DOI: https://doi.org/10.52756/ijerr.2024.v40spl.008 *Int. J. Exp. Res. Rev.*, Special Vol. 40: 104-116 (2024)

Original Article | Peer Reviewed | Open Access

International Journal of Experimental Research and Review (IJERR)
© Copyright by International Academic Publishing House (IAPH)
ISSN: 2455-4855 (Online)
www.iaph.in

ISSN 2455-4855

Security and Privacy for Smart Transportation Management using Big Data Analytics

Govindasamy R.[1*], Shanmugapriya N.[1] and Gopi R.[2]

[1]Department of Computer Science & Engineering, School of Engineering & Technology, Dhanalakshmi Srinivasan University, Trichy - 621112, Tamil Nadu, India; [2]Department of Computer Science & Engineering, Dhanalakshmi Srinivasan Engineering College, Perambalur - 621212, Tamil Nadu, India

E-mail/Orcid Id:

GS, ggnavaneeth@gmail.com, https://orcid.org/0009-0009-3429-5648; *SN,* shanmugapriyan.set@dsuniversity.ac.in, https://orcid.org/0000-0002-4620-1807; *GR,* gopircse@gmail.com, https://orcid.org/0000-0003-4957-1843

Article History:

Received: 26th Nov., 2023
Accepted: 05th Feb., 2024
Published: 30th June, 2024

Keywords:

Data encryption, security, privacy, big data, smart transportation

How to cite this Article:
Govindasamy R., Shanmugapriya N., and Gopi R. (2024). Security and Privacy for Smart Transportation Management using Big Data Analytics. *International Journal of Experimental Research and Review*, *40(spl.)*, 104-116.
DOI:
https://doi.org/10.52756/ijerr.2024.v40spl.008

Abstract: Security and privacy are vital aspects of smart transportation management with big data analytics because they assure the security of sensitive information, prevent unwanted access to essential systems, and retain public trust in the safety and dependability of the transportation infrastructure. Protecting data from cyber threats, ensuring secure communication and data transmission, protecting passengers' personal information and addressing privacy concerns related to data collection and usage to maintain transparency and accountability in data handling practices are all obstacles to smart transportation management using big data analytics. This paper proposes a Secure Data Encryption Control based Big Data Framework (SDEC-BDF) to strike a middle ground between data analytics and privacy protection, establishing the way for more private and secure transportation systems that benefit everyone involved. The intention of this approach is to offer strong security while simultaneously safeguarding people's privacy. The method has many potential uses in the Intelligent transportation sector (ITS), including traffic control, passenger security, fleet management, preventative maintenance, and road network design. It ensures privacy and security while facilitating effective data analysis. Furthermore, it protects the public confidence in the security and dependability of the transportation system, protects sensitive passenger data, and stops hackers from breaking into vital systems. The simulation analysis is conducted on the assumption that the system can maximize its security, privacy, and efficiency to create a more trustworthy transportation network.

Importance of Security and Privacy for Smart Transportation Management

The use of big data analytics for smart transportation management places a premium on security and privacy (Neilson et al., 2019; Samadder et al., 2023). First, they safeguard information like passenger records, patterns of traffic and fleet management strategies from hackers and other cybercriminals (Al-Turjman et al., 2022). The integrity and confidentiality of this data can be preserved by the use of stringent security measures that include encryption, restricted access and secure communication routes (Soomro et al., 2019; Kumar et al., 2023 a,b). Maintaining transparency and accountability requires attending to concerns over the privacy of collected and used data (Mohanta et al., 2021; Rajak et al., 2023). People have a right to be informed about all data collection, storage, and use that occurs with or without their knowledge, and they should be given the opportunity to consent to or take action over the use of their data (Chanal and Kakkasageri, 2020). A middle ground between data analysis and individual privacy can be achieved with privacy-preserving analytics techniques like data anonymization and differential privacy (Bhattarai et al., 2019). Moreover, protecting confidentiality and safety in smart transportation management encourages people to feel comfortable and confident in the transportation system's dependability (Ding et al., 2021). It is imperative that people have faith

*Corresponding Author: ggnavaneeth@gmail.com in the safety of their transportation systems and the confidentiality of their personal data (Ding et al., 2019). The widespread implementation and acceptance of smart transportation systems rely on this level of confidence. Robust privacy and security safeguards (Gupta et al., 2020) aid efficient transportation system management. By preventing unwanted entry and data breaches, transportation operations can continue relatively smoothly (Tariq et al., 2019). Authorities in the transportation sector can benefit greatly from the analysis of big data if they can do so in a safe and private way (Atitallah et al., 2020). This includes areas such as traffic management, passenger safety, vehicle administration, maintenance scheduling, and road network planning and design. Smart transportation management is impossible without security and privacy when using big data analytics. They protect private data, address privacy issues, build public trust, and provide effective data analysis, all of which contribute to transportation networks that are more secure and dependable. Concerns about gathering and use of personal information are mitigated by privacy-preserving methods (Chang, 2021). These safeguards promote the security and dependability of transportation networks by winning over the public's confidence.

Data encryption, access control, secure communication protocols, privacy-preserving analytics, and permission mechanisms are some of the current methods in security and privacy for smart transportation management employing big data analytics (Singh et al., 2020). Encryption protects private data and access controls stop hackers from getting into sensitive networks. Protocols for secure communication ensure that information is safe while in transit (Zeadally et al., 2020). Anonymization of data and differential privacy are merely a couple of instances of privacy-preserving analytics approaches that make it possible to conduct useful analyses without compromising users' personal information. Individuals are given agency over their data and privacy issues are addressed by consent methods (Arooj et al., 2022).

The complexity and scale of transportation data, however, make it difficult to apply effective security measures. The acquisition and use of private information raise privacy concerns. One of the difficulties in the transportation sector is making sure that security measures from different systems work together (Garg et al., 2019). Data analysis and privacy protection are two areas that can be difficult to reconcile. Keeping up with the ever-changing security and privacy standards is another layer of difficulty. Finally, concerns about public trust and stakeholder participation must be addressed if the plan is to be successfully implemented.

Improved network efficiency, security, and user experience are the objectives of Smart Transportation Management (STM) systems that use big data analytics. To ensure the effective implementation and acceptance of STM, however, it is necessary to resolve the substantial privacy and security concerns created by these innovations. Data availability in the context of potential cyber-attacks, protection of Internet of Things (IoT) endpoints, and authenticity and integrity of data maintenance are the key security issues. The use of distributed systems to ensure data availability, reliable encryption to prevent unauthorized access, and device authentication constitute essential solutions. Data reduction, user permission, data anonymization, and accessibility in data utilization are important aspects of privacy. Data masking and k-anonymity are important approaches to implement, as are specific information standards and user control. Improving security is possible with big data analytics by means of real-time monitoring, anomaly detection, predictive risk management, and encrypted, highly secure data storage with strong limitations on access. The main contribution of the paper is followed as:

- The present research intends to create a Secure Data Encryption Control based Big Data Framework (SDEC-BDF) for use in cyber-protected smart transportation management using big data analytics to safeguard sensitive data.
- The intention is to find an acceptable compromise between data analytics and privacy protection, ensuring that personal information is protected while simultaneously facilitating thorough data analysis.
- The proposed framework is intended to improve the transportation network's credibility with the public by safeguarding sensitive passenger data and preventing intruders' attempts to gain access to critical infrastructure.

The rest of the paper is structured as follows: This paper is organized as follows: Section 2 provides a review of the relevant prior research; Section 3 describes the Secure Data Encryption Control based Big Data Framework (SDEC-BDF); Section 4 describes the investigation's findings; and Section 5 provides a summary and conclusion.

