, No 14, 2018, pp. 1891-1895.

[7] R. Jesuthasan, "HR's New Role: Rethinking and Enabling Digital Engagement", Strategic HR Review, Vol. 16, No. 2, 2017, pp. 60-65.

[8] R. Rathi, "Artificial Intelligence and the Future of HR Practices", International Journal of Applied Research, Vol. 4, No. 6, 2018, pp. 113-116.

[9] Q. Jia, Y. Guo, R. Li, Y. R. Li, Y. W. Chen, "A Conceptual Artificial Intelligence Application Framework in Human Resource Management", Proceedings of the 18th International Conference on Electronic Business, Guilin, China, 2-6 December 2018, pp.106-114.

[10] C. Chen-Fu, C. Li-Fei, "Data Mining to Improve Personnel Selection and Enhance Human Capital: A Case Study in High-technology Industry", Expert Systems with Applications, Vol. 34, 2008, pp.280-290.

[11] T. Kaczmarek, M. Kowalkiewicz, J. Pikorski, "Information Extraction from CV", Proceedings of the 8th International Conference on Business Information Systems, 20-22 April 2005, Poznan, Poland. pp. 185-189.

[12] H. Zeng, "Adaptability of Artificial Intelligence in Human Resources Management in this Era", International Journal of Science, Vol. 7, 2020, pp. 271-276.

[13] Y. Zhao, M. Hryniewicki, F. Cheng, B. Fu, X. Zhu, "Employee Turnover Prediction with Machine Learning: A Reliable Approach", Proceedings of the Intelligent System Conference, London, UK, 1 September 2018, pp. 737-758.

[14] H. Sri, A. Varaprasad, L. V. N. Sujith, "Early Prediction of Employee Attrition", International Journal of Scientific & Technology Research", Vol. 9, No. 3, 2020, pp. 3374-3397.

[15] P. M. Usha, "An Analysis of the Use of Machine Learning for Employee Attrition Prediction – A Literature Review", Journal of Information and Computational Science, Vol. 10, 2020, pp.1429-1438.

[16] D. Shawni, S. Bandyopadhyay, "Employee Attrition Prediction Using Neural Network Cross Validation Method", International Journal of Commerce and Management Research, Vol. 6, 2016, pp. 80-85.

[17] T. Aniket, D. Motwani, "Employee Churn Rate Prediction and Performance Using Machine Learning", International Journal of Recent Technology and Engineering, Vol. 8, 2019, pp. 824-826.

[18] R. Punnoose, P. Ajit, "Prediction of Employee Turnover in Organizations Using Machine Learning Algorithm", International Journal of Advanced Research in Artificial Intelligence, Vol. 5, No. 9, 2016, pp. 22-26.

[19] R. Jain, A. Nayyar, "Predicting Employee Attrition using XGBoost Machine Learning Approach", Proceedings of the International Conference on System Modeling & Advancement in Research Trends, 23-24 November 2018, pp. 113-120.

[20] J. Sukhadiya, H. Kapadia, M. D'silva, "Employee Attrition Prediction Using Data Mining Techniques", International Journal of Management, Technology and Engineering, 2018, pp. 2249-7455.

[21] S. Ponnuru, G. Merugumala, S. Padigala, R.Vanga, B. Kantapalli,"Employee Attrition Prediction Using Logistic Regression", International Journal for Research in Applied Science and Engineering Technology, Vol.8, 2020, pp. 2871-2875.

[22] H. Zhang, L. Xu, X. Cheng, K. Chao, X. Zhao, "Analysis and Prediction of Employee Turnover Characteristics Based on Machine Learning", Proceedings of the 18th International Symposium on Communications and Information Technologies, Bangkok, Thailand, September 2018, pp. 371-376.

[23] D. Sisodia, S. Vishwakarma, A. Pujahari, "Evaluation of Machine Learning Models for Employee Churn Prediction", Proceedings of International Conference on Inventive Computing and Informatics, Coimbatore, India, 23-24 November 2017, pp. 1016-1020.

[24] C. A. Al Mamun, M. Hasan, "Factors Affecting Employee Turnover and Sound Retention Strategies in Business Organization: A Conceptual View" Problems and Perspectives in Management, Vol. 15, No. 1, 2017, pp.63-71.

[25] D. Verma, R. Chaurasia, "A Study to Identify the Factors Affecting Employee Turnover in Small Scale Industries", International Journal of Engineering Sciences & Research Technology, 2016, pp. 639-652.

[26] N. Govindaraju, "Demographic Factors Influence on Employee Retention", International Journal of Engineering Studied and Technical Approach, Vol. 7, 2018, pp.10-20.

[27] S. Badillo, B. Banfai, F. Birzele, I. Davydov, L. Hutchinson, T. Kam-Thong, J. Siebourg-Polster, B. Steiert, J. Zhang, "An Introduction to Machine Learning", Clinical pharmacology & therapeutics, Vol. 107, No 4, 2020, pp. 871-885.

[28] G. Xiang, J. Wen, C. Zhang, "An Improved Random Forest Algorithm for Predicting Employee Turnover", Mathematical Problems in Engineering, Vol 2019.

[29] I. Onuralp, H. Shourabizadeh, "An Approach for Predicting Employee Churn by Using Data Mining", Proceedings of the International Artificial Intelligence and Data Processing Symposium, Malatya, Turkey, 16-17 September 2017, pp. 1-4.

[30] F. Francesca, M. Coladangelo, R. Giuliano, E. Luca, "Predicting Employee Attrition Using Machine Learning Techniques", Computers, Vol. 9, No. 4, 2020.

Secured SDN Based Blockchain: An Architecture to Improve the Security of VANET

Original Scientific Paper

Swapna Choudhary
G H Raisoni College of Engineering,
Research Scholar, Department of Electronics Engineering, Nagpur, India
swapna.choudhari@raisoni.net

Sanjay Dorle
G H Raisoni College of Engineering,
Faculty of Electronics Engineering, Department of Electronic Engineering, Nagpur, India
sanjay.dorle@raisoni.net

Abstract *– Vehicular Ad-hoc networks (VANETs) during the communication process, nodes are always varying and the process is always under security threats like Sybil attacks, masquerading attacks, etc. In order to reduce the probability of these attacks and to regulate traffic flow in the network, a software-defined network (SDN) is used. The SDN is used for implementing protocols like OpenFlow and reducing the routing load in the network, but it doesn't provide a high level of security to the network, hence protocols like encryption, hashing, etc. are applied to the VANET. In the paper, SDN based blockchain-inspired algorithm is implemented, which coordinates network traffic and improves the overall security of the network. Security analysis of the proposed algorithm shows that the combination of blockchain with encrypted SDN is removing more than 95% of the network attacks as compared to its non-blockchain counterparts.*

Keywords*: Vehicular Ad-hoc Network, Software-Defined Networks, Encryption algorithm, network security, blockchain*

1. INTRODUCTION

Security has always been a major research issue with wireless networks. This is due to the fact that packets transmitted between wireless nodes are intercepted by adversaries, and a wide variety of malicious operations are performed on them. The malicious packets are then re-communicated via the network for affecting other nodes, thereby reducing the network's optimum per formance capability.

The design of Vehicular ad-hoc networks [1] requires that a vehicular node must do the following operations:

a. Register on the network once it comes in the range of either a network vehicle node or a network infrastructure node (hub)

b. Perform communication either directly in a peer-to-peer manner or using infrastructure hub as a hopping node

c. Periodically broadcast information regarding events that are sensed by the vehicle

d. Inform the neighboring nodes and the infrastructure once the node is leaving the ad-hoc network

e. Send heartbeat packets to the neighboring nodes and the infrastructure hub regarding the current parameters of the node (energy levels, location, etc.)

Based on these operations, each vehicular node communicates with other nodes in an effective manner. In some cases, attacker nodes without register in the network try to interact with healthy nodes. In such cases, the attacker sends the critical information outside the network or changes the information and tried to insert malicious packets in the network which decreases the efficiency of networks.

In order to protect the networks from various attacks, different control mechanisms like bandwidth & security are applied. Protocols like Open flow and SFLOW of SDN can control the traffic of the network and reduce attacks. SDN protocols are not effective for the attacks like Sybil and masquerading. To protect the networks from these attacks which change the identity of nodes' secure hashing, Public Key Infrastructure (PSK) like algorithm can be used. These algorithms can improve the probability of attacks but algorithms like blockchain peer-to-peer communication are more efficient [2]. In the next section analysis of different architecture and algorithms are given and compared with proposed

encrypted blockchain open flow SDN architecture (ABOSS).

2. RELATED WORK

Blockchains are applied for securing VANETs extensively. The self-organized secure (SOS) framework [3] is based on a peer-to-peer network. The main advantage of this network is to secure the network even if there is no roadside infrastructure exist in the network. Shamir sharing combined with a trust-based routing scheme to achieve this objective. Due to the combination of these techniques Vehicles Authority, Message Integrity, Privacy, Non-Repudiation, Traceability, Anonymity, and Availability are improved. Moreover, attacks like Impersonation, Modification, ID Disclosure, Location Tracking, Repudiation, and denial of service (DoS) are removed. Similar work with improved cryptographic primitive attribute-based encryption (CP-ABE) is defined in [4]. In this work, due to attribute-based encryption, the overall network speed is improved along with a reduction of the attack probability in the network. A verifiable hidden policy CP-ABE with a decryption testing scheme is also proposed in [4], it allows nodes to be tested and authenticated before performing network communication. This work is improved by adding true blockchain solutions, which is described in [5], wherein blockchain is used for privacy preservation along with reduction of the computational complexity of the system. This work is implemented for the Internet of Things (IoT), but it can be extended to VANETs by replacing IoT-specific blocks like IoT platform providers with RSU and cloud services with VANET infrastructure services. This work can be further extended with the help of Shor's algorithm as described in [6], wherein a lattice-based conditional privacy-preserving & authentication scheme is defined. The lattice-based scheme is able to combine data from RSUs, vehicles, application providers, and trusted-third parties as shown in Fig. 1.

