International Journal of Scientific Engineering and Research (IJSER)
ISSN (Online): 2347-3878
SJIF (2025): 8.036

Comparative Analysis of Forecasting Methods for 5G Using Real Time Mobile Network Data

Dr. Anupriya

Assistant Professor, Department of Computer Science, Govt. College for Women Gharaunda (Bastara) Karnal, Haryana, India

Abstract: ***The increasing complexity and heterogeneity of modern network traffic present significant challenges for accurate traffic classification in Software-Defined Networking (SDN) environments. Conventional classification methods repetitively struggle to comply with expeditiously evolving traffic behaviors, resulting in suboptimal resource management. To address this issue, an enhanced AGBFM (Adaptive Gradient Based Forecasting Model) and EWOFM (Enhanced Whale Optimization Based Forecasting Model) is proposed for efficient and accurate SDMN traffic classification. The incorporation of optimization algorithms improves the efficiency of the standard forecasting models, making it suitable for dynamic SDMN environments. Experimental results determine that the proposed models accomplishes a classification accuracy of 96.60%, surpassing LSSVM-PSO, LSSVM-ACO and LSSVM-WOA using the similar dataset scenarios. These findings show that AGBFM and EWOFM offer a lightweight yet effective solution for real-time traffic forecasting in SDMN.***

Keywords: Forecasting, Adaptive, Gradient, Optimization, traffic

1. Introduction

Performance analysis is utmost important for driving the advancements of software and hardware systems. An in-depth investigation can discover system and architectural bottlenecks, provide crucial information for selecting frameworks and platforms, and lead to performance improvement. Performance analysis is a major step towards performance optimizations. Various methods of optimization exist in order to achieve better performance of various machine learning algorithms. In this paper, performance of proposed forecasting models is compared with various existing models. The performance of proposed models Enhanced Whale Optimization Based Forecasting Model (EWOFM) [1] and Adaptive Gradient Based Forecasting Model (AGBFM) [2] is compared after the final iteration of the forecasting models with LSSVM-PSO, LSSVM-ACO and LSSVM-WOA using the similar dataset scenarios.

In recent years, network traffic prediction has emerged as a critical topic in the telecommunications industry. The exponential growth in mobile device usage and data services has posed significant challenges in managing and optimizing network capacity [3]. Accurate prediction models enable network operators to efficiently manage resources, prevent congestion, and enhance service quality for end-users [4]. Traffic forecasting has become a cornerstone of telecommunication network management and optimization, particularly in the context of Long-Term Evolution (LTE), which underpins modern mobile data communications. Understanding traffic patterns and characteristics, such as seasonality, trends, and the continuous rise in mobile cellular traffic, is essential for effective resource allocation and ensuring quality of service for users [5]. The value of traffic forecasting lies in its ability to provide insights into future network demand, allowing operators to anticipate and adapt to traffic pattern changes. By analyzing historical data and identifying trends and seasonal variations, operators can make informed decisions regarding network expansion, capacity upgrades, and resource allocation. This proactive approach mitigates network congestion, minimizes service disruptions, and enhances the overall user experience. The proposed forecasting models (AGBFM and EWOFM) are compared with three other models namely LSSVM-ACO, LSSVM-PSO and LSSVM-WOA. Although the meta-heuristic algorithms work really well in many domains of application, however, it is observed that LSSVM-ACO, LSSVM-PSO and LSSVM-WOA does not perform well as compared to the proposed ones, in forecasting traffic bursts. This is due to the fact that the first proposed model, adaptive gradient based optimizer (AGBO) can be directly integrated with proposed prediction model using parameter optimization (gamma and sigma). The second proposed model, IWOA enhances the exploration ability by using inertia weight factor whereas ACO and PSO cannot do optimization and proper exploration using LSSVM. Moreover, the search criteria of ACO and PSO include exploitation instead of inherent gradient based optimization in AGBFM. Also, the usage of input data for tuning of parameters makes the proposed algorithms (AGBFM and EWOFM) adaptive and efficient. The extra searches make the ACO and PSO loosely coupled optimizers. In case of PSO the convergence did happen faster, but it did not achieve good fitting cost.

