Hybridization of KF and GA as a Predictive Algorithm for Reduction of Power Consumption in a Cloud Computing Data Centre

Afolabi Rotimi - **afolabi.rotimipgs@stu.cu.edu.ng** **,** *Department of Electrical & Communication Engineering, Covenant University, Ota, Nigeria*

Bamidele Adebisi - **b.adebisi@mmu.ac.uk** **,** *Department of Electrical & Communication Engineering, Covenant University, Ota, Nigeria*

Anthony U. Adoghe - **anthony.adoghe@covenantuniversity.edu.ng**
Department of Engineering, Manchester Metropolitan University, Manchester, UK

Abstract: Data Centre (DC) has becoming a major and important component of a cloud computing to meet up with the rapid increase in the demand of telecommunication services. However, the cost of maintaining a DC is very high due to high power consumption of the unit in a telecommunication industry. The situation is further exacerbated in a country like Nigeria where there is highly unstable power supply from the national grid. Kalman Filter (KF) which is one of the energy consumption optimization technique through power consumption prediction model used to reduce the power consumption is characterized with high prediction error due to random selection of KF parameters' value during the prediction period. Hence, in this paper, Hybridization of KF and GA for power reduction through accurate power consumption prediction is proposed. Data were collected from Four different servers in Nigeria, named BSC 13, BSC 14, RNC 05 and RNC 06 using power analyzer, multimeters and thermometer. The historical assessment of data collected were carried out for the DC for two years (January to December of 2019 and 2020). The GA was used to obtain best possible values for the KF parameters and KF was then used to predict the future power consumption value on hourly basis for each day of the week. The proposed PCoKFGA model gave better performance with accurate prediction, higher power usage effectiveness and lower energy consumption than the existing KF model. Therefore, the PCoKFGA model proposed would be most useful where accuracy in prediction is of utmost importance, and for run-time application.

Keywords: Kalman Filter (KF), Genetic Algorithm (GA), Data Centre (DC), Energy Consumption (EC) and Power Usage Effectiveness (PUE)

1. INTRODUCTION

Data Centre (DC) is a physical infrastructure that houses network servers and energy storage facilities in which telecommunication companies use for processing, organizing, disseminating and storing large quanta of data in a cloud computing [1, 2]. Nowadays, telecommunication sector depends majorly on the services, applications and processed in data centres, making it an important factor in the sector's ability to maintain robust day-to-day operations [3, 4]. DC has become an essential part of the smart world being the core centre of information processing and data storage for all telecommunication industries. It is generally associated with high energy demands which represents a major impact on the power system as well as a high cost of powering data centres [5, 6]. Nowadays, the demand for data related services continues to rise exponentially resulting in continuous increase in the load demand. However, DC is characterized with high power consumption that result in high cost of maintaining data centre. Therefore, the installed DC must be intelligent and efficient to ensure system reliability, protection, robustness by reducing the power consumption [7, 8, 9]. The total energy consumption in DC is further expected to grow in the coming years because of an increase in data demand. Little reduction in power consumption by a small percentage will result in a huge amount of savings in the operation cost of data centres [10, 11]. Several algorithms such as Kalman Filter (KF) have been proposed to optimized the power consumption of DC through power

3001

Hybridization of KF and GA as a Predictive Algorithm for Reduction of Power Consumption in a Cloud Computing Data Centre

Section A-Research paper

consumption predictive model. However, the existing KF algorithm is characterized with high prediction error due to random selection of KF parameters' value during the prediction period [12, 13, 14]. Hence, in this paper, Hybridization of KF and GA for power reduction through accurate power consumption prediction is proposed.