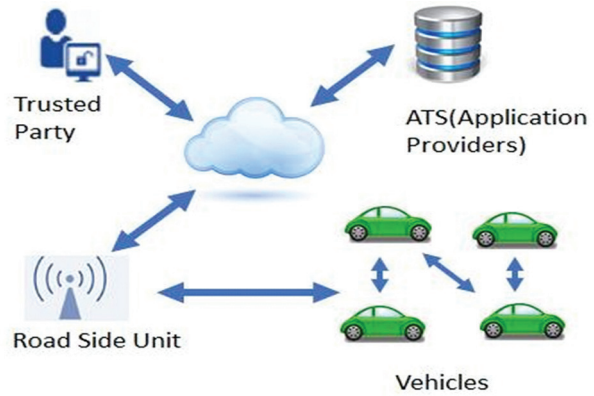


Fig.1. Lattice-based security mechanism

Due to lattice computations, the delay of the system is reduced, thereby improving the speed of communication along with providing security to the network. But, all these schemes mentioned in [3-6] suffer from inherent drawbacks, which are given as follows:

a. Limited area of applicability, because each of these protocols requires either the presence of a control unit or a high-powered computational unit.

b. Limited security performance due to the lack of decentralized control.

In order to eliminate these problems, the work in [7] presents efficient decentralized management mechanism with Blockchain. The solution employs managing the security of VANET by using a decentralized key-management mechanism. Bi-variate polynomial for a key agreement which is based on light-weight authentication is used. This technology can manage user identity and public key material which will improve the efficiency and cost as compared to traditional schemes of VANET. An example of the network is shown in Fig.2, wherein vehicles are connected to each other with the help of a decentralized blockchain model.

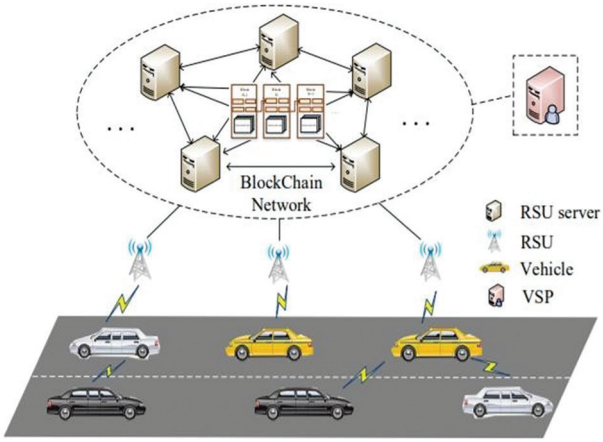


Fig. 2. Blockchain-based VANET

Due to the use of distributed blockchain-based authentication and communication, the overall communication cost is reduced by 50% of the cost of a standard public key infrastructure (PKI) system. This enables the network to be used for a larger set of users without increasing system cost. This work is modified in [8], in which trust-based routing and location privacy schemes are added to the VANET. Due to the addition of these schemes, the quality of service (QoS) performance of the network reduces, therefore there is a need to improve it with the help of machine learning models. The location privacy is maintained using k-copy scheme while the trust-based routing is maintained with the help of the same decentralized scheme as mentioned in [7].

The work in [8] can be further optimized in terms of QoS parameters with the help of an incentive scheme as mentioned in [9]. In this scheme, the node decisions regarding communication, channel selection, packet rate, etc. are monitored, and the parameters combination responsible for improving the QoS is incentivized using a scoring mechanism. Due to this, there is a balance between the security offered by the network and

the overall QoS of the network. The parameter combination which has the best score can communicate events in the network to the nearby nodes with the highest security and QoS. The event is mitigated with the help of the best possible route and parameter combination to each of the nearby nodes.

A combination of these schemes [7-9] is given in [10], wherein privacy-preservation is maintained using blockchain-powered trusted authorities. These authorities are responsible for node registration, node communication, and node termination from the network. All this data is stored in the form of the Merkle Patricia tree (MPT) for faster storage and retrieval performance. The system also supports conditional privacy by allowing each vehicle to have multiple certificates, and each certificate is responsible for a particular communication scenario in the network. A sample of this process is shown in Fig.3, the certificate for every special access for segregation is given to the vehicle in the network. This enables the vehicles to get better security performance, as the memory and computational requirement of these certificates is very low, therefore the QoS of the network is also maintained at an optimum level when compared to single-certificate computation systems.

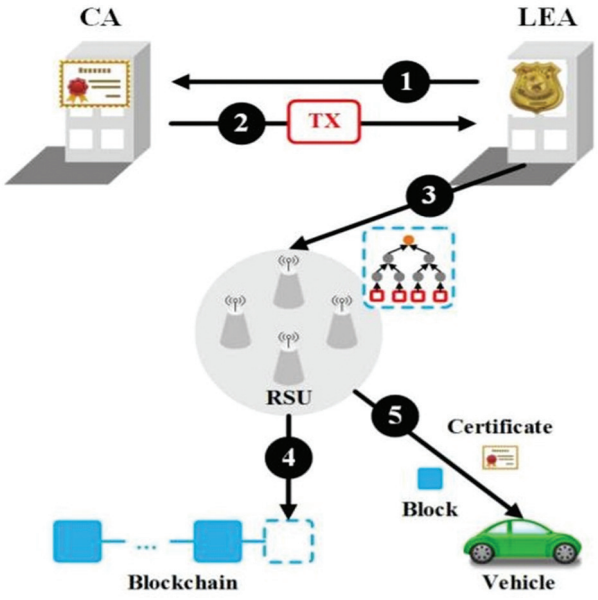


Fig. 3. Blockchain-based individual certificate scheme

In order to add features like bandwidth control, dynamic addressing, etc. VANETs are coupled with SDN. SDN is responsible for controlling the communication between nodes in the network. In order to eliminate these problems, the work in [7] presents efficient decentralized management mechanism with Blockchain. The solution employs managing the security of VANET by using a decentralized key-management mechanism. Bi-variate polynomial for a key agreement which is based on light-weight authentication is used. This technology can manage user identity and public key material which will improve the efficiency and cost as compared to traditional schemes of VANET. An example of the network is shown in Fig.2, wherein vehicles are connected to each other with the help of a decentralized blockchain model. An application of [10] and [11] can be observed in [12], in which blockchain and distributed ledger systems along with SDN are used for securing high-performance cyber-physical systems.

A similar trust-based model based on blockchain is also described in [13]. In this paper, a blockchain-based anonymous reputation system (BARS) is used to protect distribution of fake messages and privacy of the vehicles. The privacy preserving mechanism Lexicographical Merkle tree (LMT) is used to provide linkability between public key and identity of the vehicle through the certificate authority without disclosing private information of the vehicles. Law Enforcement Authority (LEA) is used to store the public key and identities of the vehicles. The reputation evaluation algorithm safeguards the vehicles from exposed behaviors thereby improving efficiency , robustness and security of the system.Work in [14] uses a blockchain-based Trust conditional privacy preservation announcement scheme (BTCPS]. It allows the vehicle to send messages to non-trusted environments. RSUs calculate reliability of the messages as per the reputation values of the vehicles. It can trace the malicious vehicle identity with associated public address . Proof of Work algorithm is used to improve the efficiency and QoS performance of the network.Due to conditional privacy, each node-to-node communication can be traced back to its source. It allows the system to trace the attacking node and eliminate it.

This scheme can be extended to include Group Mobility Management as given in [15]. In this work, handover latency and signalling costs during authentication can be reduced using aggregate message authentication code and one-time password authentication. Due to these techniques, the proposed scheme is not only fast but also supports faster handoffs whenever nodes are shifting between different internal mini-networks. This results in reducing the signalling overhead and handover latency of the system, which improves the QoS of the proposed system. In order to evaluate the system under different security threats, possible solutions are discussed in [16]. In this survey, it is observed that blockchain technology is the most useful when it comes to removing attacks from VANETs. It is used as the base security model for this underlying research. While security is a major aspect of any VANET, an effective data collection process must also be taken into consideration while designing networks. The work in [17] reviews a wide variety of data collection mechanisms for VANETs, and observe that topology-based methods have the best performance in terms of reducing delay and increasing the overall throughput performance. Using this mechanism, the QoS performance of VANETs can be improved. While using blockchain with topology-based methods, it is required that not more

than 51% of nodes must be clubbed together for communication. This results in a highly secure blockchain network, which is given in [18]. Due to topology issues if the blockchain network is attacked with more than 51% of users, then the network rules can be changed. These changes can allow an attacker to inject malicious rules into the network, and make backdoors in the system. It is also observed that when less than 51% of nodes are active then proper data dissemination can take place in the network. The overall performance and storage capabilities of the network depend on the memory and computational power needed per-node basis. The higher number of dissemination nodes will require a larger memory for storage, and will also require a higher computational power when compared to a network with a moderate number of participating dissemination nodes. This can be observed from [19], in which nodes within a radius of 1-hop distance are considered ideal candidates for data dissemination. This indicates that all kinds of topology constraints must always keep less than 51% nodes in close vicinity with each other.

Another attribute-based blockchain algorithm with privacy preservation and authentication is indicated in [20]. In this work, due to the decentralized nature of blockchain, there is no need to perform pre-authentication in the network. Moreover, the network provides high security without the presence of any roadside unit or infrastructure components. The network also demonstrates trust management considerations, which are improved with the help of SDN as observed in [21]. Trust based Deep Enforcement Learning Framework with SDN uses deep enforcement learning algorithm to find the highest routing path of the network. The trust model is used to evaluate behavior of neighboring nodes of the forwarding packets which helps to improve QoS parameters. These techniques can be extended to 5G networks as given in [22], wherein it is observed that network security is improved if SDN architectures are applied to high-speed 5G networks. While each network type has its own design requirements, the work in [23] indicates that SDN-based VANET networks require a lot of integrations before real-time deployments. For instance, to support Privacy violations the SDN-based VANET must implement 'Disclosing sensitive information module from SDN and 'Revealing the identity of vehicles module from VANET. In order to improve the security performance of stand-alone SDN systems, the work in [24] indicates the usage of broadcast encryption mechanisms. These mechanisms allow the system to secure new and existing SDN networks. The broadcast encryption mechanism is based on the Advanced Encryption Standard (AES) in 256-bit mode (AES-256). The security performance of this network is found to be far superior to other SDN networks, and thereby the same AES-256 implementation is used by the underlying research. Network safety can also be improved with the help of cooperative communication similar to P2P networks, this is given in [25]. To improve the network performance ,roadside Open Flow switch (ROFS) is used. SDN based Medium Access Control protocol is classified into two levels as follows:

a. Controller and management of vehicles is used to control Road Side Units .

b. Controller is used to schedule the cooperative time slot sharing between Road Side Units. Slots are allotted based on sharing information between control and data plane.

SDN based blockchain can improve vehicle density fluctuation which helps to improve agility and speed of the network which further improves the security. To improve network performance and better security, The Open Flow protocol is connected in tandem with ROFS, and a distributed communication architecture is implemented.