Compared to ACO, the PSO achieved lower average fitting cost i.e. lower efficiency but not better than AGBFM and EWOFM.

The evaluation parameters, namely, MSE, accuracy, TPR, FPR, precision and F1-score are also calculated for all the existing and proposed forecasting models.

2. Related Work

Seasonality in traffic patterns refers to predictable, repetitive variations in traffic volume occurring at regular intervals, such as hourly, daily, weekly, or monthly [6]. Identifying and forecasting seasonal patterns optimize network resource utilization and enhance cost efficiency. Similarly, analyzing trends in traffic data is essential for long-term network planning [7]. Trends reflect the overall directional movement of traffic volume over time, capturing gradual changes driven by user behavior, technological advancements, or market dynamics. By recognizing and forecasting these trends, operators can anticipate future traffic demand, strategically plan network expansion, and implement upgrades to address evolving user needs. The continuous increase in mobile cellular traffic presents significant challenges for network operators. The proliferation of smartphones, IoT devices, and high-bandwidth applications has resulted in a consistent surge in data consumption, exerting immense pressure on existing network infrastructure [8]. Addressing this growing demand requires innovative forecasting methods and strategic planning to ensure sustainable and efficient network operations. Accurately forecasting mobile cellular traffic growth is critical to ensuring network scalability and maintaining quality of service. Advanced forecasting methods enable operators to predict future traffic levels and implement proactive measures to scale network capacity and optimize resource utilization. In the context of Base Station traffic, accurate forecasting is foundational for optimizing performance and resource allocation. As 5G technology continues to support high-speed data services and multimedia applications, the importance of precise traffic predictions becomes increasingly significant [9]. By anticipating traffic patterns and demand dynamics, operators can equip 5G Base Stations to provide seamless connectivity, high data rates, and low latency, meeting the growing demand for superior mobile broadband services. To identify the most effective prediction method for 5G network traffic, a comparative evaluation of AGBFM and EWOFM is essential. Each model offers distinct advantages in handling various data characteristics, and selecting the appropriate model can significantly enhance prediction accuracy and operational efficiency.

3. Research Methodology

Figure 1 illustrates the processes involved in the methodology of this work. Traffic forecasting follows a structured approach to predict future traffic patterns based on historical data. The first step is to collect relevant historical traffic data, which is then pre-processed to ensure its quality and consistency. Next, a model is selected according to the characteristics of the data. For instance, if the data exhibits strong seasonal patterns, AGBFM may be the preferred model. Alternatively, if the data has a complex structure with many variables, EWOFM may be more suitable. Once the model is selected, it is trained using the historical data, with its parameters adjusted to best fit the data. The model is then evaluated using metrics such as Mean Absolute Percentage Error (MAPE), Root Mean Squared Error (RMSE), and Coefficient of Determination (R^2) etc. which provide insights into the model's performance. These metrics help identify the most appropriate model for the task. After evaluation, the best model is selected to forecast future traffic patterns.

The forecasted values are then post-processed to ensure quality and consistency. Finally, the results are visualized to facilitate interpretation and decision-making. This structured approach ensures that the most accurate and reliable traffic forecasts are generated, supporting informed decisions regarding traffic management and infrastructure planning.