Related works

This section introduces some research papers that examine the effects of technology in a variety of fields. Using tools like big data analytics, intelligent logistics management, and machine learning, Liu et al. (2020)

DOI: https://doi.org/10.52756/ijerr.2024.v40spl.008

presented a novel IoT based E-commerce business model (IoT-E-CBM). Hotel guests in the area can take advantage of cloud laundry's flexible, scalable, and user-friendly services because of the platform's ability to give the best washing solutions depending intelligently and dynamically on the current condition spaces of the laundry terminals. Companies that do their laundry on the cloud typically have greater capital turnover, liquidity, and profitability than their traditional laundry business counterparts.

Liu et al. (2021) proposed the Faster Region Convolutional Neural Networks (RCNN) model to segment a traffic image into multi-regions with different importance levels; then, multi-threshold image extraction schemes were designed based on progressive secret image sharing schemes to extract images containing key traffic information, such as reg numbers and human faces. However, there is always sensitive information in traffic photos, including license plate numbers or people's faces. The misuse of such information poses a risk to the privacy of drivers, passengers, pedestrians, and anybody else who uses the roads. The suggested techniques offer a reliable and clever means of extracting images for additional research in ITS.

The objective of the Security and Privacy Issues in Intelligent Transportation Systems (SPI-ITS) suggested by Hahn et al. (2021) is to improve mobility, comfort, safety, and efficiency by integrating sensing, control, analysis, and communication technologies into travel infrastructure and transportation. In this research, they categorize all the different ways in which ITS can compromise users' personal information. Finally, we point out areas where more study is needed to make ITS even more secure and private. Block chain-based solution for enhancing security and privacy (BC-SESP) was developed by Wan et al. (2019) for building dispersed networks, which fundamentally altered the conventional IIoT design. Then, to further enhance and perfect the new structure, they implement various relevant security technologies. The architecture's data-interaction mechanism and algorithms are then designed. Therefore, the suggested design is a major upgrade over the baseline architecture, ushering in a fresh approach to IIoT growth.

Gifty et al. (2019) first introduced Cyber physical systems with Weibull distribution (CPS-WD). Massive streams of data with varying characteristics are the focus of big data analysis in CPS. These data come from various independent origins and are processed using a decentralized database. The research discusses the recent issues in data privacy and evaluates the security and privacy implications of managing big data for CPS. Machine Learning-Based Big Data Analytics (ML-BDA), as presented by Li et al. (2021), has made it possible for these wearables to collect data at a previously unimaginable scale because of developments in the IoT. Recently, there has been a lot of interest in using big data analytics for data mining, knowledge extraction, and inference-making. In this paper, we provide an in-depth look at how machine learning may be used to analyze large amounts of healthcare data. Our research will help medical professionals and government bodies stay abreast of developments in machine learning-based big data analytics for intelligent healthcare.

Musa et al. (2023) provided a long-term solution to smart city transportation issues by developing a system that makes use of IoT and ITS. The traffic management in smart cities, which is made up of a combination of human-driven vehicles (HDV) and connected automated vehicles (CAV), requires an integrated strategy that considers both the decision-making challenges of previous research and the modeling and analysis of traffic. In addition, both techniques made use of ITS-based devices and artificial intelligence sensors to gather data on vehicles and road users in real-time. Management of traffic, regulations for traffic decision-making, and preservation for potential utilization can all profit from this data's processing and transmission using cloud computing and machine learning algorithms.

Public transportation networks in enormous Indian cities can be optimized for efficiency, customer experience, and sustainability by using big data analytics and the IoT (Rajbhandari et al., 2024). Here, they examined the state of public transportation in India and pointed out some of the challenges that the country faces. The following subject is about the IoT and big data as they pertain to solving urban transportation problems. Based on the results of this study, the IoT has the potential to greatly increase efficiency, safety, and rider satisfaction via data-driven optimization of scheduling, vehicle tracking, ticketing systems, ridership analysis, and maintenance schedules.

These research papers eliminated light on the current state of the art and the challenges faced by researchers in a wide variety of technological fields. They additionally propose novel approaches to these problems and highlight the importance of factors like privacy, security, and analysis in several contexts, for which finally, SDEC-BDF provides a solution.

Proposed method

Transportation management using big data analytics should prioritize security and Privacy. Data authentication is critical to ensure that only authorized individuals can access sensitive information. The rules governing information acquisition must specify precisely what data will be acquired, for what purpose, and by whom. In addition, data encryption is necessary to prevent unauthorized parties from reading the data. The analytics infrastructure should be routinely checked and upgraded to avoid and identify security vulnerabilities so that user data is always safe. Finally, data backups should be performed regularly in case any data is lost or stolen from its storage place. Security must also be considered for intelligent transportation management. Safety measures must be taken to prevent third parties from accessing users' personal information. Any personally identifiable information should be removed from the data.

Understanding user concerns and actions around data sharing in smart transportation management (STM) systems requires data privacy theories like contextual integrity and privacy calculus. When assessing and guaranteeing data security, security concepts, especially the CIA trinity (transparency, reliability, and availability), are used. Based on systems theory which provides an extensive perspective on how STM components are interdependent, we have created an integrated method. We use a mixed methods approach when conducting research, gathering data from both primary and secondary sources. Primary data comes from surveys and interviews with stakeholders, which help to comprehend the problem at concern. Secondary data comes from literature reviews and case studies, which help to set the findings into context. Other analytical approaches are quantitative, such as real-time monitoring and anomaly identification using machine learning and statistical analysis, whereas other methods are qualitative, like subject and content analysis. Various approaches for risk assessment, privacy impact evaluations, blockchain frameworks, and encryption/access control are used to handle and ease privacy and security issues effectively.

In addition, users should be able to participate in or decline certain types of data collecting. It is also crucial to know whether the data acquired will be utilized in ways other than those for which it was first gathered. When employing big data analytics for smart transportation management, it is crucial to ensure the safety and confidentiality of all collected data. Developing a clear set of regulations and safety procedures is important to provide consumers with a safe, private, and reliable experience. Using big data analytics may compromise the security and efficiency of smart transportation initiatives if suitable security and privacy policies are not in place.

The proliferation of private automobiles on the road has made intelligent vehicular technology an integral part of ITS. As the industry evolves, smart car innovations have the greatest influence on travel and transportation infrastructure. Using Vehicle-to-Vehicle (V2V) communication technology, autonomous cars may form convoys to increase efficiency, as seen in Figure 1. Vehicles can now share information and form groups because of studies on vehicular ad hoc networks (VANETs). To improve the security and efficiency of transportation networks, cars may form a network using wireless communication devices to share information about hazards, travel times, and more. Various distributed control tasks need communication between individual smart cars and components inside vehicles. These days, public transportation networks provide a crucial service in many cities worldwide.

$$\int Rr(D_0, sa(A_d)), \partial_0(D_r) * dA_0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

$$\int Rr(D_0, At(A_d)), \partial_0(D_r) * \sum_{i=0}^{1} ef + \sum_{i=0}^{1} cp \frac{dA_0}{Rq} \quad \ldots\ldots\ldots(2)$$

Receiver encrypts Rr the written content and sends it to the agent sa such that the data object D_0 can be

