

Nature or nurture: Revealing intrinsic privacy in the electricity consumption data

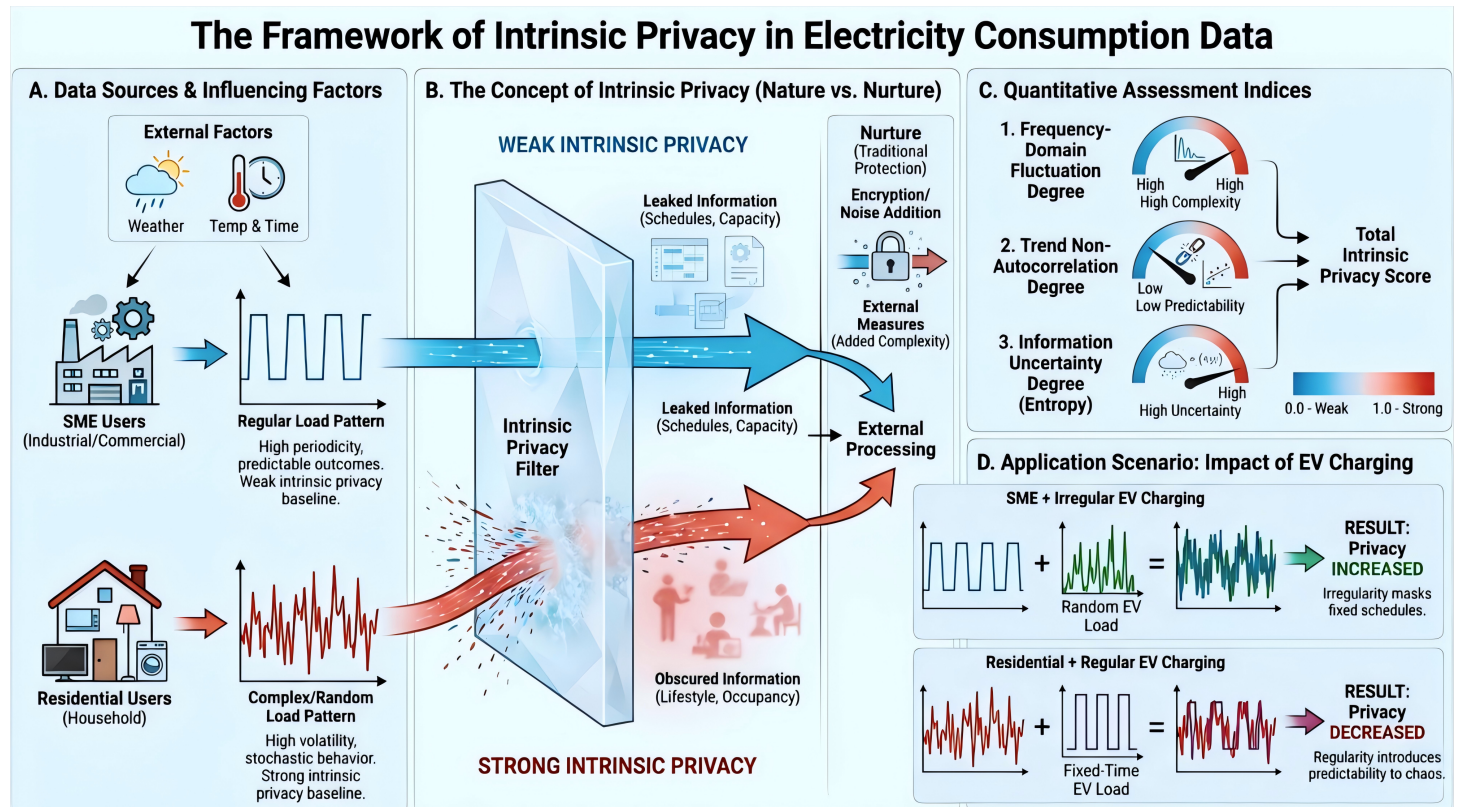
Zhiyi Li,^{1,*} Zehao Song,² Xutao Han,¹ Yue Xu,¹ Renjie Luo,¹ and Zhen Xu¹

*Correspondence: zhiyi@zju.edu.cn

Received: ; Accepted: January 20, 2026; Published Online: January 22, 2026; <https://doi.org/10.59717/j.xinn-energy.2026.100142>

© 2026 The Author(s). This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

GRAPHICAL ABSTRACT



PUBLIC SUMMARY

- Intrinsic privacy is defined as the innate capacity of raw electricity consumption data to conceal user behavior.
- Intrinsic privacy is quantified using indices for frequency fluctuation, non-autocorrelation, and information uncertainty.
- Residential load data shows higher intrinsic privacy than Small and Medium Enterprise (SME) data due to greater volatility.
- Intrinsic privacy helps users design personalized protection strategies, such as optimized smart charging for EVs.

Nature or nurture: Revealing intrinsic privacy in the electricity consumption data

Zhiyi Li,^{1,*} Zehao Song,² Xutao Han,¹ Yue Xu,¹ Renjie Luo,¹ and Zhen Xu¹

¹College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China

²State Grid Hangzhou Power Supply Company of Zhejiang Electric Power Co., Ltd., Hangzhou, Zhejiang 310000, China

*Correspondence: zhiyi@zju.edu.cn

Received: ; Accepted: January 20, 2026; Published Online: January 22, 2026; <https://doi.org/10.59717/j.xinn-energy.2026.100142>

© 2026 The Author(s). This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

Citation: Li Z., Song Z., Han X., et al. (2026). Nature or nurture: Revealing intrinsic privacy in the electricity consumption data. *The Innovation Energy* 3:100142.

Electricity consumption data collected by smart meters, while valuable for grid management, raises privacy concerns regarding user behavior. A concept, intrinsic privacy, is introduced in this study to identify the privacy characteristics that are tightly coupled with the user's electricity usage patterns. Intrinsic privacy represents the innate capacity of electricity consumption data to obscure or avoid disclosing information that the user does not wish to make public. A methodology including three key indices is developed to assess and quantify intrinsic privacy and is validated with data from two user groups: small and medium-sized enterprise and residential users. Examining electric vehicle charging as a representative scenario, we analyze how charging behaviors influence intrinsic privacy. The results show that intrinsic privacy provides electricity users with deeper insights into their privacy strengths and needs, facilitating the development of precise and personalized privacy protection strategies, such as optimized smart charging for electric vehicles.

INTRODUCTION

The rapid development of smart grids has led to an increase in the variety and complexity of user electricity consumption patterns.^{1–3} There are significant differences in electricity consumption types across users with different social attributes. Small and medium enterprise (SME) users rely primarily on fixed, large-scale production equipment, resulting in relatively stable electricity usage patterns.^{4,5} In contrast, residential users primarily utilize household appliances, such as air conditioners, washing machines, and microwave ovens, which have more fluctuating power usage patterns.^{6,7} Despite these differences, all these load types provide highly valuable user electricity information, providing operational insights that support power companies in optimizing grid scheduling and management, thereby increasing power system efficiency and reliability.^{8–11} For example, by analyzing electricity consumption data, retailers utilize this data to forecast demand and optimize energy purchasing strategies, while Distribution System Operators (DSOs) rely on it for grid stability analysis and congestion management. Electricity consumption data also facilitates the development of personalized service plans that are tailored to meet diverse user needs and increase overall user satisfaction.

While electricity consumption data is invaluable for optimizing grid management and increasing energy efficiency, it also involves sensitive private information. Detailed records of appliance types and usage times can indicate daily routines and business activities, creating unique privacy concerns for different user groups.^{12–16} Notably, different user groups have unique privacy concerns and requirements. Small and medium enterprises (SMEs) are particularly wary that high-resolution data might expose critical business secrets and competitive advantages. For instance, the start and stop times of heavy machinery in a factory can reveal shift patterns and production volumes, while the duty cycles of ovens in a bakery could disclose proprietary recipe processes or business intensity. These patterns allow external parties to infer production capacity and operational efficiency. Residential users, in contrast, primarily worry that electricity consumption data may reveal personal habits and activity patterns. Specifically, prolonged periods of low energy usage can signal that a home is vacant and vulnerable to burglary, while distinct load signatures from high-end electronics or medical devices may expose sensitive lifestyle details and health status.^{17–20} If attackers gain access to this data, they can analyze it to invade users' privacy. This vulnerability was starkly demonstrated during a data breach in Germany, where attackers accessed and analyzed the electricity consumption data of

an industrial user, obtaining production schedules and equipment operations. This breach facilitated industrial espionage and the theft of trade secrets and technical information.²¹ Similarly, in a California criminal case, hackers acquired detailed electricity usage data from thousands of households. Through data analysis, they inferred residents' schedules and identified valuable appliances, leading to multiple targeted burglaries and fraud.²² To mitigate these risks, defining and safeguarding the privacy of such data is essential.²³

Intrinsic privacy, a novel concept proposed in this study, refers to the capacity of long-term electricity consumption data to obscure or prevent the disclosure of information or matters that the electricity user does not wish to make public; this is achieved through the inherent characteristics of the data, such as its shapes and states, which are closely coupled with the user's electricity usage behavior. It should be noted that, shapes, which refer to the morphological features of the load curve such as peaks, valleys, and ramp rates; and states, which represent the operational modes (e.g., on/off/standby) or distinct statistical regimes of power consumption. Unlike traditional privacy protection, which is achieved through external protective measures,²⁴ intrinsic privacy focuses on the inherent and objective strength of privacy protection derived from the natural shapes, states, and other characteristics of the electricity consumption data itself. This suggests that, even in the absence of external privacy protection, the inherent characteristics of electricity consumption data may provide some natural safeguards against privacy breaches. Therefore, in this paper, intrinsic privacy is defined here not merely as a statistical complexity measure, but as a semantic framework that quantifies the inherent capability of the raw load data to conceal user behavior without external encryption. Although we used some signal analysis techniques in the subsequent quantization, such as entropy-based metrics, to compute this value, the concept specifically links signal complexity to the difficulty of inferring private behavioral attributes from the energy trace.