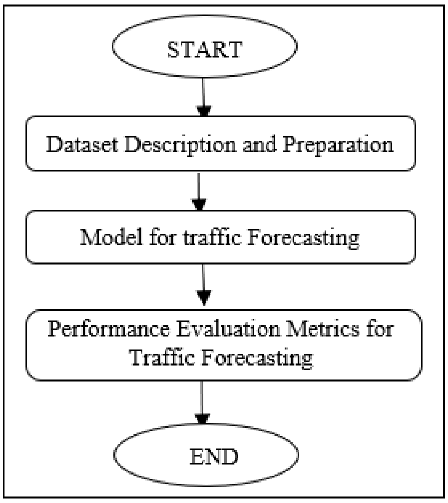
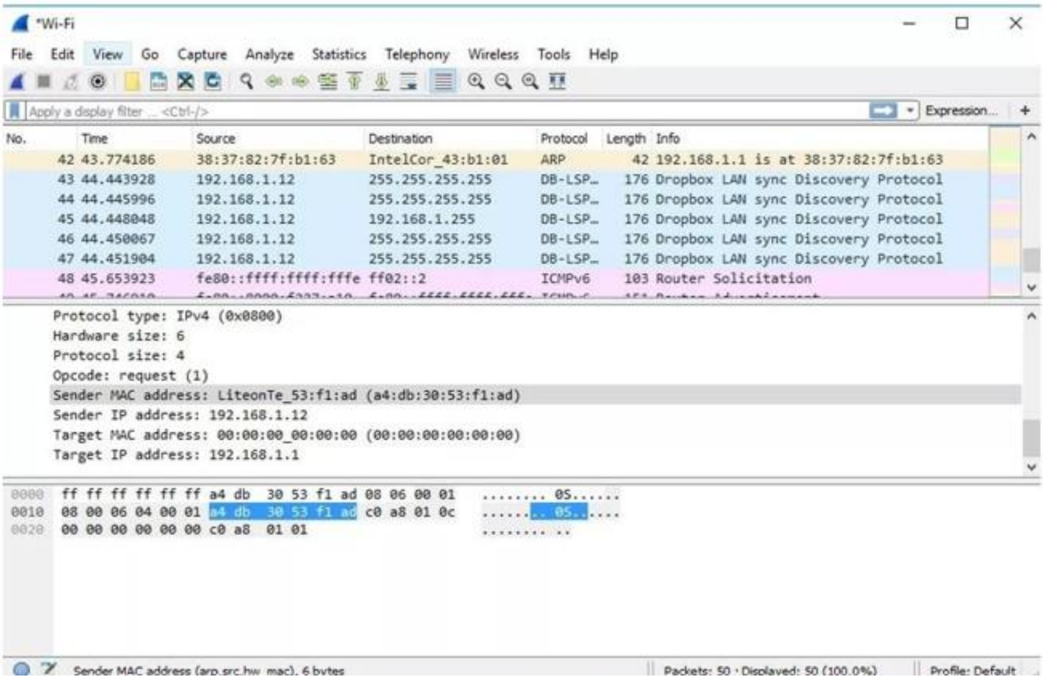


Figure 1: Flowchart of the research methodology

3.1 Dataset Description and Preparation

The data which is used for network traffic forecasting is taken using two different scenarios: Scenario I: Raw data is generated by flooding various virtual machines with network traffic. The network traffic could be generated by running various commands at command prompt of the host machine. Here "ping command" and "apache server" 50 commands are run again and again to flood network with traffic. These bursts of data are captured by the host machine by using Wireshark [10]. The Wireshark capture engine captures live network data simultaneously from multiple network interfaces. The subset of captured traffic is shown in figure 2.

When capturing of data gets completed for a certain time interval, it is saved in .csv form. In the present study, firstly, normal data is captured without any traffic burst and plotted in a graph and then VMs (virtual machines) are flooded with data packets by running commands on prompt very frequently and these packets are captured. The normal and exponential bursts seen in captured traffic is shown in figure 3 and 4 respectively.



