Reinforcement Learning-Based Chaotic Communication System for Secure Transmission of Encrypted State Information in the Smart Grid

Shahram Hosseinzadeh*[1] and Ruhollah Jafari Lagran[1]

***Abstract*—This paper proposes a system for transmitting and receiving encrypted state information via a communication channel in the presence of noise and interference. The proposed system uses a chaotic signal generated by a Lorentz oscillator to encrypt the state information, which is then transmitted through the communication channel. At the receiver end, the received signal is decrypted using a similar key generated by a forced Lorentz oscillator. The accurate determination of the force signal is essential for synchronizing chaotic signals, and this paper proposes the use of reinforcement deep learning agents to train and determine the force signal. The proposed communication scheme involves the use of a state estimator, a master chaotic oscillator, two slave oscillators, and two RL agents. The proposed system was simulated using MATLAB Simulink, and the results show that the errors exhibit a repetitive nature, with low and high values corresponding to the input signal. The proposed system provides a reliable and secure system for transmitting sensitive information over communication channels.**

***Index Terms*- Chaotic Communication System, Synchronize-tion, Reinforcement Learning.**

I. INTRODUCTION

A smart grid is an advanced electrical power system that uses modern communication and information technologies to improve the efficiency, reliability, and sustainability of electricity generation, distribution, and consumption. The traditional power grid was designed to deliver electricity from large centralized power plants to consumers through a one-way flow of electricity. However, with the increasing penetrationof renewable energy sources, electric vehicles, and energystorage systems, the power grid needs to be more flexible, resilient, and interactive to accommodate these changes. A smart grid integrates various technologies such as sensors, meters, automation, and control systems to monitor and manage the flow of electricity in real time. It also enables two-way communication between the utility and the customers, allowing them to actively participate in energy management and conservation. The smart grid can optimize the use of renewableenergy, reduce carbon emissions, enhance grid security, and improve the overall quality of electricity service. However, theimplementation of a smart grid requires significant investmentsin infrastructure, technology, and policy frameworks [1]–[5].

State estimation is a crucial process in a smart grid that estimates the state variables of the power system based on measurements obtained from various sensors and meters. In traditional grids, state estimation is performed using a centralized system, but in a smart grid, it is distributed, which reduces communication overhead and improves scalability and reliability. Advanced Metering Infrastructure (AMI) and Phasor Measurement Units (PMUs) are two novel structures used for state estimation in smart grids. AMI provides real-time data on energy consumption and production, enabling the utility to monitor energy usage patterns and optimize energy distribution. PMUs measure the voltage, current, and phase angle of the power system at high speeds, providing accurate and synchronized measurements for real-time monitoring and control. Smart grids require more advanced state estimation techniques and communication technologies to enable real-time monitoring and control of the power system. The useof AMI and PMUs in smart grids enables more accurate and reliable state estimation, improving the overall performance and efficiency of the power system [6]–[9].

Communication protocols are essential for the implementation of a smart grid as they enable the exchange of information between different components of the power system. The communication protocol used in a smart grid should be reliable, re-silent, and secure. Smart grid applications require high-speed data transmission to enable real-time monitoring and controlof the power system. The communication protocol should alsobe scalable, flexible, and cost-effective to accommodate the growing number of devices and sensors in the power system and support different types of data and applications [10].

The implementation of smart grid communication systems introduces several security issues that are not present in traditional networks. Smart grids are complex systems that rely on a large number of interconnected devices and systems, which makes them vulnerable to cyber-attacks. Unauthorized access, data integrity, denial of service attacks, malware and viruses, and physical security are some of the security issues that arise in the implementation of smart grid communication systems. Smart grid communication systems require robust security mechanisms to protect against cyber threats and ensure the confidentiality, integrity, and availability of data [11], [12].

Symmetric protocols, such as Advanced Encryption Standard (AES), are commonly used in smart grid communication networks for encrypting data that is transmitted between devices. Asymmetric protocols, such as RSA, are also usedin smart grid communication networks for tasks such as key exchange and digital signatures.