2. METHODOLOGY

2.1. POWER CONSUMPTION OPTIMIZATION USING KF WITH GA

The structure of the proposed model is shown in Fig. 1, the model measures the total traffic inflow for the current hour (h) and then forwards the value $A_r(h)$ to a logger. The logger supplies the previous hour's traffic size $A_r(h-1)$ to be multiplied with the value of k representing the current day of the week to estimate the power usage, P_{usage}, for the server. The estimated P_{usage} was forwarded to the power supply controller, which computes the total power to be consumed by the server in the current hour $P_S(h)$. The process was carried out every five minutes to the next hour of the day. Power consumption optimization technique using KF with GA technique is named as PCoKFGA and the structure for PCoKFGA technique is shown in Fig. 2. The GA optimizer is used to obtain best possible values for the KF parameters and KF model is used to predict the future power consumption value on hourly basis for each day of the week.
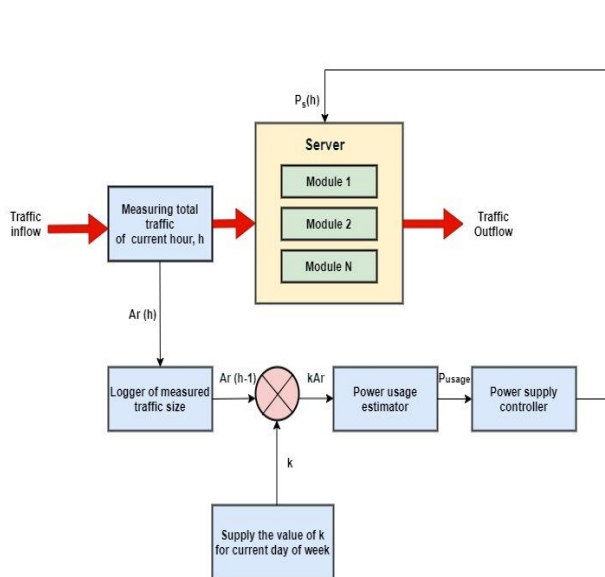

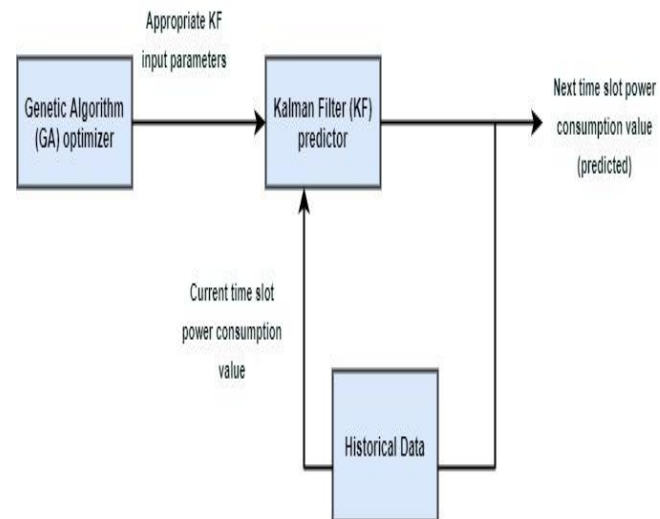Fig. 1: Server's power consumption prediction model

Fig. 2: The PCoKFGA technique for power consumption prediction model

2.2. KALMAN FILTER-BASED POWER CONSUMPTION ESTIMATION MODEL

A server's power consumption at a given time t, P_{S_t}, can be expressed as:

$$P_{S_t} = P_{S_{t-1}} + a_t \tag{1}$$

where $P_{S_{t-1}}$ is the power consumption as time $t-1$ and a_t is the variation of power consumption at time t. The parameters P_{S_t} and a_t are referred to as process state and process noise, respectively, in KF terminology. The measured power consumption at time t, Z_t, is expressed as

$$Z_t = P_{S_t} + b_t \tag{2}$$

where b_t is the measurement noise. The prediction of power consumption for the next time slot $t+1$ is obtained from the current time slot t as

3002

Hybridization of KF and GA as a Predictive Algorithm for Reduction of Power Consumption in a Cloud Computing Data Centre

Section A-Research paper

$$\tilde{P}_{S_{t+1}} = A\hat{P}_{S_t} \qquad (3)$$

where $\tilde{P}_{S_t}$ is a priori estimate of P_{S_t}

$\hat{P}_{S_t}$ is a posteriori estimate of P_{S_t}

A is the transition matrix

The state covariance matrix is given as

$$\tilde{C}_{t+1} = A\hat{C}_t A^T + Q_t \qquad (4)$$

where C is the state noise covariance matrix

Q is the process noise covariance matrix

The next stage is the measurement update which begins as

$$K_t = \tilde{C}_t H^T S_t^{-1} \qquad (5)$$

with

$$S_t = H\tilde{C}_t H^T + R_t \qquad (6)$$

where K is Kalman gain matrix.