Scenario II: In this case, the data used to evaluate the present study is taken from CRAWDAD (Community Resource for Archiving Wireless Data) iitkdg/apptraffic datasets of a smartphone app collected using tcpdump. Tcpdump is a command-line interface-based data-network packet analyzer computer software. It allows the user to see TCP/IP and other packets that are being sent and received across a network to which the machine is connected. The desired traffic for evaluation came from Google Hangout of smartphone app. An application called google hangout traffic (GB/micro-sec) time (micro-seconds) traffic (GB/micro-sec) time (micro-seconds) 52 facilitates its users to do chats, carry out VoIP calls and video calls. Google hangout is not a completely peer-to-peer service platform, although it has features of a peer-topeer application as it permits two users to interconnect in real-time using a session server which is selected dynamically. In the form of .pcap files, data was collected with only the Google Hangouts app running in the foreground and only required system functions running in the background. A subset of dataset form google hangouts of smartphone app collected from CRAWDAD community.

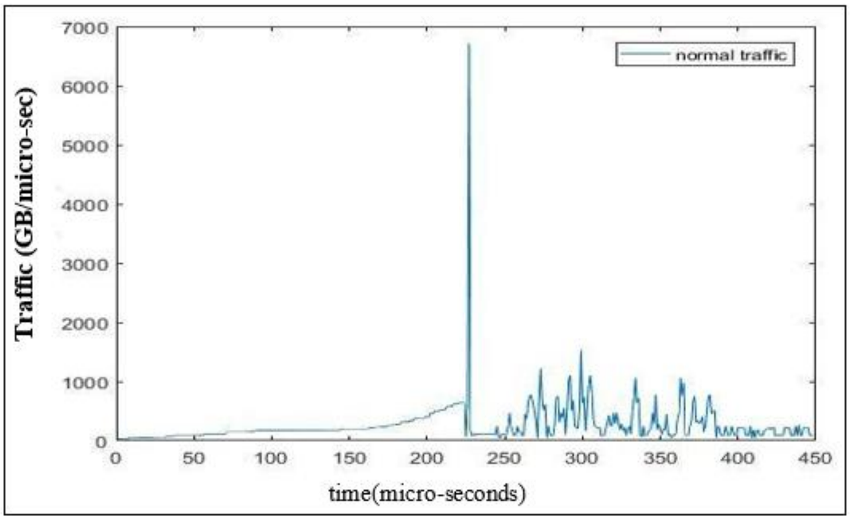


Figure 2: Wireshark engine running for network traffic capturing

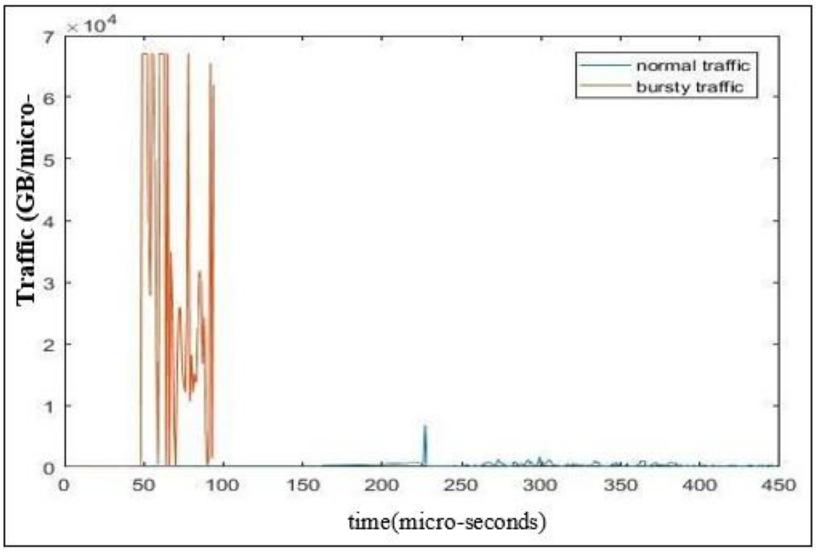


Figure 3: Normal network traffic bursts data captured through Wireshark

Figure 3, 4: Exponential (bursty) network traffic laid over normal traffic data captured through Wireshark

3.2. Models for Traffic Forecasting

Traffic forecasting in 5G base stations is crucial for efficient network management, resource allocation, and ensuring quality of service. Several models have been proposed and tested for this purpose, including traditional statistical models such as ARIMA, machine learning models using gradient methods like AGBFM, and techniques like EWOFM. This section presents a theoretical approach to traffic forecasting using these models, emphasizing their strengths and application scenarios.