1. Azarbaijan Shahid Madani University, Tabriz, Iran.

Corresponding author Email: s.hosseinzadeh@azaruniv.ac.ir

Asymmetric protocols aregenerally considered more secure than symmetric protocols, but they are also more computationally intensive and slower.

Therefore, a combination of symmetric and asymmetric pro-tools is often used in smart grid communication networks toprovide a balance between security and efficiency [13], [14].The use of chaotic signals in smart grid communication networks can enhance security by generating unpredictable and difficult-to-reproduce cryptographic keys. Chaotic signalscan protect against various types of attacks, ensuring the confidentiality, integrity, and availability of data transmittedover the smart grid communication network. The use ofchaotic signals can also improve the efficiency and reliabilityof the network by optimizing the allocation of resources andreducing interference and noise. This approach can provide a range of benefits for both security and performance, makingit a promising approach for the reliable and secure operation of the smart grid [15].

Chaotic signals are complex, irregular, and unpredictable signals. Synchronization of chaotic signals involves making two or more chaotic signals behave identically, whichis necessary for applications such as secure communication, chaos-based cryptography, and control of chaotic systems. The possibility of synchronization depends on the type of chaotic system and the synchronization method used. The condition for synchronization is that the difference between the chaotic signals' states should converge to zero asymptotically. The method for synchronization involves designing a feedback control system that adjusts the parameters of the chaotic system to match the parameters of the reference system. Various methods, such as the Pyragas method [16], the Ott- Grebogi-Yorke method [17], or the adaptive control method, can be used to design the feedback control system [18].

The paper is divided into four sections: introduction, problem formulation, results and discussion, and conclusion. Inthe introduction, a brief literature survey on strategies for secure communication in smart grids is presented. The problem formulation section outlines the proposed communication scheme. The results and discussion section presents the search findings and analyzes their implications. Finally, the conclusion of the paper is provided in section IV.

II. PROBLEM FORMULATION

In modern communication systems, the transmission and reception of encrypted state information via communication channels is a common practice. The encrypted state information is then properly decrypted to further process the data, such as bad data detection, operational decision making, and other related decisions. However, the transmission and receptionof encrypted state information can be challenging due to the presence of noise and interference in the communication channel.

To address this challenge, a system is proposed that can effectively transmit and receive encrypted state information via a communication channel. Fig. 1 depicts an overviewof the proposed system. The system consists of an encryption module, a transmission module, a reception module, and a decryption module. The encryption module encrypts thestate information using a secure encryption algorithm. The transmission module transmits the encrypted state information via a communication channel. The reception module receives the transmitted encrypted state information and performs error correction to remove any noise or interference in the communication channel. The decryption module then decrypts the received encrypted state information to obtain the original stateinformation.

Once the original state information is obtained, it can be further processed for various purposes, such as bad data detection, operational decision-making, and other related decisions. The proposed system can effectively transmit and receive encrypted state information, even in the presence of noise and interference in the communication channel. This makes it a reliable and secure system for transmitting sensitive information over communication channels. States from the smart grid are digitized and encrypted by a time-varying key. Keys are generated via free running Lorentz chaotic y parameter. The Lorentz oscillator is described by the following set of differential equations [19]:

$$\frac{dx}{dt} = \sigma(y - x) \tag{1}$$

$$\frac{dy}{dt} = x(\rho - z) - y \tag{2}$$

$$\frac{dz}{dt} = xy - \beta y \tag{3}$$

After 8-bit digitization of the oscillator signal, it is transmitted through an ideal communication channel. The received signal is decrypted using a similar key, which is generated by the forced Lorentz oscillator with the same parameters, i.e.σ, ρ and β, as those at the transmitter. The forced oscillator obeys the same coupled differential equations, only that $y$is forced by an external time-varying parameter, i.e. u [20]

$$\frac{dy}{dt} = x(\rho - z) - y + u \tag{4}$$

To properly retrieve the signal y same as the free-running oscillator in the transmitter, and consequently generate the key from the received signal. This process allows for the original signal to be achieved for further processing. The accurate determination ofthe force signal is essential for synchronizing chaotic signals. This paper proposes the use of reinforcement deep learning agents to train and determine the force signal. Reinforcement learning (RL) is a type of machine learning where an agent learns to make decisions by interacting with an environment. The agent receives feedback in the form of rewards or penaltiesfor its actions, and its goal is to maximize the cumulative reward over time.