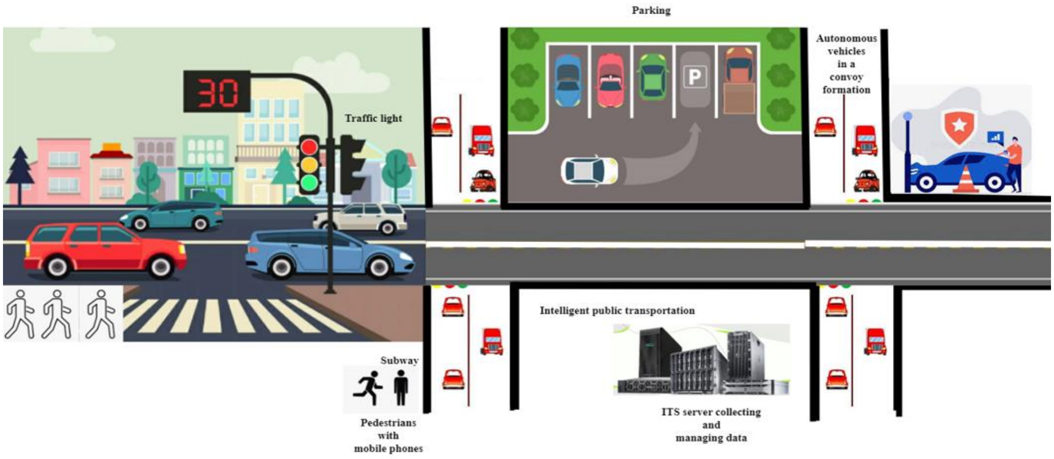
Parking
Autonomous vehicles in a convoy formation
30
Traffic light
P
Intelligent public transportation
Subway
Pedestrians with mobile phones
ITS server collecting and managing data

Figure 1. Smart transportation management system.

protected. Distributor(D_r) satisfies all agents' requests by verifying the availability of data A_d and whether they have a history of bad weather. The agent At contacts the distributor with a request Rq for the necessary data objects. Near-constant service of trains and buses ensures that people can go where they need to go quickly and affordably by applying Equations (1) and (2). The efficiency ef and capacity cp of public transportation networks may be improved with the help of ITS.

Smart bus stops may notify waiting passengers of any changes to the bus timetable, including arrival and departure timings and delays, and taking into consideration real-time traffic circumstances (such as congestion from other cars and accidents). At the same time, optimizing routes may improve passenger comfort and save trip time. The controllers of an intelligent transportation system are its administrative, controlling, or altering parts. The controllers in ITS show that the system has seen, processed, and acted upon the data gathered.

In the transportation industry, controllers may be found in traffic signals, public address systems, electronic road signs, and railroad switches. Decisions in ITS are made by controllers using sensor data. Extremely simplistic state machines, such as an if-then-else structure, are often used as controllers in ITS. Some controller devices contain built-in vulnerabilities that may enable attackers to alter traffic patterns. A bus's controller may get its control input from a pedestrian waiting at the bus stop, or it may receive its control input from parking spot sensors, allowing it to update the parking availability sign.

Smart Traffic management

The analytical and intelligent parts of ITS inform the decisions made by smart traffic controllers. The data collected from connected cars, public transportation systems, and other IoT devices is used to improve routing algorithms and traffic schemes used by smart traffic controllers. Intelligent traffic controllers react to and contribute to the data collection process. Camera sensor systems, for instance, are often used in today's traffic lights to identify the presence of approaching automobiles. Another method of increasingly common intelligent traffic management on U.S. roads and interstates is the use of changeable message signs. Modifying these signs may warn drivers of bad weather, accidents, dangers, speed limits, and more.

$$sc_i^* = td_i - \sum \alpha_{ij} C * td_j \ldots\ldots\ldots(3)$$

$$\frac{dC_{ni}}{dt} = a_i df_i(\{dy_i - d_i - \sum \alpha_{ij} sf_j\})/\sum \alpha_{ij} C * td_j \ldots (4)$$

where, C_i^* - ratio of strong competition, Z_i - novel technological developments, α_{ij} – coefficient of competitiveness time-dependent i and$j$, C_j – competition in several different arenas $a_i df_i$. With a communication network $\frac{dC_{ni}}{dt}$ linking many smart traffic controllers in Equation (3), each controller's observations may be shared with the others$\sum \alpha_{ij} sf_j$, improving the system's responsiveness to traffic dynamics $\sum \alpha_{ij} C * td_j$ and emergencies in Equation (4). In the case of an emergency, for instance, notifying the rest of the system would allow first responders to get to the scene faster and have more time to take the necessary measures$(\{dy_i - d_i - \sum \alpha_{ij} sf_j\})$. It's traffic algorithms and artificial intelligence (AI) components rely on data collected and

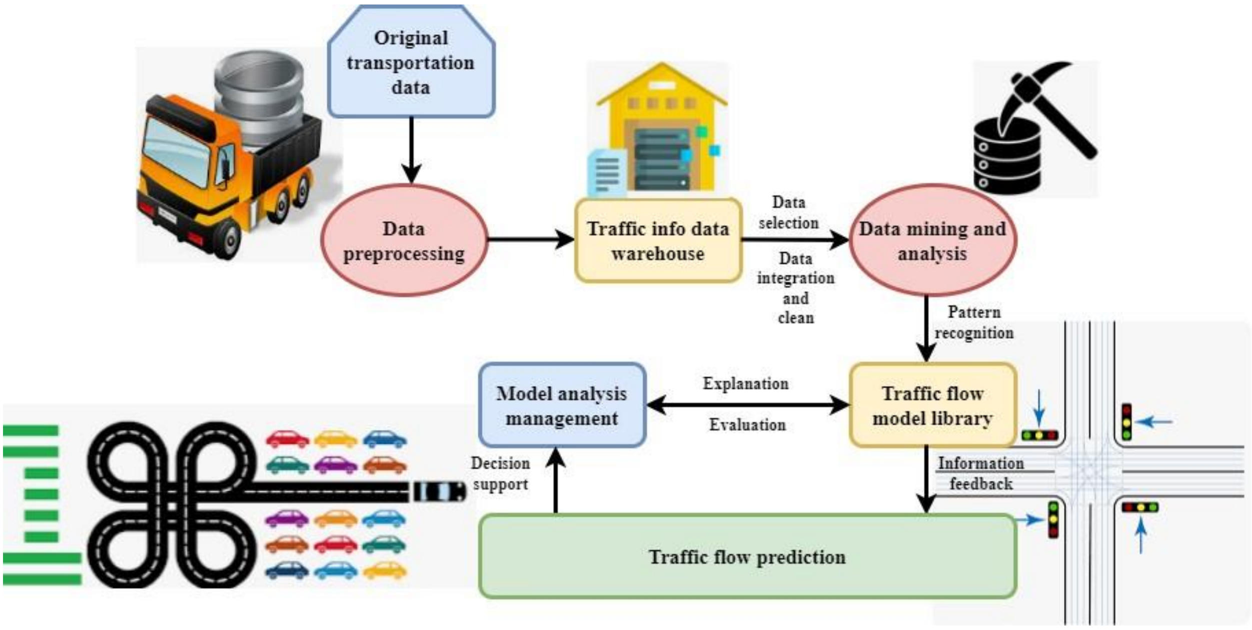

Figure 2. Data collection for traffic management.