3.3. Performance Evaluation Metrics for Traffic Forecasting

Various parameters are presented in literature which is considered for performance analysis in this study. Choice of an appropriate parameter is very important to do the comparative analysis of the proposed model with the existing ones. A number of metrics exist to express forecasting accuracy. A brief overview of the parameters explained by authors in their papers [11] . Figure 6 shows the confusion matrix consisting of the count of number of true positives, true negatives, false positives and false negatives. Equations to be used for evaluating the performance of forecasting model in terms of sensitivity or true positive rate, specificity, accuracy and precision is shown below.

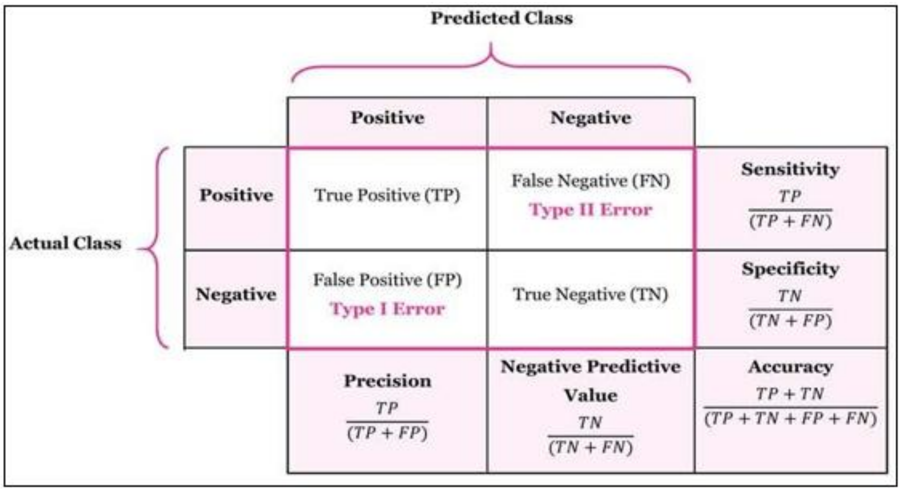


Figure 6: Confusion Matrix

4. Result and Discussion

4.1 Total Traffic Forecast

The comparison between actual total traffic data and the forecasting results generated by the proposed and existing models over the testing period is done using a number of parameters. The x-axis represents time in hours, and the y-axis indicates the traffic volume. Each model's forecast is plotted alongside the actual traffic data to visualize their performance. The graph highlights discrepancies between predicted and actual values, providing a clear visual representation of each model's accuracy. Forecasting models comprises of training and optimization algorithms. LSSVM

(Least Square Support Vector Machine) [12] is used for training and ACO (Ant Colony Optimization) [13], PSO (Particle Swarm Optimization) [14], WOA (Whale Optimization Algorithm) [15], IWOA (Improved Whale Optimization), AGO (Adaptive Gradient Optimization) are used as optimization algorithms. Convergence curve for existing and proposed models is as shown in figure 7.

