A Big Data–Driven Data Modeling Framework for Smart Energy Platforms

P. Hemalatha
Assistant Professor
Adaikalamatha College, Vallam, Thanjavur

Dr. J. Lavanya
Research Supervisor, Department of Computer Science
Soka Ikeda College of Arts And Science for Women, Chennai

OPEN ACCESS

Volume: 13

Special Issue: 3

Month: February

Year: 2026

P-ISSN: 2321-788X

E-ISSN: 2582-0397

Citation:
Hemalatha, P., and J. Lavanya. "A Big Data–Driven Data Modeling Framework for Smart Energy Platforms." *Shanlax International Journal of Arts, Science and Humanities,* vol. 13, no. 3, 2026, pp. 139–47.

DOI:
https://doiorg/10.34293/sijash.v13iS3-i1-Feb.10273

Abstract

The intensive development of smart energy systems has resulted in the continual production of high volumes of heterogeneous data of smart meters, sensors, grid infrastructure, and distributed energy resources. The management of this data must be effective to achieve real-time monitoring, predictive analytics and intelligent decision-making in the smart energy platforms. In this paper, a data modeling framework using big data is introduced to facilitate data mapping, integration, and management of high volume, velocity and variety energy data. The suggested framework follows a layered architectural design in order to process structured, semi-structured, and unstructured data and guarantee the scalability and flexibility of the system. The big data technologies are used to offer distributed storage, fault tolerance and parallel processing features, which increases efficiency of data access and lowers processing latency. The framework is tested with representative energy datasets and performance analysis is performed in terms of scalability, data retrieval performance as well as processing performance. According to the results of the experiment, the suggested data modeling framework proves to be much more efficient in the organization of the data and the provision of sophisticated analytics to serve smart energy exploitation applications (demand forecasting, energy monitoring, and operational optimization).

Keywords: Big Data, Data Modeling, Smart Energy Platforms, Energy Data Management, Data Mapping, Scalable Architecture, Energy Analytics

Introduction

In an effort to conceptualize the term smart energy, some of the scholars carried out analyses rooted in the energy concepts adopted before 2012 which included non-renewable energy and renewable energy. According to Lund et al. [1] there are three stages which can be used in the implementation of renewable energy: the introduction stage, the large-scale integration stage and the 100 percent renewable energy stage. There is greater consumption of the non-renewable energy in the initial phase, which is more than that of the renewable energy, hence addition of renewable does not influence the current distribution systems. During the second stage, renewable energy will be incorporated in raising the supply hence the distribution networks should be stabilised. The third phase is based on the assumption that the only sources that produce energy are renewable ones; thus, it is necessary to reach the balance of power. The conversion change challenges that must be considered are conversion changes among energy generators,infrastructure, storage and stabilisation of the frequency and voltage in the current distribution networks.

Literature Survey

Abella et al. [1] observe that in the case of the electricity industry, smart energy is conceptualised as a series of technologies, applications and services, which allow the prosumer to actively become a part of it. A prosumer refers to the consumer and producer of energy.

In his publication, Sanchez [2] indicates that smart energy is viewed as a new method of energy management process, both in production and consumption.

In his book on intelligent energy systems, Aichele [3] argues about the implementation of digital electricity meters that track the consumption and performance of electricity in real time. These devices relay information to the energy companies to be analyzed and are deployed as a tool to improve their services; they can also find and bring on board consumers as well as producers of energy. This, according to the author, will be one of the steps towards the design of self-regulating networks that will make up the intelligent energy system.

In their article about energy storage and smart energy systems, Thellufsen et al. [4] explain the significance of diversifying renewable sources of energy and how the latter affects the development of intelligent storage systems.