H is design matrix

S is the residual covariance

R is the measurement noise covariance matrix

The process state is corrected as

$$\hat{P}_{S_t} = \tilde{P}_{S_t} + K_t(Z_t - H\tilde{P}_{S_t}) \text{ and} \qquad (7)$$

$$\hat{C}_t = (1 - K_t H)\tilde{C}_t \qquad (8)$$

where Z is the measurement vectors.

The design matrix, H, is a 2x4 matrix such that $h_{ij} \in \{0,1\}$, and the set of H can be given as

$$H = \begin{pmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \end{pmatrix} \qquad (9)$$

Similarly, the transition matrix, A, is a 4x4 matrix such that $a_{ij} \in \{0,1\}$, and the set of A is given as

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \qquad (10)$$

Furthermore, the measurement noise, R, is a Gaussian random variable with possible value as $0 < R \leq 1$. It is worth noting that the elements of R have the same value.

2.3 OPTIMIZING THE KF MODEL PARAMETERS USING GA

The values of the KF parameters: H, A and R have significant effect on the estimation or prediction accuracy of the KF model; hence, appropriate values must be chosen for the parameters. This is to be achieved by the help of Genetic Algorithm (GA). The GA is an optimization tool applicable to non-linear problems. The GA consists of five processes namely; Initialization, Objective function, Selection, Crossover and Mutation.

i. **Initialization process:** This process involves the initialization of the population. Each (chromosome) in the GA population is a possible solution set of the KF parameters H, A and R. Each element of H and A takes a bit 0 or 1, while each element of R takes a random value that is greater than 0 but less than or equal to 1.

ii. **Objective function process:** The objective function for the PCoKFGA technique is formulated as

$$Minimize \ \frac{\sum_{t=1}^{N}(P_{S_t} - \hat{P}_{S_t})^2}{N} \qquad (11)$$

$$subject \ to: h_{ij} \in \{0,1\}$$

$$a_{ij} \in \{0,1\}$$

$$0 < r \leq 1$$

where P_{S_t} is the actual power consumption value at t-th time

$\hat{P}_{S_t}$ is the predicted power consumption value t-th time

N is the number of times (or samples)

Hybridization of KF and GA as a Predictive Algorithm for Reduction of Power Consumption in a Cloud Computing Data Centre

Section A-Research paper

ISSN 2063-5346

h_{ij}, a_{ij} and r are the elements of KF parameters H, A and R, respectively

The objective function is to minimize the Mean Squared Error (MSE) between the actual and predicted values. The objective function is evaluated on each chromosome in the GA population, and the chromosome that gives the lowest MSE is the fittest in the population.

iii. **Selection process:** The process involves the selection of fit chromosomes for reproduction. The chromosomes are sorted by their fitness values and the fit chromosomes are selected to replace the weak parent chromosomes in the population.

iv. **Crossover process:** In this process, some of the genes between two parent chromosomes are exchanged to produce offspring (a child chromosome). The genes to crossover are randomly selected with a specified crossover probability. Only the selected parent chromosomes from the selection stage are used to reproduce two new offspring for the next generation.

v. **Mutation process:** This process performs little changes to some genes of a chromosome. A boundary is chosen within the chromosome and some genes can be mutated with a specified probability of mutation. This helps to ensure that all the chromosomes in the population do not have the same genes in a given generation.