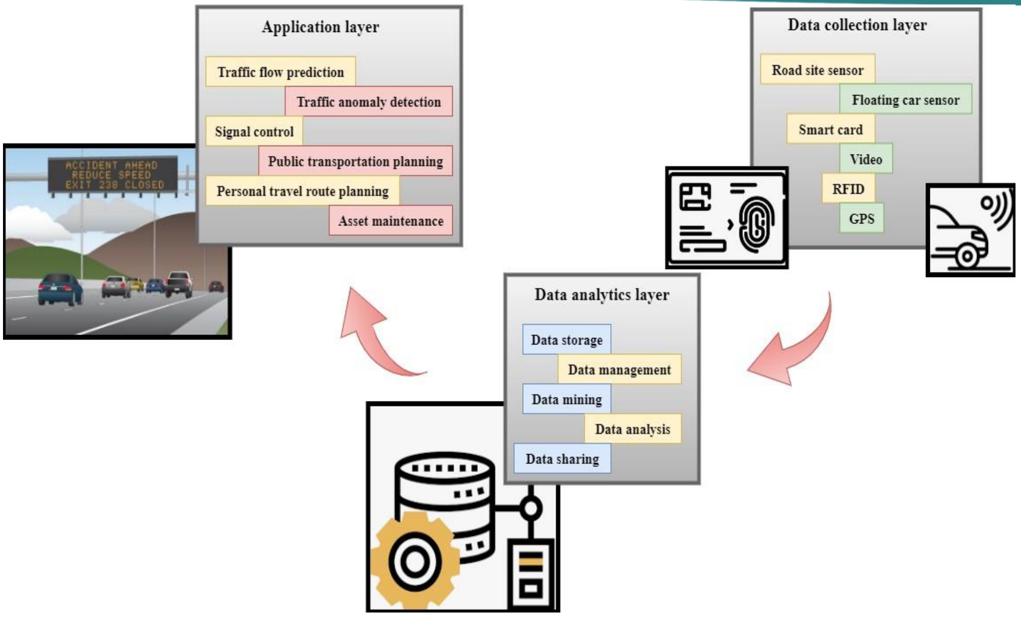

Figure 3. Data Management process using BDA.

processed by traffic management centers, which serve as the nerve center of smart traffic control.

Big Data might provide useful technical support for ITS research, development, and deployment. By efficiently gathering, evaluating, and processing data in the road and rail transportation system in a timely way, Big Data applications have the potential to supply the general public with convenient and highly effective means of transportation.

$$\int_{tp=i}^{1} pm + ct\frac{dA_0}{Rq} * dtf_{\mathrm{n}} = \int_{i=0}^{1} \partial_0(D_r) * dpp \quad \ldots\ldots(5)$$

To address problems, improve ITS performance pm, reduce costs ct, and gain useful understanding. Transportation planning relies heavily on up-to-date and accurate data on traffic state as $\frac{dA_0}{Rq}$. Predicting traffic flow is one area where intelligent transportation systems using big data analytics excel. Traditional traffic flow prediction dtf_{n} utilizing big data analytics is seen in Figure 2. To get the final useful data set, the raw data from the ITS must first undergo preprocessing pp. A traffic flow model is built from the cleaned and prepared data using a specific data mining or analytic technique stated in above Equation (5). The data from real traffic flows is used to calibrate the traffic flow model, which in turn aids the traffic management department in making decisions.

An intelligent transportation system integrates cutting-edge technologies into the transportation network, including electronic sensor technology, data transmission technology, and intelligent control technology. All users stand to gain from the enhancements made possible by implementing intelligent transportation systems. A transportation management system, traveler information system, vehicle control system, commercial vehicle management system, public transit system, and urban transportation system are the six essential components of an ITS system.

Figure 3 depicts the three layers that comprise the big data analytics-based data management process. These layers are, respectively, the data collection layer, the data analytics layer, and the application layer.

- The data collection layer is the starting point of the design since it is responsible for gathering all the data that subsequent levels will use. Information is gathered by various technologies such as radio frequency identification, global positioning systems, induction loop detectors, microwave radars, video surveillance, remote sensing, and radio frequency identification.

$$rt_{op} = \frac{AR*S*md*\tau_{op}}{Q_{op}} \quad \ldots\ldots(6)$$

$$\delta = \int_{i=0}^{1} \partial_0(D_r) * dpp \quad \ldots\ldots(7)$$

- Likewise, route optimization rt_{op} might potentially boost traveler satisfaction and cut down on time spent in transit Q_{op}. Administrative, controlling, or modifying components $AR * S * md$ of a smart transportation system are known as controllers. The ITS controllers demonstrate that the network has observed, analyzed, and responded upon the collected data with the help of Equations (6) and (7).
- The data analytics layer should be the starting point for any design. This layer's major responsibility is to take data from the data collection layer and complete storing, managing, mining, analyzing, and disseminating the data using various Big Data

DOI: https://doi.org/10.52756/ijerr.2024.v40spl.008

analytics methodologies and the corresponding platform using Equations (6) and (7).

- The application layer operates atop this architecture. Data processing results from the analytics layer are used in various transportation-related scenarios, including but not limited to traffic flow forecasting, traffic directing, signal management, emergency rescue, etc.

$$mt = \begin{bmatrix} ad_{1,1} & \cdots & ad_{1,n} & \cdots & pd_{1,n-1} \\ \vdots & & \vdots \quad \vdots \quad \vdots & & \vdots \\ ad_{2,1} & \vdots & ad_{2,n} & \vdots & pd_{2,n-1} \\ \vdots & & \vdots \quad \vdots \quad \vdots & & \vdots \\ ad_{n,1} & \cdots & ad_{n+1,n} & \vdots & pd_{n+1,n-1} \end{bmatrix}_n \quad \ldots (8)$$

The data-gathering layer employs cutting-edge techniques to track traffic, vehicular activity, and environmental variables through the matrix mt. Raw traffic data through wires or wireless is sent into the data analytics layer as ad. This information might be fully or partially organized. The data analytics layer receives the original traffic data, where it is then sorted, duplicates are eliminated, data is cleansed, and the findings are presented using Equation (8). Applying mathematics and engineering theory, it then employs descriptive and predictive analysis to uncover hidden information. The application layer uses the analytical findings to make decisions that help the city's management team forecast future traffic and passenger flow, pinpoint areas prone to traffic accidents, fine-tune signal distribution, and enforce traffic control.

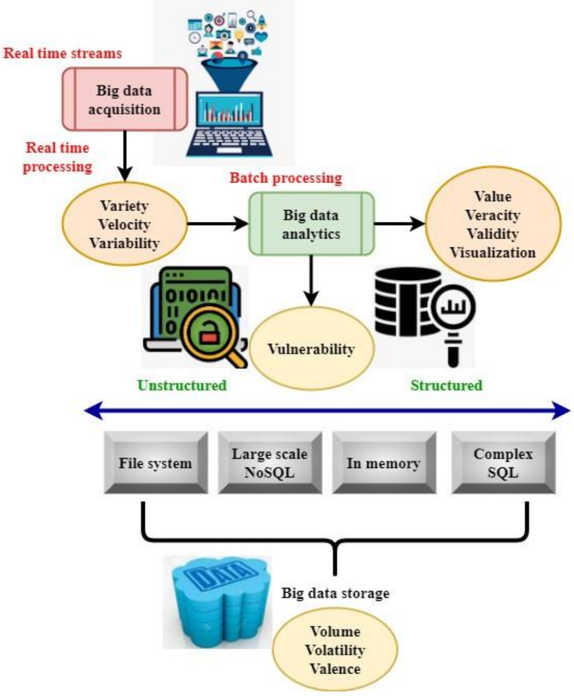

Figure 4. Secure Data Encryption Control-based Big Data Framework.