All these views highlight the complexity and the multi-dimensionality of the necessity to transition to the smart energy systems, which not only will ensure efficient use of renewable resources but also will combine the advanced technologies to enhance the efficiency, reliability and the Quality of Energy Services in order to define the way towards the sustainable and highly interconnected energy future. Smart energy refers to the overall fusion of the generation, transmission and consumption systems using real-time monitoring, two-way communication and adaptive control to maximise efficiency, reliability and sustainability; smart energy is informed by such standards as IEEE 2030 on smart grid interoperability, IEC 61850 on substation communication protocols, and NIST SP 1108 on demand response and distributed resource architecture to enable smooth data flow and dynamic response of the system.

The conceptual framework diagram (Figure 1) below shows the general structure and logical sequence of this review, outlining its main parts and their interrelations with each other to help a reader navigate through the analysis.

Although significant progress has been made regarding Smart Grid Algorithms, three main gaps remain: (1) lack of universal taxonomy to integrate traditional and ensemble machine learning methods; (2) little benchmarking of algorithm performance on applications of demand forecasting and battery storage management and (3) discussion of policy and ethical implications of large-scale application of Smart grids. The present paper will close these gaps by (i) suggesting a twelve-category classification scheme and (ii) by describing directions of future research.

We present in our review aspects that have not been considered in the analyses that have been made previously: we include the use of algorithms other than traditional Random Forest - namely applied in load-distribution issues, system failure prediction and real-time power-supply adjustments to achieve high quality of service and grid stability, and we formulate mathematically the optimisation problem in general in the smart energy, specifying the objective functions, the decision variables and the system-specific constraints. Further, we pre- introduce revised concepts in Quality of Energy Service, including harmonic distortion indices, voltage fluctuation response protocols, and these concepts are not found in the previous literature. Lastly, a practical case study to integrate the idea of Random Forest, XGBoost and LSTM model to optimize grid management, predict energy demand and improve operational efficiency is presented, thus showing how different methodologies can be synergically implemented to improve smart grid technologies.

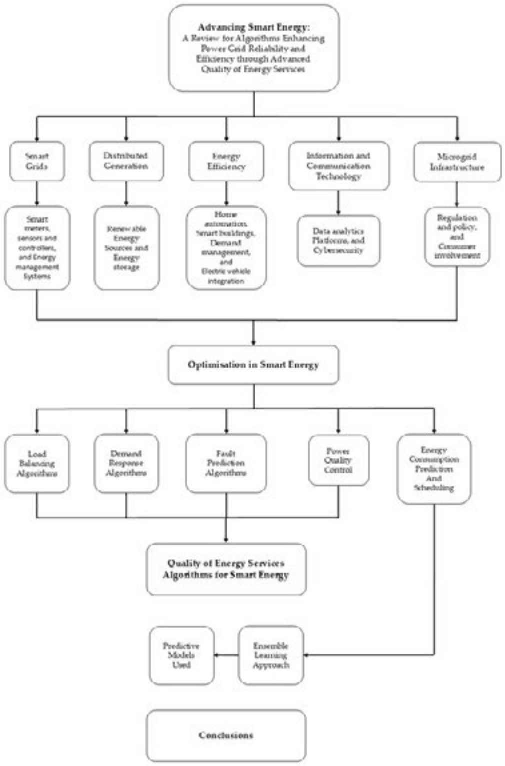

Figure 1. Conceptual framework diagram of the review

Smart Energy

Smart energy denotes efficient and sustainable consumption of energy by introducing smart technologies to the energy generation, distribution, and consumption systems that reduce energy wastage by ensuring optimal use in residential, industrial, and transportation sectors. It entails energy optimization to ensure high efficiency and also promoting the sustainability of the environment and grid stability, improving the user experience by utilizing intelligent technologies and activities.

Smart energy enhances the ability of the grid to survive and recover after shocks and attacks to ensure there is a stable and secure supply of energy. It gives consumers an enhanced power on their energy consumption by enabling them to monitor and adjust their consumption through intelligent devices and mobile applications. Smart energy promotes efficiency and use of clean energy thus reducing greenhouse gas emissions and carbon footprint. Demand-response programs and dynamic pricing can be used to encourage the consumer to consume energy during the off-peak periods which will help to balance the load in the grid.

