IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 27 July 2023, accepted 30 August 2023, date of publication 6 September 2023, date of current version 15 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3312552

RESEARCH ARTICLE

A QoS-Aware Data Aggregation Strategy for Resource Constrained IoT-Enabled AMI Network in Smart Grid

ASFANDYAR KHAN[1], SYED HAMAD SHIRAZI[1], MUHAMMAD ADEEL[2], MUHAMMAD ASSAM[3], YAZEED YASIN GHADI[4], HEBA G. MOHAMED[5], AND YONG XIE[6]

[1]Department of Computer Science and Information Technology, Hazara University Mansehra, Mansehra 21120, Pakistan
[2]The Hong Kong Polytechnic University, Hong Kong
[3]Department of Software Engineering, University of Science and Technology at Bannu, Bannu, Khyber Pakhtunkhwa, Pakistan
[4]Department of Computer Science, Al Ain University, Al Ain, United Arab Emirates
[5]Department of Electrical Engineering, College of Engineering, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia
[6]Department of Computer Technology and Applications, Qinghai University, Xining, Qinghai 810016, China

Corresponding authors: Syed Hamad Shirazi (syedhamad@hu.edu.pk) and Heba G. Mohamed (hegmohamed@pnu.edu.sa)

This work was supported by Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia, through the Princess Nourah bint Abdulrahman University Researchers Supporting Project under Grant PNURSP2023TR140.

ABSTRACT Emerging Internet of Things (IoT) technologies and applications have enabled the Smart Grid Utility control center to connect, monitor, control, and exchange data between the smart appliances, smart meters (SMs), data concentrators (DCs) and control center server (CCS) over the Internet. In particular, DC receives different Advanced Metering Infrastructure (AMI) applications data from multiple SMs for processing, queuing, aggregation, and forwarding onward towards the CCS over the things networking. However, DCs are expensive component of the AMI network. Recently, SMs are used as relay-devices to accomplish a cost-effective AMI network infrastructure to avoid the DC placement and bottleneck problem. However, SMs are recourse constrained (limited CPU, RAM, storage, and network capacity) intelligent devices which faces numerous communication challenges during outage conditions and summer peak hours where bulk amount of data with different traffic rates and latency are exchanged with the Utility control center. Therefore, an efficient data aggregation is required at relay-devices to deal with high volume of data exchange rates in order to optimize the constrained-resources of the AMI network. In this article, we propose a hybrid data aggregation strategy implemented on an aggregator-head (AH) in the clustering topology which performs data aggregation on the Interval Meter Reading (IMR) application data. AH induction greatly reduces the workload of the cluster-heads (CHs), and efficiently utilizes the constrained-resource of AMI devices in a cost effective-manner. The proposed strategy is evaluated for different existing approaches using the CloudSim simulation tool. Experimental and simulation results are obtained and compared which show the effectiveness of the proposed strategy such that limited resources are optimized, CH workload is minimized, and QoS of AMI applications are maintained.

INDEX TERMS Advanced metering infrastructure, data concentrator, data aggregation, interval meter reading, Internet of Things, quality of service, RESTful APIs, smart grid, smart meter.

I. INTRODUCTION

In recent years, advancement in Information Communication Technology (ICT) along with intelligent services has revolutionized and enabled the Internet to transmit approximately 2.5 quintillion bytes of data per day among different networked environments such as smart homes, smart city, and Smart Grid. IoT [1], [2], [3] has become an essential computing technology of Internet which enables everyday life things (smart devices) to remotely collect

The associate editor coordinating the review of this manuscript and approving it for publication was Ayaz Ahmad.

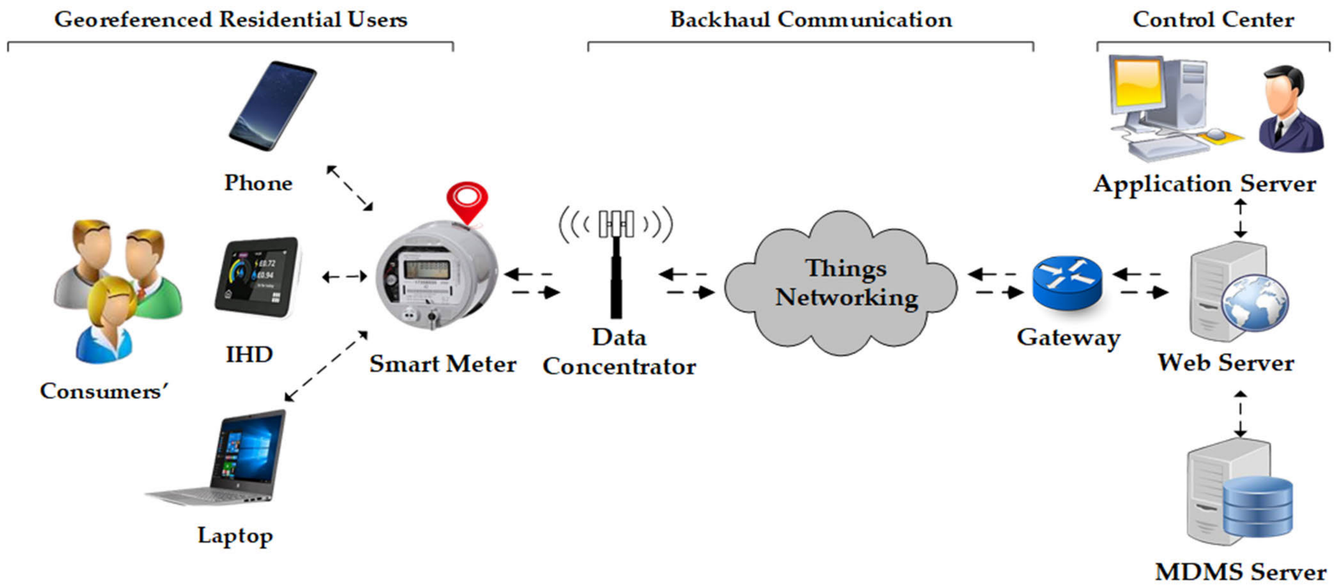

FIGURE 1. Simplified IoT-enabled AMI network in smart grid.

and transfer data automatically using various communication technologies without any human intervention. For example, prominent IoT applications [4], [5] include smart homes, smart healthcare, smart education, and Smart Grid etc. Similarly, IoT provides useful applications in Smart Grid infrastructure [6], [7] such as smart metering, electric vehicles (EVs) charging, battery, and solar farm monitoring, remote control, energy theft, billing, pricing rates, and much more.

In modern Smart Grid, AMI [8], [9] is an essential network part which enables two-way information flow between consumers and shared pool of power Utility resources e.g. computing servers, memory, storage, network applications, and vice-versa. Currently, various useful and innovative AMI applications are assisting Smart Grid to have a wide range of services and better remote control in contrast to traditional electric Grid system. Among these AMI applications, Interval Meter Reading (IMR) enables each household SM to transfer the electricity consumption readings of few kilobytes at every fixed time interval usually 15 to 60 minutes in a day, night, and peak hours in a season to grid operators. These electricity readings are transferred via the backhaul network to the Utility CCS for further processing and storage in the Metering Data Management System (MDMS) [10] which runs a database application to maintain a master database to store data of all SMs according to their location, residential area, and region in a city as depicted in Figure 1. Typically, a SM supports multiple communication technologies such as cellular (3G/4G/5G), WiFi, Zigbee, Wi-SUN, and Bluetooth etc to connect consumers via phone, in-home display (IHD) screen and laptop to the Utility provider systems. The support for these communication technologies is mainly based on the specific SM model and the requirements of a particular residential region in the Smart Grid infrastructure.

These electricity consumptions provide a clear picture to the Utility control center for tracking power supply, usage, demand [11] while the consumers receives real-time accurate information about their billing and power pricing via text messages, emails, and IHD screen.

In densely populated residential areas (urban cities) [12], [13], [14] where thousands to millions of SMs are installed and deployed in consumer premise e.g., homes, offices, buildings, and industrial sites. These SMs generate enormous amount of traffic in short time towards intermediate devices (here, DCs) which forwards onward through the backhaul communication to MDMS server located at the Utility control center. In such scenarios, the existing AMI network with constrained network resources especially SMs and DCs with limited processing, memory and bandwidth (BW) capacity make it difficult to handle and control the enormous data collection and sharing with the MDMS server of Utility control center within a short span of time (recommended latency). One solution to this problem is to increase the number of DC devices in the AMI network. However, these AMI devices are too expensive [15] and increase the overall AMI infrastructure cost. Moreover, to eliminate the need of redundant DCs and reduce the network cost, SMs are utilized to perform as relay-devices at the neighborhood area network (NAN) level in the AMI network. Since, SMs are resource constrained [16], [17] in nature, data and traffic handling is a critical challenge to be solved in AMI networks.

This study extends our past works [18], [19] in order to minimize the data traffic workload and optimize the limited resources of cluster-heads (CHs) in the IoT-enabled AMI network. We propose an efficient hybrid data aggregation strategy employed at the aggregator-head (AH) such that the data aggregation operation of IMR application data is shifted from CH to AH in each cluster in the NAN topology. The AH receives periodic electricity consumption (IMR application) data from the cluster-members, perform data aggregation using both combining and manipulating method and then transmit to the corresponding CH. The CH forwards the aggregated data towards the MDMS server via the DC using IoT communication technologies. The primary purpose of the proposed work in this article is to reduce the workload at the CH level via data aggregation such that congestion and queue contention is eliminated, desired QoS [20] (e.g., latency, throughput, and data's priority) is maintained in the AMI network.

In the context of AMI network, following are the significant contributions of our proposed research work in this article:

1. We propose a hybrid data aggregation strategy employed at the AH in clustering topology by developing an efficient algorithm to aggregate the consumption data of SMs.
2. The functionality of each CH is limited to only traffic classification, queueing, and relaying AMI applications data between DC and CCS.
3. We formulate the optimization problem by mathematical modelling an objective function to optimize the constrained resources of CHs.
4. Finally, we evaluate and validate the performance accuracy of our proposed strategy through experimental and CloudSim simulations. The obtained results show the effectiveness of the proposed strategy which successfully optimize the limited resources (CPU processing, memory, and BW) while reducing the traffic load with QoS guarantee as compared to existing strategies.