Figure 4 explains secure data encryption control-based Big Data framework. The 3Vs, or volume, velocity, and variety, are often used to characterize large databases and data warehouses, and they inspired the phrase big data to describe the immense volumes of data in today's digital world. Due to technological advancements, however, big data is now also characterized by other, more semantically relevant characteristics. The reliability or accuracy of information is known as veracity, and it is one of the most prominent extra dimensions. Therefore, big data's three most important characteristics are Validity, volatility, and value.

$$bd_{vi}^{vr} = \begin{cases} dm_{val} + sec\left(bd_{ql} - bd_{ch}\right), mt < pd_{n+1,n-1} \\ \qquad\qquad pd_{n+1}, 0 \end{cases} \quad \ldots (9)$$

$$Pv = \begin{cases} dm_{vul} + bd_{vi}^{vr}, mt < pd_n \\ \quad bd \qquad\quad , otherwise \end{cases} \quad \ldots\ldots(10)$$

Additionally, aspects of big data bd, such as variability vr and visualization vi, have changed, and new technological obstacles have been recognized. In particular, the dimensions dm of Valence dm_{val} and Vulnerability dm_{vul} have been linked to the privacy Pv and security concerns sec that now plague today's big data. To prepare for the security concerns that will inevitably accompany future big data endeavors, the qualities bd_{ql} and characteristics bd_{ch} of all eleven dimensions in this section as big data continues to expand beyond the conventional 3Vs has been identified from the above-mentioned Equations (9) and (10).

Volume

The sheer magnitude of the data being gathered is responsible for the first characteristic of big data: Volume. The vast bulk of today's data has just recently been gathered during the last few years, as commercial websites have been coupled with social networking and mobile apps. The quantity of information created and stored every day is expanding at an exponential rate. YouTube and other social networks see an average of 300 hours of video and other huge multimedia assets uploaded per minute. Many billions of records have been amassed due to businesses' usage of social media sharing platforms and standard transactional data. Petabytes, exabytes, and even zettabytes of data might one day be collected from a single source. The proliferation of big data threatens security and privacy in at least two keyways:

- When data is spread across several servers, it becomes hard for conventional databases and software tools to keep continual vigil and make sure proper security precautions are implemented.
- Security flaws and dangers are introduced whenever a cluster or node failure impacts data transactions or performance within the tolerance time constraints.

DOI: https://doi.org/10.52756/ijerr.2024.v40spl.008

Velocity

The second component, velocity, refers to how quickly organizations generate and consume data, as well as how urgently they need to handle this data in real time.

The effect may be seen in big data analytics, where the rapidity of data production necessitates keeping pace with the processing power of modern computers in real-time. There is room for improvement in data storage capacity, but the pace at which information is being created should be prioritized.

Variety

The Variety dimension of big data describes the fact that it comes in three distinct flavors: structured, semi-structured, and unstructured. Most of this information is unstructured, such as audio, video, picture, and sensor signal files or logs from the Internet, satellites, computers, and other devices. The Variety dimension encompasses the many data representations mentioned above and the various forms and channels through which the same information is communicated. It is crucial to distinguish not only the most frequent variety, which shows the structural diversity of data representation, but also the media diversity, or the various medium in which the same data is represented, and the semantic diversity, which indicates distinct interpretations depending on the varied circumstances in which data appears.

Veracity

Considering the increased ambiguity around data streaming and availability, the credibility or quality of data has been proposed to describe veracity in addition to the previously stated three criteria of big data (volume, variety, and veracity). Improving the trustworthiness of big data may help reduce the dangers associated with making business decisions.

Validity

The fifth component, "Validity," is like "Veracity" in that it relates to the extent to which data may be used for its intended purpose. Therefore, Validity verifies the accuracy of the data for a certain application or perspective, allowing to use big data analytics in real-world settings.

Volatility

While the dimensions of veracity and Validity are characteristic of big data quality assurance, volatility is a temporal element of data that dictates how long the data must remain valid before it is removed from storage.

Value

$$vn_i^f = Cp(\underset{R}{\mathrm{ot}}\, Q\,,\, \underset{1....R}{\mathrm{roi}})\, P \qquad .... (11)$$

$$sp = cr(Q) = sec\left(bd_{ql} - bd_{ch}\right), mt < pd_{n+1,n-1} + I\ ... \quad (12)$$

The traditional search selective has been replaced with this value of networks vn . Feature maps serve as input vn_i^f , and a set of potential candidates Cp serves as output ot . roi pooling P compresses data of varying lengths into a uniform output shown in Equation (11). The candidate region's class and its specific picture sp coordinates $cr(Q)$are output by the classification as well as regression layers, respectively shown in Equation (12).

Understanding the value of big data to diverse parts of an organization has been pinpointed as the seventh dimension of big data. It guides businesses toward an effective big data strategy, increasing the likelihood that big data analytics will provide the rich troves of information needed to address difficult operational

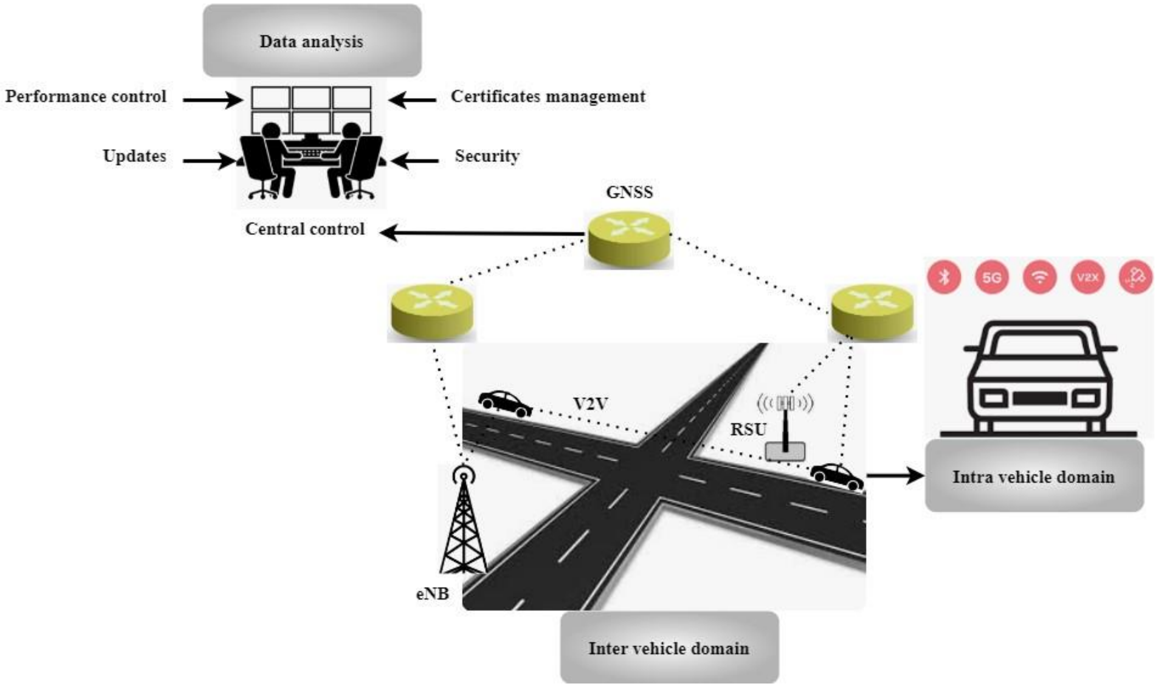

Figure 5. Privacy and security for managing transportation.

challenges. The importance of analytics contributing to action in organizations means that effective access controls and permissions over analytical assessments are essential. Finding the right security measures is also essential for developing such data insights.

Variability

Variability, the eighth dimension of big data, relates to discrepancies caused by the fact that different data sources may load data into the data storage at different speeds, formats, or types, affecting the other seven Vs.

$$f(ip)\int dt = \frac{1}{2}(li^{-db}) * \sum_{0}^{t}(lg^{-mp} + \exp(upd)) \quad .... (13)$$

From the above inclined Equation (13), it might be simply finding interesting patterns $f(ip)$in the integrated data$\int dt$. Generate actionable insights from big data by recording and linking information li about big data's variability to data already stored in a database db. IT operations need to modify their audit log gathering lg and monitoring practices mp using summation function to account for the unpredictability of big data upd as an exponential term.