In addition, intrinsic privacy also differs from the conventional concept of data randomness, as it emphasizes the inherent structural privacy characteristics of electricity consumption data rather than the information-hiding effects of randomness or fluctuation alone. Randomness typically refers to the unpredictability in the time or amplitude of the data, which mainly indicates the difficulty faced by attackers attempting to infer user behavior from data volatility. In contrast, intrinsic privacy focuses on the overall shapes and states of the consumption data, which are features that are determined by user electricity usage patterns and independent of random variations. In other words, intrinsic privacy arises from the inherent coupling between the data and the user's lifestyle or production patterns and represents an objective, innate capacity for privacy protection. Understanding intrinsic privacy facilitates a deeper comprehension of the unique, inherent privacy characteristics of electricity consumption data.^{25,26} This awareness can not only help users better understand their own privacy needs but also provide a benchmark for evaluating the effectiveness of external privacy protection methods. (Note that, unless otherwise specified, the concept of intrinsic privacy mentioned in this paper refers to the intrinsic privacy of users' electricity consumption data.)

The privacy characteristics of a user's electricity consumption data are shaped by the user's electricity consumption behavior and are closely tied to core elements that define the data's overall form, thereby influencing its intrinsic privacy.²⁷ Environmental conditions act as external factors influencing user electricity consumption behavior. For example, winter heating demands and summer air conditioning needs generate peaks in electricity

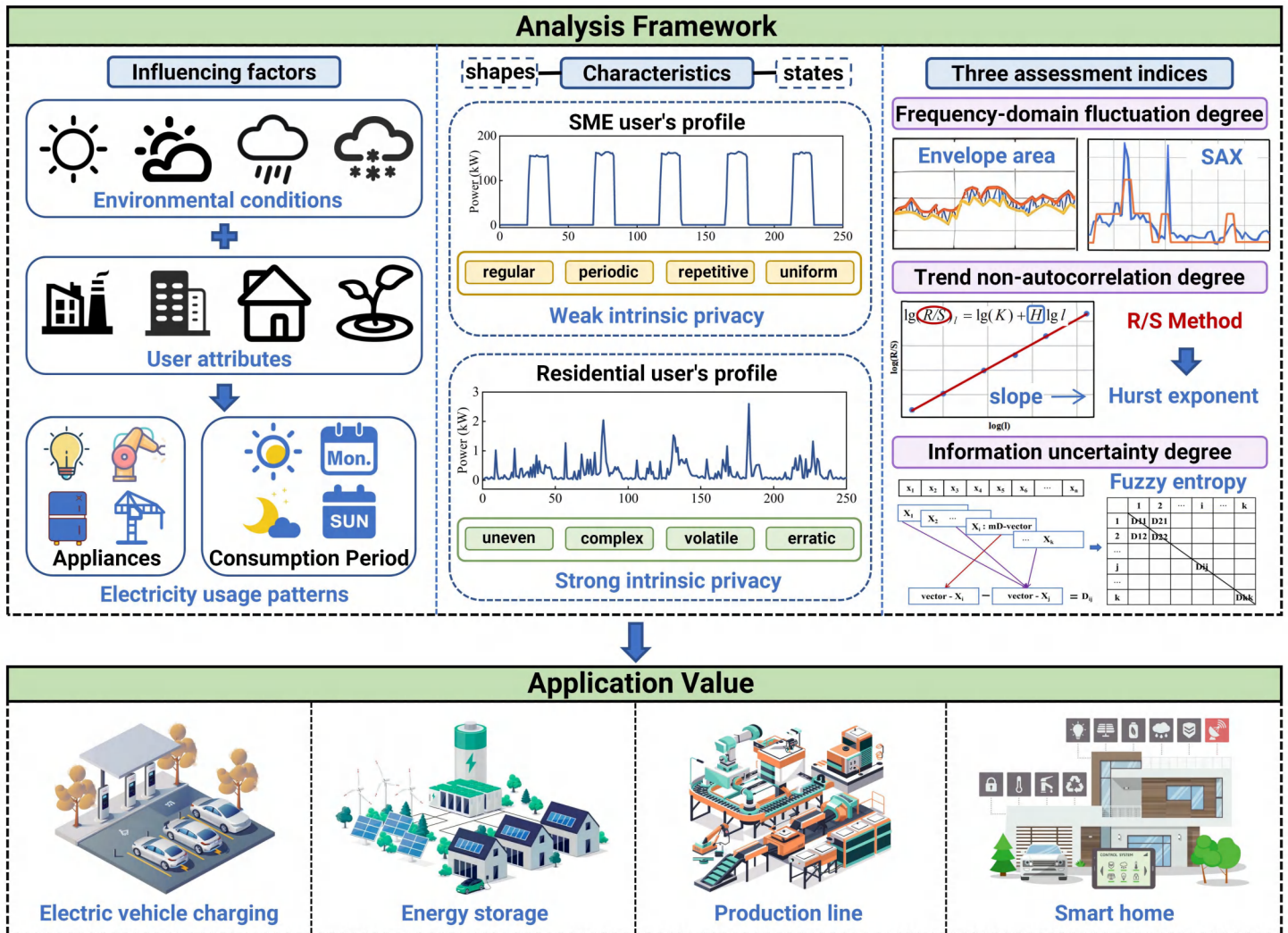


Figure 1. Analysis framework and application value of intrinsic privacy We analyze the factors influencing intrinsic privacy, its key characteristics, and the indices used for assessment. Environmental conditions, such as weather, temperature, and humidity, along with users' social attributes within industrial, commercial, agricultural, and urban sectors, jointly shape users' electricity usage patterns. These factors contribute to the varying shapes and state characteristics of power consumption data. SME users' power consumption data tends to exhibit weaker intrinsic privacy due to their regularity and periodicity. In contrast, residential users' data has stronger intrinsic privacy, as they are characterized by greater fluctuations, complexity, and irregularity. Three key indices, frequency-domain fluctuation degree, trend non-autocorrelation degree, and information uncertainty degree, are used to assess and quantify intrinsic privacy. Intrinsic privacy is applicable to privacy assessments across a variety of power consumption scenarios, such as electric vehicle charging, energy storage, production line operations, and smart home environments.

consumption. Extreme weather conditions, such as rain and snowstorms, can affect appliance operation frequency, leading to increased heat usage or decreased outdoor appliance use.²⁸ Moreover, the electricity consumption of specific equipment can vary depending on the temperature and humidity. User attributes serve as internal factors that also affect electricity usage behavior. Industrial users generally operate high-power equipment during the day, resulting in daytime electricity peaks, whereas residential users typically use more appliances in the mornings and evenings. Overall, environmental conditions and user attributes jointly influence electricity usage habits, leading to differences in appliance use and usage times. These variations contribute to distinct electricity consumption data shapes and states, ultimately resulting in different levels of intrinsic privacy.

As technology advances rapidly and society continues to develop, user electricity consumption patterns are becoming increasingly intricate, and diverse electricity consumption scenarios have emerged, such as smart home devices and distributed energy systems. These evolving patterns and scenarios not only alter the temporal distribution of user electricity consumption but also introduce greater diversity in the range of electrical devices used, serving as internal factors that influence changes in intrinsic privacy. Since intrinsic privacy is defined on the basis of the characteristics of the electricity consumption data itself, it is applicable to privacy assessments across various consumption scenarios. Here, we consider EV charging behavior as a

representative type of electricity consumption to investigate its specific impact on intrinsic privacy across user groups with differing social attributes. The diversity of EV charging behavior, whether at home, at the workplace, or at public charging stations, shapes users' daily routines and travel patterns, subsequently impacting intrinsic privacy. For example, regular home charging at fixed times can reveal users' work and rest schedules, whereas the frequent use of public charging facilities can reveal travel habits and commonly visited locations.^{34,35} Therefore, studying the impact of EV charging behavior on intrinsic privacy and developing effective measures to protect user privacy are of significant practical relevance.

In this work, as shown in Figure 1, we introduce the novel concept of intrinsic privacy and analyze it from three perspectives: influencing factors, structural characteristics, and quantitative assessment indices. External factors, such as environmental conditions, and internal factors, such as user attributes, jointly determine electricity usage behaviors, including appliance types and usage time distributions, which in turn affect intrinsic privacy. SME users have weaker intrinsic privacy, as their electricity consumption data is more regular and cyclical; in contrast, residential users have stronger intrinsic privacy because of their more complex and volatile consumption patterns. Three key indices, the frequency-domain fluctuation degree, trend non-autocorrelation degree, and information uncertainty degree, are proposed for assessing and quantifying intrinsic privacy. The results confirm the effective-

ness of these indices. Intrinsic privacy is applicable to privacy assessments across various electricity consumption scenarios. We select EV charging behavior as a representative electricity consumption scenario and analyze its impact on intrinsic privacy.