For improving security, the energy consumption of the network is considered in [26]. In this work, a small change in the learning function with the inclusion of energy consumption results in routing solutions that have greater energy efficiency than the ones which do not include energy consumption into the equation. The energy consumption can be further improved by offloading all the security and related computations on fog devices. This is observed in [27], wherein mobile edge nodes are utilized for performing complex encryption calculations, while the main communication and event-triggered processing are done on the main vehicular node. A device-to-device clustering (D2DC) method is described, which provides coverage to nodes that are not in the coverage radius of the main infrastructure node. This D2DC method does not only provides better coverage but also improves the overall energy efficiency of the network. The nodes which are not in coverage range do not require sending unnecessary communication packets in search of the infrastructure nodes. Due to this, more than 74% of the unserviceable nodes come under proper service of the network, which improves the QoS performance of the network. A similar approach that uses multi-agent architecture is given in [28]. A hybrid SDN based geographic routing protocol allows selection of reliable nodes to avoid communication problems between source and destination. By using load balancing criteria allows to form hierarchical topology of the network by creating group and selecting the group leader. Routing protocol provides the better network flexibility and resource management which helps to improve QoS parameters. The geographic routing protocol can be further modified as per the survey done in [29], in which it is suggested that VANET routing can be best adopted with key management-based trust & secure routing protocols for better routing efficiency.

The future of VANETs is the integration of the network with cloud-based approaches [30].Integration of Fog computing with SDN enhances the flexibility and programmability of the network. It will help minimize

future challenges in VANET. Blockchain- based SDN in combination with different protocols can improve inherent security of the network. The proposed Encrypted Blockchain based open Flow SDN architecture is explained and the performance evaluation of the given protocol and comparison with other standard algorithms is described in the next section.

3. PROPOSED ENCRYPTED BLOCKCHAIN OPEN FLOW SDN ARCHITECTURE (ABOSS)

The proposed encrypted blockchain open flow SDN architecture is described by dividing the entire VANET traffic flow into 3 different parts which are given as:

a. Securing node to node communication using AES- 256 & ad-hoc on-demand distance vector (AODV) routing protocol.

b. Improving security for the entire network using blockchain-based data transfer.

c. Adding QoS improvement layer with network control using Open Flow SDN.

A Block diagram of the entire system is shown in Fig.4, where node-to-node communications are shown.

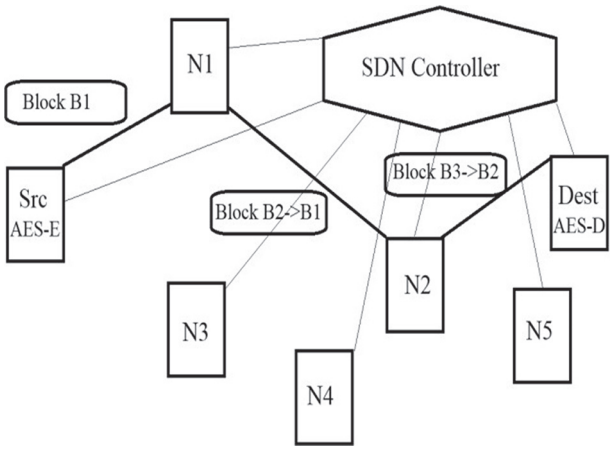


Fig. 4. Proposed encrypted blockchain open flow SDN architecture (ABOSS)

The input data originates from the source node, is encrypted with the AES protocol. The private key of AES is shared with the source and destination nodes. This key is moved into a secure block using public-key cryptography and is sent in the network. The block diagram of AES is shown in Fig.5. AES follows the given steps for encryption of data:

AES is a standard encryption algorithm that follows the given steps for encryption of data:

a. Add round key

b. Substitute bytes

c. Shift rows

d. Mix columns

e. Add round key

For decryption the same process used in reverse order using AES, the input data is altered to cipher text , and is kept ready for broadcast from the source node. Once the data is encrypted, to set the best routing paths between the source and destination packet is send to nearby nodes. These packets sent are known as Route Request (RREQ) packets as shown in Fig.6, where node 'A' is the source and node 'F' is the destination.

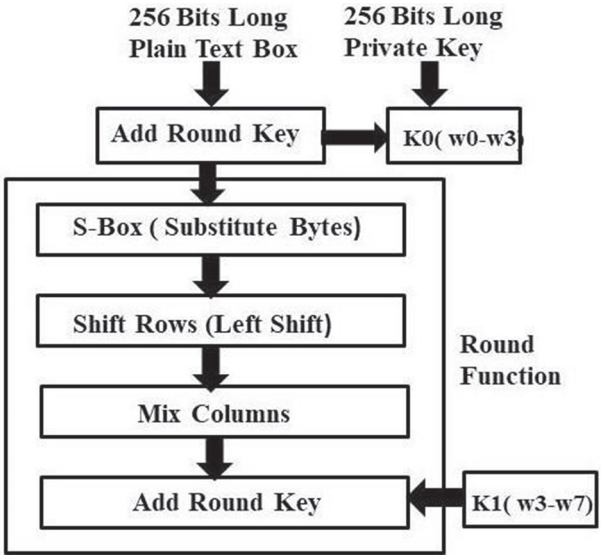


Fig. 5. AES block diagram

The nodes which are near, send Request reply (RREP) packets. Based on the reception of RREP packets, a path is selected between source 'A' and destination 'F' as A—B—D—F. This path is selected and kept stored on the SDN node. The source node then applies a blockchain based data transformation protocol. Using this protocol, the input data is converted into the following block structure which is given in Table 1.

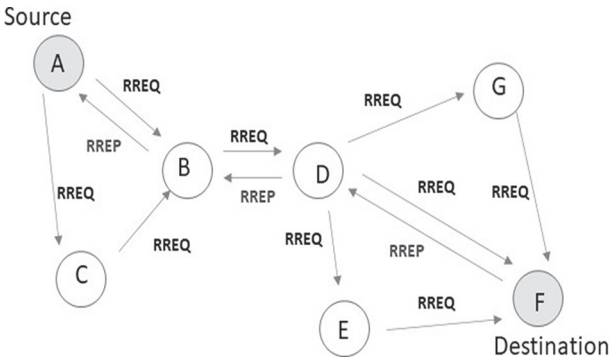


Fig. 6. Routing process

Table 1. Structure of the blocks in the blockchain

Previous Hash	Source	Destination	Current Node
Timestamp	Nonce	Data	Hash

The above structure gives the information of source node and destination nodes along with the timestamp of the transfer, current hash value and previous blocks

hash values and nonce value saved in the respective blocks. How the nonce value is responsible for creating the unique hash value. How to find the nonce number in blockchain the steps are shown below:

a. Initialize a random nonce value
b. Store the value in a block
c. Find the SHA256 hash of block
d. Check whether the hash value is repeated, if yes then discard and restart.
e. Check whether the hash value is following blockchain rules, if not then discard and restart.
f. If both rules (d, e) are followed, then store the nonce in the blockchain.

With above pseudo steps, the nonce number is evaluated and put in the blockchain. Which update the hash table's data and transferred it from source node to the next nearby nodes. As the data transmit to the next hopping nodes, whole blockchain is checked for authentication.

The blockchain, checking process is done using the following steps:

a. Check the current hash of the block
b. Check the previous hash of the next block
c. If these hashes match, then continue with the Checking,
d. If these hashes do not match, then discard the blockchain and re-start the communication
e. Once the blockchain is verified, then the communication proceeds to the next node.

Each blocks should follow the blockchain rules once the hash values of previous and current blocks is checked and verify. Proof –of –Work consensus algorithm is used to verify all these process. Once the verification done, packets send from any nodes using Open Flow SDN Protocol. This protocol has the following rules:

a. Remove the nodes which are not found in registered node list during communication process.
b. Set of 'N' packets allow to be transmitted within the nodes, where 'N' is link of maximum capacity to handle the packets.
c. k is the maximum hopes between the source and destination nodes decided. During the communication only k hoping is allowed in the network.

All these rules are applied to each of the communication packets. After application of SDN rules, the network will have the following advantages:

a. The network will be resilient to denial of service (DoS) attacks, due to SDN rules.
b. The network will not be affected by Spoofing (Masquerading) attacks, because of a combination of SDN rules and the blockchain verification process.
c. The network will not be affected by spying attacks due to blockchain verification, due to which there will be no communication of any node with unwanted nodes, thereby removing any chances of spying or spoofing.

Once all these verifications are completed, then data communication proceeds on a node-to-node basis. A comparison of the results obtained for these protocols with other standard methods is done and conclusions are derived from these results in the next section.

4. RESULTS EVALUATION & ANALYSIS

In order to evaluate the results for the given protocol, the network is simulated under similar conditions as directed in the standard VANET simulation networks in [10]. Due to the use of blockchain, it is observed that the proposed model has high security, and is able to identify Sybil, Masquerading, DDoS, and Smurf attacks with 100% efficiency. The reasons for this high efficiency are traceability, immutability, and improved trust levels of blockchain. Thereby making the network 100% efficient in terms of attack detection. The QoS parameters were evaluated by changing the number of communications, and averaging the values. The number of communications was varied between 10 to 100. The following network parameters are decided for simulating the network,

Channel Type:	Wireless Channel
Propagation Mode:	Two Ray Ground
Network interface:	Wireless Physical
MAC Protocol:	Mac/802.11
Interface Queue type:	Drop Tail Priority Queue
Antenna Type:	Omnidirectional
Routing:	AODV
Network X Size:	300
Network Y Size:	300
Packet Size:	1000 bytes per packet
Packet Interval:	0.01 seconds per packet

Parametric values for the end-to-end delay, throughput, energy consumption, and packet delivery ratio are evaluated based on the following formulas,

$$D = Tr - Tt \tag{1}$$

$$E = Et - Er \tag{2}$$

$$\text{Thr} = \frac{Psr}{D} \tag{3}$$

where, D is the end-to-end communication delay, E is the energy consumed during communication, Thr is Troughput and T_t are the reception and transmission time for the packet, E_t and E_r are the transmission and reception energies in the network, P_{sr} are the number

of packets successfully received and P_t are the number of packets transmitted. Using these values, the following parameters were evaluated. Delay performance of the network using different protocols are given in Table 2 where NSV is Non-Secured VANET, ASV is AES Secured VANET, ASS is AES Secured SDN, ABOSS is AES Blockchain Open flow Secured SDN.