Smart grids (comprising of smart meters, sensors, controllers, and energy management systems), distributed generation (including renewable sources and energy storage), energy efficiency (through home automation, smart buildings, and demand management), electric vehicle integration (using smart charging stations), information and communication technology (including data analytics platforms and cybersecurity), microgrids infrastructure, energy industry regulations and policies, and consumer interaction supported by smart technologies like sensors, the Internet of Energy, smart meters, and automated control systems to monitor and manage energy flows are major parts of smart energy.

Another concept is the Internet of Energy (IoE), a new concept in the energy industry, which seeks to highly integrate the information and communication technologies with traditional energy networks. This convergence enables effective and sustainable energy management, including distributed generation, smart grids, energy efficiency and active consumer engagement. The author concludes that through the IoE, the energy sector will be able to use digital technologies to their fullest to enhance the resilience, efficiency, and sustainability of the global energy system.

The optimal point of an optimization problem is to find the set of decision variables x that optimize or minimize the objective function f(x) subject to certain given constraints. The general optimization formulation of smart energy issues generally contains:

- Objective: Minimizing or maximizing an objective function that is a meaningful metric of a system, e.g. total cost, energy usage, or efficiency.
- Decision Variables: The power that can be modified to attain maximum objective, which includes power generation, distribution, storage and demand.
- Constraints: The restrictions and conditions to be met within the system including capacity limits, energy conservation laws, and regulations among others.

The mathematical model of the general optimization problem in smart energy can be written in the form of the following:

Objective Function

$f(x) = \min \Sigma (c_i * x_i)$ (1)

Subject to:

$g_i \leq G_i, i = 1,..., n$ (Generation capacity)

$\Sigma g_i = \Sigma d_j$ (Energy balance)

$s_i \leq S_i, i = 1,..., k$ (Storage limitations)

$l \leq L$ (Grid constraints)

$r_i \leq R_i, i = 1,..., p$ (Restrictions on renewable energy integration)

In this case, c i is the cost or benefit coefficient, x i is the decision variable, g i is the current generation, G i is the maximum generation capacity, d j is the demand at node j, s i is the current storage, S i is maximum storage capacity, l is the flow in the line, L is the capacity of the line, r i is the current store generation and R i is the maximum store generation capacity. Other limitations that can be added are connected with CO 2 emission and the governmental regulations.

Areas of Optimization in Smart Energy

1. Smart Grid Optimization: Concentrates on the optimization of the smart meters to reduce the number of errors and enhance the extent of the communication process and optimization of the sensors and controllers to balance the load and control the power flow. It also incorporates optimization of energy management systems in the efficient planning of generation, distribution as well as consumption.
2. Distributed Generation Optimization: The objective is to maximize the amount of energy captured and reduces the costs of renewable energy through optimizations of the sources. It also entails maximization of energy storage to maximize the capacity and minimized losses.
3. Energy Efficiency Optimization: Can be done by ensuring smart buildings and automated systems are optimized to use less energy and the use of algorithms and demand management that will ensure demand is reduced depending on the demands of the grid.

4. Electric Vehicle Integration: Refers to the optimization of smart charging stations to make them optimally efficient in terms of charging and grid integration.
5. Information and Communication Technologies (ICT) Optimization: It involves optimization of data analytics systems to be used in data analysis and to optimize cybersecurity to reduce vulnerabilities and risks.
6. Microgrid Infrastructure Optimization: Refers to the design and operation of the microgrids in order to make them as resilient as possible and as cost-effective as possible.
7. Regulation and Policy Optimization: Involves coming up with regulation systems and policies that promote the use of smart technologies.
8. Consumer Engagement Optimization: Concerns itself with creating interfaces and systems that result in consumer participation.
9. Smart Technology and Renewable Energy Optimization: This is done through a combination of sensors, IoT, and smart meters, which can optimize the efficiency of monitoring and controlling the flow of energy, and the combination of renewable sources that can optimize their usage and reduce their use of fossil fuels.