2.4   Energy Consumption (EC)

The Energy Consumption (EC) of a data centre (DC) should be relatively low, depending on the traffic load, for energy efficiency [15]. The EC of the DC is a function of the total power consumed by the DC. The EC of the DC for the proposed model is obtained as

$$E_{DC} = \sum_{n=1}^{N}\left(P_{n\_avg} \times T\right) \qquad (12)$$

where P_{n_avg} is the average power consumption of the nth server

T is the total operating time of the DC

N is the number of active servers in the DC

2.5   Power Usage Effectiveness (PUE)

The PUE is a metric that gives the information of the level of effective utilization of the Information and Communication Technology (ICT) equipment in a DC. The higher the value of PUE of a DC the lower the efficiency of the DC, and the ideal value for PUE is 1 [16, 17]. The power usage effectiveness for the proposed model is obtained as

$$PUE = \frac{Total\ availably\ DC\ power}{Total\ power\ consumed\ by\ ICT\ equipment} \qquad (13)$$

3.   RESULTS AND DISCUSSION

3.1. Prediction of power consumption of DC using the proposed PCoKFGA model

The prediction of power consumption was optimized for more accuracy which was investigated in this paper using the proposed PCoKFGA model. Fig. 3 shows the PCoKFGA prediction for Sunday 2019 and 2020. For the year 2019, the predicted and actual average DC power consumption for the day was 12.9698 kW and 12.9474 kW, respectively, which gives an absolute error of 0.0017 (0.2%). For the year 2020, the predicted and actual average DC power consumption for the day was 14.6146 kW and 14.5727 kW, respectively, which gives an absolute error of 0.0029 (0.3%). Fig. 4 shows the PCoKFGA prediction for Wednesday 2019 and 2020. For the year 2019, the predicted and actual average DC power consumption for the day was 13.0528 kW and 13.0323 kW, respectively, which gives an absolute error of 0.0016 (0.2%). For the year 2020, the predicted and actual average DC power consumption for the day was 14.5496 kW and 14.5140 kW, respectively, which gives an absolute error of 0.0025 (0.25%). Fig. 5 shows the PCoKFGA prediction for Friday 2019 and 2020. For the year 2019, the predicted and actual average DC power consumption for the day was 13.0384 kW and 13.0179 kW, respectively, which gives an absolute error of 0.0016 (0.2%). For the year 2020, the predicted and actual average DC power consumption for the

3004

Eur. Chem. Bull. 2023, 12 (Special Issue 6), 3001-3008

Hybridization of KF and GA as a Predictive Algorithm for Reduction of Power Consumption in a Cloud Computing Data Centre

Section A-Research paper

day was 14.5385 kW and 14.5036 kW, respectively, which gives an absolute error of 0.0024 (0.2%). The average DC power consumption predictions by the PCoKFGA model revealed that for all the days of the week, the calculated absolute errors between the predicted and actual were less than 0.0029 (0.3%). This implies that the PCoKFGA model is suitable for predicting the average daily and hourly power consumption.
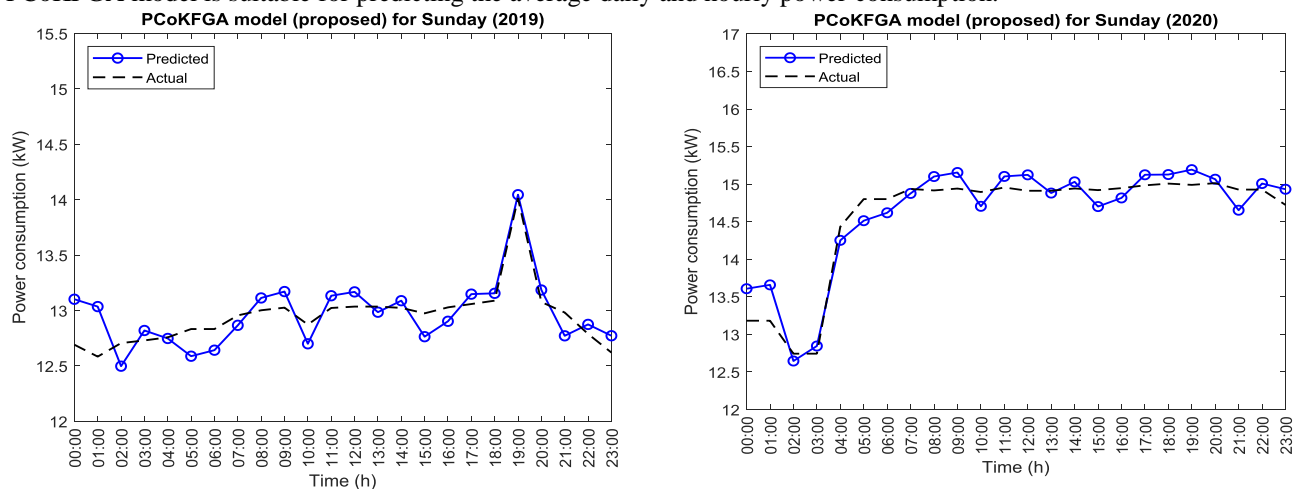