Table 2. Delay Performance

No. of Nodes	Delay (ms) NSV	Delay (ms) ASV	Delay (ms) ASS	Delay (ms) ABOSS
10	34.09	45.80	25.02	21.40
20	31.09	44.80	32.02	22.40
30	34.31	44.82	27.79	25.72
40	37.94	42.04	47.56	24.31
50	54.93	45.33	28.31	27.47
60	57.02	35.06	26.39	22.30
70	47.45	33.66	25.39	24.71
80	32.69	29.54	29.89	21.89
90	49.29	44.80	32.02	21.40
100	61.24	51.27	22.20	21.69

From Table 2, it is observed that the Average delay of NSV is 44ms, ASV is 41.7ms, ASS is 29.6ms, and delay of ABOSS is 23.2ms. Thus, the delay using ABOSS is improved by 44% when compared to VANET secured systems, and the delay using ABOSS is improved by 22% when compared to SDN secured systems. Similar comparisons are made for energy and throughput. These values are given in Table 3 and 4 respectively.

From Table 3, it is observed that the Average Energy consumption of NSV is 25.48mJ, ASV is 39.73mJ, ASS is 18.3mJ, and energy consumption of ABOSS is 15.85mJ. Thus, the energy consumption using ABOSS is improved by 60.1% when compared to VANET secured systems, and the energy consumption using ABOSS is improved by 13.4% when compared to SDN secured systems.

Table 3. Energy performance

No. of Nodes	Energy (mJ) NSV	Energy (mJ) ASV	Energy (mJ) ASS	Energy (mJ) ABOSS
10	20.3	38.3	19.2	16.2
20	21.9	39.1	19.4	16.4
30	22.6	42	20.2	16.3
40	23.6	41.8	18.8	15.5
50	29.5	42.5	17.4	13.2
60	28.5	26	17.3	17.5
70	26.7	40.1	18.3	16
80	22.2	41.5	17.8	14.2
90	27.5	42.5	17.6	16.9
100	32	43.5	17	16.3

From Table 4, it is observed that the Average Throughput of NSV is 12.74Tbps, ASV is 16.34Tbps, ASS is 262.6Tbps and the Throughput of ABOSS is 442.6Tbps. Thus, the Throughput using ABOSS is improved much more than 100 % when compared to VANET secured systems, and the Throughput using ABOSS is improved by 68.5 % when compared to SDN secured systems.

Table 4. Throughput performance

No. of Nodes	Thr (Tbps) NSV	Thr (Tbps) ASV	Thr (Tbps) ASS	Thr (Tbps) ABOSS
10	19.8	19.3	163	423
20	18.8	17.3	173	435
30	15.3	8.83	240	620
40	12.3	8.78	287	354
50	5.4	8.5	244	429
60	7.73	14.2	261	372
70	16.6	22.1	256	447
80	7.91	40.2	466	469
90	4.8	6.94	363	442
100	18.8	17.3	173	435

5. CONCLUSION AND FUTURE SCOPE

The QoS parameters are improved by combining SDN with Open Flow with blockchain, and AES encryption 256 standards of the network. The network is secured from DOS, Masquerading, and Spying attacks due to the SDN rules and blockchain verification process. Due to the incorporation of AES, there is a further improvement in the security performance of the network in terms of data confidentiality. The work can be carried out on the cloud by unloading computations associated with security and blockchain and thus improving the parameters of the network. By the addition of machine learning in the routing process, the routing algorithm can be improved, which will further improve the QoS and security performance of the network.

6. REFERENCES:

[1] Y. Yang, D. He, H. Wang, L. Zhoc, "An efficient blockchain-based batch verification scheme for vehicular ad hoc networks", Transaction on Emerging Telecommunications Technologies, 2019, pp. 1-12.

[2] S. V. Akram, P. K. Malik, R. Singh, G. Anita, S. Tanwar, "Adoption of blockchain technology in various realms: Opportunities and challenges", Security Privacy, Vol. 3, No. 5, 2020.

[3] F. M. Salem, A. S. Ali, "SOS: Self- organized secure framework for VANET", International Journal Communication System, Vol. 33, No. 7, 2020.

[4] Y. Zhao, X. Zhang, X. Xie , Y. Ding, S. Kumar, "A verifiable hidden policy CP-ABE with decryption testing scheme and its application in VANET", Transaction on Emerging Telecommunications Technologies, 2019. (in print)

[5] M. D. Firoozjaei, R. Lu, A. Ghorbani, "An evaluation framework for privacy-preserving solutions applicable for blockchain-based internet-of-things platforms", Security Privacy, Vol. 3, No. 6, 2020.

[6] Dharminder, D. Mishra, "LCPPA: Lattice-based conditional privacy-preserving authentication in vehicular communication", Transaction on Emerging Telecommunications Technologies, Vol. 31, No. 2, 2019.

[7] Z. Ma, J. Zhang, Y. Guo, Y. Liu, X. Liu, W. He, "An Efficient Decentralized Key Management Mechanism for VANET with Blockchain", IEEE Transaction on Vehicular Technology, Vol. 69, No. 6, 2019, pp. 5836-5849.

[8] B. Luo, X. Li, J. Weng, J. Guo, J. Ma, "Blockchain-Enabled Trust-based Location Privacy Protection Scheme in VANET", IEEE Transaction on Vehicular Technology, Vol. 69, No. 2, 2020, pp. 2034-2048.

[9] M. S. Iftekhar, N. Javaid, O. Samuel, M. Shoaib, M. Imran, "An Incentive Scheme for VANETs based on Traffic Event Validation using Blockchain", Proceedings of the International Wireless Communications and Mobile Computing Conference, Limassol, Cyprus, 15-19 June 2020.

[10] Z. Lu, Q. Wang, G. Qu, Senior Member, IEEE, H. Zhang, and Z. Liu, "A Blockchain-Based Privacy-Preserving Authentication Scheme for VANETs", IEEE Transactions on Very Large Scale Integration Systems, Vol. 27, 2019, pp. 2792-2801.

[11] D. Zhang, F. R. Yu, and R. Yang, "Blockchain-Based Distributed Software-defined Vehicular Networks: A Dueling Deep Q-Learning Approach", IEEE Transactions on Cognitive Communications and Networking, Vol. 5, No. 4, 2019, pp. 1086-1100.

[12] M. Wagner, B. McMillin, "Cyber-Physical Transactions: A Method for Securing VANETs with Blockchains", Proceedings of the IEEE 23rd Pacific Rim International Symposium on Dependable Computing, Taipei, Taiwan, 4-7 December 2018.

[13] Z. Lu, W. Liu, Q. Wang, G. Qu, Z. Liu, "A Privacy-preserving Trust Model based on Blockchain for VANETs", IEEE Access, Vol. 6, 2017, pp. 45655-45664.

[14] X. Liu, H. Huang, F. Xiao, Z. Ma, "A blockchain-based trust management with conditional privacy-preserving announcement scheme for VANETs", IEEE Internet of Things Journal, Vol. 7, No. 5, 2019, pp. 4101-4112.

[15] C. Lai, Y. Ding, "A Secure Blockchain-Based Group Mobility Management Scheme in VANETs", Proceedings of the IEEE/CIC International Conference on Communications in China, Changchun, China, 11-13 August 2019.

[16] S. Tanwar, J. Vora, S. Tyagi, N. Kumar, M. S. Obada, "A systematic review on security issues in vehicular ad-hoc network", Security and Privacy, Vol. 1, No. 5, 2018.

[17] B. Pourghebleh, N. J. Navimipour, "Towards efficient data collection mechanisms in the Vehicular ad hoc networks", International Journal Communication System, Vol. 32, No. 5, 2019.

[18] R. Shrestha, S. Yeob, "Regional Blockchain for Vehicular Networks to Prevent 51% Attacks", IEEE Access, Vol. 7, 2019, pp. 95033-95045.

[19] R. Shrestha, R. Bajracharya, S. Yeob Nam, "Blockchain-based Message Dissemination in VANET", Proceedings of the IEEE 3rd International Conference on Computing, Communication and Security, Kathmandu, Nepal, 25-27 October 2018.

[20] Q. Feng, D. He, S. Zeadally, K. Liang, "BPAS: Blockchain-Assisted Privacy-Preserving Authentication System for Vehicular Ad-Hoc Networks", IEEE Transactions on Industrial Informatics, Vol.6, 2019.

[21] D. Zhang, F. R. Yu, R. Yang, "A Machine Learning Approach for Software-defined Vehicular Ad Hoc Networks with Trust management", Proceedings of the IEEE Global Communications Conference, Abu Dhabi, United Arab Emirates, 9-13 December 2018.

[22] A. Hussein, I. H. Elhajj, A. Chehab, A. Kayssi, "SDN VANETs in 5G: An Architecture for Resilient Security Services", Proceedings of the 4th International Conference on Software Defined Systems, Valencia, Spain, 8-11 May 2017.

[23] W. B. Jaballah, M. Conti, C. Lal, "Security and Design Requirements for Software-Defined VANETs", Computer Networks, Vol 169, 2020.

[24] J. S. Weng, J. Weng, Y. Zhang, W. Luo, W. Lan, "BENBI: Scalable and Dynamic Access Control on the Northbound Interface of SDN-based VANET", IEEE Transactions on Vehicular Technology, Vol. 68, No. 1, 2019, pp. 822-831.

[25] G. Luo, J. Li, L. Zhang, Q. Yuan, Z. Liu, F. Yang, "SDN MAC: A Software-Defined Network Inspired MAC Protocol for Cooperative Safety in VANETs", IEEE Transactions on Intelligent Transport Systems, Vol. 19, No. 6, 2018, pp. 2011-2024.

[26] J. Joshi, K. Renuka, P. Medikonda, "Secured and Energy Efficient Data Transmission in SDN-VANETs", Proceedings of the 22nd International Computer Science and Engineering Conference, Chiang Mai, Thailand, 21-24 November 2018.

[27] A. Muthanna, R Shamilova, A. Ateya, A. Paramonov, M. Hammoudeh, "A mobile edge computing/software-definednetworking- enabled architecture for vehicular networks", Internet Technology Letters, Vol. 3, No. 6, 2019.

[28] L. Alouache, N. Nguyen, M. Aliouat, R. Checotah, "HSDN-GRA: A hybrid software-defined networking-based geographic routing protocol with the multi-agent approach", International Journal Communication System, Vol. 33, No. 15, 2020.

[29] L. Alouache, N. Nguyen, M. Aliouat, R. Chelouah, "Survey on IoV routing protocols: Security and network architecture", International Journal Communication System, Vol. 32, No. 2, 2019.

[30] R. Shrestha, R. Bajracharya, S. Y. Nam, "Challenges of Future VANET and Cloud-Based Approaches", Wireless Communication and Mobile Computing, 2018.