In RL, the agent learns a policy, which is a mapping from states to actions. The policy determines what action the agent takes in each state. The goal of the agent is to learn a policy that maximizes the expected cumulative reward.

The value of a state is the expected cumulative reward starting from that state and following the current policy. The policy iteration algorithm alternates between policy evaluation and policy improvement.

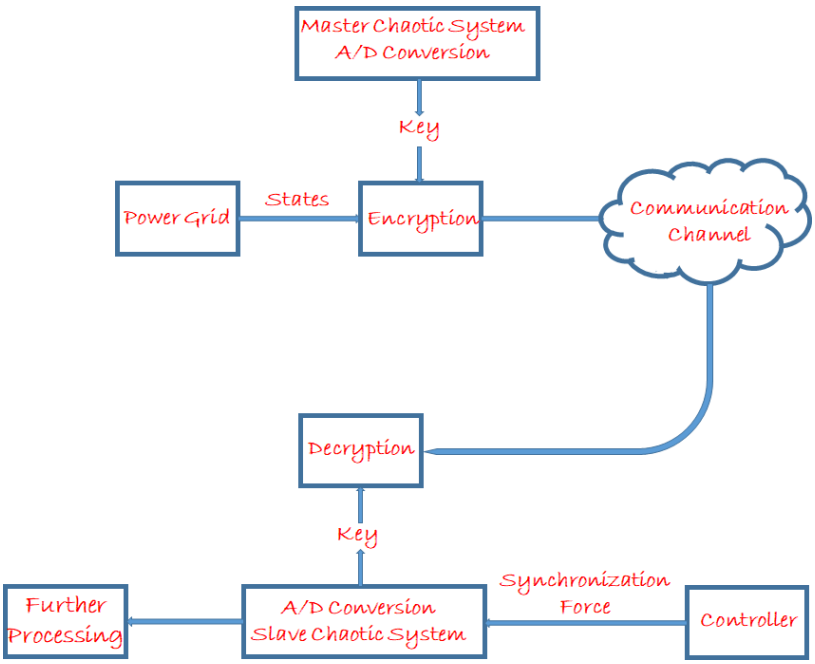

Fig. 1. Overview of the proposed system for transmitting and receiving encrypted state information via a communication channel.

In policy evaluation, the value function is computed for the current policy. In policy improvement, the policy is updated to be greedy concerning the value function. The Bellman equation is a recursive equation that expresses the value of a state in terms of the values of its successor states:

$$V(s) = \max_{a}\left[R(s,a) + \gamma \sum_{s'} P(s'|s,a)V(s')\right] \quad (6)$$

where $V(s)$ is the value of state s, $R(s, a)$ is the reward for taking action a in state s, $P(s'|s, a)$ is the probability of transitioning to state s' from state s after taking action a, and

γ is a discount factor that determines the importance of future rewards [21].

Q-learning is a popular RL algorithm that learns the optimal action-value function, $Q(s, a)$, which is the expected cumulative reward starting from state s, taking action a, and following the optimal policy thereafter. The Q-learning update rule is:

$$Q(s,a) \leftarrow Q(s,a) + \alpha[r + \gamma \max Q(s',a') - Q(s,a)] \quad (7)$$

where r is the reward for taking action a in state s, α is the learning rate, and s' is the next state [22]–[24].

In this paper, the states are determined using the error, error derivative, integral of error, and received signal y. The reward is determined based on the absolute values of the error and force.

The actor-critic model is a deep reinforcement learning approach that combines value-based and policy-based methods. The actor learns a policy that maps states to actions, while the critic learns the value function to evaluate the policy and provide feedback to the actor.