The remainder article is organized as follows. In Section II, we review related research works by focusing on existing data aggregation strategies in AMI network. Section III briefly presents the problem definition, formulation of objective function, and give insight details about the proposed QoS-aware hybrid data aggregation strategy. Section IV describes the design, performance evaluation metrics, and analyze both the experimental and simulation results. Lastly, we present our concluding remarks and offer some useful future directions for the proposed research area in Section V.

II. RELATED WORK

In recent past, IoT-enabled AMI network in Smart Grid has got attention of the research community to revamp its performance and make it convenient in practice. AMI network faces numerous communication issues such as network congestion and queuing delays due to limited network capacity in backhaul links and higher arrival rate of traffics at the intermediate devices. Therefore, managing high volume of traffics and network congestion in AMI networks have been extensively studied in literature. One of the solution to these problems is data aggregation carried out via centralized, cluster-based, peer-to-peer, and tree-based techniques. Hence, data aggregation has been chosen as a modest tactic in several research articles in the field of AMI network as follows.

In our recent research works [18], [19] we emphasized on AMI applications traffic handling with QoS provisioning in IoT-enabled Smart Grid network coupled with cloud computing. In [18], SMs (relay-nodes) are clustered using modified K-Means algorithm to extend the AMI network coverage such that to eliminate the DC hotspot problem and reduce additional cost required in the communication network topology. Similarly, a hybrid queue scheduling (HQS) scheme is proposed in [19] for AMI applications traffic in Smart Grid network. Further, AMI traffic are classified and scheduled using the priority metrics of these traffics in order to lower the cost of cloud service and ensure QoS in the Smart Grid network. However, both works lack to incorporate the data aggregation method in order to investigate its impact on the overall AMI network traffic load and resources utilization in the clustering topology.

The authors in [21] proposed a reliable AMI network planning solution based on machine learning for residential grids (urban, rural, and sub-urban) areas. Their work intend to optimize the data aggregator point (DAP) placement problem employing K-Medoid clustering to assist the transfer of metering data (e.g., voltage profile and power quality) between the power Utilities and the consumers. Results show that proposed clustering topology is deployed to ensure appropriate coverage, network device connectivity, and cost of network topology is minimized for the NAN zones.

In [22], authors proposed an effective method of smart metering data aggregation at the concentrator device (data relay point) in the Smart Grid network topology. The basic aim was to reduce message transfer size and processing time at the server end in the network. Simulation results obtained through ns-3 show that message volume and server utilization is reduced due to data aggregation in the Smart Grid network.

The research work in [23], proposed an autonomous and distributed electricity usage data collection mechanism based on clustering scheme in order to transfer the aggregated electricity usage data of consumers to the Utility control center. However, as SMs (relay nodes) have limited resources capabilities due to which time consuming operations are shifted from online to offline state. The proposed work ensures data privacy and integrity of consumers aggregated consumption data such that the communication overhead is reduced. However, workload of each CH due to routing and data aggregation is so heavy i.e. requires traffic handling and resource optimization.

In [24], authors proposed a secure routing and data aggregation (SRDA) for wireless smart metering networks. The wireless network is managed via domains each having a domain controller (here, SM) while intra and inter domain proxies are selected by the controller to perform data aggregation. The proposed work achieve scalability, efficient energy usages, and provide relative security. However, the trusted third party (TTP) node and controller requirement in each domain makes the network architecture complex.

To reduce the number of DAPs installation required in rural and suburban areas (neighborhood network), authors proposed [25] a grid-based scheme using corresponding algorithms to minimize the needs of DAPs and eliminate its impact on relay location such that communication quality is improved in the wireless network. Similarly, the authors in [26] investigated the DAP point placement problem in terms of guaranteeing real-time communication and shortening latency in the smart metering networks. Their work uses clustering DAP placement (CDP) approach using K-Means method to solve the network partitioning problem such that maximum propagation latency is reduced.

The authors in [27] proposed a novel multi-level data collection trees approach via data aggregation policy at the DC units (DCUs) i.e. forwarding nodes to handle high volume of AMI traffic such that network congestion and delays are eliminated in the AMI network. Further, it guarantee QoS of Smart Grid applications.

In [28], authors proposed a compression algorithm that compresses excessive volume of metering data more efficiently which causes congestion and exchange this data with MDMS server where it is decompressed in the IoT-based AMI network. However, in this approach decompression increases the processing time at the server end. The authors in [29] proposed a secure multidimensional data aggregation based on Horner rule for multidimensional data. Further, certificate-based aggregate short signature is used to protect data leakages and ensure privacy of consumers to overcome differential attack problem. The proposed approach lowers the communication overhead and computational cost.

In [30], a power consumption profile model is proposed for consumers using two level clustering in heterogeneous grid topology. The local power consumption profiles are drawn at first level while the global power consumption profiles are derived at second level which provides efficient pricing prediction models and helps to identify power consumption outliers. The proposed model reduces communication and computational complexity with better accuracy. A genetic algorithm based routing approach is proposed in [31] for efficient routing of alarm messages for fault and/or outage detection and localization in the monitoring electric grid. The proposed data aggregation method uses both combined and manipulating method for alarm messages at the sink level. The proposed method enhances to receive redundant messages from sensors and reduces network congestion at the sink. The work proposed in [32] investigate the limited capabilities of electric meters and network as AMI architectural problem. The authors proposed a hierarchical multi-tiered agent based AMI design using fog computing in order to achieve scalability and real-time performance for periodic and on-demand metering data in the distributed Smart Grid. The proposed three-tier architecture accommodates Local Meter Concentrator (LMC) fog node at tier-1, Transformer Station Concentrator (TSC) fog node at tier-2 that has area specific functionality, while the fog node at tier-3 named as Metering Data Collection (MDC). The performance analysis validates the data acquisition, availability, and communication in one year pilot project. Data aggregation is one of the approaches to preserve the privacy of electricity consumer's data. A comprehensive review has been presented in [33] based on different aggregation schemes using different cryptographic techniques to preserver privacy of consumptions data in AMI network. The survey highlighted open research issues of privacy challenges in Smart Grid.

The authors in [34] addressed the constraints of routing and AMI traffic flow demands at the network layer for smart metering in wireless heterogeneous networks (WHNs). They used the column generation approach to provide scalable solution to the optimization problem called capacitated multicommodity flow for AMI (CMCF-AMI) such that more SMs are covered, infrastructure cost of NAN is reduced, and capacity constraint of short range wireless technology is solved via multi-hop fashion in the wireless mesh network. In the multi-hop routing tree, the SMs data is aggregated at the DAPs which are connected via base station to communicate with the Utility using cellular network.

In [35], the authors proposed a heuristic model based on evolved network architecture that consider various aspects of WHN. The proposed optimization model utilizes geo-reference model consist of elements such as base station, group of residential SMs, and universal DAPs connected via LTE cellular network to the Utility provider systems. Results show that the optimization model achieves the target SMs coverage considering the transmission range and network capacity of the communication technology and minimizes the cost of resources by efficient resource utilization in the WHN.

The authors in [36] have proposed a novel algorithm for smart metering data of electricity, gas, and water meters at specific time interval exchanged with the Utility provider. The proposed method optimizes the resource efficiency in the cognitive mobile virtual network operator (C-MVNO) through opportunistic channel allocation and/or secondary channels and virtualization for smart metering in order to ensure coverage with lower maintenance and deployment cost.

The authors in [37] proposed a new approach based on network partitioning technique to tackle the multi-controller placement problem in software-defined network (SDN). An optimized K-Means method is applied to partition the network into sub-networks. As compared to regular K-Means, the proposed optimized technique significantly reduces the response latency between switches and centroids. Further, objectives like load balancing, reliability, and energy saving can be achieved which greatly minimize complexity in the SDN.

In [38], authors proposed a smart energy meter based on LoRaWAN technology (SEM-LoRaWAN) to collect the energy consumption of a photovoltaic (PV) system and transmit the real-time data to IoT-enabled consumer/Utility for billing and monitoring. The proposed technology monitors related parameters such as current, voltage, energy, temperature, light intensity, and humidity of PV system and exchange these information via LoRaWAN gateway to the cloud-enabled IoT server which can be accessed via web and mobile applications. Results show the efficacy of the proposed system in terms of accurate energy consumption measurement and monitoring the environmental conditions of PV system in real-time.

The authors in [39] proposed an IoT and SM data processing framework based on Edge-Fog-Cloud computing environment in order to extract meaningful patterns from metering data to monitor and control the SMs, IoT appliances, and develop applications for consumers, prosumers, aggregators, retailors, and grid operators. The computational and processing applications are distributed among the Edge, Fog, and Cloud layers such that the communication latency, response time, resource utilization, and load are minimized. The data is exchanged via MQTT (message queuing telemetry transport) messages between the edge nodes (DC node) and cloud nodes. Simulation results obtained show better performance as compared to other methods.

In the aforementioned discussion, various network models such as clustering ([18], [19], [21], [23], [26], [30], [31], [37]), star ([22], [25], [36], [38]), tree-based ([27], [28], [34]), and hierarchical ([24], [29], [32], [39]) were proposed for AMI network design problem using different data aggregation and data collection techniques as needed for different AMI applications data. In particular, cluster-based network topology uses CH (here, SM and/or DCs) for data collection, processing, data aggregation, and relaying purposes in the AMI network. As these relay-devices are resource constrained and there is a need of an efficient data aggregation strategy to minimize their workload, optimize limited resources (CPU, RAM, and BW etc), avoid network congestion that leads to hotspot, and ensure QoS-provisioning for latency bounded AMI applications data in the IoT-enabled AMI network. Table 1 presents insight review of recent research works with intended objectives regarding the network topology for AMI design, aggregation approaches for different AMI applications data which are closely related to our proposed data aggregation strategy for intended IMR application data in this article.