BCSDN-IoT: Towards an IoT security architecture based on SDN and Blockchain

Original Scientific Paper

Younes ABBASSI
Hassan 2 University,
Faculty of Sciences Ben Msik, Computer Sciences Department
Casablanca, Morocco
younes.abbassi@univh2c.ma

Habib Benlahmer
Hassan 2 University,
Faculty of Sciences Ben Msik, Computer Sciences Department
Casablanca, Morocco
h.benlahmer@gmail.com

Abstract *– The Internet of Things (IoT) aims to create a digital world where any information system can expose, discover, understand and consume data and services for analysis, diagnosis, decision support and task automation in various domains such as healthcare, transportation, energy, industry, agriculture, etc. Faced with this diversity of applications and rapid evolution, infrastructures must be able to achieve high levels of security and confidentiality while being open, sustainable, and agile to adapt to the multiple requirements of applications.*

To meet these needs, new paradigms are emerging. These include the Software Defined Networks (SDN) paradigm, which offers the ability to dynamically program different applications and devices to provide end-to-end service chains. In parallel, the Blockchain paradigm is increasingly used in the Internet of Things, making distributed transactions between connected objects such as financial transactions or "smart contracts" possible.

Although the combination of these two paradigms (Blockchain/SDN) is a major issue for the success of the Internet of Things, paving the way for new business models and management/control of communication networks, there is not yet a specified/formalized architecture allowing the use of the "Blockchain" in SDN. In this research, a new architecture for a system combining blockchain and SDN for IoT security is proposed.

Keywords*: Internet of Things (IoT), Software-Defined Networking (SDN), OpenFlow, Blockchain, BCSDN-IoT Architecture*

1. INTRODUCTION

In 2030, it is announced that there will be more than 500 billion devices connected to the Internet with a variety of uses leading to security problems and an increase in traffic on the networks that will be estimated in Zeta (1021) bytes [1].

However, currently, the security architectures deployed in networks are mainly based on experience and work on wired networks. These architectures are mainly based on centralized equipment, whose main role is to control the information that is exchanged between the company's network and the outside world. It is therefore not possible to control the information exchanged between a terminal equipment that a user will connect to his computer. On a corporate network, users can connect their phone to their computer, via Bluetooth for example, and thus the computer becomes a new entry point to the network. With the Internet of Things (IoT), we have sensors, thermostats, webcams, watches connected to our phones, themselves eventually connected to the Internet or to our computers [1]. So how can we control the information coming from this large mass of heterogeneous devices?

With the increase in the number of these heterogeneous devices, the complexity in their administration is growing. This requires a verification of the coherence of the configurations of all the network devices of a company, for example the security rules and the user rights [2].

With the support of a great combination of modern technologies such as IoT, SDN, and Blockchain, as the number of connected things to the internet grows these days, managing and controlling IoT has become a very difficult task. SDN steps in to provide the IoT

network's adaptability and programmability without requiring existing implementations to change their design. It may also assess how the network affects the overall performance and efficiency of the network system, which is very useful when dealing with real-time transactions [4]. SDN is utilized in IoT applications to reduce response time and security concerns. In SDN, many controllers have recently been used instead of a centralized controller. The fundamental purpose of using multiple controllers is to balance the load between devices and controllers while minimizing packet loss. When the user of the SDN-IoT network need resources, they will be available immediately. In addition, utilizing an SDN controller, a network can be configured dynamically. One of the most common protocols used by SDN is OpenFlow [4].

Some other advanced technology is blockchain, a decentralized, emergent technology that can be combined with SDN-based IoT applications. The hash value is used to link various blocks together, and each block of the transaction is saved forever [5]. Combining this Blockchain technology will boost security and privacy. Several academics have proposed numerous clarifications to increase the network's performance, but they are unable to entirely cure the problem.

Although IoT, SDN, Blockchain technologies are combined to provide a better solution for any smart technology such as intelligent building, smart homes, smart cities, and smart grids [5]. These technologies can also provide reliable data transmission as well as communication in the networks [3].

However, the potential use of this disruptive technology spawn to each and every application that need to evolve from a centralized authorization entity acting as a trusted intermediary or sometimes a third-party verifiable trust anchor, towards a purely distributed authentication model.

Our goal in this paper is to give the reader particularly interested in IoT security, a proposal for a new security architecture combining SDN and Blockchain technologies, with the aim of improving, and simplifying the deployment of IoT security.

The remainder of the paper is laid out as follows: In section 2, we go over some background information before introducing IoT, SDN, and Blockchain, as well as their designs. Then, in section 3, we describe our proposed BCSDN-IoT architecture, its operation, and analysis and alert generation. In section 4, we conduct an implementation of the BCSDN-IoT solution in virtual through the open source solution OpenDayLight, starting with its installation, the realization of the BCSDN-IoT architecture and the simulation of some attacks. And finally in section 5 we conclude our article with some perspectives.

2. BACKGROUND & RELATED WORKS

2.1 INTERNET OF THINGS

The Internet of Things (IoT), as well as the Internet of Everythings (IoE) in a larger sense, is a relatively new concept. It is considered a major technological and economic innovation in the industry of new information technologies and communication.

The IoT does not have a unique definition but generally speaking, It is characterized as a broadening of the current Internet to include all objects that can communicate directly or indirectly with electronic equipment that is also linked to the Internet.

The International Telecommunication Union [7] defines the Internet of Things as: "A global infrastructure for the information society, which enables advanced services by interconnecting objects (physical or virtual) through existing or evolving interoperable information and communication technologies".

IoT devices are typically sensor nodes, RFID (Radio Frequency IDentification) tags and wireless communication devices connected to the Internet in a smart environment [2]. These devices are very diverse (phone, watch, refrigerator...) and are now widely used in everyday life.

With the exponential development of these connected objects with heterogeneous characteristics, the networks of the future must evolve towards new architectures to adapt to the increase in traffic and ensure their security. Security is one of the issues of today's Internet, as there are more and more intelligent security attacks to deal with. In addition, security attacks for IoT are more difficult to handle due to the minimal energy storage, data and processing capacity that are not suitable for existing network security mechanisms based on firewall and IDS/IPS [21]. The concept of IoT is relatively simple but there are many problems because these connected devices do not have enough capacity to handle the communications and processing associated with the applications.

Architecture: The most commonly used IoT architecture for SDN solutions, and as shown in Figure 1, is made up of three layers: the perception layer, the network layer and the application layer [8].

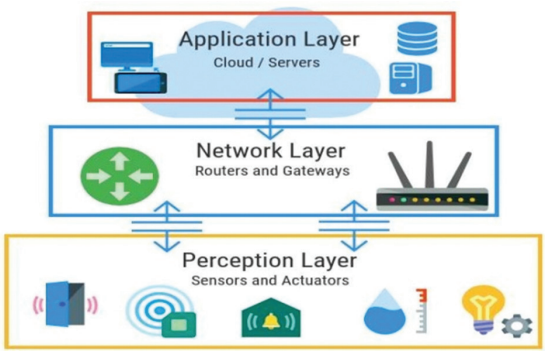


Fig. 1. The architecture of IoT

Perception layer: it is at this level that the collection of information takes place. Various devices and devices help it in this, such as smart cards, readers and sensors, RFID tags, etc.

It is personalized by a function that allows it to detect the whole object in order to acquire information about it at any time and place, through the RFID system. (EPC) Electronic Product Code is a unique identifier that distinguishes each object in the IoT infrastructure, it generates a sequence of numbers giving an idea about the producer of the object, its production date and the expiry date.... [8].

Network layer: This layer allows the sending of the required information from the previous layer to the internet through machines, wired or wireless network equipment. As a transport layer, digital data is transported reliably [8].

Application layer: or also known as the process layer analyzes the information received and makes control decisions to perform its intelligent processing function by connecting, identifying and controlling objects and devices. Intelligence assets use intelligent computing technologies such as cloud computing and process information for intelligent control, as well as the tasks that must be completed and when they must be completed [8].

2.2 SOFTWARE-DEFINED NETWORKING

SDN (software-defined networking) is a new network architecture paradigm that describes a control plane that is completely separate from the data plane. According to the ONF (Open Network Foundation) [9] SDN is an architecture that separates the control plane from the data plane, and centralizes all network intelligence [27] in a programmable entity called "Controller", in order to manage several elements of the data plane (e.g. switches or routers, etc.) via APIs (Application Programming Interface).

More concretely, we can say that a network architecture follows the SDN paradigm if, and only if, it verifies the following points:

- The control plane is completely decoupled from the data plane; this separation is materialized through the definition of a programming interface (Southbound API)
- All network intelligence is externalized in a logically centralized point called the SDN controller, which offers a global view on the entire physical infrastructure.
- The SDN controller is a programmable component that exposes an API (NorthboundAPI) to specify control applications.

Architecture: A traditional network is generally composed of interconnection equipment such as switches and routers. This equipment incorporates both the transmission and control parts of the network. In this architecture model, it is difficult to develop new services because of the strong coupling between the control plane and the transmission plane.

In order to open the network equipment to innovations, the SDN architecture was born. It allows decoupling the control part from the transmission part of the interconnection equipment [29]. As depicted in Figure 2, the SDN is made up of three layers and communication interfaces.

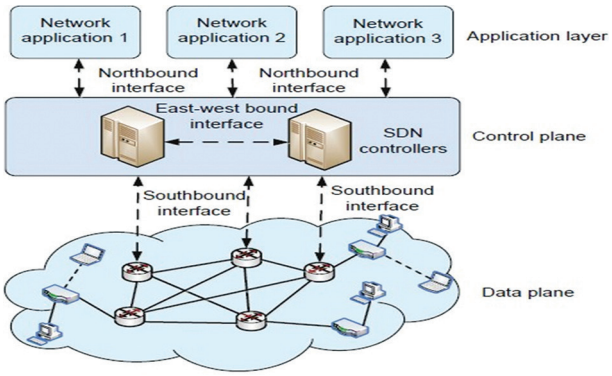


Fig. 2. The architecture of SDN

We describe in the following these layers, as well as the communication interfaces:

- **The transmission layer:** also called "data plane", it is composed of routing equipment such as switches or routers, its main role is to transmit data and collect statistics.
- **The control layer:** also called "control plane", it is mainly composed of one or more SDN controllers, its role is to control and manage the infrastructure equipment through an interface called 'south-bound API'.
- **The application layer:** represents the applications that enable the deployment of new network functionalities, such as traffic engineering, QoS, security, etc. These applications are built through a programming interface called 'northbound API'.

Communications Interfaces

There are three main types of interfaces, which allow controllers to communicate with their environment: South, North and East/West interfaces

Southbound APIs: are the interfaces that allow the SDN controller to communicate with infrastructure layer devices like switches and routers.