Q-learning, a value-based RL algorithm, is used in this paper to train the critic. A neural network is utilized to implement both the critic and actor, with the critic estimating Q-values using the state as input and the actor outputting action probabilities based on the state. The actor and critic are trained simultaneously using the Q-learning algorithm.

Our proposed communication scheme is designed for base- band 2-level communication, where the receiver must be able to distinguish between the levels. To achieve this, we have composed a system consisting of a state estimator, a master chaotic oscillator, 2 slave oscillators, and 2 RL agents.

To digitize the critical state information and master informa- tion, we use an 8-bit A/D block. The state information is then encrypted using the XOR logic gate from the master oscillator and transmitted through the communication channel.

At the receiver end, the received signal is applied twice to two XOR gates with the 0 and 1 logics as the other gate. The output from both gates is differentiated from two master oscillators. The state equation from both oscillators and differentiators are fed to RL agents to derive the proper forces. Our anticipated outcome is that the oscillator, by the proper input of the logic, is well-synchronized, and by measuring the error, the proper signal can be retrieved. Fig. 2 illustrates the communication scheme we propose.

While our proposed communication scheme involves the use of two RL agents in the receiver end, it is not necessary to train both agents separately. Instead, we can train one agent and then copy the results to the other

agent. This approach saves timeand effort, as training RL agents can be a cumbersome task.

By copying the results from one agent to the other, we ensure that both agents have the same knowledge and expe- rience. This allows them to work together more efficientlyand effectively. Additionally, it reduces the risk of errors or inconsistencies that may arise from training both agents separately.

Using a single trained RL agent and copying the results to the other agent is a practical and efficient approach that can simplify the implementation of our proposed communication scheme.

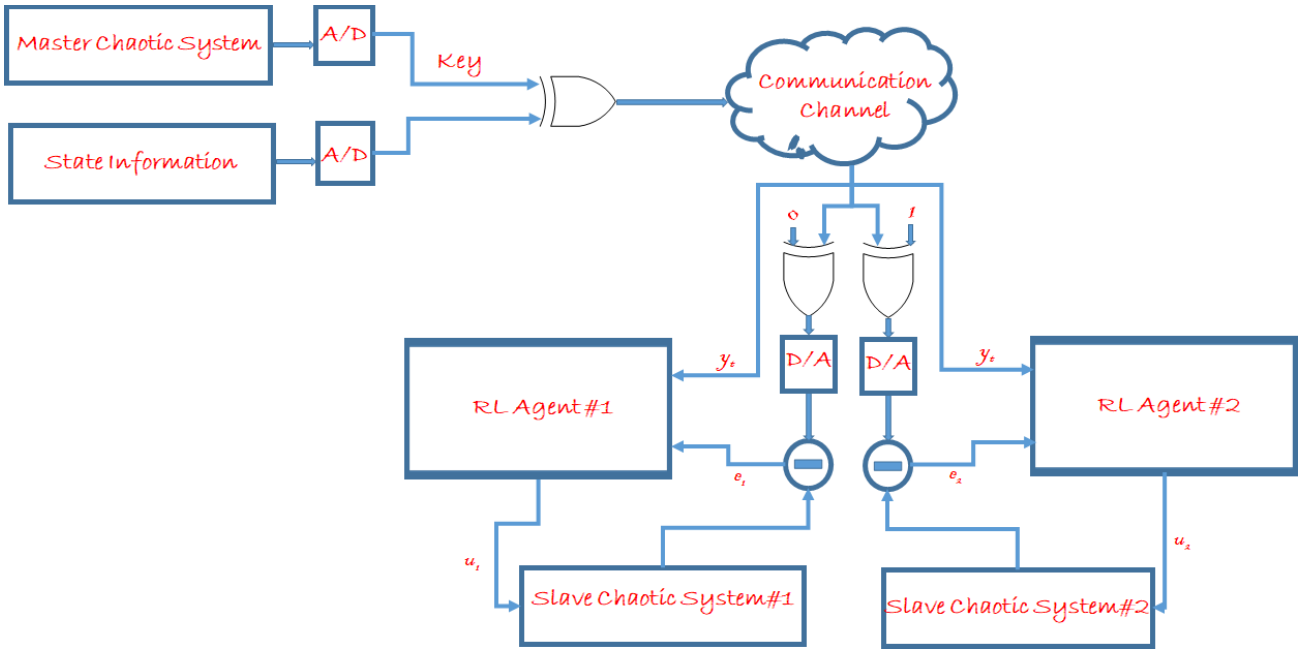

Fig. 2. Communication scheme composed of state estimator, master chaotic oscillator, 2 slave oscillators, and 2 RL agents.