Fig. 3: Prediction of power consumption of data centre by the PCoKFGA model for Sunday
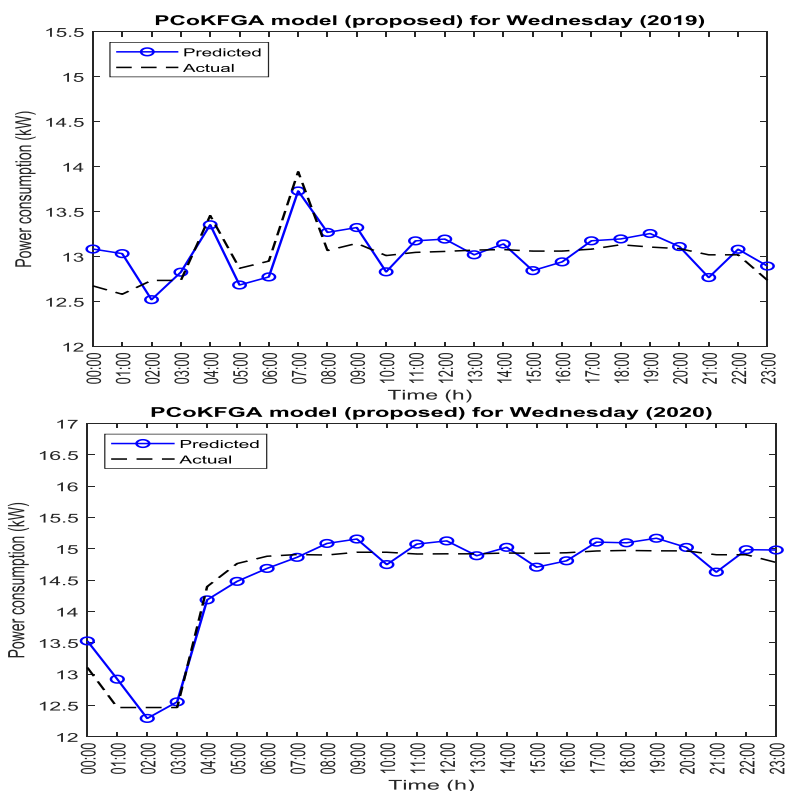

Fig. 4: Prediction of power consumption of data centre by the PCoKFGA model for Wednesday
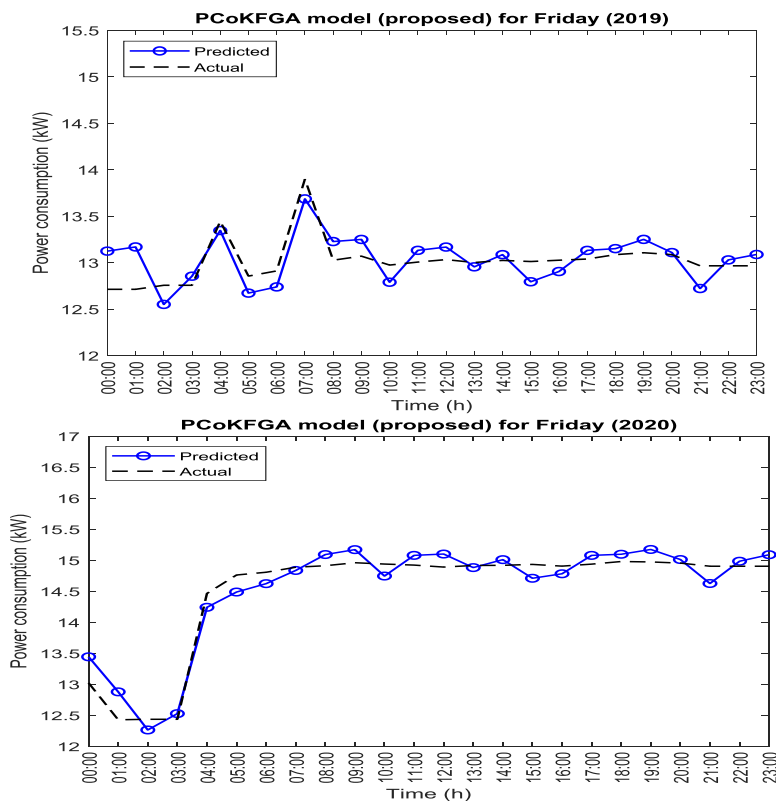
3005