Figure 5 shows the scenario of privacy and security for managing transportation. Functionality and communication linkages between ITS nodes are described highly in the ITS architecture. It comprises several parts that work together and are split into the intra-vehicle and extra-vehicle categories. The data journey across ITS applications begins with data collecting. It allows for the collection of all observable measurements from various sources to be shared between moving vehicles and stationary infrastructure.

Intra Vehicle Space

$$IVS = \int_{-\infty}^{\infty}\beta(ECU) - (\log aU * \frac{1}{dni^{2}}) \quad .... (14)$$

There is a steadily growing trend toward more electrical complexity and embedded gadgets in today's automobiles explained using Equation (14). Electronic operate Units ECU are embedded computer systems often by logarithmic function$\log aU$ used in automobiles to operate various aspects of an automobile, such as the vehicle's engine IVS or its in-car entertainment system, over the twice of dispersed network dni^{2} . Several different types of bus communication networks inside a vehicle allow electronic control units (ECUs) to talk to one another. Each one may be useful in different situations due to differences in criticality, cost, bandwidth, and time needs.

Inter Vehicle Space

$$Rd(nm) = \frac{1}{2}(ivp^{-e}) * \sum_{i=0}^{1}(sr^{-\alpha} + \exp(2de)) \quad ...(15)$$

Vulnerable Road Users (VRU) refers to non-motorized nm road users Rd including walkers and bicyclists $Rd(nm)$, as well as regional infrastructure, and the inter-vehicle paradigm ivp encompasses all forms of communication between cars and their surroundings sr. Any vehicle having an OBU installed may join the network and exchange data $(2de)$for a wide range of uses, including but not limited to safety, traffic control, and information dissemination with Equation (15).

This paper provides an overview of the analysis conducted by Framework SDEC-BDF in the context of smart transportation management, covering the topics of security, privacy, efficiency, interoperability, resource scarcity, and energy management.

Results and Discussion

This section offers an overview of the research findings relating to the topics of security, privacy, efficiency, interoperability, resource scarcity, and energy management analysis of the Secure Data Encryption Control based Big Data Framework (SDEC-BDF) in the context of smart transportation management. The review additionally places an emphasis on the framework's effectiveness, its interoperability with a variety of data formats, its capacity to deal with limited resource availability, and its ability to successfully manage the consumption of energy.

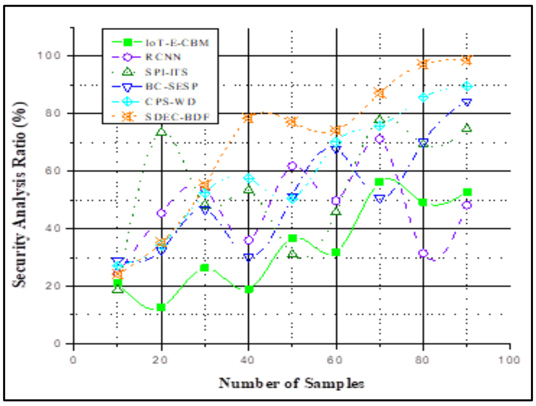

Figure 6. Security Analysis.

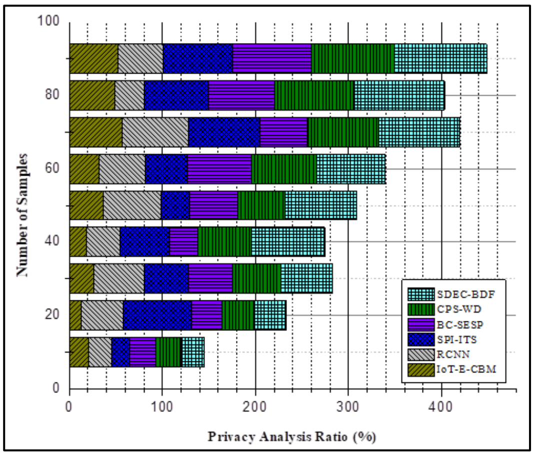

Figure 7. Privacy Analysis.

The above Figure 6 reveals that the algorithms for encryption utilized within the SDEC-BDF framework are capable of protecting sensitive transportation data in an

DOI: https://doi.org/10.52756/ijerr.2024.v40spl.008

efficient manner. They have a high level of resistance against the known attack vectors, which guarantees the data's integrity and confidentiality. The access control techniques provide a fine-grained level of control over data access while incurring a low amount of overhead. This helps to ensure that authorized individuals may access the system and stops unauthorized breaches from occurring.

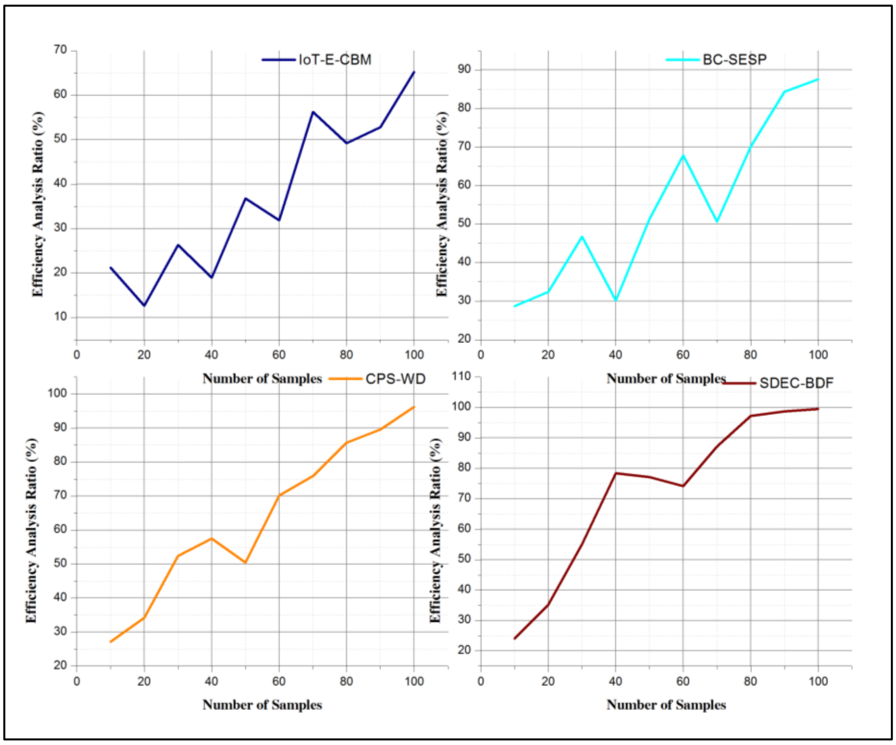

Figure 8. Efficiency Analysis.

The effectiveness of the anonymization and pseudonymization mechanisms that have been applied in the SDEC-BDF framework has been highlighted in the privacy evaluation. The above Figure 7 shows that there is a low chance of re-identification in the data that has been anonymized. This ensures that the privacy of individuals may be protected while still enabling data analysis to take place. The architecture exhibits resistance to intrusions into users' privacy, displaying durability against de-anonymization and inference attacks.

In Figure 8, according to the findings of the efficiency research, the SDEC-BDF architecture presents exclusively a marginal increase in computational burden when it comes to the implementation of encrypting it, control of access enforcement, and privacy protection measures. Both the processing speed and the response times continue to fall within the allowed range, which ensures that transportation data will be handled effectively. The framework has a high degree of scalability, which allows it to effectively manage growing data loads without causing a major drop in performance.

