Journal of IoT and Cybersecurity Systems

Research Article

Probability Based Traffic Analysis focus on Intersection between Student Boarding House President University

Phyo Thiha Aung, and Thandar Swe

Abstract. Traffic flow analysis is essential for improving transportation services, campus accessibility, road safety, and congestion management. Previous studies have widely applied statistical analysis, probability models, machine learning, and traffic visualization to understand vehicle movement and predict congestion; however, limited small-scale studies have examined campus-related traffic routes using simple probability-based methods that can directly support student safety and service operations. This study aims to analyze traffic flow and estimate congestion probability on the road section between the Jl. Ki Hajar Dewantara and Mekarmukti–Lemah Abang Junction and the Center Point Jababeka roundabout, specifically along the route from the Student Boarding House to President University. Vehicle count data, mainly cars, were manually collected over time and analyzed using Python. Descriptive statistics, including mean, variance, standard deviation, maximum, and minimum values, were calculated, while congestion probability was determined using a threshold of 12 vehicles per minute. Matplotlib was used to visualize traffic flow patterns, vehicle arrival frequency, and probability distribution. The results showed an average traffic flow of 7.78 vehicles per minute, variance of 13.85, standard deviation of 3.72, maximum of 15 vehicles, minimum of 1 vehicle, and a congestion probability of 16.67%. These findings indicate moderately stable traffic conditions with occasional congestion. The study contributes a practical probability-based approach for campus traffic monitoring, although it is limited to one road section and one observation period. Future research should use broader datasets and predictive modeling to improve traffic management and student safety services.

Keywords: Python, President University, Probability, Student Safety, Traffic flow.

Author(s):

Phyo Thiha Aung
Department of Electrical Engineering, Universitas Presiden, Cikarang, Indonesia
E-mail: phyo.aung@student.president.ac.id

Thandar Swe
Department of Electrical Engineering, Universitas Presiden, Cikarang, Indonesia
E-mail: thandar.swe@student.president.ac.id

Corresponding Author: Phyo Thiha Aung (phyo.aung@student.president.ac.id).

How to Cite This Article: Aung, P. T., & Swe, T. (2026). Probability Based Traffic Analysis focus on Intersection between Student Boarding House President University. Journal of IoT and Cybersecurity Systems, 1(1), 31–42. https://doi.org/10.58291/jicss.v1i1.6

Received : 23 May, 2026; Revision: 2 July, 2026; Accepted: 8 July, 2026; Publication: 9 July, 2026.

1. Introduction

Transportation is one of the most fundamental necessities for human beings. It is closely connected to people's daily lives through travel and is continuously related to various aspects of livelihood. Therefore, improving convenience in daily life, reducing travel time, and ensuring safety are among the priority considerations. This study is a small-scale study that takes these factors into consideration and presents several findings and recommendations. One of the fundamental methods used in analyzing traffic which is the focus of this study is the application of statistics. Therefore, to obtain accurate data, the field observations were conducted, and data were collected manually. Despite significant advances in short-term traffic flow forecasting techniques, particularly in prediction accuracy and applicability (Gomes et al., 2023; He & Yao, 2026; Osho et al., 2026; Wang et al., 2026). These measurements mainly focus on average traffic demand and variations in traffic flow over time and further studies are required to obtain more comprehensive information. Probability theory also plays an important role in the field of traffic engineering. It is used to represent the uncertainty and randomness of vehicle arrivals by developing models based on probability distributions. In many traffic studies, the probability of traffic congestion is determined by comparing actual measured results with a predetermined threshold value.