In this section, we present Smart Energy Algorithms as a general term describing a category of computation, which is aimed at improving the efficiency, sustainability, and reliability of energy systems. This includes smart grid management and optimization, construction energy efficiency, demand-side management, and integration of electric vehicles, and optimization of energy production and consumption. Smart energy Smart Grids Algorithms are a specific type of Smart Energy Algorithms, focusing on the control, monitoring, and optimization of the electricity grid to provide higher resilience, quality, and real-time flexibility of the service.

Some of the classical algorithms that can be applied to the smart energy issues with the aim of enhancing the Quality of Energy Service are:

- The Linear Programming is able to optimize the distribution of energy, load, and identifying the optimal result in a model that has linear relationships: it is essential to make sure that the smart energy systems are available and perform optimally.
- Dijkstra and the Algorithm are generally applied in power networks in order to determine the shortest and optimum power transmission path between nodes, which is crucial in terms of reliability and speed of power supply.
- The Convex Optimization is especially applicable in the case where the solution space is convex so that whatever local minimum is achieved is also a global minimum. This is required to stabilize frequency and voltage in distribution networks.
- Simplex Algorithms are applied to a problem of linear programming in the organization of power systems planning and functioning, where the resources are provided as efficient and optimized.
- The NewtonRaphson Method is a method that solves non-linear equations and is often used in the state estimation of power systems, which is some of the most important in ensuring quality and reliability of supply.
- Greedy Algorithms obtain efficient and sub-optimal solutions selecting an optimal local solution at each step which can be used in the allocation of resources and responding to real-time demand changes.
- Tabu Search Algorithm(g): Tabu search algorithm is a way of local search of solutions in solution space to identify optimal or near-optimal solutions, significant in continuous improvement in the energy grid management.
- Monte Carlo Methods are used to find the solution of complicated problems by using random variables, which can be used in predicting system failures and assessing plan mitigation of risk.

- Finite Element Method finishes the partial differential equations in the field of engineering, among which is the problem of power transmission and distribution, which is required to design and maintain smart energy infrastructure.

The Nature-Inspired Algorithms can be particularly effective when adopted in the context of smart energy systems to enhance the efficiency and effectiveness of managing the grid and optimizing the resources. The next section provides a description of the manner in which each of the types of these algorithms can be applied to the environment of smart energy systems.

Natural algorithms [6] and their implementation in smart energy systems in particular, can be useful to enhance efficiency and effectiveness of grid management and resource optimisation. The following summarizes the application of each of these types of algorithms to the smart energy systems:

Swarm algorithms: Swarm algorithms are applicable in optimisation of power transmission routes, load distribution during real time. The principal ones are the Ant Colony Optimisa- tion (ACO) algorithm to identify the most efficient paths through the power distribution network and reduce losses and enhance demand response (see the Algorithm 1 (pseudocode)) and the Firefly algorithm which solves multi-objective optimisation problems in grid management, including cost minimisation and energy efficiency maximisation.

Algorithm. # Pseudocode of an Ant Colony Optimisation (ACO) method for Smart Grid Dispatch Optimisation
Input:
Number_of_ants, Number_of_iterations
Graph representing decision paths (e.g., energy dispatch options) Pheromone_initial, Evaporation_rate, α (alpha), β (beta)
Energy demand profile, generation/storage capabilities and constraints
Initialise pheromone trails τ[i][j] on all graph edges with Pheromone initial
For iteration = 1 to Number_of_iterations do:
For each ant k = 1 to Number_of_ants do:
Initialise empty path P[k]
For each decision step t in the energy dispatch sequence:
For each possible move (i → j), calculate transition probability:
prob[i][j] =i][j] = 1/cost[i][j] (heuristic: inverse of local cost)
Select next decision j using probabilistic rule Append j to path P[k]
Evaluate the fitness (objective function) of the complete path P[k]:
Cost of energy dispatch;
Renewable energy usage;
Storage efficiency;
Penalties for constraint violations.
1. End For (ants)
2. Update pheromone trails for all edges:
(a) Apply evaporation:
(b)τ[i][j] = (1 − Evaporation_rate)•τ[i][i]
For each ant k that found a good solution: For each edge (i → j) in path P[k]:
τ[i][j] = τ[i][j] + Δτ[k][i][j]
1. End For (iterations)
2. Output the best solution path (lowest cost, constraint-compliant dispatch)