MATERIALS AND METHODS

Definition and quantification of intrinsic privacy

Intrinsic privacy is defined as the extent to which sufficiently long-term electricity consumption data, owing to inherent characteristics such as its original shapes and states that are tightly coupled with the user's electricity usage behavior, can obscure or prevent the disclosure of information or activities that the user prefers to keep private. Unlike traditional privacy protection, which relies on external measures such as encryption, intrinsic privacy focuses on the natural, objective privacy protection capabilities inherent in the electricity consumption data itself.

Assessing and quantifying intrinsic privacy requires multiple dimensions to be considered, as many features influence it. Fluctuation, a crucial time series metric, plays a significant role in intrinsic privacy. Electricity consumption data with greater fluctuations is typically more complex and irregular, making it more difficult to predict and thus increasing intrinsic privacy. Autocorrelation measures the relationship between power values at different time points in the data, primarily through periodic and trend-based patterns. Strong autocorrelation indicates high periodicity and information continuity, facilitating easier prediction and thus weakening intrinsic privacy. Furthermore, information uncertainty, which reflects the absence of clear patterns or regularities over time, is another critical factor. High information uncertainty signifies a lack of repetitive patterns and the emergence of new patterns, indicating greater complexity and stronger intrinsic privacy.

Thus, we propose a set of three key indices to quantify intrinsic privacy: the frequency-domain fluctuation degree, trend non-autocorrelation degree, and information uncertainty degree, as mentioned previously. Specifically, for the second index, the non-autocorrelation degree is used instead of autocorrelation to ensure consistency in the directions of the three indices (i.e., higher index values correspond to stronger intrinsic privacy). To facilitate the visual observation of changes in intrinsic privacy, each index is assigned a weighting coefficient that sums to a comparable value, as defined in Equation (1).

$$\text{Intrinsic privacy} = \gamma_1 \text{FdF} + \gamma_2 \text{TACor} + \gamma_3 \text{InfUn} \quad (1)$$

where FdF , TACor , and InfUn are the values of the three indices; γ_1 , γ_2 , and γ_3 represent the corresponding weight coefficients. When the scenario changes, the coefficients should also change accordingly. Detailed algorithms for the indices can be found in the supplementary information.

Frequency-domain fluctuation degree. The first index is the frequency-domain fluctuation degree, which is measured by both high- and low-frequency fluctuation rates. The high-frequency fluctuation rate is determined by calculating the area of the electricity consumption data envelope, whereas the low-frequency fluctuation rate is computed after dimensionality reduction via symbolic aggregate approximation (SAX). The final frequency-domain fluctuation degree index is obtained by weighting these two components via wavelet decomposition. Given that the data envelope is a discontinuous curve, the Lebesgue integral, which effectively handles discontinuities, is used to calculate the area of the envelope.

Trend non-autocorrelation degree. The second index quantitatively characterizes the autocorrelation of the electricity consumption data; for this index, the R/S analysis method is used to examine the scaling behavior of the rescaled range and calculate the Hurst exponent. This method assesses the temporal correlation and trend persistence of the electricity consumption data. A larger Hurst exponent indicates stronger long-term dependence in the electricity consumption data, implying a certain degree of information continuity and stronger periodicity. This makes the series easier to predict accurately, thus weakening its intrinsic privacy. Specifically, to ensure consistency in the directions of all three assessment indices, the trend non-autocorrelation degree is considered as one of the indices in measuring intrinsic privacy. The trend non-autocorrelation degree measures the unpredictability of the electricity consumption data, which contributes to its intrinsic privacy.

Information uncertainty degree. The third index measures the probability

of new patterns emerging in the electricity consumption data in terms of the fuzzy entropy. A higher fuzzy entropy value indicates a greater probability of new patterns, resulting in greater complexity and uncertainty in the electricity consumption sequence. This makes it more challenging to accurately extract information from user electricity consumption data, thereby increasing information uncertainty. Fuzzy entropy can thus quantitatively characterize the information uncertainty of user electricity consumption data. Fuzzy entropy evaluates complexity by calculating the probability distribution of similar patterns in the electricity consumption data.

Validation tests for the intrinsic privacy indices

To validate the effectiveness of the intrinsic privacy assessment indices, the real-world Ireland CER dataset, which records half-hourly electricity consumption data for SME and residential users, was selected as the data source for the following analysis. After removing low-quality data (e.g., permanently uninhabited users or users clearly lacking specific social attributes) and data that did not meet the required sequence length, 120 groups of electricity consumption data (including 60 SME users and 60 residential users) over a 16-day period were randomly selected to test whether data with higher intrinsic privacy level is more difficult to predict and decompose and whether data with lower intrinsic privacy level is easier to analyze and explore; this would be compelling evidence of the effectiveness of the indices. The strength of intrinsic privacy was determined by the three key indices, with the assumption that for weak-intrinsic-privacy data, all three index values are between 0 and 0.5, whereas strong-intrinsic-privacy data has values between 0.5 and 1.

Experiments on load prediction and load decomposition were conducted to test the observability and analyzability of user electricity consumption data. In the load prediction experiment, four distinct prediction algorithms were considered to test the load prediction accuracy for all users, measured by the MSE. Specifically, H-W and SARIMA were selected to represent classical linear statistical methods, while BPNN and LSTM were employed to represent modern non-linear machine learning approaches. This selection covers a broad spectrum of forecasting capabilities to minimize the impact of different methods on various prediction scenarios, ensuring that the evaluated intrinsic privacy is not an artifact of specific model limitations but a fundamental property of the data. In the load decomposition experiment, three users from each social attribute group were selected to construct a decomposable load dataset on the basis of their total electricity consumption data. The residential user loads included LCD TVs, energy-saving lamps, and microwave ovens, whereas the SME user loads included air conditioners, product lines, and servers. Load decomposition was performed via the noninvasive load monitoring method to determine the operating status of electrical equipment.

Note that data scaling and normalization were used in this study solely to ensure the comparability of the results for different groups of user electricity consumption data with varying lengths and peak-valley values, thereby eliminating any potential biases. In this validation test, data normalization uses the maximum and minimum index values of all the groups of users' electricity consumption data as the upper and lower limits, respectively. This approach differs in the subsequent EV charging tests.

EV charging tests: Charging load models

EVs include hybrid electric vehicles, plug-in hybrid electric vehicles, extended-range electric vehicles, and battery electric vehicles. In this study, we consider only battery electric vehicles. The method of EV charging, as well as the time and location of charging, affects the charging load model. Three charging methods—regular slow charging, fast home charging, and superfast charging—with constant rated powers of 7 kW, 20 kW, and 100 kW, respectively, were introduced in the analysis to consider different charging scenarios. The charging durations were 10 hours for regular slow charging, 4 hours for fast home charging, and 0.5 hours for superfast charging, as users typically choose superfast charging for emergencies rather than for fully charging the vehicle. The choice of charging method is closely related to the type of equipment at the charging location, which is the site where the vehicle is connected for charging. The charging locations were considered to be workplaces (e.g., SMEs) or residential areas. SME users have self-built charg-

ing stations capable of supporting all types of charging methods. However, owing to hardware constraints, residential users are limited to choosing between regular slow charging and fast home charging, as they have only home charging piles.

Moreover, batteries are charged via various techniques, including constant current, constant voltage, constant power, and taper charging. Typically, unidirectional charging is considered, meaning that power flows only from the grid to the vehicle (G2V). In this study, we make appropriate simplifications under certain conditions to reduce the computational complexity; for example, we assume constant charging power throughout the entire charging session, as users typically do not intentionally adjust the charging rate. However, in practice, the charging power gradually increases at the beginning and decreases toward the end of the charging process. Therefore, the individual EV charging load is approximated as a translatable trapezoidal wave superimposed on the user's original electricity consumption data.

EV charging tests: Simulation overview

The electricity consumption data of 6 SME users and 6 residential users from the CER dataset over a one-week period (starting at midnight on Monday) was used in these tests. Four key factors were identified for considering the impact of EVs on intrinsic privacy: the charging method, charging time, charging location, and number of EVs. These factors combine to create various scenarios of EV charging behavior. Depending on these factors, EV integration introduces new elements into the overall electricity consumption data that may increase load peaks. By examining the differences between the users' original electricity consumption data and that under various charging scenarios with the assessment indices, we can analyze the impact of EV charging behaviors on intrinsic privacy. All the scenarios involving randomness are repeated 100 times, with the average value taken as the final result for each scenario.

For SME users, EV charging behaviors occur mainly during weekdays, with a possibility of multiple (up to five) EVs being charged at different times of the day. However, considering real-world circumstances, a residential user typically has only one EV. In the simulation tests, since SME users prioritize private information such as device power parameters, which are closely tied to fluctuations in the electricity consumption data, the weight of the frequency-domain fluctuation degree is increased in the calculation. Conversely, residential users are more concerned about private information related to their daily travel habits, which is highly correlated with the trend non-autocorrelation index, so the weight of this index is set to a higher value. Various parameter combinations need to be configured during the tests to ensure generalizability. All other conditions are identical in the two tests, and all the simulated charging scenarios comply with the previous assumptions and real-world conditions, ensuring the reliability of the experimental results.