III. RESULTS AND DISCUSSION

The proposed communication scheme was simulated usingMATLAB Simulink, with the reinforcement learning toolboxof MATLAB used to implement and train the RL agent. Tosimplify the analysis, a normalized time scale was considered. Fig. 3 shows the error from both oscillators compared to the input signal. A repetitive 0 and 1 signal with positive logic wasused as the input. It can be observed that the errors exhibit arepetitive nature, with low and high values corresponding to the input signal.

To retrieve the received signal appropriately, the errors wererectified and filtered. Fig. 4 shows the results of this process. The retrieved signals from both oscillators were then comparedto derive the final output, as shown in the Fig.5.

The simulation results demonstrate the effectiveness of the proposed communication scheme. The use of RL agents in the receiver end allows for efficient and accurate signal retrieval, even in the presence of noise and other disturbances. The rectification and filtering of errors further improves the accuracy of the received signal.

It seems that the proposed communication scheme shows promise for practical implementation in real-world scenarios. Further research can explore the use of different input signals and noise levels to evaluate the robustness of the scheme.

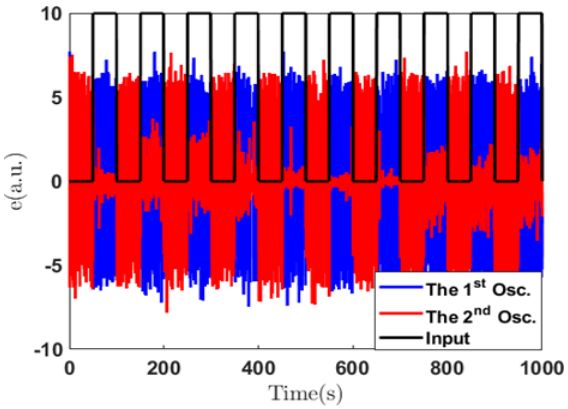

Fig. 3. Error from both oscillators compared to the input signal.

IV. CONCLUSION

In this paper, we proposed a reinforcement learning-based chaotic communication system for the secure transmission of encrypted state information in the smart grid. The proposed system uses a chaotic signal generated by a Lorentz oscillator to encrypt the state information, which is then transmitted through the communication channel. At the receiver end, the received signal is decrypted using a similar key generated bya forced Lorentz oscillator. The accurate determination of the force signal is essential for synchronizing chaotic signals, and this paper proposes the use of reinforcement deep learning agents to train and determine the force signal.

The proposed communication scheme involves the use of a state estimator, a master chaotic oscillator, two slave oscillators, and two RL agents. The proposed system was Simulated using MATLAB Simulink, and the results show that the errors exhibit a repetitive nature, with low and high values corresponding to the input signal. The proposed system provides a reliable and secure system for transmitting sensitive information over communication channels.

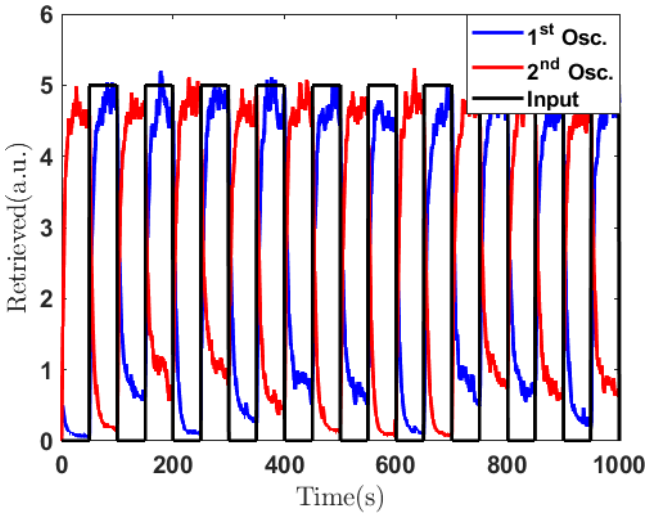

Fig. 4. Retrieved signals after rectification and filtering.