III. PROBLEM FORMULATION AND METHODOLOGY

Existing schemes in literature have employed different clustering architecture ([18], [19], [21], [23], [26], [30], [31]) to group SMs and DCs into clusters with designated CH (SM) which performs similar functionality of allocating resources to cluster members as DC do in a NAN topology as shown in Figure 2. However, SM as relay-device has less resources (4)-12 KB RAM, 16 MHz CPU, 64 KB-1MB Flash memory, and leased network capacity) [40], [41]. For example in certain AMI scenarios, AMI applications (SMs) generate heavy traffic rates towards corresponding CHs. If traffic arrival rates exceeds (i.e. more data arrives than CH can handle) at the output capacity link (i.e. BW), data will wait longer in queues for their turn or being dropped if no free buffers (RAM) available. Similarly, longer waiting times (i.e. queueing delay) at output link increases queue contention and congestion level at CHs leads to hotspot(s). Moreover, end-to-end latency of AMI applications traffic must be fulfilled to maintain QoS requirements in the AMI network. It implies that reducing average nodal delay (CH delay) during data transmission is a critical parameter in order to achieve fairness, QoS, and throughput in delay-intensive AMI network. Mathematically the nodal delay (D_{CH}) can be expressed as below.

$$D_{CH} = D_{Queue} + D_{proc} + D_{agg} * N_{SM} + D_{tran} + D_{prop} * h_{DC} \tag{1}$$

where, CH delay is denoted as D_{CH} that mainly consists of queueing delay (D_{Queue}), processing delay (D_{proc}), aggergation delay (D_{agg}), transmassion delay (D_{tran}), and propagation delay (D_{prop}) which occurs for one hop (h_{DC}) transmission towards the DC. Specifically, D_{agg} denotes the aggregation delay for each CH to perform data aggregation on the IMR application data received from the total cluster-members (N_{SM}) which heavily depends on the employed aggregation technique. Among these, D_{Queue} is a direct function of network congestion which occurs in M/M/1 model due to low capacity of the backhaul network and can be computed in Eq. (2).

$$D_{Queue} = \frac{PktSize}{BW} * N \tag{2}$$

where, *PktSize* is the packet size in bytes, *BW* is the allocated *BW* in bps, and *N* represents the average packets queued-in and can be expressed as follows:

$$N = \frac{\rho}{1-\rho} \tag{3}$$

where, ρ is the probability that the server is working i.e. server utilization of queuing system and can be expressed in Eq. (4).

$$\rho = \frac{\lambda_{CH}}{\mu_{CH}} \tag{4}$$

where, λ_{CH} denotes the packet arrival rate (Poisson distribution) and μ_{CH} represents the service rate (Exponential distribution) at the CH. In the context of AMI networks, cloud-based control centers plays a crucial role and provides insights about the important optimization system parameters like λ_{CH} and μ_{CH} which determines the overall behavior (efficiency and stability) of a queueing system in terms of queue length and average waiting time. For example, if ($\rho < 1$), then the server capacity utilization is low (underutilized), when ($\rho > 1$), then the server utilization is high

TABLE 1. Summary: comparison of state of the art studies against the proposed work for different AMI applications.

Scheme (s)	Publish Year	Network Topology	Data Aggregation (Level)	AMI Application (s)	Objective (s)
[18]	2021	Clustering	n/a	AMI	- Maximizes SMs coverage - Lowers DCs cost and achieves QoS
[19]	2022	Clustering	n/a	AMI	- Cloud services cost minimization - Improves QoS
[21]	2021	Clustering	n/a	Quality and Voltage profile	- Extends coverage - Increases connectivity
[22]	2015	Star	DC	Electricity consumption	- Reduces messages volume and server processing time
[23]	2016	Clustering	CH	Electricity usage data	- Ensures privacy and integrity of aggregated usage metering data
[24]	2016	Hierarchical	Inter/Intra Domain Proxy	Electricity usage data	- Provide scalability, energy efficiency and security
[25]	2022	Star	DAP	Metering data	- Minimize number of DAPs
[26]	2017	Clustering	DAP	AMI	- Reduces latency
[27]	2018	Tree	DCU	Smart grid	- Reduce congestion and delays
[28]	2021	Tree	SM	Metering data	- Reduce congestion and improves QoS
[29]	2018	Hierarchical	Gateway	Consumer data	- Minimize communication overhead and computational cost
[30]	2017	Clustering	DC	Electricity usage data	- Reduce communication and computational complexity
[31]	2018	Clustering	Sink	Alarm message	- Reduces network congestion
[32]	2022	Hierarchical	LMC	Periodic & on-demand	- Improves scalability and performance
[33]	2019	Multiple	n/a	Consumption data	- Ensures privacy
[34]	2019	Tree	DAP	AMI	- Increase Coverage, reduce cost, and minimize calculation time
[35]	2017	Multiple	DAP	Consumption & VoIP	- Achieves target coverage and reduce number of DAPs
[36]	2023	Star	DC	Consumption data	- Achieves coverage with lower cost
[37]	2016	Clustering	Controller	Multiple	- Reduces maximum latency
[38]	2022	Star	IoT server	Energy consumption	- Accurate billing and real-time monitoring
[39]	2023	Hierarchical	DC	Meter readings	- Minimize latency, execution time
Proposed work	2023	Clustering	AH	IMR, ODMR & ODMRR	- Resource Optimization - Ensures QoS

(overutilized) and if ($\rho = 1$) indicates that the server is utilized at its full capacity (also called critical point in the queuing theory).

Motivated by the aforementioned discussion, it is imperative to tackle the issues associated with resource constrained CHs in cluster-based NAN topology in AMI networks such that number of transmission towards CHs are reduced, workload is fairly distributed, throughput is increased, constrained resources are optimized, and QoS of diverse AMI applications are ensured. The notations used onward in the article are described in Table 2.

To solve the optimization problem, the main objective is to efficiently utilize and allocate the limited resources of CHs i.e. to find an optimal solution for each cluster in the NAN topology. Therefore, the optimization problem becomes a minimization problem of the requested resources usage at the CH which can be formulated as an objective function in Eq. (5) as follows:

$$\textbf{Minimize} \sum_{k=1}^{K} \sum_{i=1}^{N_{SM}} (CH_{CPU} CCH_{RAM} + CH_{BW})_i^k \tag{5}$$

Subject to the following constraints :

$$\bigcup_{i=1}^{K} C_i = \text{NAN} \tag{5a}$$

$$\sum_{k=1}^{K} \sum_{i=1}^{N_{SM}} SM_{i,j}^k = 1 \tag{5b}$$

$$\sum_{k=1}^{K} \sum_{i=1}^{N_{SM}} SM_{i,j}^k . \lambda_i \leq \mu_j \tag{5c}$$

$$\sum_{k=1}^{K} \sum_{i=1}^{N_{SM}} SM_{i,j}^k . r_{CPU} < CH_{CPU} \tag{5d}$$

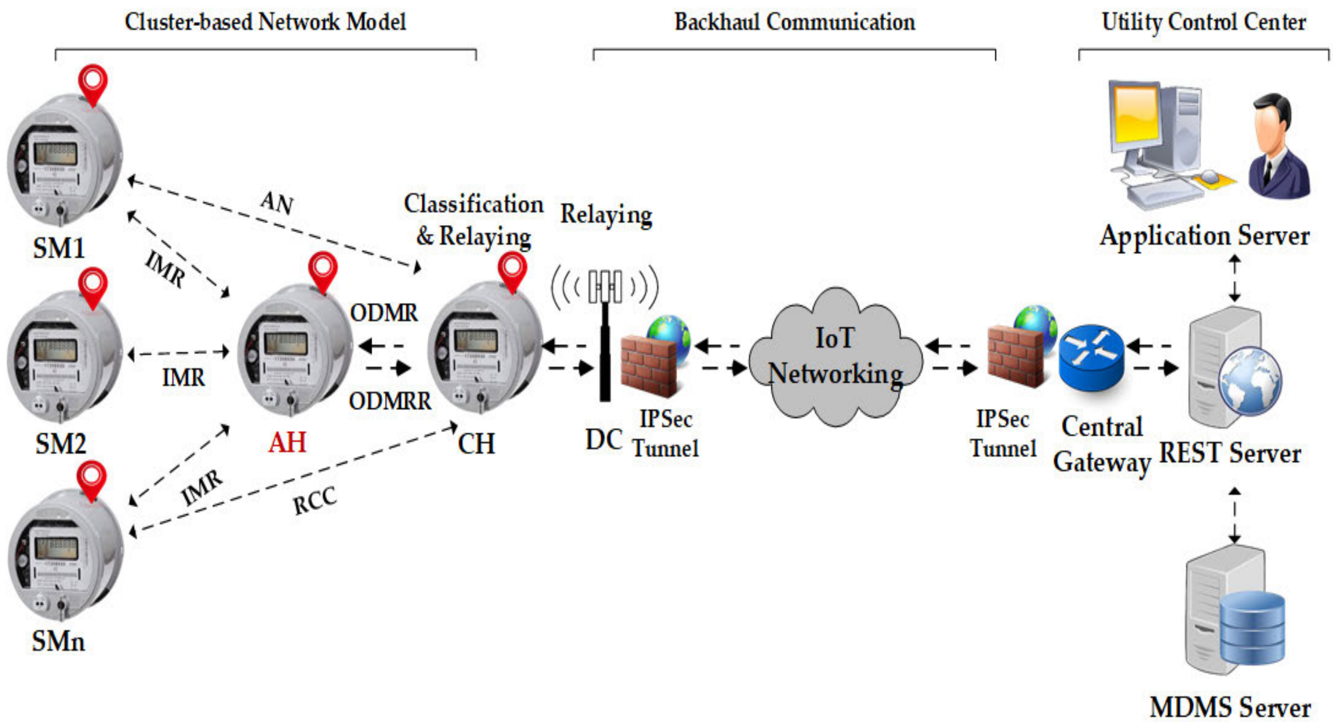

FIGURE 2. Schematic diagram of cluster-based network model [18].