The most widely used protocol, and the most deployed as a Southbound interface is the OpenFlow protocol, which has been standardized by the ONF, its latest version is 1.5 [10], more details on this protocol will be given in the next section. There are now other southern interface alternatives, such as ForCes [11], or

Open vSwitch Database (OVSDB) [12], but the OpenFlow protocol is currently the de facto standard, which is widely accepted and spread in SDN networks.

Nord interfaces: are used to program transmission devices, exploiting the network abstraction provided by the control plane. It is noted that unlike the Southbound API which has been standardized, the North interface still remains an open question.

While the need for such a standardized interface is a considerable debate within the industry, the advantage of an open Northbound API is also important, as an open Northbound API allows for more innovation and experimentation. Several implementations of this interface exist, each of which offers very different functionality. The RESTful [13] is considered the most widespread North API in SDN.

East/West interfaces: are communication interfaces that allow communication between controllers in a multi-controller architecture to synchronize the network state. These architectures are very new and no inter-controller communication standard is currently available.

2.3. BLOCKCHAIN

Blockchain technology emerged in early 2009 with the crypto-currency Bitcoin (BTC). Bitcoin users use a variable public key (PK) [14] to generate transaction information and broadcast it to the network for transferring funds. Transaction information is stored by all users in its own block. Once the block is full, a network mining process is performed; the hash value of the block is calculated, and the encrypted information and blocks are added to the blockchain [15]. To mine the cryptographic hash value of a block, certain nodes in the network, known as miners, compete to solve a proof of work called the cryptographic resource consumption puzzle (POW) [28]. The node that solves the puzzle first and gets everyone's approval is considered to have mined the block. This is because blockchain technology maintains all transaction data counts among all members, and all members update the counts simultaneously to maintain completeness when new transactions occur [16] [23]. The Internet and encryption technologies are the underlying technologies that allow all members to verify the reliability of each transaction to resolve a single point of failure caused by a traditional third-party authorized transaction. Because the blockchain is a peer-to-peer (P2P) network [17], the transaction is free of unauthorized third-party charges. As everyone keeps their transaction information up to date, the hacking effect of single point records is very limited, and it frequently fails. In addition, users of a blockchain system can openly access transaction records and reduce the costs of monitoring transactions. Since the hash value stored in each block peer is affected by the block peer is affected by the value of the previous block, forgery and modification of data requires modification of the entire chain [18] and the amount of computation at one point is far behind the computation of the entire network. As a result, forgery is almost impossible.

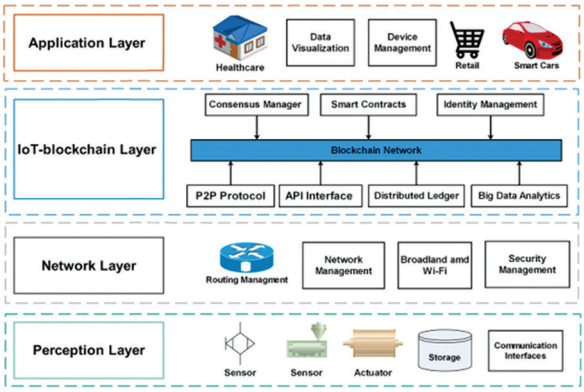


Fig. 3. The architecture of IoT with Blockchain

2.4 RELATED WORKS

Several researchers have recently addressed emerging leading technology such as IoT, SDN, Blockchain, and other smart technologies in today's world [19, 20, 21, and 22]. In this section, some literature reviews of recent works have been mentioned which are given below:

Table 1. State of the art.

Research Work	Used			Summary contributions and features
	IoT	SDN	Block chain	
M. J. Islam, M. Mahin, S. Roy, B. C. Debnath, and A. Khatun.[19]	✓	✓	*	Presented a distributed black net with SDN-IoT architecture for smart cities and addressed the cluster head selection scenario
M. A. Ferrag, M. Derdour, M. Mukherjee, A. Derhab, L. Maglaras, and H. Janicke [20]	✓	*	✓	Provided several overviews of the Blockchains application domain in IoT, e.g: Vehicle Internet, Energy Internet, Cloud Internet
P. K. Sharma, S. Singh, Y.-S. Jeong, and J. H. Park, [21]	✓	✓	✓	Proposed a literature combination between Blockchain and SDN for IoT networks and presented flow rule table for validation of blocks as well
C. Qiu, F. R. Yu, F. Xu, H. Yao, and C. Zhao [22]	✓	✓	✓	Proposed an imminent permitted blockchain-based consensus in distributed SDIoT and also efficiently used a novel dueling deep Q-learning approach.

3. BCSDN- IoT PROPOSED ARCHITECTURE COMBINING SDN AND BLOCKCHAIN FOR IoT SECURITY

Based on the analysis in the previous section for the rapidly growing IoT networks created by new communication paradigms, we observed that the current distributed network architecture, protocols, and techniques are not designed to meet the design principles required for future challenges and satisfy new service requirements. Today, organizations need a unique distributed security architecture that includes powerful network security devices that provide real-time proactive protection and high performance to address the design principles analyzed. In this section, we provide the distributed secure SDN architecture,BCDSN-IoT architecture, its workflow, and a mechanism for updating high-performance availability flow rule tables play an important role in a distributed blockchain network.

3.1. BCSDN-IoT APPROACH

BCSDN-IoT adopts distributed secure network control in the IoT network using the concept of blockchain technology to improve security, scalability, and flexibility without the need for a central controller. Figure 5 shows the overall view of the proposed architecture. In the proposed architecture, all controllers in the IoT network are interconnected in a distributed blockchain network fashion so that each IoT transmitting device in the network can communicate easily and efficiently. Each local network view includes an IDS/IPS (Intrusion Detection System/Intrusion Prevention System) service. By putting an IDS module on each controller, the BCSDN-IoT architecture not only enables operational flexibility, but also proactive and reactive incident prevention based on the repeating threat environment, which is fast evolving, dynamic and high performing. It provides an agile, modular and secure network infrastructure. Protections must dynamically adapt to the threat landscape without requiring security administrators to manually process large numbers of notifications and approvals. These assurances must be well-coordinated across the broader IoT environment, and the architecture must adopt a protection posture that uses both internal and external sources constructively.

Our solution is inspired by the security grid concept and our intelligent firewall approach to improve security in a conventional network and extend it to the IoT.

In this approach, we propose a collaborative security solution with a distributed controller architecture coupled with IDS. We have opted for a distributed SDN architecture distributed SDN architecture because a centralized architecture with a single controller increases the danger of network in the event of a denial of service (DoS) attack, there will be a service outage. For example, if the threat is only on one machine, it is not critical and isolating the machine can be a solution, but if the single controller is compromised, the whole network is at risk. The use of multiple controllers therefore creates redundancy, ensures high availability and reduces network latency.

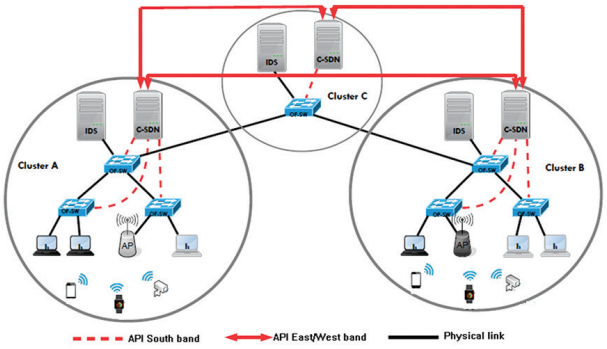


Fig. 4. Distributed Routing Cluster for SDN

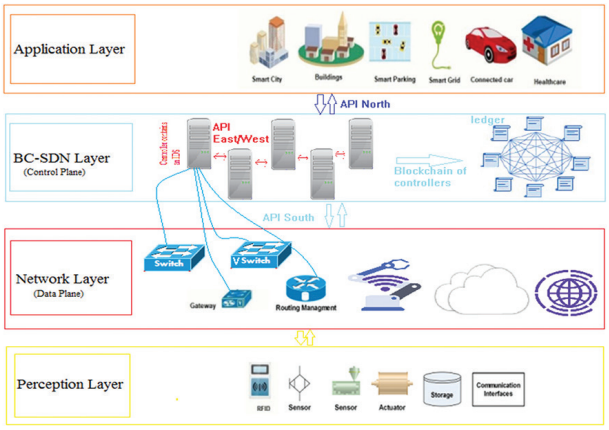


Fig. 5. The BCSDN-IoT architecture proposed

As shown in Figure 4, our solution consists of one or more clusters. Each cluster is composed of one or more network devices that are responsible for the interconnection of devices including connected objects. Within each cluster, an SDN controller manages the OpenFlow network. Each SDN controller is coupled to an IDS. The IDS is responsible for detecting intrusions into the network perimeter of each cluster. In other words, a cluster is an SDN domain in which we use an OpenFlow network with an SDN controller and an IDS to manage the security domain, which we call the zone of trust. To form a trust zone, all the equipment and devices in this zone must be fully secured. This security work is done by the controller and IDS pair. The SDN controller acts as an intelligent firewall for the trust zone and has security rules specific to the security needs of the cluster. These rules are programmed by an administrator. They can be distributed to other trust zones if the security need is the same through the East-West API.

Our approach allows not only to manage security in a totally decentralized way through a local management of security by the SDN/IDS controller couple, but also that the controllers exchange information on the threats detected in their respective clusters.

The SDN controller is the central element for security management in each trust zone. It has a global view of the network, manages traffic and distributes security policies to the network devices in its own cluster.

Before the SDN controller can isolate the threat in each cluster, it must be detected. That is why we use an IDS to solve this problem. In practice, this can be done by setting up an IDS like snort or other.

3.2 ANALYSIS, DETECTION AND ALERT GENERATION

To achieve this, we used an IDS to listen to all network traffic, analyze and detect malicious flows.

The IDS analyzes the network data and detects anomalies or attack patterns predefined by the blockchain network administrators. This detection is mainly based on the analysis of the network and transport layer packet headers but also on the packet content. To detect a malicious flow, the IDS mainly uses two analysis methods, namely the signature-based detection method which allows to detect known patterns in the analyzed data, or the behavioral detection method which detects deviations of a behavior from a normal profile. In both cases, the IDS compares the analyzed data to a reference described either by a signature or by a normal profile. Once the data is analyzed, the IDS can generate an alert in the form of a log file in case of malicious flows.

4. IMPLEMENTATION OF THE BCSDN-IoT SOLUTION

Our implementation model is entirely realized in a virtual environment with open source tools.