Additionally, data scaling is necessary to ensure that the three assessment indices can be weighted and summed. The lower limit is set using the index value of a straight line, as a straight line has the lowest intrinsic privacy and is easy to predict. The upper limit is established via the index value of white noise, as white noise has the highest intrinsic privacy, making it impossible to analyze. This scaling ensures that the indices can be effectively summed to obtain a comprehensive privacy assessment.

RESULTS

Intrinsic privacy of SME and residential users and validation of the assessment indices

To gain a deeper understanding of intrinsic privacy, it is first necessary to establish methods for assessing and quantifying it. According to our definition, the inherent shapes and states of user electricity consumption data are the most fundamental determinants that affect intrinsic privacy.³⁶ We propose a set of three key indices that are closely correlated with these factors to quantify intrinsic privacy. These three key indices are (1) the frequency-domain fluctuation degree, (2) the trend non-autocorrelation degree, and (3) the information uncertainty degree.

It is crucial to emphasize why basic statistical measures, such as the mean or variance of power consumption data, are insufficient for this purpose. While mean and variance describe the global magnitude and dispersion of data, they fail to capture the temporal structures and sequential dependen-

cies where privacy leakage actually occurs. For instance, a user with a strictly periodic routine and one with completely random behavior could yield identical variances, yet they represent opposite extremes of privacy preservation. Hence, we use the proposed three indices to specifically target the dynamic dimensions overlooked by basic statistics. The frequency-domain fluctuation captures the complexity of behavioral events, the trend non-autocorrelation measures the lack of predictable routines, and the information uncertainty quantifies the entropy of operational states. Together, these metrics provide a comprehensive and rigorous assessment of the difficulty involved in extracting private information from raw power consumption data.

In general, user electricity consumption data with strong intrinsic privacy is more challenging to analyze and predict accurately, as it has higher level of randomness and fluctuation.^{37,38} This structural complexity significantly hinders attempts to extract meaningful patterns or trends, which strengthens the protection of sensitive user information against unauthorized inference or exploitation. The opposite is true for electricity consumption data with weak intrinsic privacy. Therefore, to demonstrate the effectiveness of the assessment indices, we conduct numerical tests to determine whether these indices can distinguish user electricity consumption data with varying levels of intrinsic privacy. Additionally, we examine whether data with higher intrinsic privacy scores is more challenging to predict and decompose, providing compelling evidence of the indices' effectiveness.

The data used in the tests was sourced from the Ireland CER dataset, which records half-hourly consumption data. After the data was cleaned to remove low-quality entries, 120 groups of electricity consumption data for 60 SME users and 60 residential users over a 16-day period was selected for analysis. In the load prediction experiment, four algorithms, Holt–Winters (H-W), seasonal ARIMA (SARIMA), the backpropagation neural network (BPNN), and long short-term memory (LSTM), were employed, and performance was measured by the mean squared error. In the load decomposition experiment, three users from each social attribute group were chosen to construct a decomposable load dataset on the basis of their total electricity consumption data. The nonintrusive load monitoring method was applied to assess the operational status of electrical appliances for each user category.

Figure 2 shows the three-dimensional spatial distribution of intrinsic privacy strength for electricity users with different social attributes. In this 3D space, the 1/8 cube nearest the origin (blue region) represents the weak-intrinsic-privacy region, and the 1/8 cube farthest from the origin (red region) represents the strong-intrinsic-privacy region, with the intermediate area indicating moderate intrinsic privacy. If the intrinsic privacy value of a user's electricity consumption data falls within the blue region, it is classified as weak-intrinsic-privacy data. Conversely, if it falls within the red region, it is labeled strong-intrinsic-privacy data. We find that most groups of residential user electricity consumption data fall within the strong-intrinsic-privacy region, whereas most groups of SME user data are located in the weak-intrinsic-privacy region, aligning with expectations on the basis of the privacy characteristics of these user groups. These results indicate that the proposed assessment indices effectively distinguish between SME and residential user electricity consumption data, likely because of the differing social attribute characteristics of these user groups.⁴⁻⁷

Load prediction experiment results are shown in Figure 3. It is evident that residential user electricity consumption data with strong intrinsic privacy generally has a higher prediction mean squared error (MSE) than SME user electricity consumption data with weak intrinsic privacy, a trend that becomes particularly apparent as the number of the groups of electricity consumption data increases. This indicates that SME user electricity consumption data is easier to predict accurately with various prediction algorithms, which may be attributed to its relatively simple data structure, characterized by regularity and periodicity;^{4,5} this structure enables prediction algorithms to effectively capture its patterns and trends. Conversely, strong-intrinsic-privacy residential user electricity consumption data, with its complex data structures and greater randomness and fluctuation,^{6,7} poses challenges for prediction algorithms in achieving accurate forecasting. This finding also confirms the obstructive role of strong intrinsic privacy in load prediction, further supporting the study's hypothesis regarding the relationship between intrinsic privacy and prediction accuracy.

Similarly, the corresponding load decomposition accuracy was obtained

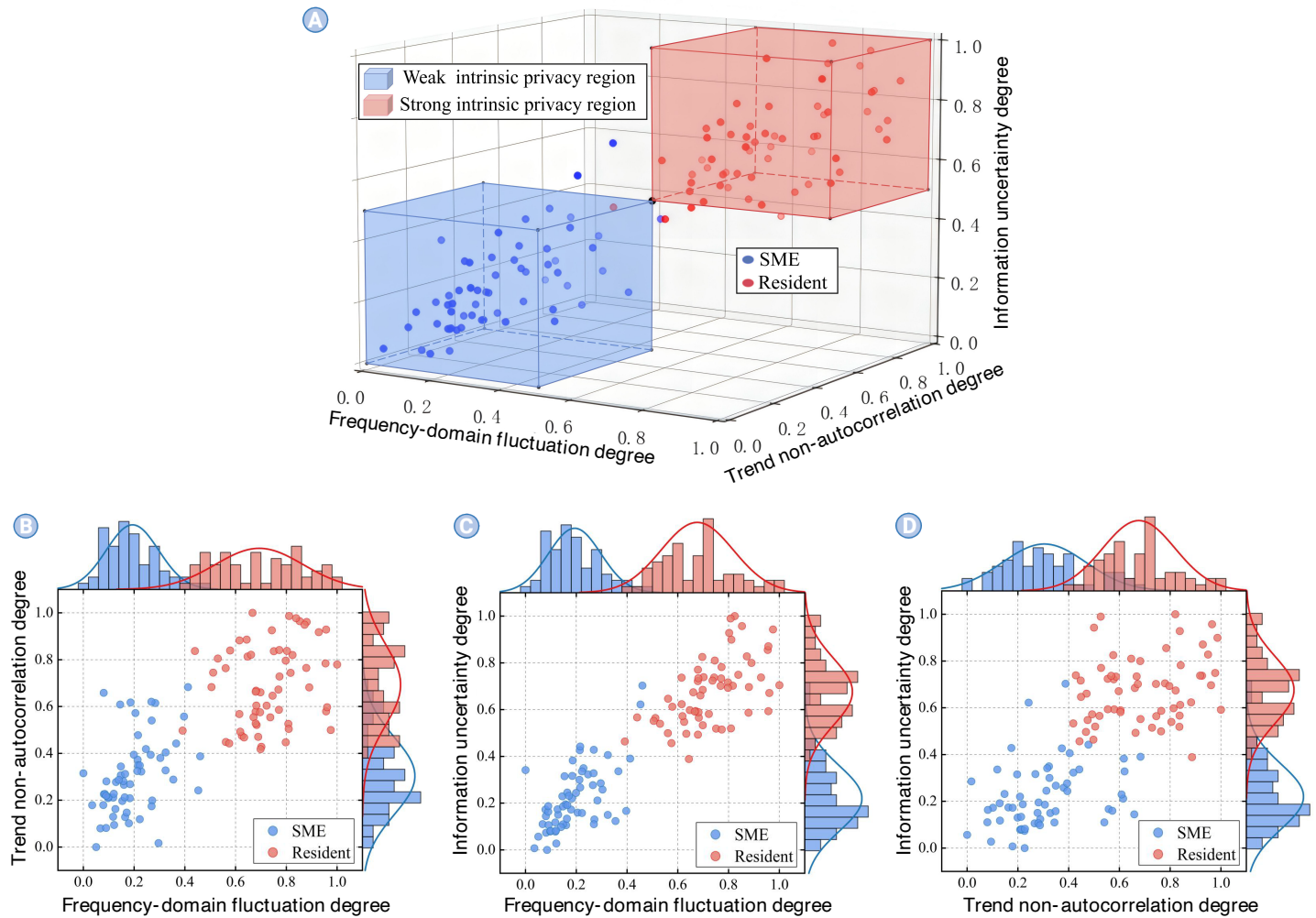


Figure 2. The distribution of two typical users' intrinsic privacy levels The blue region represents a predefined area of weak intrinsic privacy, while the red region corresponds to a predefined area of strong intrinsic privacy. Most groups of the intrinsic privacy index values (represented by blue dots) of SME users' power consumption data fall within the blue region, indicating weak intrinsic privacy for these users' data. In contrast, the intrinsic privacy index values (represented by red dots) across most groups of residential users' power consumption data fall within the red region, indicating strong intrinsic privacy for these users' data. (a) 3D space; (b) projection along the z-axis; (c) projection along the y-axis; (d) projection along the x-axis.

from the load decomposition experiment (see Table 1), along with the ROC curves (see Figure 4). The distinct discrepancy in decomposition accuracy suggests that the strong intrinsic privacy of the residential user electricity consumption data creates a more challenging scenario for accurate load decomposition. The ROC curves in the figure further support these findings. Each curve represents the true positive rate (TPR) versus the false positive rate (FPR) for the load decomposition results across different user groups.³⁹ The areas under the ROC curves (AUCs) serve as a measure of decomposition model performance, with values ranging from 0.5 (no discrimination) to 1 (perfect discrimination). A higher AUC indicates a stronger ability of the model to distinguish between the positive and negative classes. The AUCs for SME users are notably larger than those for residential users, quantitatively confirming the superior load decomposition performance of SME user electricity consumption data.

