Journal of Computer Sciences and Applications, 2019, Vol. 7, No. 1, 31-36
Available online at http://pubs.sciepub.com/jcsa/7/1/5
Published by Science and Education Publishing
DOI:10.12691/jcsa-7-1-5

Efficient Data Privacy and Security in Autonomous Cars

Rushit Dave[1,*], Evelyn R Sowells Boone[2], Kaushik Roy[3]

[1]Department of Computer Science, North Carolina A&T State University, Greensboro, USA
[2]Department of Computer System Technology, North Carolina A&T State University, Greensboro, USA
[3]Department of Computer Science, North Carolina A&T State University, Greensboro, USA
*Corresponding author: rrdave@ncat.edu

Received March 06, 2019; Revised April 17, 2019; Accepted May 10, 2019

Abstract As the advancement and testing of self-driving auto innovation has advanced, the possibility of exclusive self-ruling vehicles working on open streets is nearing. Industry specialists foresee that self-governing vehicles will be financially accessible inside the following five to ten years. As automation becomes more prevalent in the transportation industry, driverless vehicles are appearing more frequently in the news. Asymmetric algorithms have shown their impact on large amount of data that have been secured by generating a public key. Cyber security is the major concern for any autonomous vehicle. The main contribution of this research is to secure data using an Asymmetric algorithm technique which will be saved in cloud storage.

Keywords: *automated cars, privacy, security, RSA algorithm*

Cite This Article: Rushit Dave, Evelyn R Sowells Boone, and Kaushik Roy, "Efficient Data Privacy and Security in Autonomous Cars." *Journal of Computer Sciences and Applications*, vol. 7, no. 1 (2019): 31-36. doi: 10.12691/jcsa-7-1-5.

1. Introduction

Driverless vehicles are currently gaining momentum not only in industry, but also in the public forum. The media coverage on driverless transportation, especially cars, has increased over the past decades. A growing number of news outlets are not only focusing on companies developing automated driverless technologies, but also consumer experiences with driverless cars. For example, recent news articles have reported self-driving cars with topics ranging from the technology advancements for self-driving vehicles [1], to the benefits of self-driving vehicles for the elderly and those with disabilities [2] , and of course to the catastrophic – a self-driving car killing a pedestrian [3]. Driverless cars are frequently reported on as new developments happen, whether they are technological developments, social developments, or policy developments. However, not all development is positive. With the increasing capability of driverless technologies (i.e. automated braking systems, lane change assist) there is also a certain level of accountability companies' products are expected to work properly and safely. In the case of automated driving technologies, it is immediately clear to the designers, engineers, and the consumers when these products are unreliable or unsafe. This is due largely in part to the role of the media in shaping consumer perceptions of driverless cars. News headlines may be biased a certain way, which can benefit or hinder public perception of these cars; therefore, they will likely play a large role in swaying consumer perceptions and behaviors.

Self-driving cars, or autonomous vehicles, may be the greatest disruptive innovation to travel that we have experienced in a century. A fully-automated, self-driving car is able to perceive its environment, determine the optimal route, and drive unaided by human intervention for the entire journey. Self-driving cars have the potential to drastically reduce accidents, travel time, and the environmental impact of road travel. However, obstacles remain for the full implementation of the technology including the need to reduce public fear, increase reliability, and create adequate regulations.

Of particular concern with regard to self-driving cars are Data privacy and Cyber Security risks. To date, five states and the District of Columbia have enacted laws that address autonomous vehicles or autonomous technology, but none of these state regulations address key areas of data privacy and security, such as the collection, use, choice, and security of consumer data gathered from these autonomous vehicles or autonomous technology. As vehicles become more computerized and begin to generate huge amounts of data, the potential for unwanted third-party access of that data and the risk of cyber threat increases. Hackers could potentially access the personal data of a driver, such as the vehicle's location, the identity of others in the car, and whether the driver is home at any particular time. Additionally, cyber -attacks could have potentially fatal consequences, not just for the driver and passengers inside the vehicle, but for anyone or anything physically surrounding the self-driving car. The model is based on cloud computing and for security with RSA algorithm and using ORAM for more privacy.

Oblivious Random-Access Memory (ORAM) is a security mechanism to hide data access patterns [4].