In Figure 9, the interoperability evaluation sheds light on the framework's capacity to manage the varied data formats and standards that are typical of situations, including smart transportation. The SDEC-BDF architecture integrates data standardization techniques, which make it possible to combine data coming from a variety of sources. This means that different sorts of data, such as information regarding traffic flow, data regarding vehicle tracking, and passenger records, may be handled and evaluated in an efficient manner inside the framework.

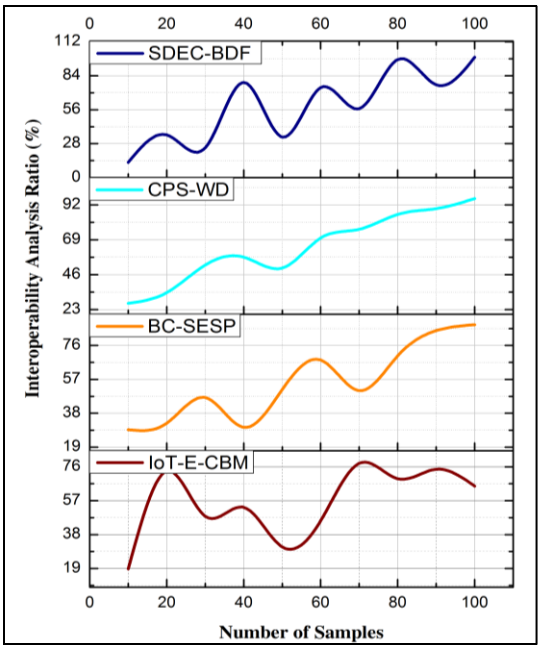

Figure 9. Interoperability Analysis.

DOI: https://doi.org/10.52756/ijerr.2024.v40spl.008

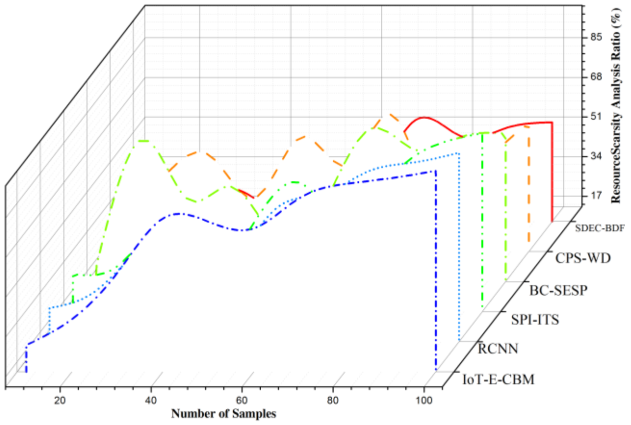

Figure 10. Resource scarcity.

The analysis investigates the scalability of the framework as well as its ability to manage resources, paying particular attention to situations in which there may be a shortage of resources because of rising data volumes or system expansion. From the above Figure 10, as the size of the system expands, the SDEC-BDF architecture exhibits its scalability by effectively allocating and managing the system's resources. The evaluation shows that the framework's resource management algorithms adequately handle resource scarcity, which ensures that operation can continue without interruption while preserving the necessary security and privacy protections.

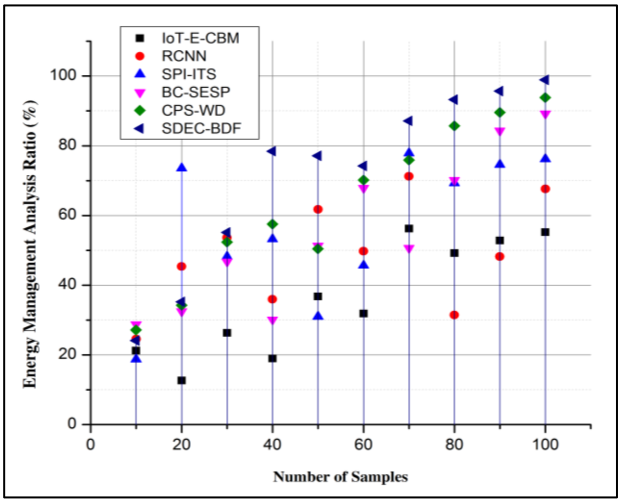

Figure 11. Energy management.

From Figure 11, the analysis considers the amount of energy that is consumed by the framework as well as the impact that this consumption has on the availability of resources, particularly in settings where resources are restricted or in the case of devices that have energy restrictions. The evaluation demonstrates that the SDEC-BDF framework employs energy-saving techniques to reduce the amount of energy that is consumed. These mechanisms include efficient data processing algorithms and low-power modes. Even in situations with a limited supply of resources, the framework can continue to function, attributable to its energy-efficient design.

Encryption techniques protect sensitive transportation data from known attack vectors, according to the research. Data access is fine-grained and efficient using access control techniques. While facilitating data analysis, the framework's anonymization and pseudonymization methods protect privacy. Its computing performance, scalability, and resource consumption are efficient. The interoperability analysis shows the framework's data format compatibility and system integration. In resource-constrained contexts, the resource scarcity analysis stresses the framework's resource management tactics and scalability. Finally, the energy management analysis reveals the framework's energy-saving techniques for sustainable operation. The SDEC-BDF system for secure, privacy-preserving smart transportation management utilizing big data analytics is proven effective and suitable by the research.

Conclusion

In the context of smart transportation management utilizing big data analytics, security and privacy are of the utmost importance. By balancing data analytics and privacy protection, the proposed Secure Data Encryption Control-based Big Data Framework (SDEC-BDF) safeguards critical information and keeps the public's faith in the transportation system. The framework protects against cyber threats and illegal access to vital systems by integrating comprehensive security mechanisms such as data encryption, access control and secure communication protocols. Concerns about data gathering and usage are addressed by privacy-preserving analytics methods, which safeguard personal information without compromising analytical power. The framework safeguards critical passenger data, prevents hackers from gaining access, and increases public trust in the transportation system. Building a reliable transportation system requires relentlessly optimizing for safety, privacy and productivity. The simulation analysis performed in the present research requires the system's capability to attain these goals, laying the groundwork for further advancements in safe and confidential smart transportation management. The benefits of big data analytics can be realized in the transportation sector by resolving the problems and implementing the proposed framework; doing such will protect privacy, ensure security and lead to a more dependable and efficient transportation network.

Conflict of interest

The authors declare that there is no conflict of interest.