The combined results of these three experiments demonstrate that the proposed assessment indices can effectively differentiate the intrinsic privacy levels of SME and residential user electricity consumption data, confirming that residential user data, which has higher intrinsic privacy, consistently results in lower accuracy in both load prediction and decomposition tasks. This reduced accuracy is observed across different prediction algorithms and nonintrusive load monitoring methods. Therefore, the relationships between the intrinsic privacy strength and the corresponding load prediction and decomposition accuracies underscore the effectiveness and reliability of these indices as quantitative metrics for intrinsic privacy.

Impact of EV Charging Behavior on Intrinsic Privacy

EV charging behavior is analyzed here as a representative component that significantly alters the shape of household power consumption data. While smart meters typically measure the net demand, the mechanism by which EV charging introduces irregularity serves as a valid case study to demonstrate how modifying the composition of power consumption data impacts intrinsic privacy.

The experiments used 12 groups of electricity consumption data from 6 SME users and 6 residential users in the CER dataset, covering a one-week period starting at midnight on Monday. Four key factors—charging methods, charging time, charging location, and number of EVs—were considered in modeling various scenarios of EV charging behavior. The impacts of these behaviors on intrinsic privacy were assessed by comparing the original electricity consumption data with that under different charging scenarios via the established intrinsic privacy indices. For SME users, multiple EVs (up to five) charge at different times on weekdays, whereas residential users typically have a single EV charging at night. During the simulations, privacy parameters such as power fluctuations and travel habits were weighted differently for SME and residential users, reflecting the specific privacy concerns of each group.

For SME users, privacy concerns primarily involve the potential exposure of critical information, such as operational schedules, rated power, and the energy consumption behaviors of key equipment. Typically, the original electricity consumption data of SME users exhibits high periodicity and regularity,

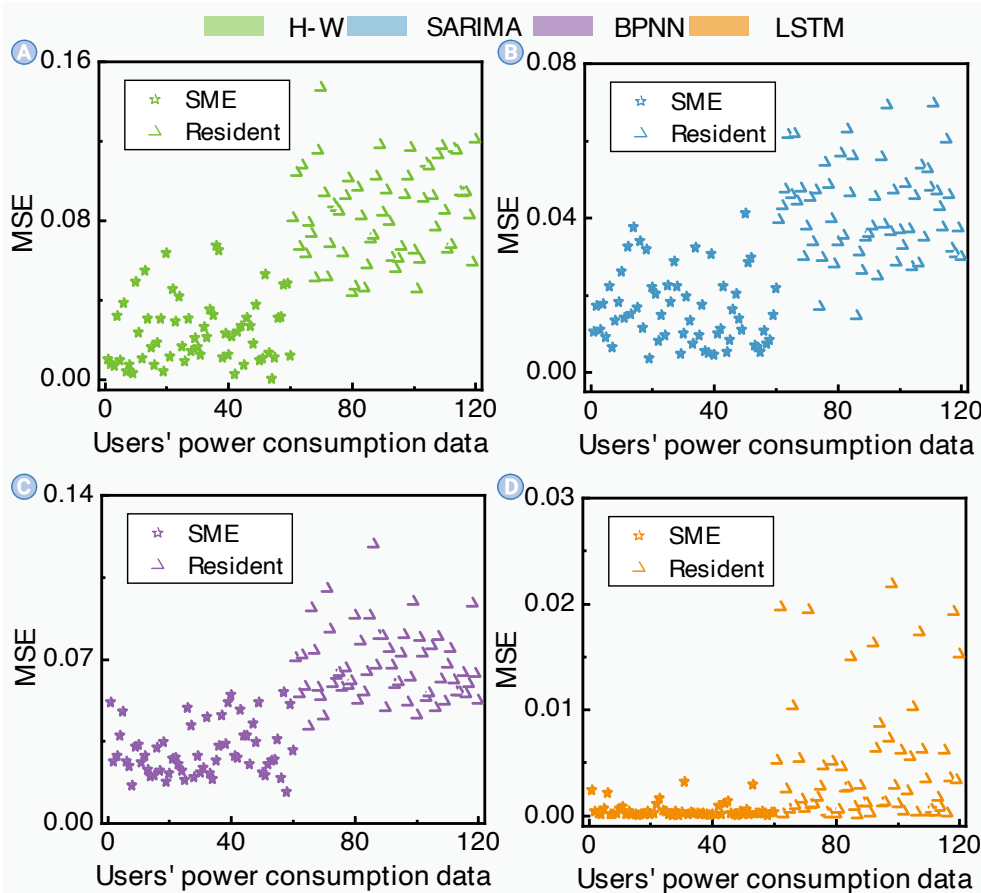


Figure 3. Load prediction accuracies for all users under four classic prediction methods Across four different prediction algorithms, the prediction accuracy for SME users' power consumption data consistently surpasses that of residential users' data. (A) H-W; (B) SARIMA; (C) BPNN; (D) LSTM.

Table 2 shows descriptions of different EV charging scenarios for SME users, where the EV charging behaviors become increasingly irregular. Figure 5 illustrates the changes in intrinsic privacy for the SME user electricity consumption data. The experiment was conducted three times, each with a different set of weighting coefficients. The highest weighting coefficient (SME: frequency-domain fluctuation degree; residential: trend non-autocorrelation degree) was set to 0.5, 0.6, and 0.7 for each iteration, and the other two coefficients remained the same.

In scenarios II and III, where regular charging behaviors are introduced, there is no significant increase in intrinsic privacy because of the high periodicity and regularity in the charging load. Although the additional charging load can obscure certain information, such as the rated powers of production lines, attackers can still employ filtering techniques to remove the regular charging patterns, thereby extracting useful information from SME user electricity consumption data. Furthermore, the peak increase from fast home charging in scenario III is greater than that from slow charging in scenario II, resulting

potentially leading to the inadvertent disclosure of such sensitive information. However, the introduction of EV charging behavior adds a layer of complexity to the electricity consumption data by introducing varying load patterns.^{35,36} This probably acts as a form of noise, obscuring the identifiable characteristics of the original SME user electricity consumption data. In contrast, the privacy concerns of residential users differ significantly from those of SME users, focusing primarily on the potential exposure of personal mobility patterns and daily routines. Unlike SME user electricity consumption data, which has high periodicity and regularity, residential user electricity consumption data is shaped by individual behaviors and various household electrical devices, resulting in strong intrinsic privacy.^{18,19} When EV charging behaviors are introduced into residential user electricity consumption data, these charging activities are often governed by work schedules and home routines, resulting in more regular and predictable charging patterns. This regularity likely introduces consistent charging waves into the original residential electricity consumption data, which may inadvertently reduce intrinsic privacy.

To simplify the comparison of intrinsic privacy across different charging scenarios, the three key indices were weighted and summed to produce a single effective value in the EV charging test (see Experimental procedures). Note that the data scaling method used in the test differs from that in the index validation test; it is expected that the index values for SME users may generally exceed 0.5, whereas those for residential users may fall below 0.5. Since the same data scaling standards were applied to both SME and residential users across all charging scenarios, this does not affect result interpretation, as the experiment focuses on horizontal comparisons of scenarios for users with similar social attributes.

in greater fluctuation, which accounts for the higher intrinsic privacy observed in scenario III. Scenarios IV and V indicate a notable increase in intrinsic privacy, attributed to the introduction of irregular and nonperiodic charging behaviors. These behaviors disrupt the daily repetitive patterns of SME user electricity consumption data, and the pronounced peaks introduced by superfast charging further increase this fluctuation. These findings indicate that while regular EV charging behaviors can introduce some level of increased privacy, irregular and highly random charging behaviors significantly increase the intrinsic privacy of SME users. The added stochastic noise from these charging patterns effectively masks sensitive information, such as operational schedules and the electricity consumption behaviors of key equipment, providing SMEs with better protection against potential privacy breaches.