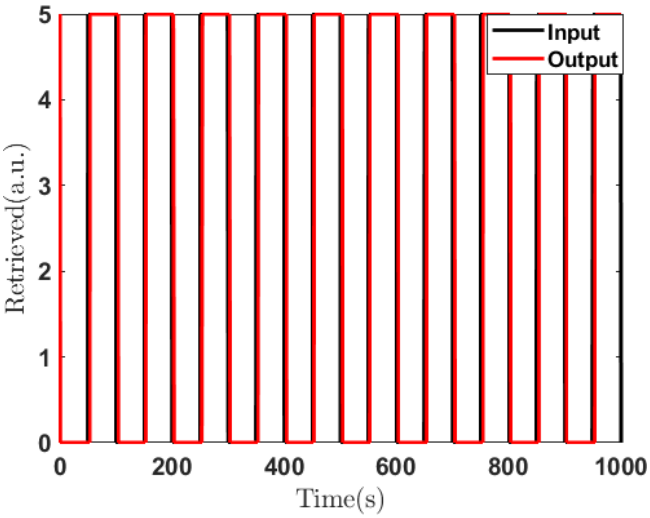

Fig. 5. Comparison of retrieved signals from both oscillatorsto derive final output.

REFERENCES

[1] M.K. Hasan, A.A. Habib, Z. Shukur, F. Ibrahim, S. Islam, M.A. Razzaque, Review on the cyber-physical and cyber-security system in smart grid: Standards, protocols, constraints, and recommendations, *Journal of Network and Computer Applications*, vol. 209, p. 103540, Jan. 2023.

[2] M.K. Hasan, A.A. Habib, S. Islam, M. Balfaqih, K.M. Alfawaz, andD. Singh, Smart Grid Communication Networks for Electric Vehicles Empowering Distributed Energy Generation: Constraints, Challenges, andRecommendations, *Energies*, vol. 16, no. 3, p. 1140, Jan. 2023.

[3] X. Xiang and J. Cao, An efficient authenticated key agreement scheme supporting privacy-preservation for smart grid communication panel, *Electric Power Systems Research*, vol. 203, p. 107630, Feb. 2022.

[4] P. Haji Mirzaee, M. Shojafar, H. Cruickshank, and R. Tafazolli, "Smart Grid Security and Privacy: From Conventional to Machine LearningIssues (Threats and Countermeasures)," *IEEE Access*, vol. 10, pp. 52922-52954, May 2022.

[5] Y. Kim, S. Hakak, and A. Ghorbani, "Smart grid security: Attacks and defense techniques," *IET Smart Grid*, vol. 6, no. 2, pp. 103-123, Apr. 2023.

[6] M. Zhang, C. Shen, N. He, S. Han, Q. Li, Q. Wang, and X. Guan, "False data injection attacks against smart grid state estimation: Construction, detection and defense," *Science China Technological Sciences*, vol. 62, no. 11, pp. 2085-2102, Sep. 2019.

[7] T. Nguyen, S. Wang, M. Alhazmi, M. Nazemi, A. Estebsari, and P. Dehghanian, "Electric Power Grid Resilience to Cyber Adversaries: Stateof the Art," *IEEE Access*, vol. 8, pp. 86905-86922, May 2020, doi:10.1109/ACCESS.2020.2993233.

[8] Linge Sagar Gajanan, Mukesh Kirar, and More Raju, "Wide area Monitoring and Protection in Cyber-Physical Infrastructure," published in*2023 IEEE Renewable Energy and Sustainable E- E-Mobility Conference (RESEM)*, Bhopal, India, 17-18 May 2023, pp. 1-6.