ORAM with RSA algorithm has given more successful output for multipurpose cloud architectures of online shopping and banking systems.

Cloud computing has become one of the most exciting platforms to use when it comes to the everyday tasks of an end user. The reasons for such are various due to its high scalability, enablement of ubiquitous capabilities, and the feature of having on-demand access to a shared pool of computing resources. It has become a vital part to usability as well as rapid access over interconnected networks in fields such as web development, big data analytics, and IoT.

Though cloud computing has made a substantial contribution to the technical field [5], there are risks and issues associated with it of which include, but not limited to, security, privacy, and secure management of sensitive information in transmission.

Autonomous vehicles are the future of transportation, offering a variety of social benefits such as improving mobility for the elderly, disabled and children, reducing accidents caused by driver errors, decreasing congestion linked to selfish driving behaviour, increasing fuel efficiency by reducing unnecessary braking, and enhancing human productivity by freeing us from driving [6]. To acquire real-time information about their environment, autonomous vehicles use an array of sensing technologies, e.g., sonar devices, cameras, lasers, and radars. However, because sensor data may be inaccurate or incomplete due to limited range, field of view, and obstructions. Vehicles can retrieve information that assist autonomous driving (e.g., map publishers) through vehicle-to-infrastructure (V2I).

2. Literature View

This paper describes an important Novel method to secure data of autonomous cars from cyber-attacks. Throughout the analysis data generated and stored by cameras, different censors, location tracking are secured using cryptography algorithm. It will reduce chances of cyber-attacks and will be able to prevent data more effectively.

Data privacy and security issues in autonomous cars at various levels:

2.1. Location Tracking

Location Data is fundamentally gathered and utilized in self-ruling vehicles for route purposes [7] – e.g., goal data, course data, speed, and time travelled. Area highlights are likewise utilized in existing customary vehicles to recall areas; give extra data applicable to the excursion, for example, constant movement information and purposes of-enthusiasm along the planed route; and to set steering inclinations, for example, staying away from interstates or toll streets.

An informational collection that corresponds area and travel information (e.g., current area, goal, speed, course, date and time) with extra data about the proprietor and traveller, could give different advantages. For instance, this sort of informational collection may help in rush hour gridlock arranging, decreasing clog, and enhancing wellbeing. In any case, this kind of consolidated informational index may likewise uncover delicate data about people, especially if this data is kept up after some time. These protection concerns exist both at the individual and societal level [8]. The Supreme Court has officially perceived in U.S. v. Jones that area data "creates an exact, far reaching record of a man's open developments that mirrors an abundance of insight about her familial, political, expert, religious, and sexual affiliations." The New York Court of Appeals noticed that a person's chronicled area and goal data would uncover "unquestionably private" trips, for example, to a specialist, plastic specialist, fetus removal facility, AIDS treatment focus, strip club, criminal resistance lawyer, by-the-hour motel, association meeting, place of love, and gay bar. Beside the disclosure of private data, data about one's present area or travel examples may make a danger of physical mischief or stalking if that data fell into the wrong hands.

2.2. Sensor Data

Autonomous vehicles (and existing human-driven vehicles) contain sensors that gather information about the vehicle's activity and its environment. For instance, sensors in Google's self-driving auto incorporate cameras, radar, warm imaging gadgets, and light location and going (LIDAR) gadgets [10] that gather information about the earth outside the vehicle. Among different utilizations [7], this information enables self-ruling vehicles to decide the articles it experiences, make forecasts about the earth, and make a move dependent on these expectations.

Organizations that consistently gather information about a vehicle's surroundings may wish to hold up under at the top of the priority list exercises from past requirement activities. For example, in 2012, the FCC authorized Google for social occasion Wi-Fi arrange data and the substance of information transmitted over Wi-Fi systems (counting touchy data, for example, email content, sites visited, and passwords) as its Google Street View vehicles accumulated pictures of open streets. Self-ruling vehicles are fit for gathering driving propensities, goals and other uncovering data about different drivers without their insight or assent. Extra concerns could emerge dependent on the utilization of symbolism caught by the vehicle, including proprietorship question and potential intrusion of protection claims, contingent on the conditions in which the pictures are caught [9]. Hence, organizations taking part in expansive information accumulation may wish to actualize shields to ensure singular protection.