Table 2 also provides descriptions of various charging scenarios for residential users, with EV charging behaviors becoming increasingly irregular. The changes in intrinsic privacy for residential users are shown in Figure 6. The weighting coefficient settings are identical to those in the previous experiment. In residential scenario I, where no EV charging behavior is introduced, intrinsic privacy is higher than that in SME scenario I because of the inherent fluctuation and irregularity of residential electricity consumption data. However, in scenario II, where slow charging occurs at a fixed time each night, there is a significant decrease in intrinsic privacy. This reduction occurs because regular EV charging behavior introduces repetitive patterns into otherwise volatile residential user electricity consumption data, increasing load regularity and autocorrelation. These regular patterns make the electricity consumption data easier to predict, thereby significantly reducing intrinsic privacy. In the other three scenarios, the index values of intrinsic privacy

Table 1. Load Decomposition Accuracy

User	SME 1	SME 2	SME 3	Re 1	Re 2	Re 3
Accuracy	0.863	0.804	0.815	0.714	0.696	0.661

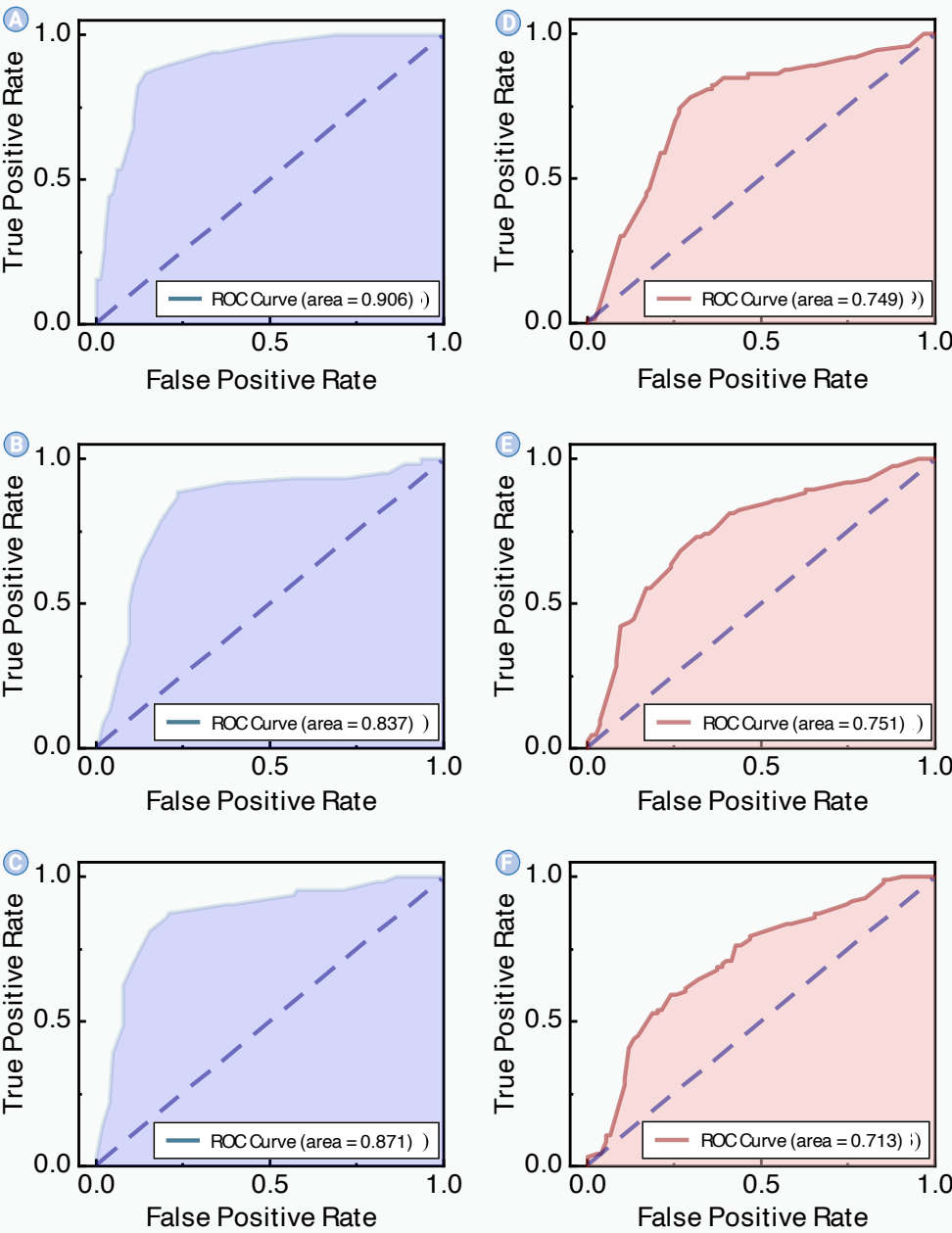


Figure 4. ROC curves of load decomposition The blue curves represent the ROC curves generated by SME users in the load decomposition task, whereas the red curves represent the ROC curves for residential users in the same task. The AUC values for the three groups of SME users' power consumption data in the load decomposition task are consistently higher than those for residential users (A) SME user 1; (B) SME user 2; (C) SME user 3; (D) residential user 1; (E) residential user 2; (F) residential user 3.

rebound somewhat because the increase in regularity from the repetitive patterns is counterbalanced by added fluctuations from the new peaks. The analysis of residential users suggests that while regular EV charging behaviors at fixed times can significantly reduce intrinsic privacy by introducing predictable patterns, random charging behaviors and varying charging speeds can mitigate this effect by adding complexity and fluctuation to electricity consumption data.

DISCUSSION

The novel concept of intrinsic privacy is introduced here to capture the unique privacy characteristics arising from the inherent shapes and states of a user's electricity consumption data, reflecting its innate capacity to conceal or withhold details that the user prefers to keep private. Three key indices are used to assess and quantify intrinsic privacy. To validate the effectiveness of these indices and analyze the impact of EV charging behaviors on intrinsic privacy across different social attributes, half-hourly data from both SME and residential users in the Ireland CER dataset is utilized. Although current measurement equipment limits the granularity of electricity consumption data, the selected duration is sufficient to capture short-term daily routines and weekly cycles, which are the primary targets for real-time privacy attacks such as occupancy detection. Consequently, these limitations do not affect the generalizability of the results, as representative users and examples were selected for the experiments to illustrate a

Table 2. EV Charging Scenarios for SME and Residential Users

Scenarios	Descriptions
SME	I No EV charging behavior introduced
	II All EVs charging slowly at a fixed time
	III All EVs charging quickly at home at a fixed time
	IV All EVs charging at random times with a mix of fast and slow charging
	V Random number of EVs charging at random times with a mix of fast and slow charging
Resident	I No EV charging behavior introduced
	II Slow charging at a fixed time each night
	III Fast home charging at a fixed time each night
	IV Fast home charging at random times each night
	V A mix of fast and slow charging at random times each night

process by which broader insights are drawn from specific cases. With advancements in technology, future research and analysis may utilize more fine-grained data, such as minute-level or second-level data, and extend the analysis to seasonal datasets to evaluate the long-term evolution of intrinsic privacy across different climate conditions. Furthermore, additional user attributes and a broader range of scenarios can be explored within the same framework, while future work could also focus on optimizing and updating the algorithms for the three key indices to further increase the accuracy and applicability of the intrinsic privacy assessment method.

The intrinsic privacy differences between SME and residential users, as calculated via the proposed indices, are substantial. These differences arise from the distinct power usage behaviors and habits associated with the social attributes of users, which lead to unique electricity consumption data characteristics and varying levels of intrinsic privacy.^{4-7,25,26} Specifically, SME users typically operate large equipment in predictable cycles aligned with workdays, resulting in a highly predictable load pattern and, consequently, lower intrinsic privacy. In contrast, residential users have greater intrinsic privacy due to the randomness and variability in household appliance usage, which complicates load analysis. Notably, different EV charging loads indeed