The common and frequent machine learning techniques that have been applied for traffic flow prediction are Convolutional Neural Network and Long-Short Term Memory (L. Li & Chen, 2017). Traffic flow modelling has traditionally focussed on simulation (i.e., 'traffic assignment') approaches based on origin–destination matrices (i.e., trip generation). Reserchers help transportation planners identify periods of high traffic demand and evaluate road capacity utilization. In this case, data visualization techniques are also an important component of traffic flow analysis (Razali et al., 2021). Therefore, graphs such as line charts, histograms, and bar charts are used to present traffic flow patterns and statistical results. These visualizations are mainly intended to help researchers and decision-makers better understand traffic behavior. In addition, they help provide clearer and more accurate conclusions from numerical data and assist in identifying road conditions that may not be easily observed (Q. Li et al., 2026; Silalahi et al., 2021; Simanjuntak et al., 2024). Consequently, how to better utilize the inherent structural features of transportation systems for traffic flow prediction has become a topic worthy of in-depth exploration (Boddepalli et al., 2026; Budiyanto et al., 2021; Ni et al., 2026). Improvements in programming, have also made it easier to process and analyze traffic data efficiently.

Grounded in traffic flow theory, cooperative vehicle control, and probabilistic decision modeling, Safe Pass addresses a problem of direct relevance to urban traffic planning and emergency response management (Zong & Chen, 2026). Python, which is widely used by engineers, has become one of the most popular programming languages for data analytics because of its simplicity and the availability. Furthermore, libraries such as Pandas make data management easier, while Matplotlib helps create clear and effective visualizations of data. (Wang et al., 2026) By using these tools, researchers can easily perform statistical analysis, probability calculations, and graphical presentations (Alanazi et al., 2026). Based on previous studies, statistical and probability analyses have contributed to a better understanding of traffic flow characteristics. These findings continue to serve as effective methods for estimating the risk of traffic congestion (He & Yao, 2026; Q. Li et al., 2026). Therefore, this study uses Python to (Kamalifar et al., 2026; Z. Li et al., 2026; Silalahi et al., 2023) analyze the collected data and calculates the probability of traffic congestion using probability methods. In addition, graphs are used to present the behavior of traffic flow.

Data-driven methods have recently emerged as an increasingly viable solution, and transfer learning provides a backbone for building reasonable predictive models, even for cities without sufficient data to train dedicated models (Hoogendoorn & Bovy, 2001; Simanjuntak et al., 2023; Yu et al., 2013). Traffic operations on roadways can be improved by field research and field experiments of real-life traffic flow[4]. Creating and predicting general traffic indicators, such as traffic flow, density, and mean speed, is crucial for effective traffic control and congestion prevention (Gualda et al., 2026; Morandi et al., 2026; Sheehan et al., 2026).

2. Methods

The proposed traffic flow analysis follows a systematic process consists of seven steps. These steps include surveying data, checking of missing/error data, congestion analysis, probability calculation, displaying results and adding recommendations. The complete procedure of this study is shown in Figure 1.

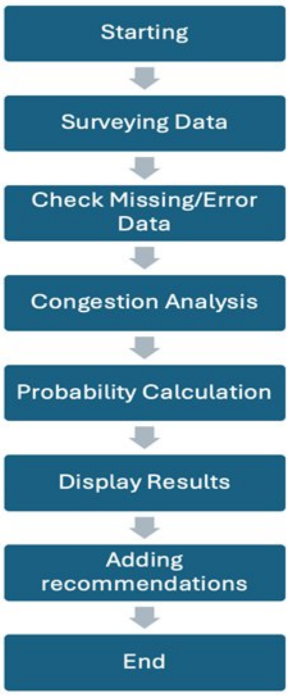

Figure 1. Architecture of the system (Source: Self Documentation)

Minute	Vehicles	Minute	Vehicles	Minute	Vehicles
16:01	8	16:21	2	16:41	11
16:02	8	16:22	5	16:42	13
16:03	11	16:23	14	16:43	6
16:04	13	16:24	13	16:44	9
16:05	13	16:25	11	16:45	5
16:06	9	16:26	11	16:46	5
16:07	10	16:27	8	16:47	3
16:08	13	16:28	7	16:48	7
16:09	13	16:29	5	16:49	8
16:10	11	16:30	4	16:50	5
16:11	15	16:31	4	16:51	5
16:12	9	16:32	3	16:52	1
16:13	9	16:33	2	16:53	4
16:14	5	16:34	4	16:54	9
16:15	8	16:35	7	16:55	12
16:16	4	16:36	5	16:56	11
16:17	3	16:37	13	16:57	7
16:18	5	16:38	11	16:58	10
16:19	4	16:39	10	16:59	9
16:20	2	16:40	8	17:00	7