4.1 INSTALLING OPENDAYLIGHT

The OpenDaylight controller is an open source network operating system software developed in Java and managed by the Linux Foundation. It is based on a modular architecture and exchanges with SDN applications using the Northbound API. OpenDaylight communicates with network devices using its Southbound API. The most commonly used Southbound API in SDN is OpenFlow.

To experiment with our solution, we created a virtual machine with 2 CPUs and 16GB of RAM with an Ubuntu 16.04 operating system on the VMware platform. Then we installed on this machine an OpenDaylight SDN controller Beryllium- SR4 version.

Once OpenDaylight was installed, we added features such as odl-l2switch-switch, odl-dlux-all and odl-restconf to support Layer 2/3 switches, web interface and communicate with applications via the REST API. It is also important to enable OpenFlow version 1.3 by adding the -of13 option on the launch script file, as OpenFlow version 1.0 is implemented on the OpenDaylight controller by default. OpenDaylight provides several types of features to use as needed.

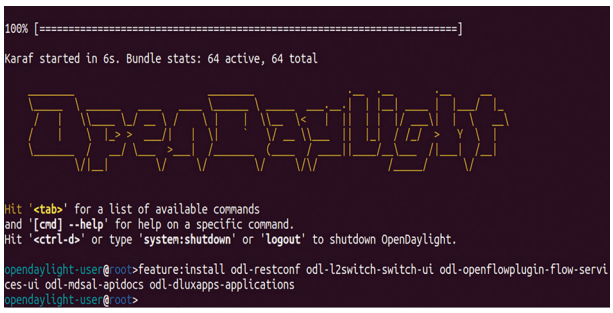


Fig. 6. Installing OpenDaylight

Once OpenDaylight was installed, we added features such as odl-l2switch-switch, odl-dlux-all and odl-restconf to support Layer 2/3 switches, web interface and communicate with applications via the REST API. It is also important to enable OpenFlow version 1.3 by adding the -of13 option on the launch script file, as OpenFlow version 1.0 is implemented on the OpenDaylight controller by default. OpenDaylight provides several types of features to use as needed.

To make a Layer 2/3 OSI routing decision, the OpenDaylight controller knows the network topology, as well as the devices that are connected with their identifiers (IP addresses and MAC addresses). Using OpenFlow 1.3, the OpenDaylight controller configures an OVS switch and manages and updates the OpenFlow network.

4.2 REALIZATION OF THE ARCHITECTURE

Most of the works in the literature use the mininet network simulator to experiment the SDN network. We have chosen to use virtual machines in a production environment on VMvare platform, to be in a real use case.

To realize our virtual network architecture, we created a second virtual machine with an Ubuntu 16.04 operating system, 2 virtual CPU and 16GB of RAM on a VMvare platform. On this machine, we installed an OpenFlow 1.3 compatible virtual switch (OVS version 2.6.0) and Qemu (Quick Emulator), an open source virtual machine emulator on x86 architecture.

The OVS is an open source software implementation of an Ethernet switch with the particularity of being multilayer and distributed. It is designed to work as an OSI level 2/3 switch in virtual machine environments supporting different protocols and standards, including the OpenFlow protocol. In our work, it has allowed us to make client virtual machines communicate with each other. Qemu is used to emulate our client machines with an Alpine Linux operating system, an ultra light distribution of Linux with 48MB of RAM. We used the basic qemu-img tool to create and manage disk images. The qcow2 format is used in this work because it integrates more features like compression and encryption.

Then, we wrote a bash script to launch several qemu client virtual machines with the possibility to manage them remotely. The same script allows to launch the OVS to interconnect the Alpine Linux virtual machines

and to create the link between the OpenFlow switch and the OpenDaylight controller, to allow the latter to control the network via the OpenFlow 1 protocol.

A dynamic allocation of IPv4 addresses in DHCP of the virtual machines clients of the network is made by the same code. This is how we set up our OpenFlow network with the possibility of scaling up just by varying the number of virtual machines and number of virtual machines and OVS desired.

4.3 SECURING THE LINK BETWEEN THE OPENDAYLIGHT CONTROLLER AND THE OVS

As discussed earlier, the communication channel between the OVS and the OpenDaylight controller is not encrypted by default, which means that encryption of OpenFlow exchange messages between these two elements of the SDN network does not run automatically. In addition, some controllers do not even support TLS for encrypting communications between the SDN switch and the controller. A hacker can exploit this lack of security on the OpenFlow channel to attack the network and conduct malicious actions. This is extremely dangerous if the hacker gains access to the controller that would give him control over the entire network. With a grip on the controller, the hacker can remove OpenFlow switches, modify OpenFlow rules in the switch, capture sensitive traffic and monitor how the controller handles OpenFlow packets. For this reason, SSL/TLS encryption of OpenFlow message exchanges on the channel between the OVS and OpenDaylight is required.

The encryption of OpenFlow messages between the OVS and OpenDaylight is done using an SSL/TLS connection, based on the Public Key Infrastructure (PKI) model.

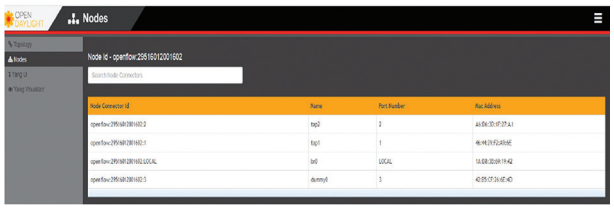


Fig. 7. Node Inventory on the OpenDaylight Beryllium-SR4

Using the OpenSSL encryption toolkit, we generated a keyStore, a file containing the controller's private and public keys. Then, the key file is imported into a JKS format key file, adapted to be configured on the OpenFlow configuration file of the OpenDaylight controller.

4. 4 SOME EXAMPLES OF SIMULATED ATTACKS

To simulate an attack and see if the IDS detects it or not, we installed the Nmap tool on one of the client virtual machines. Then, we successively launched a denial of service attack, a port scan and an IP address spoofing with the specific Nmap command on a case by case basis.

- **Denial of service**

The objective here is to detect and block attempts to saturate a target machine with DoS attacks using the ICMP protocol. We proceeded to send ICMP requests to the second machine of our network, in order to see if the Snort IDS reacted by detecting the unwanted flows. With this example, we found that after this ICMP request, our Snort IDS detected and saved a log file on the specific directory of the Snort server. This type of attack attempt can make the controller or a machine unavailable to its users. It interrupts or suspends network services temporarily or indefinitely.

With the proposed solution, it is possible to block the communication of malicious nodes in an automated way.

- **Port scan**

In this case, the goal is to detect port scan attempts on TCP and UDP protocols and to block these requests from the same source with the Nmap tool. Nmap is an open source port scanning software designed to detect open ports and, more generally, to obtain information about the operating system of a remote computer. To find out which ports are open on a machine, Nmap sends a packet to all ports on the target machine and analyzes the responses.

To simulate port scanning, we installed the Nmap tool on one of the Linux host alpine machines on our network. Then, we launched a port scan on one of the machines of the network with the specific command (nmap -p "*" Ip address target machine) and in the same way, Snort detected this attack attempt and recorded the corresponding log.

- **IP or MAC address spoofing**

MAC address spoofing is when a malicious attacker attempts to spoof a legitimate MAC or IP address in order to send packets to the network, using a trusted address. MAC/IP address replication forces systems to believe that the source is trustworthy.

In the same way as port scanning, we experimented with Nmap, and through the specific command (nmap spoof-mac target machine MAC address or target machine IP address), IP and MAC address spoofing and found that Snort detected the threat and logged the associated log.

We noticed that Snort detected all the attacks and saved the corresponding files in the log directory. This procedure can be extended to other types of more complex and intelligent threats.

5. CONCLUSION

In this paper, based on an analysis of the challenges faced by large-scale IoT networks due to new communication paradigms, BCSDN-IoT, a novel distributed secure IoT network architecture composed of an SDN backbone using blockchain technology, has been proposed to address current and future challenges and

satisfy new service requirements. BCSDN-IoT improves the performance and capacity of a system. The primary role of the BCSDN-IoT model is to generate and deploy protections, including threat prevention, data protection, and access control, and mitigate network attacks such as cache poisoning/ARP spoofing, DDoS/DoS attacks, and detect security threats. The BCSDN-IoT approach also focuses on minimizing attack window time by allowing IoT forwarding devices to check and download the most recent flow rule table if necessary. The performance evaluation is based on the scalability, defense effects, accuracy rates and performance overhead of the proposed model.

6. REFERENCES

[1] H. Hamed, D. Ali, M. Reza, A. Mohammed, K. Hadis, "A Survey on Internet of Things Security: Requirements, Challenges, and Solutions", Internet of things, Vol. 14, 2021, 100129.

[2] H.W. Rolf, "Internet of Things – Need for a New Legal Environment?", Computer Law & Security Review, Vol. 25, 2009, pp. 522–527.

[3] P. Sanghera, T. Frank, "How to Cheat at Deploying and Securing RFID", Chapter 15 - RFID Security: Attacking the Backend", How to Cheat, 2007, pp. 311-321.

[4] J. Sweta, C. Pruthviraj, S. Ayushi, "The Fundamentals of Internet of Things: Architectures, Enabling Technologies, and Applications", Healthcare Paradigms in the Internet of Things Ecosystem, 2021, pp. 1–20.

[5] O. Flauzac, C. Gonzalez, F. Nolot, "New Security Architecture for IoT Network", Procedia Computer Science, Vol. 52, 2015, pp. 1028–1033.

[6] A. I. Sanka et al. "A Survey of Breakthrough in Blockchain Technology: Adoptions, Applications, Challenges and Future Research", Computer Communications, Vol. 169, 2021, pp. 179–201.

[7] Internet of Things Global Standards Initiative, ITU 2021.

[8] A. Mayuri, T. Sudhir, "Internet of Things: Architecture, Security Issues and Countermeasures", International Journal of Computer Applications, Vol. 125, No. 14, 2015, pp. 1–4.

[9] Software-Defined Networking (SDN) Definition Open Networking Foundation, https://www.opennetworking.org/sdn-definition (accessed: 2018)

[10] Openflow-switch-v1.5.1, https://www.opennetworking.org/wpcontent/uploads/2014/10/openflow-switch-v1.5.1 (accessed: 2018).

[11] J. Halpern, S. J. Hadi, "Forwarding and Control Element Separation (ForCES) Forwarding Element Model", Internet Engineering Task Force , 2010.

[12] B. Pfaff, B. Davie, "The Open vSwitch Database Management Protocol", RFC Editor, RFC7047, 2013.

[13] F. R. Thomas, "Architectural Styles and the Design of Network-based Software Architectures", Information and Computer Science, 2000, pp. 180.