TABLE 2. Notations used in the proposed research methodology.

Notation	Meaning
K	Number of clusters
N_{SM}	Number of SMs in a cluster
CH_{CPU}	CPU utilized by ith SM in *k*th CH
CH_{RAM}	RAM utilized by ith SM in *k*th CH
CH_{BW}	Bandwidth utilized by ith SM in *k*th CH
C_i	*i*th cluster in the NAN, $i=1,2,3,\ldots, K$
$SM_{i,j}^k$	*i*th SM connected to jth CH/AH in *k*th cluster
$\lambda_{CH}/\lambda_{AH}$	Traffic arrival rates at the CH or AH
μ_j/μ_{CH}	Service rate of *j*th CH
r_{CPU}	CPU required by ith SM application
r_{RAM}	RAM required by ith SM application
$PktSize_{i,k}$	Packet generated from ith SM in *k*th cluster
PS/HS	Packet payload size or header size in bytes
po^{CH}	Empty queue probability of CH
h_i	*i*th hour in a day
Kb/s	Kilobits per second
VM_{CPU}	CPU power of a Virtual Machine
MIPS	Millions instruction per second
VM_{RAM}	RAM capacity of a Virtual Machine
VM_{BW}	Network capacity of a Virtual Machine
$Electricity_{Cons_{i,k}}$	Electricity consumption of *i*th SM in *k*th cluster
$\lambda_{Up}/\lambda_{down}$	Outgoing / incoming traffic
λ_{agg}	Aggregated packets received from AH

$$\sum_{k=1}^{K}\sum_{i=1}^{N_{SM}} SM_{i,j}^{k}.r_{RAM} < CH_{RAM} \tag{5e}$$

$$\lambda_{agg} < \lambda_{CH} \tag{5f}$$

The objective function in Eq. (5) intends to optimize the constrained resources of the CH represented as sum of utilized CPU, RAM, and BW that are the attributes of the objective function and are subject to the constraints in Eq. (5a)-(5e). Eq. (5a) indicates that the entire NAN is partitioned into clusters (C_K). Eq. (5b) ensures that cluster members are exclusively connected to one CH and/or AH. Next, Eq. (5c) assures that the average traffic arrival rate λ_i from cluster members is less than the offered service rate μ_j (i.e. network capacity constraint) of the CH. Eq. (5d)-(5e) satisfy that the resource requirements of all AMI applications traffic must be lower than the CH resource capacity i.e. CPU and RAM. Finally, Eq. (5f) indicates that the data aggregation traffic (λ_{agg}) is shifted from CH towards AH which lowers the traffic arrival rates λ_{CH} at the CH, Where, λ_{agg} represents the IMR applications data arrived for data aggregation at the AH that can be expressed in Eq. (6) as:

$$\lambda_{agg} = \sum_{k=1}^{K}\sum_{i=1}^{N_{SM}} (PS+HS)_{i,j}^{k} + \left(\frac{\sum_{k=1}^{K}\sum_{i=1}^{N_{SM}} PS_{i,j}^{k}}{N_{SM}} + HS_i\right) \tag{6}$$

Moreover, $\sum_{k=1}^{K}\sum_{i=1}^{N_{SM}}(PS+HS)_{i,j}^{k}$ presents the total IMR applications packets arrived from the cluster members (N_{SM}) at a fixed time interval towards the AH and $\lambda_{agg} = \frac{\sum_{k=1}^{K}\sum_{i=1}^{N_{SM}} PS_{i,j}^{k}}{N_{SM}} + HS_i$ denotes the aggregated packet using the aggregation technique forwarded from AH to CH. Similarly, λ_{CH} denotes the total traffic arrival from AMI applications generated from cluster members and can be computed as:

$$\lambda_{CH} = \sum_{k=1}^{K} \sum_{i=1}^{N_{SM}} \mu_{CH}(1 - po^{CH}) + \lambda_{agg} \tag{7}$$

whereas, μ_{CH} can be computed as follows:

$$\mu_{CH} = \frac{BW_{CH}}{\sum_{k=1}^{K} \sum_{i=1}^{N_{SM}} PktSize_{i,k}} \tag{8}$$

where, BW_{CH} represents the network capacity of CH and $\sum_{k=1}^{K} \sum_{i=1}^{N_{SM}} PktSize_{i,k}$ are the total arriving packets at the CH.

From Eqs. (5)-(8), we notice that the overall resource utilization of CH must be minimized i.e. considered as an optimization problem while ensuring all the constraints (Eq. (5a)-(5f)). Therefore, we propose a QoS-aware hybrid data aggregation strategy for IMR applications data (energy consumption) to find a best solution for the optimization problem.

Next, we outline the basic network model, traffic classification and the proposed hybrid data aggregation strategy in this article.

A. NETWORK MODEL

We adopt the same AMI architecture from [18] which is a three-tier framework consists of lower, middle, and upper-tier in the AMI communication network in Smart Grid as depicted in Figure 2. The lower-tier (i.e. NAN) consists of georeferenced IoT-enabled SMs which are grouped into K clusters using modified K-Means algorithm based on distance criteria. In each cluster, centroid is designated as CH which performs traffic classification and forwarding between cluster-members and DC. An AH is selected in each cluster alongside CH which perform data aggregation on periodic IMR applications data. We mainly focus in this article to implement an optimal hybrid data aggregation strategy which will extend the work of the existing clustering topology at the NAN level. The middle-tier includes DC and backhaul network. The upper-tier consists of the Utility CCS (application server, web server, and MDMS server) connected via a gateway (central router) using the backhaul network to the residential DC. The communication medium used at the lower-tier is generally Wi-SUN [42], while the DC and gateway generally uses the LoRaWAN [43] as backhaul communication technology for AMI applications traffic. Moreover, cloud computing approach is adopted at the Utility control center coupled with IoT services at the three-tier of the AMI deployment network. We incorporate the following few assumptions in the cluster-based network model:

i) All SMs have the same functionality and resources capacity (transmission range, storage, RAM, CPU, and BW so on).
ii) All communication devices are trustworthy i.e. uses IPSec protocol.
iii) Each SM is pre-programmed to transmit their electricity consumption to AH.
iv) Aggregation of IMR applications data results into one aggregated packet.

B. TRAFFIC CLASSIFICATION

AMI network has a diverse set of AMI applications traffic types in Smart Grid network. These traffic types are classified based on their QoS characteristics (latency, BW, and reliability) as tabulated in Table 3 above.

The AMI applications traffic types (deterministic and event-driven) can be classified into three traffic classes based on their latency requirements as follows:

i) **Periodic:** Traffic generated from IMR application is scheduled at regular time interval (15-to-60 minutes) corresponds to this class. These traffics can tolerate a delay up to a few minutes during transmission.
ii) **On-demand**: These type of traffics are generated on demand both from electricity consumers and Utility control center. On-demand class consists of traffics such as on-demand meter reading (ODMR), on-demand meter reading response (ODMRR), and billing information application.
iii) **Time-critical:** This traffic class includes traffics generated from remote control commands (RCC), power control commands (PCC), electric vehicles (EVs) charging, and alert notifications (AN). These traffics have higher reliability, BW, and tight latency (real-time communication) requirements.

Three AMI applications including IMR, ODMR, and ODMRR are considered in this article. The IoT-enabled SM generally measures (samples) the electricity consumption (IMR applications data) triggered automatically at scheduled time interval (periodic) which depicts the consumer's load profile that may be used in demand-response (DR), load forecasting, and billing info applications. Whereas, ODMR, and ODMRR is a request-response mechanism for exchange of electricity load pattern (e.g., minimum, maximum, average) collected on-demand from residential consumers SM whenever needed. The amount of electricity consumption of the consumer can be calculated as a collective electrical load of the smart appliances and other loads inside the house, building etc and can be mathematically modelled in Eq. (9) as follows:

$$Electricity_{Cons} = \sum_{i\mathcal{D}1}^{N_{SM}} \sum_{a=1}^{n} \sum_{h=1}^{24} kWh_{ah} \tag{9}$$

where, $a = 1, \ldots, n$ denotes the household smart appliances and kWh_{ah} expresses the electricity consumptions in Kilowatt (kW) per hour (h). The time interval (here, h) is set by the Utility provider. Terms like energy and electricity while data and traffic has the same meaning throughout this article.

C. PROPOSED HYBRID DATA AGGREGATION STRATEGY

This section details the proposed QoS-aware hybrid data aggregation strategy configured at each AH of the NAN clustering topology for IMR applications traffic. The main

TABLE 3. Classification of AMI applications traffic [18], [28].

Traffic Type	AMI Aapplication	Packet Size (Bytes)	Latency (Seconds)	Sampling Frequency (Per Day)	Bandwidth (Kb/s)	Reliability (%)
Deterministic	IMR	250-480	< 4 hours	4–6 (Residential) 12–24 (Commercial)	10–100 500 backhaul	99–99.99
Event-Driven	ODMR	50-64	< 15 sec	as needed	10–100	99–99.99
	ODMRR	100	30 sec	as needed	10–100	99–99.99
	Billing Info	100	Sec or min	as needed	9.6–56	99–99.99
	EVs charging	100	< 15 sec	2-4	14–100	99
	RCC	20	1 sec	as needed	10–100	99
	PCC	100	1 sec	as needed	10–100	99
	AN	50	3 sec	as needed	10–100	99–99.99