Figure 2. Detail collected data (Source: Self Documentation)

2.1. Surveying Data

Figure 2 descibes the data are collected from the road section between *the Jl. Ki Hajar Dewantara & Mekrmukti-Lemah Abang Junction and the Center Point Jababeka roundabout* only (on the route from the Student Boarding House-SBH to the President University Campus). The data were collected on Sunday, June 28, 2026, starting from 16:01 to 17:00 PM. This date and time were selected because they were considered to represent the peak hour within a week. Vehicle count data recorded over time were collected from traffic on the above-mentioned road section from motor vehicles especially cars**.**

2.2. Check Missing

After the data were collected, data validation was performed to make sure data quality. Since the tally system was used for recording the data, a double-check was conducted to verify the accuracy of the recorded data. The data were further examined for missing values, incomplete records and wrong entries that could affect the analysis results. Any identified errors were reviewed and corrected. This step improves the reliability and accuracy of the analysis.

2.3. Congestion Analysis

During the congestion analysis stage, congestion events were identified based on a predetermined threshold value of 12 observations per minute. In this study, if the vehicle count was greater than or equal to the specified threshold value, the condition was classified as traffic congestion. The system then calculated the total number of congestion events that occurred during the observation period. The congestion condition can be expressed as equation 1:

$$\textit{Observed Vehicle Count} \geq \textit{Threshold} \tag{1}$$

2.4. Probability Calculation

The probability of traffic congestion was calculated using the ratio of the number of congestion events to the total number of observation records. This was used to estimate the likelihood of traffic congestion occurring during the observation period as follow at equation 2.

$$P(C) = \frac{Nc}{Nt} \tag{2}$$

Where: P(C) = Congestion Probability, Nc = Number of Congestion Events, Nt = Total Number of Observations. In addition, the probability distribution of vehicle arrivals is generated by calculating the relative frequency of each vehicle count value that collected on survey.

2.5. Display Results

In the next stage, the analysis results are presented in both numerical and graphical forms. Statistical information, congestion probability, and probability distributions are presented based on the calculation results. Moreover, line graphs, histograms and bar charts are used to clearly illustrate traffic flow behavior and traffic congestion.

3. Result and Discussion

The traffic flow system was implemented using Python for analysis and was evaluated using a traffic dataset containing vehicle counts recorded over time. The analysis included statistical calculations, estimation of congestion probability, and probability distribution analysis.

3.1. Statical Analysis Result

According to the statistical analysis results, the mean vehicle arrival rate was 7.78 vehicles per minute. This represents the average traffic volume during the observation period. The calculated variance was 13.85, and the standard deviation was 3.72. These results indicate that vehicle arrivals showed a moderate level of variation over time. The maximum number of vehicles recorded during a single observation was 15, while the minimum number of vehicles recorded was 1. These results show that traffic flow during the observation time was not highly stable, with vehicle counts frequently fluctuating above and below the average value. However, the relatively low standard deviation indicates that there were no significant sudden changes in traffic flow.

3.2. Congestion Analysis Results

In this study, a congestion threshold of 12 vehicles per minute was established. Based on this threshold, the analysis identified 6 congestion events in the dataset. The calculated congestion probability was 0.167, or 16.67%. These results indicate that traffic congestion occurred in only about one-sixth of the total observation records. Since the congestion probability was relatively low, the observed traffic conditions were generally stable, and traffic congestion occurred only during a few numbers of observation periods.

3.3. Probability Distribution Analysis

According to the probability distribution analysis, the arrival of 5 vehicles per minute occurred most frequently, with a probability of 15%. And this was followed by 11 and 13 vehicles (11.8%) and the least frequent occurrences were 1, 6, 12, 14, and 15 vehicles per minute, each with a probability of only 1.8%. The probability distribution indicates that most traffic flow was concentrated between 5 and 13 vehicles per minute. Very high traffic volumes was not occurred frequently but very low traffic volumes were observed sometimes. Therefore, the dataset was primarily characterized by normal traffic conditions, while instances of high traffic density occurred only a few times.