Flexible encoding;	Genetic Algorithm (GA)	Metaheuristic High computational cost; risk of multi-modal problems	M	L	M
NSGA-II	Multi-objective GA	Generates Pareto-optimal front; Computationally intensive; parameter tuning	L	L	M
Particle Swarm	Swarm Optimisation	Fast convergence; few Can get trapped in local optima;	H	L	H
Optimisation (PSO) Control Parameters Sensitive To Parameter Selection					
Tabu Search	Local Escapes local optima via simple implementation	Needs tabu list management; parameter settings affect performance	M	M	M

Convex Programming	Convex optimisation	Guarantees global optimum;	Applicable only to convexified	H	H	H
efficient solvers available	models; may require	overfitting; handles mixed		limited interpretability		
problem reformulation	Machine learning	data types	Kernel choice complexity; scales poorly with very large datasets Complex hyperparameter	M	L	M
XGBoost	Machine learning	built-in regularisation; handles missing data	tuning; risk of overfitting if misconfigured	H	H	H
Long Short-Term Memory (LSTM)	Deep learning	Captures long-term temporal dependencies; effective for sequence data Enhances accuracy by combining	Data intensive; long training times; harder to interpret Increased complexity; risk of	M	L	M
Stacking ensemble	Ensemble learning	multiple base learners; reduces individual model biases	overfitting; challenging to interpret	H	M	H

Conclusions

The principal field of the proposed research is to carry out a review of intelligent energy systems. The present research takes into account the primary methods that are identified in the literature and connected to the issue under treatment. Thus, the bioinspired and refined algorithms include Dragonfly Algorithm, Multi-objective Genetic Algorithm, Random search algorithm based on natural selection, Backpropagation ANN, Genetic Algorithm based on chromosomes, Long Short-Term Memory Model, Hybrid Modified Grey Wolf Optimisation Sine Cosine Algorithm Crow search Algorithm (MGWOSCACSA), Non-dominated Sorting Genetic Algorithm (NSGA-II), Hybrid Multi-Objective Genetic Algorithms (HMOGA), Elitist Non- dominated Sorting

Genetic Algorithms II, According to the already mentioned work, the key arguments that have been identified by the review are as follows:

The systematic use of various methods to use energy efficiently, monitoring and processing data produced by sensors have been introduced in systematic manner to understand the advancements in the field; alongside those other contributions have also been brought to the fore.

It has been made possible by the use of sensors and/or connection via wireless networks, which have increased the standards of processing the data and incorporating it to various control algorithms and decisions when managing the energy.

Data processing techniques enable us to visualise and present techniques to help reduce or eliminate human errors and enhance decision-making. Rigorously speaking, all techniques perform better than classical techniques.

- Real environments make use of optimisation techniques and by the fact they work they can be used in real time provided that there is time available to change the system variables i.e. time the system requires to respond.
- One optimisation method cannot be adopted. Hybrid methods can be used to enhance the location of a solution to energy saving.
- The optimisation algorithms could be optimised by increasing the number of sensors and the time taken to monitor the results of electrical variables of the system.

No implementation of IoT techniques is used due to the fact there is no reference to the performance of the algorithms being enhanced by the use of such devices; however, the use of the IoT monitoring devices implies the use of cloud computing.

The current work can aid researchers to get inspired in the sphere of smart energy.