DOI: https://doi.org/10.52756/ijerr.2024.v40spl.008

References

Al-Turjman, F., Zahmatkesh, H., & Shahroze, R. (2022). An overview of security and privacy in smart cities' IoT communications. *Transactions on Emerging Telecommunications Technologies*, *33*(3), e3677. https://doi.org/10.1002/ett.3677

Arooj, A., Farooq, M. S., Akram, A., Iqbal, R., Sharma, A., & Dhiman, G. (2022). Big Data Processing and Analysis in Internet of Vehicles: Architecture, Taxonomy, and Open Research Challenges. In *Archives of Computational Methods in Engineering*, *29*(2), 793–829. https://doi.org/10.1007/s11831-021-09590-x

Atitallah, S. Ben, Driss, M., Boulila, W., & Ghezala, H. Ben. (2020). Leveraging Deep Learning and IoT big data analytics to support the smart cities development: Review and future directions. In *Computer Science Review*, *38*. https://doi.org/10.1016/j.cosrev.2020.100303

Bhattarai, B. P., Paudyal, S., Luo, Y., Mohanpurkar, M., Cheung, K., Tonkoski, R., Hovsapian, R., Myers, K. S., Zhang, R., Zhao, P., Manic, M., Zhang, S., & Zhang, X. (2019). Big data analytics in smart grids: State-of-theart, challenges, opportunities, and future directions. In *IET Smart Grid 2*(2), 141-154. https://doi.org/10.1049/iet-stg.2018.0261

Chanal, P. M., & Kakkasageri, M. S. (2020). Security and Privacy in IoT: A Survey. In *Wireless Personal Communications*, *115*(2), 1667–1693. https://doi.org/10.1007/s11277-020-07649-9

Chang, V. (2021). An ethical framework for big data and smart cities. *Technological Forecasting and Social Change*, *165*. https://doi.org/10.1016/j.techfore.2020.120559

Ding, W., Jing, X., Yan, Z., & Yang, L. T. (2019). A survey on data fusion in internet of things: Towards secure and privacy-preserving fusion. *Information Fusion*, *51*, 129-144. https://doi.org/10.1016/j.inffus.2018.12.001

Ding, Y., Jin, M., Li, S., & Feng, D. (2021). Smart logistics based on the internet of things technology: an overview. *International Journal of Logistics Research and Applications*, *24*(4), 323–345. https://doi.org/10.1080/13675567.2020.1757053

Garg, S., Singh, A., Kaur, K., Aujla, G. S., Batra, S., Kumar, N., & Obaidat, M. S. (2019). Edge Computing-Based Security Framework for Big Data Analytics in VANETs. *IEEE Network*, *33*(2), 72-81. https://doi.org/10.1109/MNET.2019.1800239

Gifty, R., Bharathi, R., & Krishnakumar, P. (2019). Privacy and security of big data in cyber physical systems using Weibull distribution-based intrusion detection. *Neural Computing and Applications*, *31*, 23–34. https://doi.org/10.1007/s00521-018-3635-6

Gupta, M., Abdelsalam, M., Khorsandroo, S., & Mittal, S. (2020). Security and Privacy in Smart Farming: Challenges and Opportunities. *IEEE Access*, *8*, 34564-34584. https://doi.org/10.1109/ACCESS.2020.2975142

Hahn, D., Munir, A., & Behzadan, V. (2021). Security and privacy issues in intelligent transportation systems: Classification and challenges. *IEEE Intelligent Transportation Systems Magazine*, *13*(1), 181-196. https://doi.org/10.1109/MITS.2019.2898973

Kumar, A., Dutta, S., & Pranav, P. (2023a). Prevention of VM Timing side-channel attack in a cloud environment using randomized timing approach in AES – 128. *Int. J. Exp. Res. Rev.*, *31*(Spl Volume), 131-140. https://doi.org/10.52756/10.52756/ijerr.2023.v31spl.013

Kumar, A., Dutta, S., & Pranav, P. (2023b). Supervised learning for Attack Detection in Cloud. *Int. J. Exp. Res. Rev.*, *31*(Spl Volume), 74-84. https://doi.org/10.52756/10.52756/ijerr.2023.v31spl.008

Li, W., Chai, Y., Khan, F., Jan, S. R. U., Verma, S., Menon, V. G., Kavita, & Li, X. (2021). A Comprehensive Survey on Machine Learning-Based Big Data Analytics for IoT-Enabled Smart Healthcare System. *Mobile Networks and Applications*, *26*(1), 234–252. https://doi.org/10.1007/s11036-020-01700-6

Liu, C., Feng, Y., Lin, D., Wu, L., & Guo, M. (2020). Iot based laundry services: an application of big data analytics, intelligent logistics management, and machine learning techniques. *International Journal of Production Research*, *58*(17), 5113–5131. https://doi.org/10.1080/00207543.2019.1677961

Liu, Y., Yang, C., & Sun, Q. (2021). Thresholds Based Image Extraction Schemes in Big Data Environment in Intelligent Traffic Management. *IEEE Transactions on Intelligent Transportation Systems*, *22*(7), 3952-3960. https://doi.org/10.1109/TITS.2020.2994386

Musa, A. A., Malami, S. I., Alanazi, F., Ounaies, W., Alshammari, M., & Haruna, S. I. (2023). Sustainable Traffic Management for Smart Cities Using Internet-of-Things-Oriented Intelligent Transportation Systems (ITS): Challenges and Recommendations. *Sustainability, 15(13),* 9859.

DOI: https://doi.org/10.52756/ijerr.2024.v40spl.008

https://doi.org/10.3390/su15139859

Mohanta, B. K., Jena, D., Ramasubbareddy, S., Daneshmand, M., & Gandomi, A. H. (2021). Addressing Security and Privacy Issues of IoT Using Blockchain Technology. *IEEE Internet of Things Journal*, *8*(2), 881-888. https://doi.org/10.1109/JIOT.2020.3008906

Neilson, A., Indratmo, Daniel, B., & Tjandra, S. (2019). Systematic Review of the Literature on Big Data in the Transportation Domain: Concepts and Applications. In *Big Data Research*, *17*, 35-44. https://doi.org/10.1016/j.bdr.2019.03.001

Rajak, R., Choudhary, A., & Sajid, M. (2023). Load balancing techniques in cloud platform: A systematic study. *Int. J. Exp. Res. Rev.*, *30*, 15-24. https://doi.org/10.52756/ijerr.2023.v30.002

Rajbhandari, S., & Sharma, R. (2024). Using Big Data and the Internet of Things to Optimize Public Transport Efficiency Across Major Cities in India. *Journal of Intelligent Connectivity and Emerging Technologies*, *9*(1), 13-24.

Samadder, M., Barman, A., & Roy, A. (2023). Examining a generic streaming architecture for smart manufacturing's Big data processing in Anomaly detection: A review and a proposal. *Int. J. Exp. Res. Rev.*, *30*, 219-227. https://doi.org/10.52756/ijerr.2023.v30.019

Singh, S. K., Rathore, S., & Park, J. H. (2020). BlockIoTIntelligence: A Blockchain-enabled Intelligent IoT Architecture with Artificial Intelligence. *Future Generation Computer Systems*, *110*, 721-743. https://doi.org/10.1016/j.future.2019.09.002

Soomro, K., Bhutta, M. N. M., Khan, Z., & Tahir, M. A. (2019). Smart city big data analytics: An advanced review. In *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *9*(5). https://doi.org/10.1002/widm.1319

Tariq, N., Asim, M., Al-Obeidat, F., Farooqi, M. Z., Baker, T., Hammoudeh, M., & Ghafir, I. (2019). The security of big data in fog-enabled IOT applications including blockchain: A survey. *Sensors (Switzerland)*, *19*(8), 1788. https://doi.org/10.3390/s19081788

Wan, J., Li, J., Imran, M., & Li, D. (2019). A blockchain-based solution for enhancing security and privacy in smart factory. In *IEEE Transactions on Industrial Informatics*, *15*(6), 3652-3660. https://doi.org/10.1109/TII.2019.2894573

Zeadally, S., Siddiqui, F., Baig, Z., & Ibrahim, A. (2020). Smart healthcare: Challenges and potential solutions using internet of things (IoT) and big data analytics. *PSU Research Review*, *4*(2), 149-168. https://doi.org/10.1108/PRR-08-2019-0027

How to cite this Article:

Govindasamy R., Shanmugapriya N., and Gopi R. (2024). Security and Privacy for Smart Transportation Management using Big Data Analytics. *International Journal of Experimental Research and Review*, *40*(spl.), 104-116.

DOI: https://doi.org/10.52756/ijerr.2024.v40spl.008