2.3. Third Party Collection and Use of Data

In 2014, Jim Farley, the Global Vice-President of Marketing and Sales at Ford, told participants of the Consumer Electronics Show: "We know everybody who oversteps the law, we know when you're doing it. We have GPS in your auto, so we realize what you're doing. Incidentally, we don't supply that information to anybody." Although he later withdrawn the announcement, Farley's remarks feature the protection ramifications of information gathering and utilize. In reality, individual

data around a self-sufficient vehicle client's areas and on-street conduct might be significant to different government and private segment substances including law requirement, news media, private specialists, and insurance agencies. From a law authorization point of view, allowing access to sensor information presents a considerable lot of a similar security worries as gadgets like movement cameras, including the capacity to track a person's area or, as recommended by Farley's remarks, the capacity to recognize when people may have damaged activity or different laws. Also, sensor information may give critical data to insurance agencies. For instance, back up plans could screen driving propensities and modify premiums, like an intentional program as of now offered by one auto safety net provider. Back up plans may likewise be keen on utilizing both inward sensors (e.g., those that track speed and direction of the vehicle) and outside sensors (e.g., cameras or LIDAR) regarding mishap examinations. The issue of responsibility for information may affect whether the safety net provider has a privilege to secure or utilize this information. Self-ruling vehicles likewise present testing addresses identified with another protection thought—risk. Different analysts have recognized a large number of these lawful issues, including whether the driver, producer or programming would be held obligated in case of an impact.

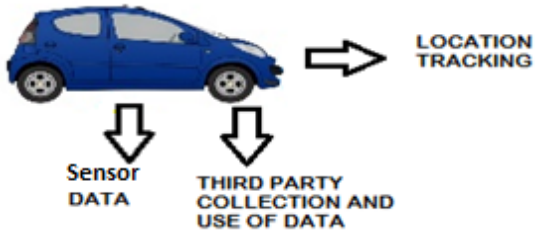

Figure 1. Security and privacy issues

3. Proposed Architecture

DES is a symmetric key algorithm which was developed by IBM in 1977 [11]. It uses blocks size 64 bits and key size 56 bits. It always operates on blocks of equal size. It always operates on blocks of equal size and uses both permutations and substitutions in the algorithm. It used 16 rounds of transposition and substitution to encrypt each group of 8 (64 bit) plaintext letters and produced output from each round is one by one. The number of rounds is exponentially proportional to the amount of time. **AES** is also a symmetric key algorithm based on the feistel structure. Security of Rijndael depends on its variable nature key size allowing up to a key size of 256-bit [12]. **ECC** stands for Elliptical curve cryptography. It is a public key encryption technique based on elliptic curve theory. It can be used to create faster, smaller, and more efficient cryptographic keys [12]. Elliptic curves are applicable for encryption, digital signatures, pseudo-random generators and other tasks. They are also used in several integer factorization algorithms such as Lenstra elliptic curve factorization. The attraction is the same level of security provided by keys of smaller size.

RSA stands for Ron Rivest, Adi Shamir and Leonard Adleman, who first publicly described the algorithm in 1977. It is the most commonly used public key encryption algorithm.

Key Generation

Select p, q	p, q both prime, p≠q
Calculate $n = p \times q$	
Calculate $\phi(n) = (p-1) \times (q-1)$	
Select integer e	$gcd(\phi(n), e) = 1; 1 < e < \phi(n)$
Calculate d	
Public key	$KU = \{e, n\}$
Private key	$KR = \{d, n\}$

Encryption

Plaintext:	$M < n$
Ciphertext:	$C = M^e \pmod n$

Decryption

Ciphertext:	C
Plaintext:	$M = C^d \pmod n$

Figure 2. RSA key generation

After surveying the previous four types of cryptographic algorithms, we decided to focus on one algorithm in particular: RSA. We began by analysing its algorithmic complexity, which we express using Big-O notation. Big O notation is used to classify algorithm runtimes by their dominant parts. Coefficients and polynomial terms of lesser degree are often dropped when using Big-O notation. For example, if an algorithm executes a loop of length n three times, instead of describing it as having a runtime of length 3n, we would say it runs in O (n) time, or linear time. If an algorithm can be described as having a runtime of n 3 + 2n2, we would say it runs in O (n3) or cubic time.