3.4. Graphical Analysis

The Traffic flow versus Time graph illustrates the fluctuations in vehicle arrivals throughout the observation period. At certain times, peaks in vehicle count were observed, while at other times, lower traffic volumes were recorded. Overall, the traffic flow remained at an average of approximately 8 vehicles per minute. The Traffic Flow versus Time graph shown in figure 3.

The histogram shows that most vehicle counts were concentrated within the middle range, indicating that the traffic flow distribution was relatively uniform. The histogram of vehicle arrivals illustrates in figure 4.

The Probability Distribution Bar Chart clearly shows that the arrival of 5 vehicles per minute occurred most frequently. Likewise, the arrival of 11 and 13 vehicles per minute was the second most frequent, indicating that these traffic conditions may occasionally lead to traffic congestion or traffic delays. However, the chart also indicates that the probability of severe traffic congestion was low. The probability distribution bar chart in figure 3.

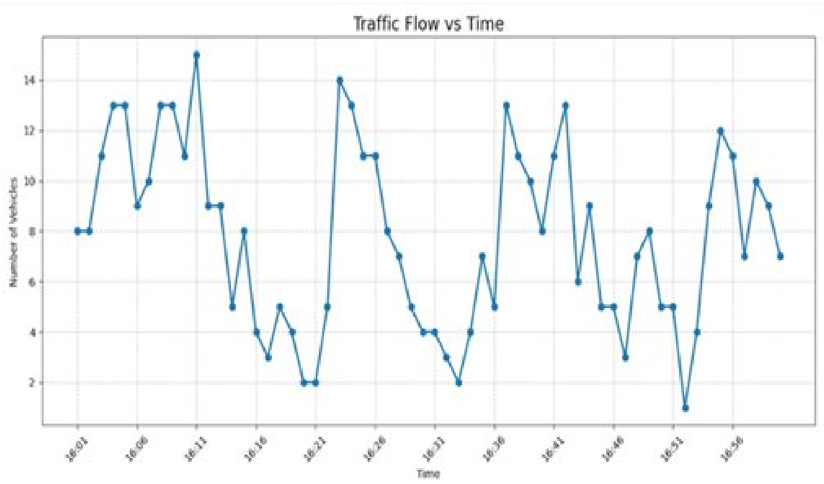

Figure 3. Traffic flow vs time of the system (Source: Self Documentation)

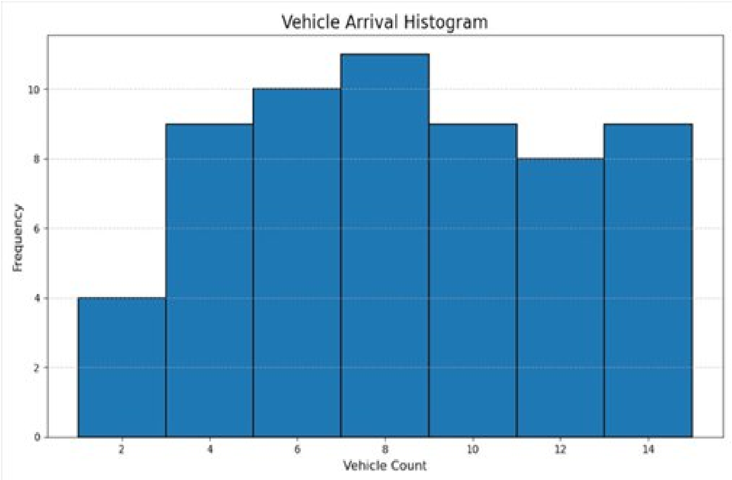

Figure 4. Vehicle arrival in the system (Source: Self Documentation)