objective in this article is to limit the functionality of the resource constrained CHs only to processing and relaying AMI applications traffic i.e. CHs will not do any data aggregation while satisfying the QoS requirements of all AMI applications in the AMI network. Each Cluster-member ($\sum_{k=1}^{K} \sum_{i=1}^{N_{SM}} SM_{i.k}$) generates traffic accordingly as tabulated in Table 3. While, Cluster-members are programmed to transfer traffic types belongs to Time-critical class directly to their corresponding CH_k. Whereas, traffic from Periodic class are transmitted towards their assigned AH_k for data aggregation. These traffics arrives over a scheduled time interval $h_i \in \{h_1, h_2, \ldots, h_{24}$ at the corresponding AH_k. Similarly, the Utility CCS generate traffic types from On-demand class exchanged via DC and CH_k towards AH_k which forwards the On-demand traffic to the requested cluster-member and the response is routed back in the corresponding flow (route) to the Utility CCS for further storage in MDMS server and other decision making processes. Finally, our proposed hybrid data aggregation strategy is based on coupling the existing combining and manipulating method performed for aggregating incoming traffics received from cluster-members, as each AH is programmed to do so. In the combining method, AH concatenates the IMR applications traffic into one aggregated packet with a common header send out to corresponding CH which relay it to the Utility control center. This method effectively reduce the traffic size and the aggregated packet includes energy consumptions information of all cluster-members that will be processed at the Utility CCS. For example, in an IP-based AMI communication network, when packet header is removed from an IP packet (i.e. IMR packet) the actual data remains unchanged in the payload (JSON string payload with key attributes as depicted in Figure 4). To ensure that the data integrity is maintained and information is not lost during the header removing and concatenating process e.g., higher level protocol such as TCP (transmission control protocol) employ checksum mechanism to verify the data integrity of the payload data. If a loss or corruption is detected during communication, TCP make request for the retransmission of corrupted or lost segment (packet). Whereas, manipulating method is applied on the On-demand traffic class. Manipulating method includes status information of local area and can effectively visualize the total electricity consumption by calculating (e.g. average, maximum, minimum, and summation) of all cluster-members. Additionally, proposed hybrid data aggregation strategy is based on using a subset of SQL query interface to execute an aggregation function command (request) to a single, multiple, and/or all cluster-members. The requests are efficiently processed at the AH and responses (retrieval) are routed via CHs for data management (retrieval and storage) that will be performed at the Utility MDMS server accessible via backhaul network. Thus, the proposed strategy effectively addresses the optimization problem by reducing the incoming traffic at the CH such that RAM, CPU, and BW are optimized in the AMI network.

One of the potential benefit of the proposed research methodology is that it will provide support to heterogeneous environment (WHN) if the network topology employs different wireless communication technologies (e.g., WiFi, ZigBee, LTE etc) and protocols. In addition, numerous applications traffic will be supported if traffics are classified, queued, prioritized, and scheduled at the CH level in the IoT-enabled AMI network.

The process of the proposed hybrid data aggregation strategy is presented in Algorithm 1 with a time complexity $\boldsymbol{O}(|\mathrm{K}| * |\mathrm{N_{SM}}|)$ to address and solve the optimization problem where K represents the number of clusters (CH) and N_{SM} represents the total deployed SMs in the residential area.

In Algorithm 1, input consists of variable *topo* (network topology), number of clusters (K) in the clustering topology, total number of cluster-members (N_{SM}) in each cluster, generated AMI application traffic with packet size (*PktSize*) according to the traffic characteristics in Table 3, and allocated network capacity (*BW*). Algorithm 1 starts and proceeds in steps as follows: First, necessary variable are initialized in Line 2. Using the two for loops in Lines 3 to 7 ensure that the network topology (*topo*) is generated as considered from [18] such that each cluster-member ($SM_{i,k}$) participate in the clustering topology and have exclusive connectivity to a single DC via dual-head (CH and AH) in order to satisfy con-

Algorithm 1 Process of Hybrid Data Aggregation Strategy

Input: $Topo, K, N_{SM}, PktSize, h, BW$

1 **Start**
2 **Initialize** $i = k = 1$; $Topo = PktSize = \varphi$; $\rho = Total = CH_{BW} = CH_{CPU} = CH_{RAM} = 0$
3 **for** $k \in K$ **do**
4 **for** $i \in N_{SM}$ **do**
5 $Topo \leftarrow SM_{i,k}$
6 **end for**
7 **end for**
8 **for** $k \in K$ **do**
9 **for** $i \in N_{SM}$ **do**
10 **Read** $PktSize_{i,k} \leftarrow SM_{i,k}$
11 **Compute** ρusing **Eq.**(6)
12 **if** $\rho < 1$**then**
13 **if** $PktSize_{i,k}$.AMI application $= =' $ RCC$'$ or $'$PCC or $'$AN$'$or $'$EVsCharging$'$ **then**
14 **if** $SM_{i,k}.Con_{ST} \rightarrow CH_k =' $ Disconnect$'$or !**Receive** $(SM_{i,k}.Con_{RP}) \leftarrow CH_k$**then**
15 **Send** $\left(SM_{i,k}.Con_{RQ}\right) \rightarrow CH_k$
16 **Receive** $(SM_{i,k}.Con_{RP}) \leftarrow CH_k$
17 **Transmit** $SM_{i,k}.PktSize_{i,k} \rightarrow CH_k$ using **Eq.**(7)
18 **Allocate** BW to CH_k using **Eq.**(8)
19 **Transmit** $\mathrm{PktSize_{i,k}.CH_k} \rightarrow CCS$
20 **else if** $PktSize_{i,k}$.AMI application $= =' $ IMR$'$**then**
21 **if** $SM_{i,k}.Con_{ST} \rightarrow AH_k =' $ Disconnec t$'$or ! **Receive** $(SM_{i,k}.Con_{RP}) \leftarrow AH_k$**then**
22 **Send** $(SM_{i,k}.Con_{RQ}) \rightarrow AH_k$
23 **Receive** $(SM_{i,k}.Con_{RP}) \leftarrow AH_k$
24 **Compute** $PktSize_{i,k}.h_i \leftarrow Electricity_{Cons\,i,k}$ using **Eq.**(9)
25 **Transmit** $\mathrm{SM_{i,k}.PktSize_{i,k}} \rightarrow \mathrm{AH_k}$
26 **Process** $\lambda_{\mathbf{agg}} + = \bigcup SM_{i,k}.PktSize_{i,k}$ using **Eq.**(6)
27 **Receive** $CH_k \leftarrow \lambda_{\mathbf{agg}}$
28 **Allocate** BW to CH_k using **Eq.**(8)
29 **Transmit** $CH_k.PktSize_{i,k} \rightarrow CCS$
30 **else if** $PktSize_{i,k}$.AMI application $= =' $ ODM R$'$ or $'$ODM RR$'$or $'$Bill Info$'$ **then**
31 **if** $CH_k.Con_{ST} \rightarrow AH_k =' $ Disconnec t$'$or !**Receive** $(CH_k.Con_{RP}) \leftarrow AH_k$ **then**
32 **Send** $(CH_k.Con_{RQ}) \rightarrow AH_k$
33 **Receive** $(CH_k.Con_{RP}) \leftarrow AH_k$
34 **Manipulate** MIN or MAX or AVGorSUM $(AH_k.PktSize_{i,k}.Electricity_{Cons}) \rightarrow CH_k$
35 **Allocate** BW to CH_k using **Eq.**(8)
36 **Transmit** $CH_k.PktSize_{i,k} \rightarrow CCS$
37 **else**
38 **Discard** $\mathrm{SM_{i,k}.PktSize_{i,k}}$
39 **end if**
40 **else if** $\rho = 1$ *or* $\rho > 1$ **then**
41 **Display** "**Resource utilization is at full capacity or overutilized**"
42 **end if**
43 $\mathrm{CH_{BW}} + = \mathrm{CH_k.BW}; \mathrm{CH_{RAM}} + = \mathrm{CH_k.PktSize_{i,k}}$
44 $CH_{CPU} + =$ CPU instrcutions per$CH_k.PktSize_{i,k}$
45 $Total + = (CH_{CPU} + CH_{RAM} + CH_{BW})_i^k$ using **Eq.** (5)
46 **end for**; **end for**
47 **Return** $Topo,\ PktSize_{i,k}, Total,\ CH_{BW}, CH_{RAM}, CH_{CPU}$
48 **End**

straint (5a) and (5b) of the objective function (Eq. (5)). Next, the AMI application traffic is read from each cluster-member into packet of size ($PktSize_{i,k}$) in Lines 8-10. Lines 11-12 are used to check the CH resource utilization to ensure constraint (5c) and (5f). To ensure the QoS of AMI traffic, the generated traffic type is checked (i.e. classified), if it belongs to Time-critical class, then an exclusive TCP/IP connection is established between cluster-members and a corresponding CH (CH_k) in Lines 14-16. Next, Lines 17-19 are used to transfer the Time-critical traffic from cluster-members to CH which further forwards it to the IoT-enabled CCS by allocating BW to satisfy constraint (5c). Lines 20-36 are used to execute the hybrid data aggregation strategy for both Periodic and On-demand traffic classes to satisfy constraint (5d)-(5e). For IMR application traffic in Line 20, an exclusive TCP/IP connection is established between cluster-members and AH (AH_k) using Lines 21-23. In Line 24, total electricity consumption ($Electricity_{Consi,k}$) at periodic time interval h_i is calculated and stored in the packet payload and transmitted towards the AH via Line 25. In Line 26, combine method is applied to concatenate the incoming IMR traffic into one aggregated packet with a common header. The incoming aggregated packet (λ_{agg}) is received at the CH in Line 27 which is further exchanged with the CCS using Line 29. In Line 30, if traffic types belong to On-demand class then a connection is established between CH and AH using Lines 31-33. On the received request (ODMR), AH apply the manipulating method on the electricity consumption ($Electricity_{Cons}$) extracted from the packet and response (ODMRR) is shared with CH using Line 34 which is further transmitted towards CCS using Lines 35-36. While the generated raw traffic is discarded in Lines 37-39. If the resource utilization probability is equal or higher than 1, then a message is displayed to show that the resource(s) are at full capacity or overutilized in Lines 40-42. Lines 43-44 are used to compute the constrained resources (BW, RAM, and CPU) usage at the CH. Line 45 is used to compute the main objective function expressed in Eq. (5). Finally, Algorithm 1 returns the network topology, total resources utilized, and AMI applications traffic as output and ends in Lines 46 and 48 respectively.

The process of the main objective function (see Algorithm 1) is depicted in Figure 3 which illustrate the optimum solution of the optimization problem and is briefly explained as follows. We considered the network topology with dual-head cooperative strategy in [18] using Wi-SUN (short range) and LoRaWAN (long range) communication technologies in the deployment of the AMI network. In addition, Algorithm 1 is used to implement the proposed hybrid data aggregation strategy in order to verify the main objective function and find an optimum solution of the optimization problem.