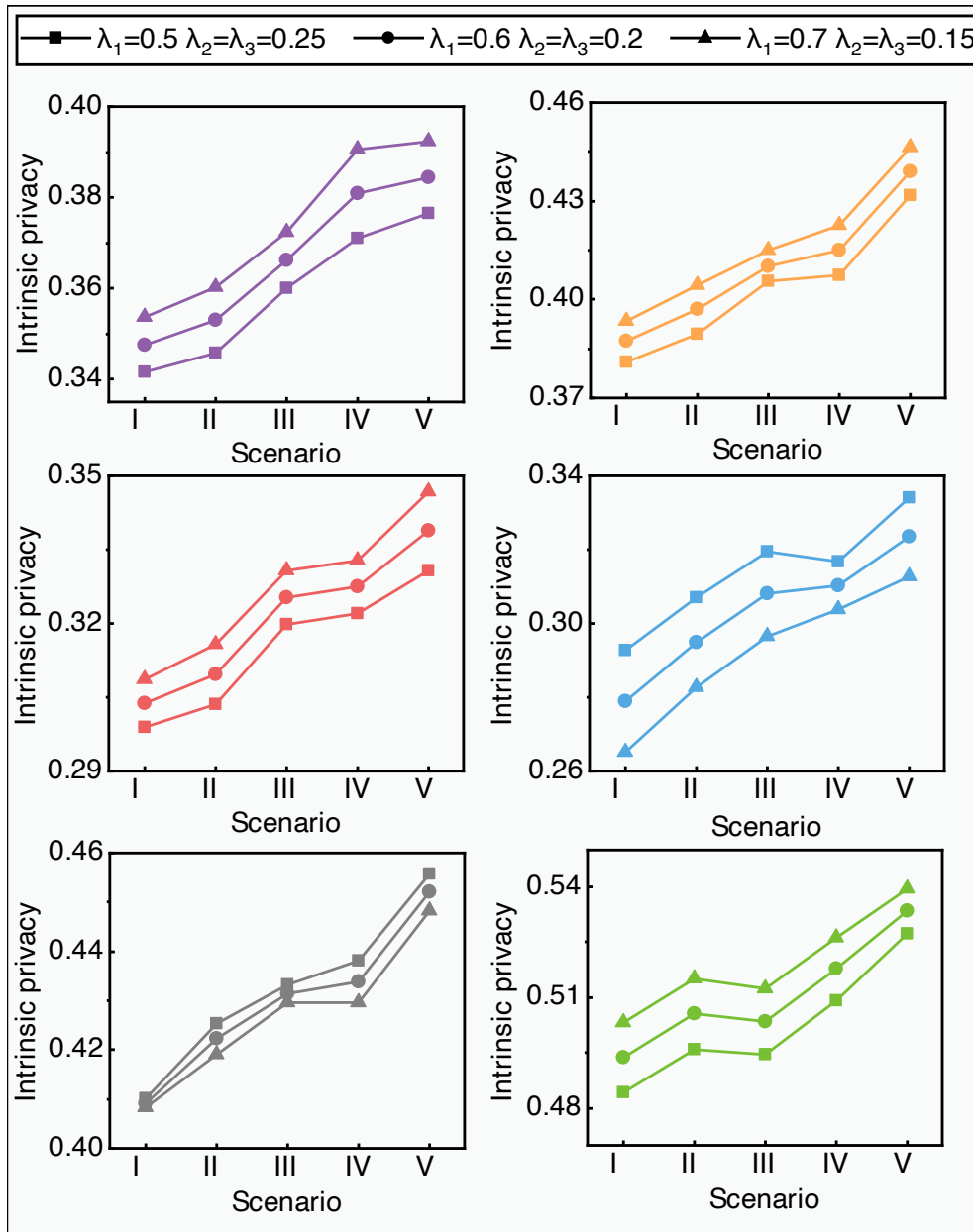


Figure 5. Changes in the intrinsic privacy of SME users under different EV charging scenarios The three distinct shapes in the line charts represent different coefficient combinations, and the six different colors represent different SME users. As the load patterns in the EV charging scenarios become more complex and irregular, the intrinsic privacy of the six SME users involved in the experiment significantly improves. Additionally, the more irregular the EV charging patterns are, the higher the intrinsic privacy of the users.

electricity consumption data for electricity users with different attributes is prone to be treated equally, and the use of untargeted and undifferentiated protection measures leads to excessive privacy protection, resulting in unnecessary resource waste. Intrinsic privacy helps electricity companies and users gain a deeper understanding of their privacy needs, allowing for the design of more precise and personalized privacy protection strategies. This approach not only safeguards privacy but also reduces resource costs. It is also observed that a high level of Intrinsic Privacy correlates with lower prediction accuracy. While this implies a challenge for grid forecasting, the core value of this metric lies in empowering the user. Users have the right to know the intrinsic privacy level of their power consumption data, i.e., to what extent their behavior is naturally exposed or concealed. This knowledge enables informed decision-making: users with naturally high intrinsic privacy might share data more freely, while those with low intrinsic privacy can decide to implement active masking strategies.

In addition, intrinsic privacy can guide users in managing the privacy characteristics of their electricity consumption data by adjusting the smart charging patterns of EVs. Users can flexibly adjust the timing and frequency of EV charging to modify the shapes of their consumption data, thereby increasing privacy protection. Furthermore, intrinsic privacy allows users to evaluate the privacy level of their electricity consumption data at various time scales and

influence intrinsic privacy to varying degrees. For SME users with regular electricity usage patterns, EV charging can introduce additional noise, effectively masking parameters such as production line equipment specifications and increasing intrinsic privacy. However, for residential users with more variable electricity usage, EV charging introduces regular load patterns, thus reducing intrinsic privacy. If the residential user considers such regular patterns to threaten privacy, smart charging or other effective privacy protection methods could be explored and applied. Note that we consider only the impact of EV charging behavior on intrinsic privacy in the EV experiment. However, as EVs and their charging stations continue to evolve rapidly, there is potential for EVs not only to absorb power from the grid but also to feed power back into it. Since EVs can be viewed as home energy storage devices, controlling their charging and discharging behaviors to add noise to user electricity consumption data could be explored as a method for more efficient privacy protection.

Overall, research on intrinsic privacy can address the limitations of existing privacy protection studies. Current studies focus primarily on external data processing and modification, such as data anonymization, encryption, and perturbation. While these algorithms enhance privacy, they often overlook the unique privacy characteristics of user electricity consumption data and the specific privacy needs of users. Consequently, in practical implementation,

granularities, reminding them that privacy depends solely on their electricity usage behavior, not measurement precision. Thus, regardless of whether users choose high-precision smart meters or specific measurement resolutions, their privacy remains unaffected. This attribute can not only reduce user concerns about the detailed data collected by smart meters but also enable grid operators to flexibly choose different data collection frequencies and resolutions while preserving user privacy. Operators can provide customized data collection plans for different users, enabling more adaptable data collection and management without reducing privacy level and thereby ultimately supporting more efficient grid operations.

REFERENCES

- Butt O.M., Zulfarnain M. and Butt T.M. (2021). Recent advancement in smart grid technology: Future prospects in the electrical power network. *Ain Shams Eng. J.* 12:687–695. DOI:10.1016/j.asej.2020.05.004
- Kataray T., Dhananjay U., Balamurugan S., et al. (2023). Integration of smart grid with renewable energy sources: Opportunities and challenges—A comprehensive review. *Sustain. Energy Technol. Assess.* 58:103363. DOI:10.1016/j.seta.2023.103363
- Khalid M. (2024). Smart grids and renewable energy systems: Perspectives and grid integration challenges. *Energy Strategy Rev.* 51:101299. DOI:10.1016/j.esr.2024.101299
- Gholian A., Mohsenian-Rad H. and Hua Y. (2015). Optimal industrial load control in

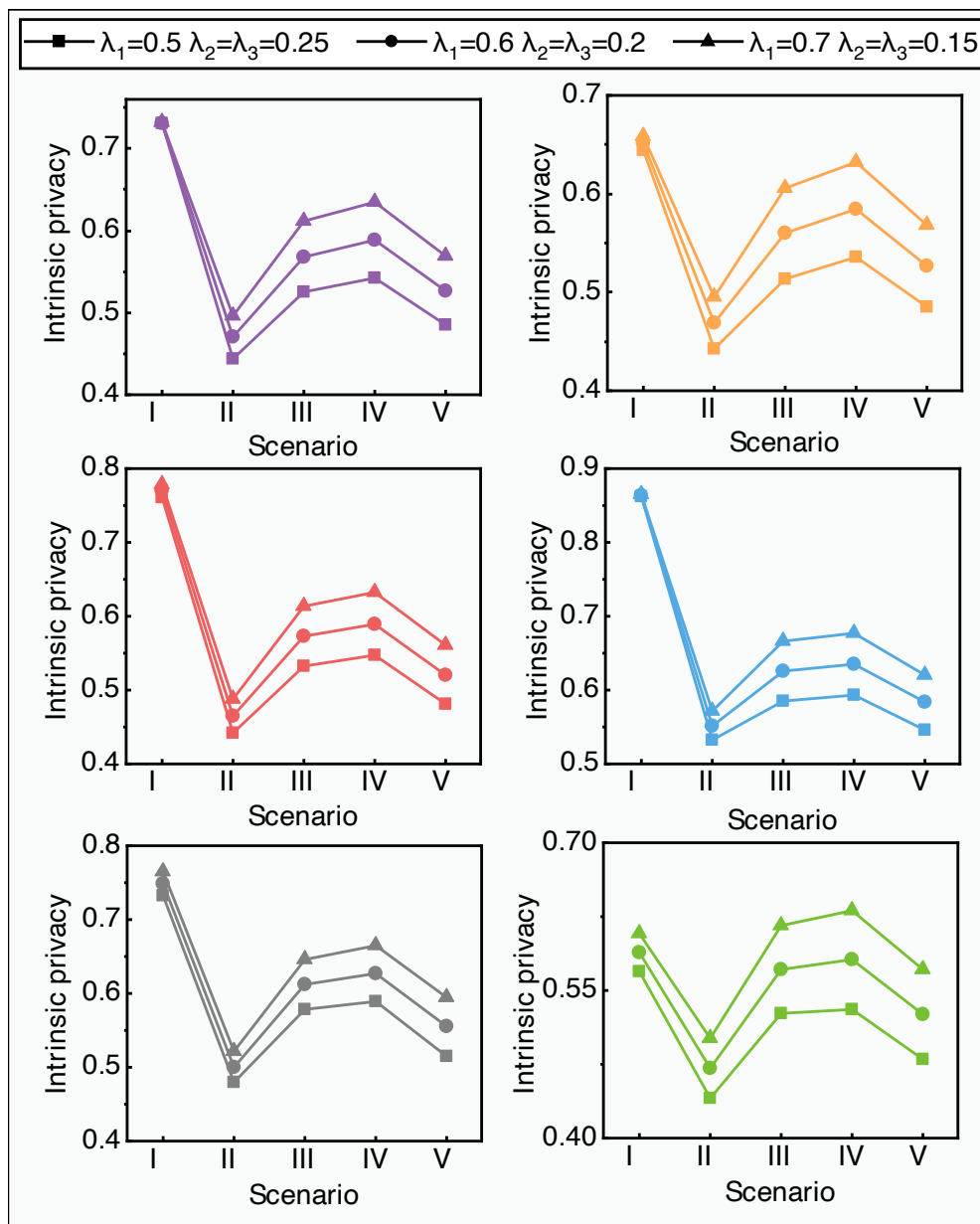


Figure 6. Changes in the intrinsic privacy of residential user loads in different EV charging scenarios The three distinct shapes in the line charts represent different coefficient combinations, and the six different colors represent different residential users. When regular EV charging behavior is introduced, the intrinsic privacy of the six residential users significantly decreases. However, as the charging behavior becomes more irregular, the intrinsic privacy of residential users begins to improve.