Discussion and Future Trends

The future of smart energy systems is going to be typified by increasingly smart machine learning and optimisation algorithms in collaboration with built in ensem- ble models. In the near future, predictive analytics will get more accurate as models like the Random Forest, XGBoost and LSTM will be trained on richer, more diverse data feeds; then, the individual strengths of each model will be used to reduce forecast error and enhance resilience in a diverse variety of operating conditions. At the same time, classical approaches to optimisation will be incorporated into real-time management systems, which will allow the automatic regulation of demand and supply, the dynamic distribution of loads and the quick elimination of failures, which will contribute to the stability of the grid and its tolerance to disruptions, including blackouts. On the infrastructural side, scalability will be essential: algorithms should be able to scale with the increasing volumes of data and with complex topologies of networks, and highly developed optimisation of battery storage will eliminate intermittency of renewables and ensure uninterrupted supply. The extensive implementation will be dependent on favorable regulatory policies and affordable investment plans, which will make the growth sustainable, without being cost-prohibitive. In the end, there is a positive side of more reliable service offered to consumers, as well as flexible pricing and increased transparency, as long as policymakers and industry stakeholders can cooperate to overcome the obstacles of infrastructure modernisation and long-term planning.

References

1. Ibrahim, H.; Aburukba, R.O.; El-Fakih, K. An Integer Linear Programming model and Adaptive Genetic Algorithm approach to minimize energy consumption of Cloud computing data centers. Comput. Electr. Eng. 2018, 67, 551–565. [CrossRef]
2. Fang, X.; Misra, S.; Xue, G.; Yang, D. Smart grid—The new and improved power grid: A

survey. IEEE Commun. Surv. Tutor. 2012, 14, 944–980. [CrossRef]
3. Lorencin, I.; And¯elic´, N.; Mrzljak, V.; Car, Z. Genetic Algorithm approach to design of multi-layer perceptron for combined cycle power plant electrical power output estimation. Energies 2019, 12, 4352. [CrossRef]
4. Dey, B.; García Márquez, F.P.; Basak, S.K. Smart energy management of residential microgrid system by a novel hybrid MG- WOSCACSA algorithm. Energies 2020, 13, 3500. [CrossRef]
5. Abdul Latif, S.N.; Shi, J.; Salman, H.A.; Tang, Y. Optimization of demand-response-based intelligent home energy management system with binary backtracking search algorithm. Information 2020, 11, 395. [CrossRef]
6. Sikorska, J.Z.; Hodkiewicz, M.; Ma, L. Prognostic modelling options for remaining useful life estimation by industry. Mech. Syst. Signal Process. 2011, 25, 1803–1836. [CrossRef]
7. Lei, Y.; Li, N.; Guo, L.; Li, N.; Yan, T.; Lin, J. Machinery health prognostics: A systematic review from data acquisition to RUL prediction. Mech. Syst. Signal Process. 2016, 104, 799–834. [CrossRef]
8. Strielkowski, W.; Vlasov, A.; Selivanov, K.; Muraviev, K.; Shakhnov, V. Prospects and challenges of the machine learning and data-driven methods for the predictive analysis of power systems: A review. Energies 2023, 16, 4025. [CrossRef]
9. Jiang, X.; Xiao, C. Household energy demand management strategy based on operating power by genetic algorithm. IEEE Access 2019, 7, 96414–96423. [CrossRef]
10. Kampelis, N.; Sifakis, N.; Kolokotsa, D.; Gobakis, K.; Kalaitzakis, K.; Isidori, D.; Cristalli, C. HVAC optimization genetic algorithm for industrial near-zero-energy building demand response. Energies 2019, 12, 2177. [CrossRef]
11. Salata, F.; Ciancio, V.; Dell'Olmo, J.; Golasi, I.; Palusci, O.; Coppi, M. Effects of local conditions on the multi-variable and multi-objective energy optimization of residential buildings using genetic algorithms. Appl. Energy 2020, 260, 114289. [CrossRef]
12. Foltyn, L.; Vysocký, J.; Prettico, G.; Beˇloch, M.; Praks, P.; Fulli, G. OPF solution for a real Czech urban meshed distribution network using a genetic algorithm. Sustain. Energy Grids Netw. 2021, 26, 100437. [CrossRef]