IV. ANALYSIS OF RESULTS

In this section, we present details about the design (experimental and simulation) of the proposed hybrid data aggregation strategy in order to solve the optimization problem in the IoT-enabled AMI network. The performance is evaluated, analyzed, and discussed based on the obtained experimental and simulation results.

A. DESIGN AND PERFORMANCE EVALUATION CRITERIA

We evaluate the proposed hybrid data aggregation algorithm based on the cluster-based network topology in the IoT-enabled AMI network.

First, we carry out different experiments in the clustering topology with QoS traffic engineering implemented at the tier-1 that strongly influence the proposed strategy in terms of workload (here, traffic) reduction and better resources utilization at the constrained CH of the AMI network. We have used a software-define approach to minimize the reliance on expensive physical hardware infrastructure. Software-defined implementation enables to program AMI components such as SMs, DCs, communication networks, and CCS which significantly enhances testing, control, monitoring, management, and troubleshooting in a cost-effective manner using standard APIs framework and graphical user interfaces making the experimental design efficient and flexible in the AMI domain as detailed below.

Since IoT communication framework is deployed to transmit and make it available the IMR application data across the IoT-enabled AMI network. Therefore, dedicated set of IoT RESTful APIs (JSON APIs) [44] are deployed over the network model which enables the periodic IMR applications data acquisition and supports device-to-device communication as web-services over the Internet. RESTful APIs are widely recognized and can be designed easily as these are lightweight, succinct to integrate the IoT and web services as well as works well with the firewalls. In the context of AMI applications, each residential SM (REST client) is periodically triggered to send average electricity consumption (i.e. both active and reactive electricity) after every 15 minute interval. The Utility CCS (REST server) obtains these real-time QoS AMI traffics from time-stamped JSON string payloads via RESTful APIs, which utilizes a request-response message format using multiple queries in the form of HTTP methods such as GET, PUT, and DELETE etc with minimum efforts in the IP-based network and store these metering data into a database application at the MDMS server.

Since most of the AMI devices (e.g., SMs) are resource constrained, REST-based service models involves different operational behaviors with QoS provisioning which requires extensive memory and processing capabilities for HTTP methods in the underline IoT-enabled AMI network. Hence, this leads to the optimization problem as formulated in Eq. (5). The metering data (electricity consumptions) exchanged between the cluster-members and AH in the experimental design is tabulated in Table 4.

Second, we use the simulation approach as an alternative to compare and evaluate the performance of Algorithm 1 with existing methods. CloudSim [45], [46], [47] simulator is used at the cloud computing system (here, CCS) to model the proposed QoS-aware hybrid data aggregation strategy in order to extend our past work [18], [19] by adding QoS module (see Algorithm 1) to support our experiments. The simulation parameters are set on a laptop whose config-

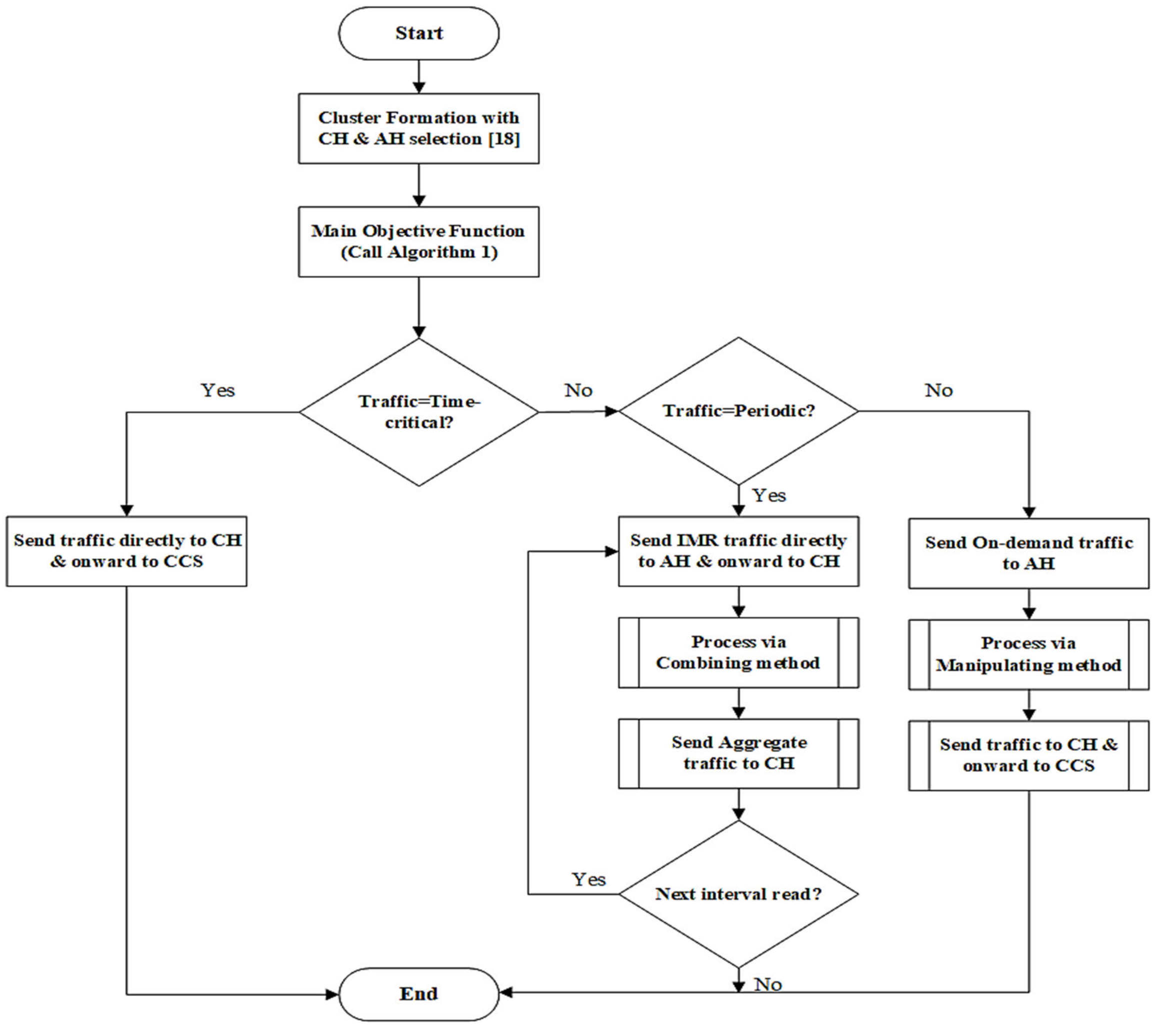

FIGURE 3. Flowchart for the proposed research methodology.

uration detail includes: Intel(R) Core™i5 8250M CPU @ 1.60 GHz 1.80 GHz, RAM (8 GB), Hard drive (500 GB), 64-bit Windows 10 Pro, and a theoretical network capacity of 2 MB. The simulated data is extracted from the JSON string payload with key attributes as depicted in Figure 4, which is generated randomly or periodically by SMs and/or on-demand queried by CCS whose traffic characteristics can be found in Table 3. These real-time collected metering data are permanently stored in a database application created in Microsoft SQL Server 2019 Express [48]. More precisely, to better simulate the proposed strategy, existing CloudSim modules are modified and configured as: Data Center (1), Host (1) has multiple CPU cores with processing speed (1000 MIPS), Broker (1), Virtual Machines (2 VMs) each sharing 2GB RAM and 2MB BW while number of cloudlets randomly varies to the number of SMs requests in the AMI network.

The main reason behind using CloudSim in the resource constrained CH (i.e. embedded system) is that it allows developers and researchers to accurately model a real AMI network


```
{
    "SMID": "198",
    "Read Date": "2023-02-10",
    "Read Time": "18:30:11",
    "Traffic Class": "On-demand",
    "AMI Application": "ODMRR",
    "Season": "Winter",
    "Time of Day": "PM",
    "Hour Type": "On-Peak",
    "Current Read": 1038,
    "Previous Read": 1024,
    "Unit": "kWh",
    "Consumer Type": "Residential Consumer",
    "Detail of Request": "Request for Electricity Consumption"
}
```


FIGURE 4. ODMRR application traffic exchange via JSON string payload.

in order to conduct multiple cost-effective experiments to test the behavior of SMs, communication networks, CCS, and cloud environment. CloudSim eliminates the need of physical hardware resources and make virtualize the available cloud resources (RAM, CPU, and BW etc) via resource optimization techniques. Such techniques improve the overall network efficiency in the AMI network. In addition, it enables the researchers to validate the performance of different algorithms like optimize resource allocation, task scheduling, and load balancing to observe underutilized resources under different workloads in different scenarios. All experiments have been simulated on the same laptop running CloudSim where each experiment has been executed for 1 hour and randomly repeated for 10 time.

Third, we use the following two performance metrics to evaluate capability of Algorithm 1 both in the experimental and simulation design. In the first experiment (scenario), we measure the impact on total traffic reduction at the CHs due to the use of the proposed hybrid data aggregation strategy in the IoT-enabled AMI network. The reason behind using this evaluation metric is to show uniform workload distribution which results to minimize traffic loads at the CH in order to control network congestion and delay in the AMI network. Second, we examine the effectiveness of the proposed strategy by computing the limited resources (RAM, CPU, and BW) of the CH utilized by each AMI applications traffic i.e. most importantly for IMR application traffic. To mimic the resource utilization in CH, the following three equations are used to quantify the resource per usage during the simulation in each VM respectively.

$$(VM_{CPU}) = \left(\sum_{i=1}^{2} \sum_{j=1}^{n} CPU\ used\ by\ Cloudlet_j\ in\ VM_i \right) \tag{10}$$

$$(VM_{RAM}) = \left(\sum_{i=1}^{2} \sum_{j=1}^{n} RAM\ used\ by\ Cloudlet_j\ in\ VM_i \right) \tag{11}$$

$$(VM_{BW}) = \left(\sum_{i=1}^{2} \sum_{j=1}^{n} BW\ used\ by\ Cloudlet_j\ in\ VM_i \right) \tag{12}$$

$$Total = \sum_{i=1}^{2} VM_{CPU} + VM_{RAM} + VM_{BW} \tag{13}$$

$$Usage(\%) = \frac{\sum_{i=1}^{2} CPU\ (MIPS) + RAM + BW}{Total} \tag{14}$$

where, Eqs. (10)-(12) show the usages of individual resources (CPU, RAM, and BW) in each VM respectively. Similarly, Eq. (13) represents the sum of resources utilized in VMs, whereas percentage of total resource usage is calculated in Eq. (14). The simulation results computed based on these equations will ensure that the objective function implemented behave as expected.