Eur. Chem. Bull. 2023, 12 (Special Issue 6), 3001-3008

Hybridization of KF and GA as a Predictive Algorithm for Reduction of Power Consumption in a Cloud Computing Data Centre

Section A-Research paper

ISSN 2063-5346

Fig. 5: Prediction of power consumption of data centre by the PCoKFGA model for Friday

3.2. EFFECTIVENESS OF THE PROPOSED PCOKFGA MODEL

The PCoKFGA model, which was also developed in this study, is applicable for forecasting the next hour's power consumption by using the knowledge of some immediate previous hour's values. The PCoKFGA model was able to achieve almost accurate predictions due to the use of the best possible KF parameters' values that were optimized by a genetic algorithm. Consequently, the PCoKFGA model gave a relatively good average prediction absolute error of 0.002 (0.2%) and 0.003 (0.3%), for the years 2019 and 2020, respectively. Furthermore, it is worthy of note that the prediction error of the PCoKFGA model is negligible; hence, the power consumption value predicted by the PCoKFGA model requires no adjustment and can be used directly in the power supply budget. Thus, the model is of practical use where a high level of prediction accuracy is required. On average, the developed PCoKFGA model (proposed) outperforms all the other prediction models investigated in this study in the area of prediction error including KF without GA, but inferior to the ASUS model in the area of implementation cost. Furthermore, the computational complexity of PCoKFGA is relatively higher than all the other models except the ARMA and Conventional method.

3.3   ENERGY CONSUMPTION (EC) AND POWER USAGE EFFECTIVENESS (PUE)

The calculated energy consumption (EC) and power usage effectiveness (PUE) of the models for the year 2019 is presented in Table 1. The lowest possible EC and the PUE of approximately 1.0 is the target of any data centre (Uster *et al.* 2020; Rezaei-Mayahi *et al.* 2019). The PCoKFGA model gave the lower EC of 313.354 kWh

3006

Eur. Chem. Bull. 2023, 12 (Special Issue 6), 3001-3008

while the existing KF model gave the highest EC of 633.600 kWh. The calculated ECs and PUEs for the year 2020 are contained in Table 2. The proposed PCoKFGA model gave the lower EC with the value of 343.410 kWh, as against 633.600 kWh obtained with the existing KF model.

Table 1: EC and PUE of data centre for the year 2019

Model	EC (kWh)	PUE
PCoKFGA (proposed)	313.354	1.68
KF (existing)	336.459	1.68

Table 2: EC and PUE of data centre for the year 2020

Model	EC (kWh)	PUE
PCoKFGA (proposed)	349.918	1.54
KF (existing)	375.060	1.54

4. CONCLUSION

In this paper, real-time data was obtained from a DC for the years 2019 and 2020 to investigate the applicability of some power consumption prediction models. The analysis carried out on the DC's data revealed that the amount of power consumed by a DC has a direct proportionality to the network traffic being handled by the DC. Furthermore, ambient temperature of the DC could also have an impact on the power consumption. In addition, it was observed in this study from the comparison between the data of 2019 and 2020 that the average daily power consumption levels in a given year may be different from that of another year. Hence, 2019 data could not be used to create a model for forecasting 2020 data. However, the data of the first few weeks of a given year is suitable for creating a prediction model for the rest of that year. The performances of the proposed PCoKFGA model was compared the existing KF model using prediction error measurements, energy consumption and Power Usage Effectiveness (PUE). The PCoKFGA model's prediction was based on estimation using Kalman filter (KF) and optimized KF parameters by genetic algorithm. The proposed PCoKFGA model gave better performance with accurate prediction, higher PUE and lower EC than the existing KF model. Therefore, the PCoKFGA model proposed would be most useful where accuracy in prediction is of utmost importance, and for run-time application.