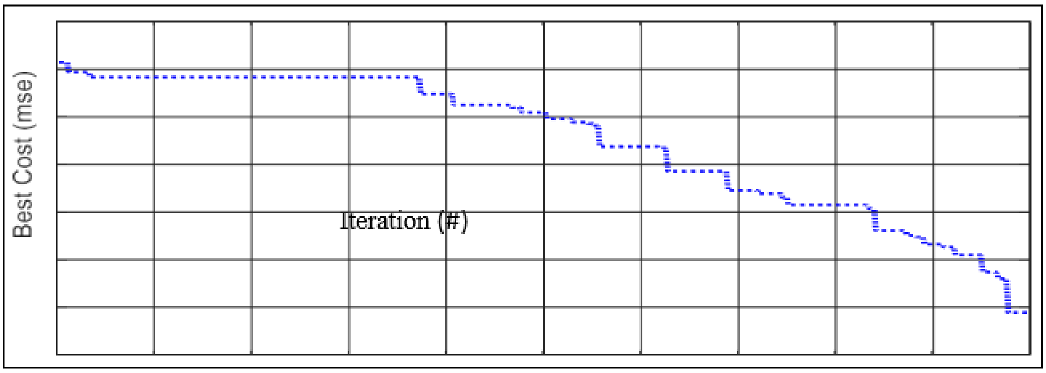


Figure 7: Convergence Curve using LSSVM-ACO

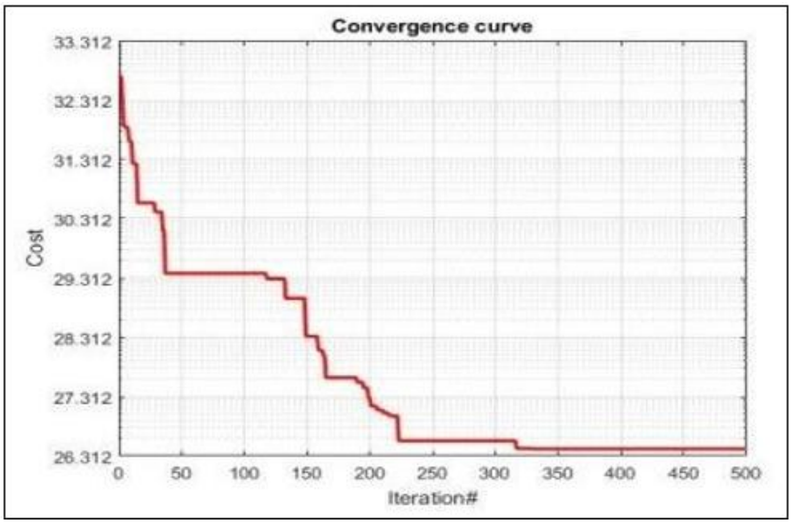


Figure 8: Convergence Curve using LSSVM-PSO

The best achievable optimization cost using ACO and PSO with LSSVM classifier is shown in figure 7 and 8. ACO Pareto optimization was not received even until 50 iterations. This makes LSSVM-ACO better in terms of optimization efficiency. This is because of the slower search criteria of ACO.

4.2 Performance Analysis

The proposed forecasting models (AGBFM and EWOFM) are compared with three other models namely LSSVM-ACO, LSSVM-PSO and LSSVM-WOA. Although the meta-heuristic algorithms work really well in many domains of application, however, it is observed that LSSVM-ACO, LSSVM-PSO and LSSVM-WOA does not perform well as compared to the proposed ones, in forecasting traffic bursts. This is due to the fact that the first proposed model, adaptive gradient based optimizer (AGBO) can be directly integrated with proposed prediction model using parameter optimization (gamma and sigma). The second proposed model, EWOA enhances the exploration ability by using inertia weight factor whereas ACO and PSO cannot do optimization and proper exploration using LSSVM. Moreover, the search criteria of ACO and PSO include exploitation instead of inherent gradient based optimization in AGBFM. Also, the usage of input data for tuning of parameters makes the proposed algorithms (AGBFM and EWOFM) adaptive and efficient. The extra searches make the ACO and PSO loosely coupled optimizers. In case of PSO the convergence did happen faster, but it did not achieve good fitting cost. Compared to ACO, the PSO achieved lower average fitting cost i.e. lower efficiency but not better than AGBFM and EWOFM. The evaluation parameters, namely, MSE, accuracy, TPR, FPR, precision and F1-score are also calculated for all the existing and proposed forecasting models.