14. Guo Z., Zhou K., Zhang C., et al. (2018). Residential electricity consumption behavior: Influencing factors, related theories and intervention strategies. *Renew. Sustain. Energy Rev.* **81**:399–412. DOI:10.1016/j.rser.2017.07.046
15. Abreu J.M., Pereira F.C. and Ferrão P. (2012). Using pattern recognition to identify habitual behavior in residential electricity consumption. *Energy Build.* **49**:479–487. DOI:10.1016/j.enbuild.2012.02.044
16. Hernández L., Baladrón C., Aguiar J.M., et al. (2012). Classification and clustering of electricity demand patterns in industrial parks. *Energies* **5**:5215–5228. DOI:10.3390/en5125215
17. Yu H., Jia H., Zhang Z., et al. (2023). Privacy-preserving demand response of aggregated residential load. *Appl. Energy* **339**:121018. DOI:10.1016/j.apenergy.2023.121018
18. McLaughlin S., McDaniel P. and Aiello W. (2011). Protecting consumer privacy from electric load monitoring. *ACM Conf. Comput. Commun. Secur.*:87–98. DOI:10.1145/2046707.2046719.
19. Lisovich M. and Wicker S. (2008). Privacy concerns in upcoming residential and commercial demand-response systems. *IEEE Power Syst. Conf. Expo.*:1–10. DOI:10.1109/PSCE.2008.4836041.
20. Palacios-Garcia E.J., Carpent X., Bos J.W., et al. (2022). Efficient privacy-preserving aggregation for demand side management of residential loads. *Appl. Energy* **328**:120112. DOI:10.1016/j.apenergy.2022.120112
21. Yusoff S., Rusli M.E., Yusoff Y., et al. (2018). Privacy protection issues in the advanced metering infrastructure. *Proc. 6th Int. Conf. Inf. Technol. Multimed.*:11–14. DOI:10.1109/ICIMU.2018.8692797.
22. Armoogum S. and Bassoo V. (2019). 8 - Privacy

smart grid. *IEEE Trans. Smart Grid* **7**:2305–2316. DOI:10.1109/TSG.2015.2490429

5. Wang Z., Li M., Li L., et al. (2023). A non-intrusive method of industrial load disaggregation based on load operating states and improved grey wolf algorithm. *Appl. Energy* **351**:121934. DOI:10.1016/j.apenergy.2023.121934
6. Lu N., Xie Y., Huang Z., et al. (2008). Direct load control for ancillary services. *IEEE Power Energy Soc. Gen. Meet.*:1–5. DOI:10.1109/PES.2008.4596786.
7. Proedrou E. (2021). A comprehensive review of residential electricity load profile models. *IEEE Access* **9**:12114–12133. DOI:10.1109/ACCESS.2021.3051459
8. Wang H., Qi L., Yan L., et al. (2020). Load photo: A novel analysis method for load data. *IEEE Trans. Smart Grid* **12**:1394–1404. DOI:10.1109/TSG.2020.3031923
9. Elahe M.F., Jin M. and Zeng P. (2021). Review of load data analytics using deep learning in smart grids: Open load datasets, methodologies, and application challenges. *Int. J. Energy Res.* **45**:14274–14305. DOI:10.1002/er.6746
10. Mehdipour Pirbazari A., Farmanbar M., Chakravorty A., et al. (2020). Short-term load forecasting using smart meter data: A generalization analysis. *Processes* **8**:484. DOI:10.3390/pr8040484
11. Wang Y., Chen Q., Hong T., et al. (2018). Review of smart meter data analytics: Applications, methodologies, and challenges. *IEEE Trans. Smart Grid* **10**:3125–3148. DOI:10.1109/TSG.2018.2818167
12. Chowdhury P.K.R., Ninsawat S. and Nitivattananon V. (2020). Electricity consumption patterns within cities: application of a data-driven settlement characterization method. *Int. J. Digit. Earth* **13**:1425–1446. DOI:10.1080/17538947.2020.1774426
13. Rahman A., Srikumar V. and Smith A.D. (2018). Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks. *Appl. Energy* **212**:372–385. DOI:10.1016/j.apenergy.2017.12.051
- of energy consumption data of a household in a smart grid. *Smart Power Distrib. Syst.*:163–177. DOI:10.1016/B978-0-12-812154-2.00008-0.
23. Cheng L., Liu F. and Yao D. (2017). Enterprise data breach: causes, challenges, prevention, and future directions. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **7**:e1211. DOI:10.1002/widm.1211
24. Kalogridis G., Cepeda R., Denic S.Z., et al. (2011). Elecprivacy: Evaluating the privacy protection of electricity management algorithms. *IEEE Trans. Smart Grid* **2**:750–758. DOI:10.1109/TSG.2011.2160876
25. Dayarathna M., Wen Y. and Fan R. (2015). Data center energy consumption modeling: A survey. *IEEE Commun. Surv. Tutor.* **18**:732–794. DOI:10.1109/COMST.2015.2481183
26. Chen W., Zhou K., Yang S., et al. (2017). Data quality of electricity consumption data in a smart grid environment. *Renew. Sustain. Energy Rev.* **75**:98–105. DOI:10.1016/j.rser.2016.10.054
27. Miah M.D., Foyzal M.A., Koike M., et al. (2011). Domestic energy-use pattern by the households: A comparison between rural and semi-urban areas of Noakhali in Bangladesh. *Energy Policy* **39**:3757–3765. DOI:10.1016/j.enpol.2011.04.003
28. Mansouri I., Newborough M. and Probert D. (1996). Energy consumption in UK households: Impact of domestic electrical appliances. *Appl. Energy* **54**:211–285. DOI:10.1016/0306-2619(95)00034-Q
29. Sanguesa J.A., Torres-Sanz V., Garrido P., et al. (2021). A review on electric vehicles: Technologies and challenges. *Smart Cities* **4**:372–404. DOI:10.3390/smartcities4010022
30. Kumar R.R. and Alok K. (2020). Adoption of electric vehicle: A literature review and prospects for sustainability. *J. Clean. Prod.* **253**:119911. DOI:10.1016/j.jclepro.2019.

119911

31. Liu W., Placke T. and Chau K.T. (2022). Overview of batteries and battery management for electric vehicles. *Energy Rep.* **8**:4058–4084. DOI:10.1016/j.egy.2022.03.016
32. Deng J., Bae C., Denlinger A., et al. (2020). Electric vehicles batteries: requirements and challenges. *Joule* **4**:511–515. DOI:10.1016/j.joule.2020.01.013
33. Dixon J. and Bell K. (2020). Electric vehicles: Battery capacity, charger power, access to charging and the impacts on distribution networks. *eTransportation* **4**:100059. DOI:10.1016/j.etr.2020.100059
34. Unterwiesing A., Knirsch F., Engel D., et al. (2022). An analysis of privacy preservation in electric vehicle charging. *Energy Inform.* **5**:3. DOI:10.1186/s42162-022-00185-9
35. Saxena N., Grijalva S., Chukwuka V., et al. (2017). Network security and privacy challenges in smart vehicle-to-grid. *IEEE Wirel. Commun.* **24**:88–98. DOI:10.1109/MWC.2017.1600318
36. Song Z. and Li Z. (2024). A review of smart grid security and privacy protection technologies. *Int. Conf. Energy Electr. Power Eng.*:954–961. DOI:10.1109/CEEPE62022.2024.10586302.
37. Mohsenian-Rad A.-H. and Leon-Garcia A. (2010). Optimal residential load control with price prediction in real-time electricity pricing environments. *IEEE Trans. Smart Grid* **1**:120–133. DOI:10.1109/TSG.2010.2055903
38. Zhao A., Mi L., Xue X., et al. (2022). Heating load prediction of residential district using hybrid model based on CNN. *Energy Build.* **266**:112122. DOI:10.1016/j.enbuild.2022.112122
39. Fan J., Upadhye S. and Worster A. (2006). Understanding receiver operating characteristic (ROC) curves. *Can. J. Emerg. Med.* **8**:19–20. DOI:10.1017/S1481803500013336

FUNDING AND ACKNOWLEDGMENTS

This work is supported by National Natural Science Foundation of China (52477132).

AUTHOR CONTRIBUTIONS

Z. S. executed the experiments and wrote the initial manuscript draft. X. H. examined and polished the manuscript. Z. L. conceived the study, designed the experiments, and revised the manuscript. P. J. contributed to discussing the data and revising the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

DATA AND MATERIALS AVAILABILITY

All original data values taken from the CER dataset in this article are available from the Irish Social Science Data Archive (Commission for Energy Regulation (CER)).

DATA AND CODE AVAILABILITY

All original code has been deposited at GitHub at https://github.com/Szzzzh826/Intrinsic_Privacy.git and is publicly available as of the date of publication. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

SUPPLEMENTAL INFORMATION

It can be found online at <https://doi.org/10.59717/j.xinn-energy.2026.100142>

LEAD CONTACT WEBSITE

Further information questions should be directed to Zhiyi Li (zhiyi@zju.edu.cn).