RSA key lengths are increased every few years to ensure that the improved factoring algorithms do not compromise the security of messages encrypted with RSA. Thus, it is important to investigate the performance of algorithms for generation of RSA key pairs larger than 1024 bits, which is one of the most used key length currently. This paper gives the performances of some discussed algorithms when generating 1024-bit primes, which correspond to 2048-bit RSA key pairs. All timing measurements are taken using the SLE66CX160S microcontroller manufactured by Infineon Technology at 3.57 MHz internal clock frequency. The SLE66CX160S microcontroller has a crypto coprocessor [13] which is used to illustrate how some of the special features of crypto coprocessors can improve the efficiency of the overall implementations.

Encryption is the process of converting original plain text (data) into cipher text (data). Steps:

- Cloud service provider should give or transmit the Public- Key (n, e) to the user who want to store the data with him or her.
- User data is now mapped to an integer by using an agreed upon reversible protocol, known as padding scheme.
- Data is encrypted and the resultant cipher text (data) C is C = me (mod n).
- This cipher text or encrypted data is now stored with the Cloud service provider.

Decryption: Decryption is the process of converting the cipher text (data) to the original plain text (data). Steps:

- The cloud user requests the Cloud service provider for the data.
- Cloud service provider verifies the authenticity of the user and gives the encrypted data.
- The Cloud user then decrypts the data by computing, m = Cd (mod n).
- Once m is obtained, the user can get back the original data by reversing the padding scheme.

RSA is an asymmetric encryption algorithm and also the first and most successful public key cryptosystem in theory. It is widely used in PKCS (Public Key Cryptography Standards) and Electronic Business. In this algorithm, there is always the computation of large numbers, so great time complexity is the greatest imperfection in either hardware or software implementation [16,17]. Therefore, an efficient implementation of RSA key generation is very important for theoretical study and practical applications. In the process of RSA key generation, the most time-consuming process is the generation of large prime number. So our main study is focused on the optimization. Generally, in order to get rid of partial odd composite numbers, pre-screening is adopted before the final Miller-Rabin algorithm [18]. Up to now, much research work has been taken to reduce key generation time.

4. Methodology

For security, focus is on data storage security. Distributed storage technique will be used as security as a service for securing cloud data. The data will be divided into several batches to achieve maximum accuracy, then encrypt the data by generating a polynomial function against each batch as discussed for RSA algorithm and store the data in separate databases. Time sensitivity will play a major role in ensuring the cloud performance. If the collision is detected, cloud server should not take more than a designated time frame to give a prompt response to the user. The performance of cloud will be increased by encrypting portion of files concurrently and uploading the encrypted data in parallel in the cloud storage. This will result in increasing the speed of message transfer. To make the cloud available, use reliable storage agreement since the hard-drive is currently the main storage media in the cloud environment, reliability of hard disks will be ensured. The privacy of the cloud data will be preserved using the oblivious RAM (ORAM) technology.

Oblivious RAM (ORAM) is a security-provable methodology for ensuring customers' entrance examples to remote distributed storage. As of late, various ORAM developments have been proposed to enhance the correspondence proficiency of the ORAM demonstrate, however little consideration has been paid to the capacity effectiveness. The best in class ORAM developments have the capacity overhead of O (N) or O (N log N) obstructs at the server, when N information squares are facilitated. As an emerging technology cloud computing provides easy access and high-performance computing on the data. Another major challenge that today software companies face, are storage of data at affordable cost and make available all the time.