[14] C. Konstantinos, D. Michel, "Blockchains and smart contracts for the internet of things", IEEE Access, Vol. 4, 2016, pp. 2292–2303.

[15] B. Mandrita, L. Junghee, R. Kim-Kwang, "A Blockchain Future for Internet of Things Security: A Position Paper", Digital Communications and Networks, Vol. 4, No. 3, 2018, pp. 149–160.

[16] G. Sangeeta, A. Kavita, "Essentials of Blockchain Technology for Modern World Applications", Materials Today: Proceedings, 2021.

[17] T. Christos, B. Konstantinos, G. Panagiotis, A. Stavros, D. Tasos, "Malicious threats and novel security extensions in p2psip", Proceedings of the IEEE International Conference on Pervasive Computing and Communications Workshops, Lugano, Switzerland, 19-23 March 2012, pp. 746–751.

[18] S. Jie, Z. Pengyi, A. Mohammed, B. Yubin, Ge YU, "Research Advances on Blockchain-as-a-Service: Architectures, Applications and Challenges", Digital Communications and Networks, 2021.

[19] M. J Islam, M. Mahin, S. Roy, B. Debnath, A. Khatun, "DistBlackNet: A distributed secure black SDN-IoT architecture with NFV implementation for smart cities", Proceedings of the International Conference on Electrical, Computer and Communication Engineering, Cox'sBazar, Bangladesh, 7-9 February 2019, pp. 1–6.

[20] M. A. Ferrag, M. Derdour, M. Mukherjee, A. Derhab, L. Maglaras, H. Janicke, "Blockchain technologies for the Internet of Things: Research issues and challenges", IEEE Internet Things, Vol. 6, No. 2, 2019, pp. 2188–2204.

[21] P. Sharma, S. Singh, Y. Jeong, J.H. Park, "DistBlockNet: A distributed blockchains-based secure SDN architecture for IoT networks", IEEE Communications Magazine, Vol. 55, No. 9, 2017, pp. 78–85.

[22] C. Qiu, F. R. Yu, F. Xu, H. Yao, C. Zhao, "Permissioned blockchain based distributed software-defined industrial Internet of Things", Proceedings of the IEEE Globecom Workshops, 2018, Abu Dhabi, United Arab Emirates, 9-13 December 2018, pp. 1-7.

[23] R. Kumar, R. Sharma, "Leveraging Blockchain for Ensuring Trust in IoT: A Survey", Journal of King Saud University - Computer and Information Sciences, 2021.

[24] S. A. Latif et al, "AI-Empowered, Blockchain and SDN Integrated Security Architecture for IoT Network of Cyber Physical Systems", Computer Communications, Vol. 181, 2022, pp. 274–283.

[25] S. Rathore, B. W. Kwon, J. H. Park, "BlockSecIoTNet: Blockchain-Based Decentralized Security Architecture for IoT Network", Journal of Network and Computer Applications, Vol. 143, 2019, pp. 167–177.

[26] P. Podili, K. Kataoka, "TRAQR: Trust Aware End-to-End QoS Routing in Multi-Domain SDN Using Blockchain", Journal of Network and Computer Applications, Vol. 182, 2021, 103055.

[27] P. P. Ray, N.j Kumar, "SDN/NFV Architectures for Edge-Cloud Oriented IoT: A Systematic Review", Computer Communications, Vol. 169, 2021, pp. 129–153.

[28] Md. A. Uddin et al. "Blockchain Leveraged Decentralized IoT EHealth Framework," Internet of Things, Vol. 9, 2020, 100159.

[29] V. Balasubramanian, M. Aloqaily, M. Reisslein, "An SDN Architecture for Time Sensitive Industrial IoT", Computer Networks, Vol. 186, 2021, 107739.

INTERNATIONAL JOURNAL OF ELECTRICAL AND COMPUTER ENGINEERING SYSTEMS

Published by Faculty of Electrical Engineering, Computer Science and Information Technology Osijek,
Josip Juraj Strossmayer University of Osijek, Croatia.

About this Journal

The International Journal of Electrical and Computer Engineering Systems publishes original research in the form of full papers, case studies, reviews and surveys. It covers theory and application of electrical and computer engineering, synergy of computer systems and computational methods with electrical and electronic systems, as well as interdisciplinary research.

Topics of interest include, but are not limited to:

- Power systems
- Renewable electricity production
- Power electronics
- Electrical drives
- Industrial electronics
- Communication systems
- Advanced modulation techniques
- RFID devices and systems
- Signal and data processing
- Image processing
- Multimedia systems
- Microelectronics
- Instrumentation and measurement
- Control systems
- Robotics
- Modeling and simulation
- Modern computer architectures
- Computer networks
- Embedded systems
- High-performance computing
- Parallel and distributed computer systems
- Human-computer systems
- Intelligent systems
- Multi-agent and holonic systems
- Real-time systems
- Software engineering
- Internet and web applications and systems
- Applications of computer systems in engineering and related disciplines
- Mathematical models of engineering systems
- Engineering management
- Engineering education

Paper Submission

Authors are invited to submit original, unpublished research papers that are not being considered by another journal or any other publisher. Manuscripts must be submitted in doc, docx, rtf or pdf format, and limited to 30 one-column double-spaced pages. All figures and tables must be cited and placed in the body of the paper. Provide contact information of all authors and designate the corresponding author who should submit the manuscript to https://ijeces.ferit.hr. The corresponding author is responsible for ensuring that the article's publication has been approved by all coauthors and by the institutions of the authors if required. All enquiries concerning the publication of accepted papers should be sent to ijeces@ferit.hr.

The following information should be included in the submission:

- paper title;
- full name of each author;
- full institutional mailing addresses;
- e-mail addresses of each author;
- abstract (should be self-contained and not exceed 150 words). Introduction should have no subheadings;
- manuscript should contain one to five alphabetically ordered keywords;
- all abbreviations used in the manuscript should be explained by first appearance;
- all acknowledgments should be included at the end of the paper:
- authors are responsible for ensuring that the information in each reference is complete and accurate. All references must be numbered consecutively and citations of references in text should be identified using numbers in square brackets. All references should be cited within the text;
- each figure should be integrated in the text and cited in a consecutive order. Upon acceptance of the paper, each figure should be of high quality in one of the following formats: EPS, WMF, BMP and TIFF;
- corrected proofs must be returned to the publisher within 7 days of receipt.

Peer Review

All manuscripts are subject to peer review and must meet academic standards. Submissions will be first considered by an editor-in-chief and if not rejected right away, then they will be reviewed by anonymous reviewers. The submitting author will be asked to provide the names of 5 proposed reviewers including their e-mail addresses. The proposed reviewers should be in the research field of the manuscript. They should not be affiliated to the same institution of the manuscript author(s) and should not have had any collaboration with any of the authors during the last 3 years.

Author Benefits

The corresponding author will be provided with a .pdf file of the article or alternatively one hardcopy of the journal free of charge.

Units of Measurement

Units of measurement should be presented simply and concisely using System International (SI) units.

Bibliographic Information

Commenced in 2010.
ISSN: 1847-6996
e-ISSN: 1847-7003

Published: semiannually

Copyright

Subscription Information

The annual subscription rate is 50€ for individuals, 25€ for students and 150€ for libraries.

Postal Address

Faculty of Electrical Engineering,
Computer Science and Information Technology Osijek,
Josip Juraj Strossmayer University of Osijek, Croatia
Kneza Trpimira 2b
31000 Osijek, Croatia

IJECES Copyright Transfer Form

(Please, read this carefully)

This form is intended for all accepted material submitted to the IJECES journal and must accompany any such material before publication.

TITLE OF ARTICLE (hereinafter referred to as "the Work"):

COMPLETE LIST OF AUTHORS:

The undersigned hereby assigns to the IJECES all rights under copyright that may exist in and to the above Work, and any revised or expanded works submitted to the IJECES by the undersigned based on the Work. The undersigned hereby warrants that the Work is original and that he/she is the author of the complete Work and all incorporated parts of the Work. Otherwise he/she warrants that necessary permissions have been obtained for those parts of works originating from other authors or publishers.

Authors retain all proprietary rights in any process or procedure described in the Work. Authors may reproduce or authorize others to reproduce the Work or derivative works for the author's personal use or for company use, provided that the source and the IJECES copyright notice are indicated, the copies are not used in any way that implies IJECES endorsement of a product or service of any author, and the copies themselves are not offered for sale. In the case of a Work performed under a special government contract or grant, the IJECES recognizes that the government has royalty-free permission to reproduce all or portions of the Work, and to authorize others to do so, for official government purposes only, if the contract/grant so requires. For all uses not covered previously, authors must ask for permission from the IJECES to reproduce or authorize the reproduction of the Work or material extracted from the Work. Although authors are permitted to re-use all or portions of the Work in other works, this excludes granting third-party requests for reprinting, republishing, or other types of re-use. The IJECES must handle all such third-party requests. The IJECES distributes its publication by various means and media. It also abstracts and may translate its publications, and articles contained therein, for inclusion in various collections, databases and other publications. The IJECES publisher requires that the consent of the first-named author be sought as a condition to granting reprint or republication rights to others or for permitting use of a Work for promotion or marketing purposes. If you are employed and prepared the Work on a subject within the scope of your employment, the copyright in the Work belongs to your employer as a work-for-hire. In that case, the IJECES publisher assumes that when you sign this Form, you are authorized to do so by your employer and that your employer has consented to the transfer of copyright, to the representation and warranty of publication rights, and to all other terms and conditions of this Form. If such authorization and consent has not been given to you, an authorized representative of your employer should sign this Form as the Author.

Authors of IJECES journal articles and other material must ensure that their Work meets originality, authorship, author responsibilities and author misconduct requirements. It is the responsibility of the authors, not the IJECES publisher, to determine whether disclosure of their material requires the prior consent of other parties and, if so, to obtain it.

- The undersigned represents that he/she has the authority to make and execute this assignment.
- For jointly authored Works, all joint authors should sign, or one of the authors should sign as authorized agent for the others.
- The undersigned agrees to indemnify and hold harmless the IJECES publisher from any damage or expense that may arise in the event of a breach of any of the warranties set forth above.

_____ _____

Author/Authorized Agent **Date**

CONTACT
International Journal of Electrical and Computer Engineering Systems (IJECES)
Faculty of Electrical Engineering, Computer Science and Information Technology Osijek
Josip Juraj Strossmayer University of Osijek
Kneza Trpimira 2b
31000 Osijek, Croatia
Phone: +38531224600,
Fax: +38531224605,
e-mail: ijeces@ferit.hr