REFERENCE

[1]   Krishan, T. and Amanpreet, K. (2021). Comparisons of Energy Efficiency Management Techniques in Cloud Based Environment, Turkish Journal of Computer and Mathematics Education, 12(2):491-499

[2]   Dhaya, R., Ujwal, U. J., Tripti, S. Prabhdeep, S., Kanthavel, R., Senthamil, S. and Daniel, K. (2022). Energy-Efficient Resource Allocation and Migration in Private Cloud Data Centre

[3]   Kazi, M. U. A., Math, H. J. B. And Manuel, A. (2021). A Review of Data Centers Energy Consumption and Reliability Modeling, IEEE Access, pp 1-28.

[4]   Emmanuel, O., Rose, O. A., Aliyu, A., Agbata, B. C. and Akpa, J. (2022). The Impact of Green Cloud-base Computing on Environmental Sustainability in Nigeria, Asian Journal of Research in Computer Science, 3(5): 51-60.

[5]   Kalyan, D., Satyabrata, D., Rabi, K. D. and Ananya, M. (2018). Survey of Energy-Efficient Techniques for the Cloud-Integrated Sensor Network, Hindawi Journal of Sensors, pp 2-15.

[6]   Imad, I., Imane, D., Ouadoudi, Z., Mohammed, R. and Driss, A. (2019) A Kalman Filter Process for Energy Optimization in WSNs, Journal of Communications Software and Systems 15 (1):9-18.

[7]   Ali, S. M., Jawad, M., Khan, M. U. S., Bilal, K., Glower, J., Smith, S. C., ... & Zomaya, A. Y. (2017). An ancillary services model for data centres and power systems. *IEEE Transactions on Cloud Computing*. https://ieeexplore.ieee.org/abstract/document/7918513

[8]   Clito, A. and João, M. (2018). Data Centers and their Energy Consumption for Climatization, Proceedings IRF2018: 6th International Conference Integrity-Reliability-Failure Lisbon/Portugal, pp 1051-1057

[9]   Nagma, K., Jagpreet, S. and Jaiteg, S. (2019) Toward energy-efficient cloud computing: a survey of dynamic power management and heuristics-based optimization techniques, The Journal of Supercomputing, Springer, pp 5-58.

[10]   Forbes (2017). Innovation: Why Energy Is A Big and Rapidly Growing Problem for Data Centers, Retrieve from https://www.forbes.com/sites/forbestechcouncil/2017/12/15

[11]   Sharmistha, P., Devpriya, P. and Brojo, K. M. (2020) Energy Efficiency for Cloud Computing Applications: A Survey on the Recent Trends and Future Scopes, IEEE Xplore, pp 1-6.

[12]   Kliazovich, D., Bouvry, P., & Khan, S. U. (2013). DENS: data centre energy-efficient network-aware scheduling. *Cluster computing*, *16*(1), 65-75.

[13]   Gupta, A. (2016). United States Data Centre Energy Usage Report. http://networks.cs.ucdavis.edu/presentation2016/Gupta-07-01-2016.pdf.

[14]   Shehabi, A., Smith, S., Sartor, D., Brown, R., Herrlin, M., Koomey, J. & Lintner, W. (2016). United states data centre energy usage report (No. LBNL-1005775). Lawrence Berkeley National Lab. (LBNL), Berkeley, CA (United States). https://www.osti.gov/biblio/1372902

[15]   Zheng, L., Mian, Z. and Xusheng, Z. and Yun, L. (2020) A Non-Intrusive, Traffic-Aware Prediction Framework for Power Consumption in Data Centre Operations, Energies, pp 1-19

[16]   Uster, F., Plocksties, F. and Timmermann, D. (2020). Decentral load control for data centers, International Conferences on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), pp. 350-359.

[17]   Rezaei-Mayahi, M., Rezazad, M. and Sarbazi-Azad, H. (2019). Future Generation Computer Systems, 94(2019), pp. 130-139.

3008

Eur. Chem. Bull. 2023, 12 (Special Issue 6), 3001-3008