4.2.1 Mean Square Error

The MSE is evaluated for the proposed and existing algorithms. MSE is a technique for determining how close estimations or projections are to actual values. For regression models, this is employed as a model evaluation metric, with a lower value indicating a better fit. The EWOFM shows minimum MSE because this model uses enhanced whale optimization model to improve the hyper parameters of training model. IWOA optimizer can be tightly coupled with the LSSVM training algorithm as compared to other models which makes this model fit for forecasting.

4.2.2. Accuracy

The accuracy is evaluated for the proposed and existing algorithms. The EWOFM shows maximum accuracy because search criteria in this model used by search agents is three dimensional which enhances the capabilities in terms of exploration of the search space and hence overall accuracy in forecasting results increases.

4.2.3. True Positive Rate (TPR) or Sensitivity or Recall and False Positive Rate (FPR)

The TPR is maximum in case of EWOFM because the number of true alarms higher as the number of iterations increases. The higher TPR indicates confirmation. The FPR is lowest in case of EWOFM. The lower value of FPR is desirable, as FPR signify wrong forecasting. The performance evaluation of EWOFM, AGBFM, LSSVM-PSO, LSSVM-ACO and LSSVM-WOA for both the scenarios of dataset is summarized in table 1 and table 2.

Table 1: Comparative analysis of existing and proposed forecasting models for scenario 1

Parameters / Forecasting Models	LSSVM-ACO	LSSVM-PSO	LSSVM-WOA	AGBFM	EWOFM
Mean Square Error	7.4%	27.9%	4.2%	3.16%	2.94%
TP alarms	21698	16931	21509	23619	24267
FP alarms	1737	6504	1926	1491	843
TN alarms	9300	7257	9219	7873	8089
FN alarms	745	2788	826	497	281
Accuracy	92.60%	72.20%	91.80%	94.10%	96.60%
TPR (Sensitivity)	0.966805	0.858614	0.963018	0.979391	0.988553
FPR	0.15738	0.47264	0.172813	0.159227	0.09438
Precision	0.92588	0.722466	0.917815	0.940621	0.966428
F1-Score	0.9459	0.784678	0.939873	0.959615	0.977365
Execution Time*	56	81	49	45	41
Computational Complexity	$\theta(n\log n)+\frac{\theta(n^2\log n)}{\rho^2}$**	$\theta(n^2\log n)$	$\theta(nlog\ n)+\theta(n^3)$	$\theta(nlog\ n)+\theta(n^2)$	$\theta(nlog\ n)+\theta(n^3)$

Table 2: Comparative analysis of existing and proposed forecasting models for scenario 2

Parameters / Forecasting Models	LSSVM-ACO	LSSVM-PSO	LSSVM-WOA	AGBFM	EWOFM
Mean Square Error	8.1%	29.2%	4.7%	3.42%	3.30%
TP alarms	15403	11973	15389	15838	17349
FP alarms	1563	4993	1577	1127	828
TN alarms	6602	5132	6596	6789	5784
FN alarms	670	2140	676	484	277
Accuracy	90.80%	70.60%	90.70%	93.40%	95.40%
TPR (Sensitivity)	0.958315	0.848367	0.957921	0.970347	0.984285
FPR	0.191427	0.493136	0.192952	0.14237	0.125227
Precision	0.907875	0.705706	0.907049	0.933569	0.954448
F1-Score	0.932413	0.770488	0.931791	0.951603	0.969137
Execution Time*	60	85	53	51	46
Computational Complexity	$\theta(n\log n)+\frac{\theta(n^2\log n)}{\rho^2}$**	$\theta(n^2\log n)$	$\theta(nlog\ n)+\theta(n^3)$	$\theta(nlog\ n)+\theta(n^2)$	$\theta(nlog\ n)+\theta(n^3)$