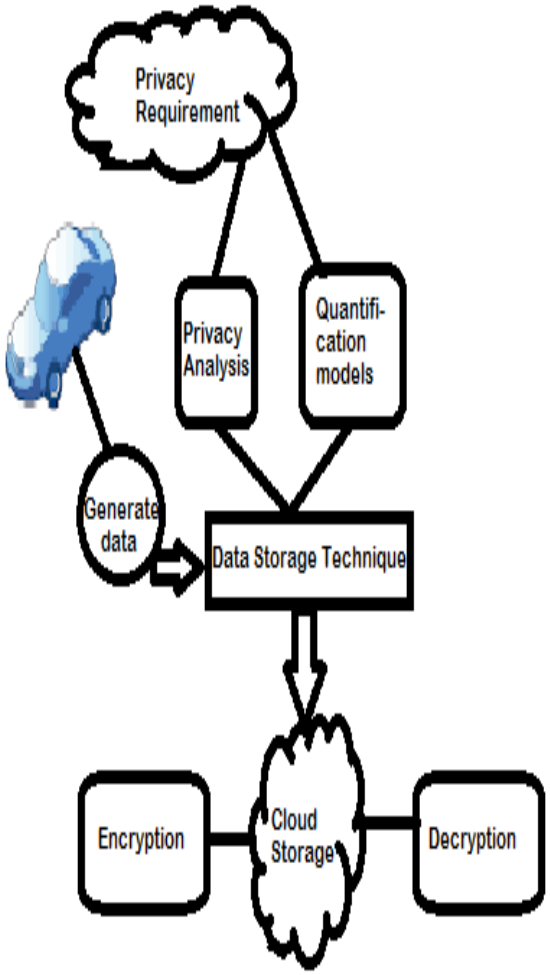

Figure 3. Basic framework

The factors that make to move cloud are [14]:

- Reduces the maintenance cost like no need of licensed software fee for each system, the purchase of new hardware and software is reduced
- Access to the application can be done anytime, anywhere provided that they should be connected to internet.
- Scalable
- Improves Flexibility
- Disaster Recovery
- As the services are based on" Pay per use", capital expenditure can be reduced
- User Friendly Environment
- Quick Deployment
- Less Energy Consumption

Table 1. Comparison of different Cryptography algorithm

Factors	DES	AES	RSA	ECC
Contributor	IBM-75	Rijman, Joan	Amazon	NealKoblitz victor S.miller
Key- Length	56 bits	128, 192 and 256	Based on number of bit in N=p*q	135 bits
Block Size	64 bits	128 bits	Variant	Variant
Security rate	Not enough	Good	Excellent	Good
Execution Time	Slow	Faster	Fastest	Fast

5. Results

Success depends on the outcome of the architecture of privacy system. Success for this system brings less risks with it and people will use it more with less fear. Data security and privacy will be increased. Due to availability the ratio of performance will increase. Less chances of accidents. This system will be more secured for any kind of cyber-attack as cyber threat is the major concern of this system. Companies of this autonomous vehicles will be less impacted due to data privacy and security. This data gathered by the autonomous vehicle will be helpful in emergency situations. Location tracking safety will increase. To check the success of this system by pointing the major acceptance part of hypothesis with more than 90% successfulness and by providing more secure system of autonomous vehicle. To check results based on accuracy of the system and to define and calculate the accuracy of hypothetical test the efficient way is to use the characteristic test using positive and negative results and their ratio.

- True Positive (TP):- The number of cases of data correctly stored and encrypted
- False Positive (FP):- The number of cases of data incorrectly stored and encrypted
- True Negative (TN):- The number of cases of data not affected using dummy cyber attack
- False Negative (FN):- The number of cases of data affected using dummy cyber attack

As per ISO 5725-1 [15], the general term "accuracy" is utilized to depict the closeness of an estimation to the genuine esteem. At the point when the term is connected to sets of estimations of a similar measure and, it includes a part of irregular blunder and a segment of precise mistake. For this situation genuineness is the closeness of the mean of an arrangement of estimation results to the real (genuine) esteem and accuracy is the closeness of agreement among an arrangement of results.

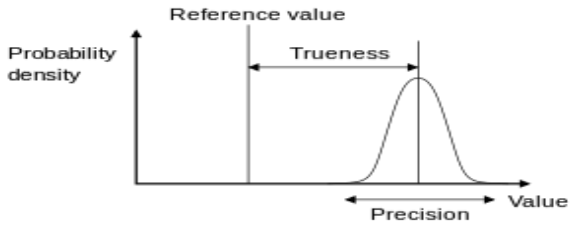

Figure 4. Accuracy factors

Accuracy: The accuracy of a test is its ability to differentiate stored and encrypted data correctly. To estimate the accuracy of a test, calculate the proportion of true positive and true negative in all evaluated cases. Mathematically, this can be stated as:

Accuracy: - TP+TN/TP+TN+FP+FN

6. Conclusion

For autonomous cars privacy and data security are major concerns. There are many states in United States of America who has autonomous cars, but they have same major issue, as autonomous cars use different sensors, cameras, information of users to collect data of surrounding, but this data is still not secured, so there are chances cyber threat or attack using this data gathered by autonomous cars. Data security and privacy after using data secured techniques with RSA algorithm to encrypt and decrypt the data and storing it in cloud it will be more secured system. As this method has shown good results for multi cloud system for online shopping and banking systems. Information security and protection will be expanded, Because of accessibility the proportion of execution will get increment and transport arrangement of self-governing vehicles. Less odds of mischances. This framework will be more secured for any sort of digital assault, digital risk is the real worry of this framework. Organizations of this independent vehicles will be less singed because of information protection and security. This information assembled by the self-sufficient vehicle will be useful in crisis circumstances.

References

[1] N. Davis. New laser technology lets driverless cars see round corners Guardian (2018). https://www.theguardian.com/technology/2018/mar/05/self-driving-cars-may-soon-be-able-to-see-around-corners.

[2] A. Halsey Driverless cars promise far greater mobility for the elderly and people with disabilities https://www.washingtonpost.com/local/trafficandcommuting/driverless-cars-promise-far-greater-mobility-for-the-elderly-and-people-with-disabilities/2017/11/23/6994469c-c4a3-11e7-84bc5e285c7f4512_story.html?utm_term=.b5549509a7a4.

[3] D. Wakabayashi. Self-driving Uber car kills Arizona pedestrian, where robots roam https://www.nytimes.com/2018/03/19/technology/uber-driverless-fatality.html.

[4] Bindschaedler, Muhammad Naveed, Xiaorui Pan, XiaoFeng Wang, and Yan Huang. Practicing Oblivious Access on Cloud Storage: the Gap, the Fallacy, and the New Way Forward Vincent. www.oblivious-storage.com.

[5] Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R., Konwinski, A., Stoica, I. (2010). A view of cloud computing. Communications of the ACM, 53(4), 50-58. Google Scholar, Crossref, ISI.

[6] Daniel J Fagnant and Kara Kockelman. 2014. Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations for capitalizing on self-driven vehicles. Transportation Research Board (2014).

[7] D. J. Glancy. Privacy in autonomous vehicles. Santa Clara L. Rev., 52:1171, 2012.

[8] The privacy implications of Autonomous vehicles By Norton Rose Fulbright on July 17, 2017. Posted in Compliance and risk management, Regulatory response.

[9] Self-Driving Cars and Data Collection: Privacy Perceptions of Networked Autonomous Vehicles Cara Bloom Joshua Tan Javed Ramjohn Lujo Bauer Carnegie Mellon University.

[10] J. Petit. Self-driving and connected cars: Fooling sensors and tracking drivers, 2015. https://www.blackhat.com/docs/eu-15/materials/eu-15-Petit-Self-Driving-And-Connected-CarsFooling-Sensors-And-Tracking-Drivers.pdf. Accessed March 2017.

[11] RituTripathi, Sanjay Agrawal. “Comparative Study of Symmetric and Asymmetric Cryptography Techniques” International Journal of Advance Foundation and Research in Computer (IJAFRC) Volume 1, Issue 6, June 2014. ISSN 2348-4853.

[12] Study on Symmetric and Asymmetric Cryptographic Techniques May Aye Chan Aung Myat Su Wai.

[13] Alfred J Menezes, Paul C Oorschot, Scott A Vanstone, Handbook of Applied Cryptography, 2005

[14] Lakshmi Neelima, M. Padma. A STUDY ON CLOUD STORAGE M. IJCSMC, Vol. 3, Issue. 5, May 2014, pg. 966-971.

[15] BS ISO 5725-1: "Accuracy (trueness and precision) of measurement methods and results - Part 1: General principles and definitions", p.1 (1994).

[16] Rivest, R.L. Response to NIST's proposal. Communications of ACM, 35, 1992, 41-47.

[17] Kaliski, B., and Robshaw, M. The Secure Use of RSA. CryptoBytes, RSA Laboratories, (Autumn 1995), 7-13.

[18] Koç, Ç. K. High-Speed RSA Implementation. Technical Report TR-201, version 2.0, RSA Laboratories, November 1994.