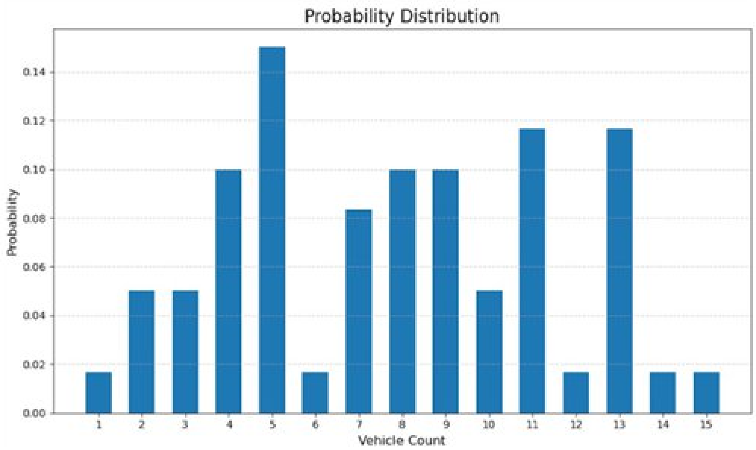

Figure 5. Probability distribution of the system (Source: Self Documentation)

These results demonstrate that the statistical and probability-based approach can effectively analyze traffic flow behavior and estimate the probability of traffic congestion. Furthermore, the findings can provide useful support for traffic monitoring, traffic congestion assessment, and transportation planning as shown at table 1.

Table 1. Summary of traffic flow analysis

Matrix	Observation
Average Traffic Volume	7.78 Vehicles/min
Most Frequent Vehicle Count	8 until 9 vehicles
Congestion Probability	16.67%
Traffic Condition	Moderately Stable
Highest Record Traffic Volume	15 vehicles
Lowest Record Traffic Volume	1 vehicle

Source: Self documentation

The overall traffic condition is moderately stable with an average flow of 6.07 vehicles per minute and relatively low congestion probability.

3.5. Recommendations

Although traffic congestion was not observed frequently during this study, it was found that vehicles continuously passed through the study focused road section. Since a traffic signal is located at one end of the route, vehicles tend to travel at relatively high speeds toward the crossing when the traffic light turns green. As a result, while traffic congestion is unlikely to cause significant travel delays or interruptions on this road section, the risk of traffic accidents involving while students and peoples crossing the road remains high. Therefore, it is highly recommended that President University Administration should assign full-time staffs to assist in crossing the road safely throughout the day and to enhance road safety along this route.

6. Conclusions and Future Directions

This study presented a traffic flow analysis system based on statistical and probability methods using Python. The proposed system successfully processed traffic data stored in a CSV dataset and performed statistical analysis, traffic congestion analysis, and probability distribution analysis. The results showed that the mean vehicle arrival rate was approximately 8 vehicles per minute, indicating a moderate level of traffic flow during the observation period. The statistical analysis produced a variance of 13.85 and a standard deviation of 3.72, suggesting that traffic conditions remained relatively stable throughout the observation period. A congestion threshold of 12 vehicles per minute was adopted in this study. Based on this threshold, the system identified 6 congestion events and calculated a congestion probability of 16.67%. These findings indicate that traffic congestion occurred only during certain periods and was not a dominant condition throughout the observed dataset. The probability distribution analysis further revealed that the arrival of 5 vehicles per minute was the most frequent traffic condition, while 11 and 13 vehicles per minute were the second most frequent. Higher traffic volumes occurred less frequently during the observation period. Overall, the proposed method effectively analyzed traffic flow behavior and provided a reliable estimation of traffic congestion probability. The findings of this study may support transportation planning, traffic monitoring, campus safety, and student service operations. In future work, collecting more detailed traffic data and incorporating predictive traffic modeling techniques may further improve the performance of the system and contribute to improving the services of President University.