5. Conclusion and Future Scope

1) A study of the various machine learning algorithms is performed in order to choose the suitable algorithm for training. LSSVM machine learning algorithm is considered for training using dataset. Various optimization schemes are also studied to optimize the parameters of machine learning algorithm for better performance.
2) The proposed models have been compared on the evaluation parameters using confusion matrix namely MSE, TPR, FPR, accuracy, execution time and computational complexity precision, f1-score.
3) The two proposed models are compared with the existing models to show the effectiveness of the proposed ones using the performances matrices. The proposed forecasting models are compared with three existing models namely LSSVM-ACO, LSSVM-PSO and LSSVM-WOA. Although the meta-heuristic algorithms work really well in many domains of application, however, it is observed that both LSSVM-ACO and LSSVM-PSO does not perform well as compared to deterministic AGBFM, in predicting traffic bursts. Also EWOFM perform best of the all models in forecasting. In future, EWOFM and AGBFM models can also be extended further, so as to apply to different real world problems like image processing, data mining and feature selection.

References

[1] Singhrova, A. (2021). Enhanced whale optimization based traffic forecasting for SDMN based traffic. *ICT Express*, *7*(2), 143-151.

[2] Anupriya, & Singhrova, A. (2022). Mobile traffic flow prediction using intelligent whale optimization algorithm. *Automated Software Engineering*, *29*(2), 48.

[3] "Cisco Annual Internet Report - Cisco Annual Internet Report (2020–2025) White-Paper-Cisco." https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html.
[4] S. Djahel, R. Doolan, G.-M. Muntean, and J. Murphy, "A Communications-Oriented on Traffic Management Systems for Smart Cities: Challenges and Innovative Approaches," Jan. 2015, vol. 17. doi: 10.1109/COMST.2014.2339817.
[5] L.-V. Le, D. Sinh, B.-S. P. Lin, and L.-P. Tung, "Applying Big Data, Machine Learning, and SDN/NFV to 5G Traffic Clustering, Forecasting, and Management," in *2018 4th IEEE Conference on Network Softwarization and Workshops (NetSoft)*, Jun. 2018, pp. 168–176. doi: 10.1109/NETSOFT.2018.8460129.
[6] Van Splunder, J. (2015). Periodicity detection in network traffic. *Technical Report, Mathematisch Instituut Universiteit Leiden*.
[7] Gijon, C., Toril, M., Luna-Ramírez, S., Marí-Altozano, M. L., & Ruiz-Avilés, J. M. (2021). Long-term data traffic forecasting for network dimensioning in LTE with short time series. *Electronics*, *10*(10), 1151.
[8] Haseeb, A., & Farooq, A. (2025). REVOLUTIONIZING TELECOMMUNICATIONS: THE IMPACT OF IOT. *Spectrum of Engineering Sciences*, 347-355.
[9] Tedjopurnomo, D. A., Bao, Z., Zheng, B., Choudhury, F. M., & Qin, A. K. (2020). A survey on modern deep neural network for traffic prediction: Trends, methods and challenges. *IEEE Transactions on Knowledge and Data Engineering*, *34*(4), 1544-1561.
[10] Nath, A. (2015). *Packet Analysis with Wireshark*. Packt Publishing Ltd.
[11] Yun, S. Y., Namkoong, S., Rho, J. H., Shin, S. W., & Choi, J. U. (1998). A performance evaluation of neural network models in traffic volume forecasting. *Mathematical and Computer Modelling*, *27*(9-11), 293-310.
[12] Mitra, V., Wang, C. J., & Banerjee, S. (2007). Text classification: A least square support vector machine approach. *Applied soft computing*, *7*(3), 908-914.
[13] Dorigo, M., Birattari, M., & Stutzle, T. (2006). Ant colony optimization. *IEEE computational intelligence magazine*, *1*(4), 28-39.
[14] Kennedy, J., & Eberhart, R. (1995, November). Particle swarm optimization. In *Proceedings of ICNN'95-international conference on neural networks* (Vol. 4, pp. 1942-1948). ieee.
[15] Mirjalili, S., & Lewis, A. (2016). The whale optimization algorithm. *Advances in engineering software*, *95*, 51-67.