B. RESULTS AND DISUSSION

In this section, we evaluate the performance of our proposed data aggregation strategy using the performance evaluation metrics. Both experimental and simulation results are quantified, compared, visualized, and discussed as follows:

We first investigate the effectiveness of our proposed data aggregation strategy in terms of the total traffic load reduction at the CHs for the related AMI applications traffic classes with QoS provisioning. In our experiment, assumed clustering topology consists of 200 SMs (cluster-members), 1 root DC connected to a central router over Internet (capacity of 2MB) to the Utility CCS. In particular, Algorithm1 extracts and collect IMR application data (electricity consumption) via RESTful APIs services from residential SMs at fixed time interval (h = 15 minutes) and AH receives and stores for short time the traffic (generated packets) for data aggregation. Then AH runs (Algorithm 1) to aggregate the electricity consumptions (meter readings) at every 1 hour (60 minutes) and send via CH towards CCS for further processing. Similarly, the Utility CCS query the AH via DC and CH to get On-demand meter readings of SMs. While the traffic generated from Time-critical class applications are directly exchanged with the corresponding CHs. We assume that the TCP/IP packet has a header size of 20 bytes and data payload size is set accordingly as mentioned in Table 3 for each AMI application. The experimental results are quantified in Table 5 and compared with AMI traffic scenario in [18] (without data aggregation strategy) and [27], [39] (with data aggregation strategy) to examine the significance of the proposed algorithm.

Table 5 shows the experimental results obtained in our proposed strategy as compared with the other three scenarios based on the first evaluation metric. To evaluate the total traffic (packets) received by the CHs is 13448 on average in the network topology based on the parameters.

Hence, we noticed that the volume of AMI traffic received at the CH in [18] is 16.90% (2274), 12.01% (1616) in [27] and 64.97% (8738) in [39] respectively. However, when we apply our proposed algorithm the total traffic received is 6.09% (820) in the IoT-enabled AMI network as shown in Figure 5.

This is because, the functionality of data aggregation particularly for IMR application data at the CH is switched towards the AH. As a result, due to our proposed strategy CH receive less AMI traffic as compared with the other three scenarios. Hence, our proposed strategy exhibits better performance (reduces total traffic load) for QoS-aware AMI applications in IoT-enabled AMI network.

Next, we use CloudSim simulation as an alternative approach to validate the performance of the proposed strategy based on second evaluation metric in terms of reducing the resource (e.g., CPU, RAM, and BW) usages for AMI applications traffic in the IoT-enabled AMI network. We create a pool of cloudlets (jobs/tasks) in response to the AMI traffics (REST requests) generated by 200 SMs (REST clients) from related traffic classes (Periodic, On-demand, and

TABLE 4. An example of metering data exchanged with AH via RESTful APIs.

SM ID	Read Date	Read Time	Traffic class	AMI application	Season / Hour Type	Current Read	Previous Read	Electricity consumption (kWh)	Aggregate (kWh)	Detail of Request (RESTful API)
198	2023-06-10	18:30:11 PM	On-demand	ODMRR	Summer / Peak	1038	1024	14	-	POST
200	2023-06-10	18:00:01 PM	Periodic	IMR	Summer / Peak	887	880	7	7	POST
200	2023-02-10	18:15:01 PM	Periodic	IMR	Summer / Peak	892	887	5	12	POST
200	2023-02-10	18:30:01 PM	Periodic	IMR	Summer / Peak	900	892	8	20	POST
200	2023-02-10	18:45:01 PM	Periodic	IMR	Summer / Peak	912	900	12	32	POST
201	2023-02-10	18:35:50 PM	On-demand	ODMR	Summer / Peak	-	-	-	-	GET

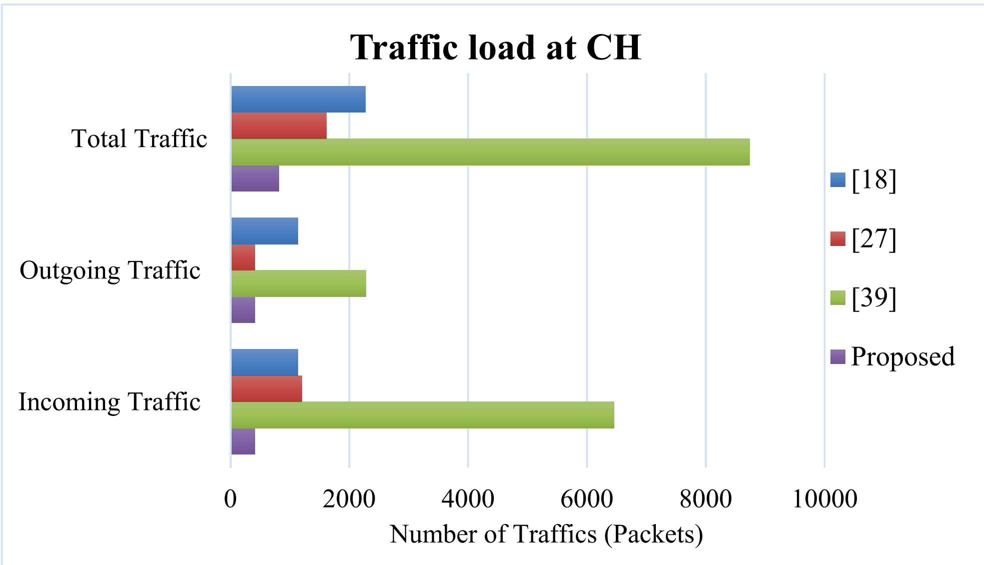

FIGURE 5. Comparison of traffic loads with existing approaches.

TABLE 5. Comparative quantification of total traffic load at the CH.

Scheme	Incoming Traffic (λ_{Down})			Outgoing Traffic (λ_{Up})	Total Traffic ($\lambda_{Down}+\lambda_{Up}$)
	Periodic	On-demand	Time-critical		
[18]	200x4=800	252	85	1137	2274
[27]	200x4=800	290	110	416	1616
[39]	1391x4=5567	510	380	2281	8738
Proposed	16	320	74	410	820

Time-critical) where each cloudlet is managed by a Broker in the CloudSim simulation environment. These cloudlets are submitted as VM requests to corresponding VMs on the physical Host in the Datacenter (detailed in Section IV subsection B). Moreover, the resources utilized by each cloudlet during execution is computed and compared in Table 6 to show that the objective function as expressed in Eq. (5) that is, minimization of constrained resources usage at the CHs is achieved using the proposed strategy for IMR application in AMI network.

During this simulation, we start by executing varying number of cloudlets submitted to VMs. The obtained simulation

TABLE 6. Comparative analysis of VM resources.

Scheme	VM Resources			Total VM Resources Usage (CPU + RAM + BW)	VM Resources Usage (%)
	CPU (MIPS)	RAM (Bytes)	BW (Bytes/s)		
[18]	21480	248784	497570	767834	27.59
[27]	14260	252800	291520	558580	20.07
[39]	44538	522360	646478	1213376	43.60
Proposed	8650	110040	124100	242790	8.72

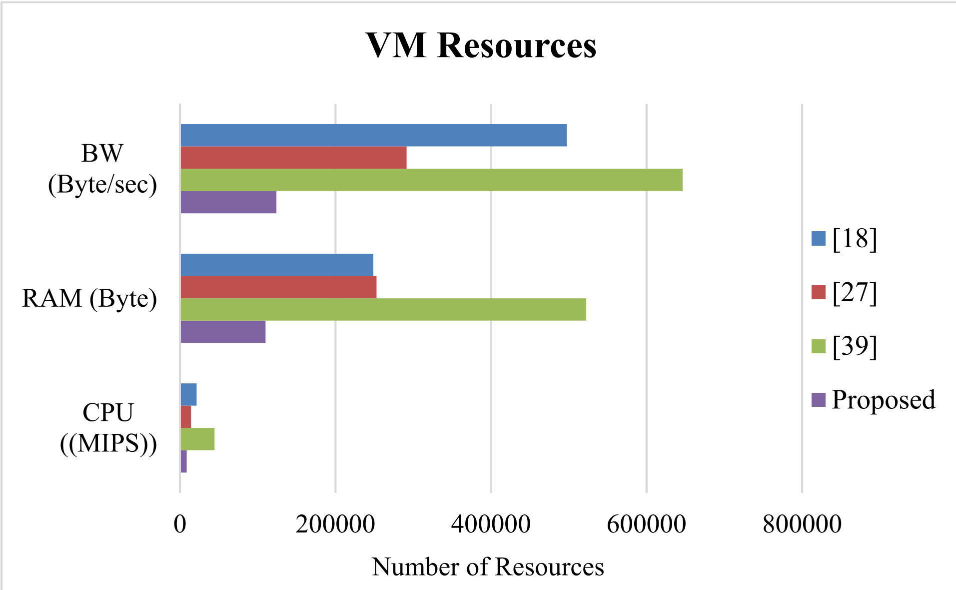

FIGURE 6. Comparison of VM resources with existing approaches.