Author Contributions: “Conceptualisation: Thandar Swe. and Phyo Thiha Aung.; Methodology: Thandar Swe.; Software: Thandar Swe. and Phyo Thiha Aung.; Validation: Phyo Thiha Aung; Investigation: Phyo Thiha Aung; Resources: Thandar Swe; Writing—original draft preparation: Thandar Swe. and Phyo Thiha Aung; Writing—review and editing: Lukman Medriavin Silalahi; Supervision: Lukman Medriavin Silalahi. All authors have read and approved the final version of the manuscript.

Acknowledgements: The author would like to express deepest thanks to Professors and Head of the Electrical Engineering, for their suggestion throughout the study. The author would like to also express thanks to the supervisor Ir. Lukman Medriavin Silalahi, A.Md., ST., MT., IPM.,

ASEAN-Eng., APEC-Eng, Supervisor of the Electrical Engineering Study Program, for giving this opportunity and his invaluable comments and guidance during this study. And finally, the author wishes to thank to all persons who helped with necessary assistance for this study.
Conflicts of Interest: The authors declare no conflict of interest

AI Declaration

During the preparation of this manuscript, the authors used AI-assisted technologies to improve language quality and readability. All intellectual content, analysis, and conclusions remain the sole responsibility of the authors.

References

Alanazi, F., Armghan, A., Abdullah Al-Gburi, A. J., & Yousef, A. (2026). Multi-modal data fusion and explainable artificial intelligence for smart traffic flow prediction in urban mobility networks. *Egyptian Informatics Journal*, *34*, 101005. https://doi.org/10.1016/j.eij.2026.101005

Boddepalli, M., Rajabi, E., & Amin, S. H. (2026). Traffic collision analysis and traffic flow simulation in Toronto using cellular automata models. *Expert Systems with Applications*, *324*, 132538. https://doi.org/10.1016/j.eswa.2026.132538

Budiyanto, S., Silalahi, L. M., Adriansyah, A., Darusalam, U., Andryana, S., & Rochendi, A. D. (2021). Development of Internet of Things Based Fertigation System for Improving Productivity of Patchouli Plantation. *2021 3rd International Conference on Research and Academic Community Services (ICRACOS)*, 230–233.

Gomes, B., Coelho, J., & Aidos, H. (2023). A survey on traffic flow prediction and classification. *Intelligent Systems with Applications*, *20*, 200268. https://doi.org/10.1016/j.iswa.2023.200268

Gualda, L., Singh, R., Abouelela, M., & Antoniou, C. (2026). Assessing a data-driven workflow for city-wide and cross-city traffic flow prediction. *Transportation Research Procedia*, *95*, 800–807. https://doi.org/10.1016/j.trpro.2026.02.101

He, P., & Yao, Y. (2026). A spatiotemporal correlation model for traffic flow based on dynamic graph attention and multi-scale feature fusion. *Alexandria Engineering Journal*, *146*, 80–94. https://doi.org/10.1016/j.aej.2026.05.019

Hoogendoorn, S. P., & Bovy, P. H. L. (2001). State-of-the-art of vehicular traffic flow modelling. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, *215*(4), 283–303. https://doi.org/10.1177/095965180121500402

Kamalifar, A., Cenedese, C., Cucuzzella, M., & Ferrara, A. (2026). A coordinated Service station enhanced Mainstream Traffic Flow Control based on METANET-s. *IFAC Journal of Systems and Control*, 100425. https://doi.org/10.1016/j.ifacsc.2026.100425

Li, L., & Chen, X. (Michael). (2017). Vehicle headway modeling and its inferences in macroscopic/microscopic traffic flow theory: A survey. *Transportation Research Part C: Emerging Technologies*, *76*, 170–188. https://doi.org/10.1016/j.trc.2017.01.007

Li, Q., Yan, Q., Qiu, C., Wang, M., Kan, R., & Liu, A. (2026). Deconstructing spatiotemporal complexity in Inland shipping: Vessel traffic flow prediction based on hexagonal grids and adaptive mixture of experts architecture. *Applied Ocean Research*, *172*, 105106. https://doi.org/10.1016/j.apor.2026.105106