[9] Gang Cheng, Yuzhang Lin, Ali Abur, Antonio Gómez-Expósito, and Wenchuan Wu, "A Survey of Power System State Estimation UsingMultiple Data Sources: PMUs, SCADA, AMI, and Beyond," *IEEE Transactions on Smart Grid*, vol. 14, no. 3, pp. 1738-1753, May 2023.

[10] L. Tightiz and H. Yang, "A Comprehensive Review on IoT Protocols' Features in Smart Grid Communication," *Energies*, vol. 13, no. 11, p. 2762, Jun. 2020, doi: 10.3390/en13112762.

[11] D. Yan, F. Liu, Y. Zhang, and K. Jia, "Dynamical model for individual defense against cyber epidemic attacks," *IET Information Security*, vol. 13, no. 3, pp. 541-551, May 2019.

[12] B.M.R. Amin, S. Taghizadeh, M.S. Rahman, M.J. Hossain, V. Varad- harajan, and Z. Chen, "Cyber attacks in smart grid – dynamic impacts, analyses, and recommendations," *IET Cyber-Physical Systems: Theory & Applications*, vol. 5, no. 4, pp. 321-329, Jul. 2020.

[13] D. He, H. Wang, M. K. Khan, and L. Wang, "Lightweight anonymous key distribution scheme for smart grid using elliptic curve cryptography,"*IET Communications*, vol. 10, no. 14, pp. 1795-1802, Sep. 2016.

[14] M. R. Asghar, G. Dán, D. Miorandi, and I. Chlamtac, "Smart Meter Data Privacy: A Survey," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2820-2835, Fourth quarter 2017.

[15] K. Sun, *Chaotic Secure Communication: Principles and Technologies*. Berlin, Germany: De Gruyter, 2019.

[16] K. Pyragas, "Delayed feedback control of chaos," *Phil. Trans. R. Soc.A*, vol. 364, pp. 2309-2334, 2006. doi:10.1098/rsta.2006.1827.

[17] J. C. Claussen, "Floquet stability analysis of Ott-Grebogi-Yorke and difference control," *New Journal of Physics*, vol. 10, no. 6, pp. 063006 (13pp), June 2008.

[18] S. Chen and J. Lü, "Synchronization of an uncertain unified chaotic system via adaptive control," *Chaos, Solitons & Fractals*, vol. 14, no. 4, pp. 643-647, Sep. 2002. doi:10.1016/S0960-0779(01)00255-1.

[19] Bo-Wen Shen, "A Review of Lorenz's Models from 1960 to 2008," *International Journal of Bifurcation and Chaos*, vol. 33, no. 10, article no. 2330024, October 2023. doi: 10.1142/S021812742330024X.

[20] Louis M. Pecora and Thomas L. Carroll, ”Synchronization in chaotic systems,” *Physical Review Letters*, vol. 64, no. 8, pp. 821-824, February 1990. doi: 10.1103/PhysRevLett.64.821.

[21] Yutaka Matsuo, Yann LeCun, Maneesh Sahani, Doina Precup, David Silver, Masashi Sugiyama, Eiji Uchibe, and Jun Morimoto, ”Deep learning, reinforcement learning, and world models,” *Neural Networks*, vol. 152, pp. 267-275, August 2022, Special Issue on AI and Brain Science: Perspective.

[22] Jesse Clifton and Eric Laber, ”Q-Learning: Theory and Applications,”

[23] *Annual Review of Statistics and Its Application*, vol. 7, pp. 279-301, 2020.

[24] B. Ramadevi and K. Bingi, ”Chaotic Time Series Forecasting Approach Using Machine Learning Techniques: A Review,” *Symmetry*, vol. 14, no. 5, p. 955, May 2022. doi: 10.3390/sym14050955.

[25] C.T. Lin and C.P. Jou, ”Controlling chaos by GA-based reinforcement learning neural network,” *IEEE Trans. Neural Netw.*, vol. 10, pp. 846-859,1999.