Li, Z., Lin, Q., Pu, F., Ahn, S., Zhang, Y., Jiang, J., & Zhou, Y. (2026). Unveiling traffic wave of linear adaptive cruise control: A second-order macroscopic traffic flow model. *Transportation Research Part B: Methodological*, *211*, 103506. https://doi.org/10.1016/j.trb.2026.103506

Morandi, V., Peirano, L., & Speranza, M. G. (2026). Real-time rerouting of traffic flows to control the risk of disruptions. *Transportation Research Part C: Emerging Technologies*, *189*, 105721. https://doi.org/10.1016/j.trc.2026.105721

Ni, Y.-C., Kouvelas, A., & Makridis, M. A. (2026). Simulating link-level interrupted flow traffic dynamics and the comparison between different models for urban road networks. *Simulation Modelling Practice and Theory*, *147*, 103252. https://doi.org/10.1016/j.simpat.2026.103252

Osho, O., Chakraborty, S., & Das, S. K. (2026). SafePass: Efficient emergency vehicle passage with minimal disruption to traffic flow. *Transportation Research Interdisciplinary Perspectives*, *37*, 102011. https://doi.org/10.1016/j.trip.2026.102011

Razali, N. A. M., Shamsaimon, N., Ishak, K. K., Ramli, S., Amran, M. F. M., & Sukardi, S. (2021). Gap, techniques and evaluation: traffic flow prediction using machine learning and deep learning. *Journal of Big Data*, *8*(1), 152. https://doi.org/10.1186/s40537-021-00542-7

Sheehan, A., Jephcote, C., & Gulliver, J. (2026). Development and evaluation of statistical models for international-scale road traffic flow estimation. *Environment International*, *210*, 110231. https://doi.org/10.1016/j.envint.2026.110231

Silalahi, L. M., Budiyanto, S., Silaban, F. A., Simanjuntak, I. U. V., Rochendi, A. D., & Osman, G. (2021). Optimizing The Performance Of The Power Station Generator Space Lighting System Performance Based On Internet Of Things Using ESP32. *2021 3rd International Conference on Research and Academic Community Services (ICRACOS)*, 219–224. https://doi.org/10.1109/ICRACOS53680.2021.9702010

Silalahi, L. M., Simanjuntak, I. U. V., Budiyanto, S., & Rochendi, A. D. (2023). Computer Hardware and Software Education for Teacher's Office of Insan Mulia Early Childhood Education School Tangerang. *Journal of Innovation and Community Engagement*, *4*(4 SE-Articles), 232–240. https://doi.org/10.28932/ice.v4i4.7313

Simanjuntak, I. U. V., Heryanto, H., Dani, A. W., Salamah, K. S., Silalahi, L. M., & Marsuki, M. (2024). Analysis Of Solar Power Public Street Lighting Optimization With Pvsyst Software In A Residential Complex Area. *International Journal of Electronics and Telecommunications*, *70*(3), 743–749.

Simanjuntak, I. U. V., Heryanto, Rochendi, A. D., & Silalahi, L. M. (2023). Simulation and Analysis of Link Failover Using Routing Border Gateway Protocol (BGP) Multi-Protocol Label Switching (MPLS) Networks. *2023 International Conference on Radar,* *Antenna, Microwave, Electronics, and Telecommunications (ICRAMET)*, 341–346.

Wang, X., Ruan, X., Casas, J. R., Zhang, M., Li, Y., Zhou, H., & Dong, Y. (2026). Extreme response scenarios for long-span bridges: Probabilistic generation and interpretable clustering under free-flow and congested traffic. *Engineering Structures*, *353*, 122346. https://doi.org/10.1016/j.engstruct.2026.122346

Yu, H., He, J., Liu, R., & Ji, D. (2013). On the security of data collection and transmission from wireless sensor networks in the context of Internet of Things. *International Journal of Distributed Sensor Networks*, *2013*. https://doi.org/10.1155/2013/806505

Zong, M., & Chen, Y. (2026). Traffic flow signal control and optimization based on multi-agent deep reinforcement learning. *Alexandria Engineering Journal*, *145*, 238–252. https://doi.org/10.1016/j.aej.2026.04.025