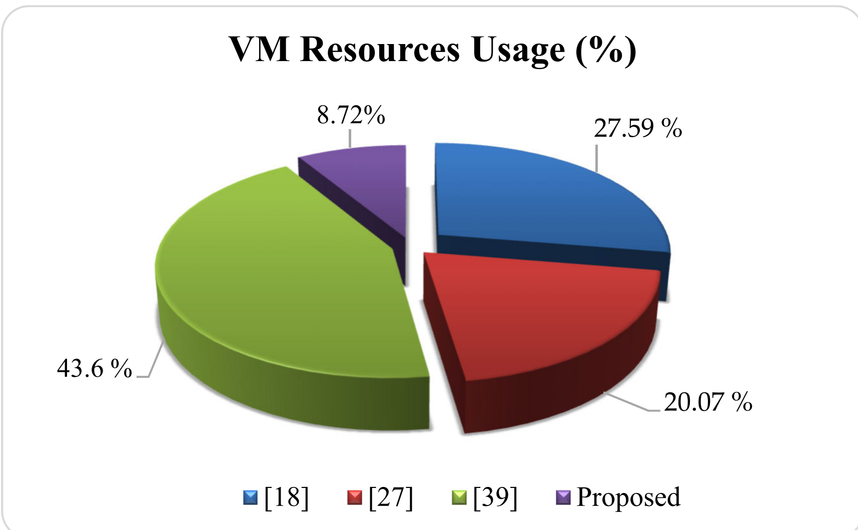

FIGURE 7. Comparison of VM resources usage (%) with existing approaches.

results in each approach are tabulated and compared in order to evaluate the objective function. For example, in [27] using data aggregation strategy at the CH level consumes total VM resources equal to 558580 (20.07%) whereas approach [18] without data aggregation strategy uses more VM resources i.e. 767834 (27.59%) than [27]. The reason behind this is that scenario [18] employs no data aggregation approach on the received IMR application traffic which leads to use more VM resources during the simulation period. Similarly, the data aggregation approach in [39] uses VM resources i.e. 1213376 (43.60%) due to the large number of SMs (5567) connected to four edge nodes (DC node) in the hierarchical network. As compared with [18], [27], and [39] approaches, our proposed strategy exhibits better performance in terms of reducing the usages of VM resources i.e., 242790 (8.72%) as depicted in Figure 6 in the considered network scenario.

The main reason behind this efficient resource optimization and utilization is that IMR application data is transmitted to the AH which employs hybrid data aggregation algorithm and hence reduces the traffic load at the CH in the clustering topology. Further, stringent QoS requirements (latency, BW etc) of all Time-critical traffics are ensured in the AMI network. Figure 7 illustrates the usage percentage (%) of the total VM resources in all research approaches.

V. CONCLUSION

IoT technology facilitate data communication amongst the IoT-enabled AMI devices (SMs, DCs, and MDMS servers) in the AMI architecture. In particular, during peak hours in a season high volume of traffic is exchanged which creates immense burden on these resource constrained devices e.g., SMs and DCs when used as relay-devices in the network topology of the AMI network. Literature investigates this high traffic volume challenge particularly caused by IMR application data in the AMI architecture. In contrast, this article proposes a novel hybrid data aggregation strategy for IMR application data in cluster-based topology in order to minimize the bulk amount of traffic with QoS guarantee and constrained resources optimization problem in the IoT-enabled AMI network. For this, we formulated the optimization problem by defining an objective function that facilitate the minimization of the constrained resources at the CHs in clustering topology. To solve the optimization problem, we developed an algorithm in the CloudSim simulator to find an optimal solution with a tradeoff between reduction of AMI traffic volume with QoS guarantee and resources optimization in the AMI network. Experimental and simulation results provide significant insights that how our proposed strategy outperforms using the two performance evaluation metrics that is, reducing high traffic volume with diverse QoS requirements and resources utilization when compared with the other three research approaches. Hence, we conclude that these contributions show the effectiveness of our proposed strategy in the constrained last-mile IoT-enabled AMI network and well help the electric Utility network managers to design a cost-effective network infrastructure by utilizing the limited available resources e.g., leased network capacity etc. Since we have assumed clustering topology at the NAN level, in future we will investigate to test the feasibility of our proposed strategy by exploring other network topologies and IoT communication technologies to tradeoff various network performance metrics (end-to-end delay, latency, throughput, and reliability). We will consider these possible limitations in future to continue this study.

REFERENCES

[1] Y. Kabalci, E. Kabalci, S. Padmanaban, J. B. Holm-Nielsen, and F. Blaabjerg, ''Internet of Things applications as energy internet in smart grids and smart environments,'' *Electronics*, vol. 8, no. 9, p. 972, 2019, doi: 10.3390/electronics8090972.

[2] D. Sun, W. Li, X. Yao, H. Liu, J. Chai, K. Xie, L. Zhu, and L. Feng, ''Research on IoT architecture and application scheme for smart grid,'' in *Proc. 9th Int. Conf. Comput. Eng. Netw.*, 2021, pp. 921–928, doi: 10.1007/978-981-15-3753-0_90.

[3] S. A. Hashmi, C. F. Ali, and S. Zafar, ''Internet of Things and cloud computing-based energy management system for demand side management in smart grid,'' *Int. J. Energy Res.*, vol. 45, no. 1, pp. 1007–1022, Jan. 2021, doi: 10.1002/er.6141.

[4] N. Ramachandran and V. Perumal, ''Delay-aware heterogeneous cluster-based data acquisition in Internet of Things,'' *Comput. Electr. Eng.*, vol. 65, pp. 44–58, Jan. 2018, doi: 10.1016/j.compeleceng.2017.03.018.

[5] K. Ab Rouf et. al., ''Data aggregation mechanisms in the Internet of Things: A study, qualitative and quantitative analysis,'' *Int. J. Comput. Digit. Syst.*, vol. 9, no. 2, pp. 289–297, Jan. 2020, doi: 10.12785/IJCDS/090214.

[6] NETL. (Feb. 2008). *NETL Modern Grid Strategy Powering our 21st-Century Economy ADVANCED METERING INFRASTRUCTURE*. U.S. Department of Energy Office of Electricity Delivery and Energy Reliability. [Online]. Available: https://www.smartgrid.gov/files/NIST_SG_Interop_Report_Postcommentperiod_version_200808.pdf

[7] I. Colak, S. Sagiroglu, G. Fulli, M. Yesilbudak, and C.-F. Covrig, ''A survey on the critical issues in smart grid technologies,'' *Renew. Sustain. Energy Rev.*, vol. 54, pp. 396–405, Feb. 2016, doi: 10.1016/j.rser.2015.10.036.

[8] R. Ullah, Y. Faheem, and B.-S. Kim, ''Energy and congestion-aware routing metric for smart grid AMI networks in smart city,'' *IEEE Access*, vol. 5, pp. 13799–13810, 2017, doi: 10.1109/access.2017.2728623.

[9] M. Shahraeini and Z. Farmani, ''Designing of communication systems in advanced metering infrastructure (AMI) using Wi-Fi offloading technology,'' in *Proc. 9th Int. Conf. Comput. Knowl. Eng. (ICCKE)*, Oct. 2019, pp. 60–66, doi: 10.1109/ICCKE48569.2019.8964924.

[10] A. K. Padhan, H. Kumar Sahu, P. R. Sahu, and S. R. Samantaray, ''Performance analysis of smart grid wide area network with RIS assisted three hop system,'' *IEEE Trans. Signal Inf. Process. Over Netw.*, vol. 9, pp. 48–59, 2023, doi: 10.1109/TSIPN.2023.3239652.

[11] M. T. Ahammed and I. Khan, ''Ensuring power quality and demand-side management through IoT-based smart meters in a developing country,'' *Energy*, vol. 250, Jul. 2022, Art. no. 123747, doi: 10.1016/j.energy.2022.123747.

[12] (2015). *Tokyo Electric Power Co*. Accessed: Mar. 9, 2023. [Online]. Available: http://www.tepco.co.jp/smartmeter/indexj.%0Ahtml

[13] T. M. Ghazal, M. K. Hasan, H. M. Alzoubi, M. Alshurideh, M. Ahmad, and S. S. Akbar, ''Internet of Things connected wireless sensor networks for smart cities,'' in *The Effect of Information Technology on Business and Marketing Intelligence Systems*, M. Alshurideh, B. H. Al Kurdi, R. Masa'deh, H. M. Alzoubi, and S. Salloum, Eds. Cham, Switzerland: Springer, 2023, pp. 1953–1968, doi: 10.1007/978-3-031-12382-5_107.

[14] T. N. Bhattarai, S. Ghimire, B. Mainali, S. Gorjian, H. Treichel, and S. R. Paudel, ''Applications of smart grid technology in nepal: Status, challenges, and opportunities,'' *Environ. Sci. Pollut. Res.*, vol. 30, no. 10, pp. 25452–25476, Feb. 2022, doi: 10.1007/s11356-022-19084-3.

[15] T. Notice, G. Disclaimer, H. I. Base, en B. Longgang. (2017). *HUAWEI AR550 Series IoT Gateway Datasheet Key Features and Values*. [Online]. Available: https://www.scribd.com/document/451491305/Huawei-AR550-Series-IoT-Gateway-Data-Sheet

[16] M. Kochański, K. Korczak, and T. Skoczkowski, ''Technology innovation system analysis of electricity smart metering in the European union,'' *Energies*, vol. 13, no. 4, p. 916, 2020, doi: 10.3390/en13040916.

[17] Department of Energy and Climate Change. (Jan. 2013). *Smart Metering Implementation Programme: Smart Metering Equipment Technical Specifications Version 2. Draft to the European Commission Jan. 2013*. [Online]. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/68898/smart_meters_equipment_technical_spec_version_2.pdf

[18] A. Khan, A. Umar, A. Munir, S. Shirazi, M. Khan, and M. Adnan, "A QoS-aware machine learning-based framework for AMI applications in smart grids," *Energies*, vol. 14, no. 23, p. 8171, Dec. 2021, doi: 10.3390/en14238171.
[19] A. Khan, A. I. Umar, S. H. Shirazi, W. Ishaq, M. Shah, M. Assam, and A. Mohamed, "QoS-aware cost minimization strategy for AMI applications in smart grid using cloud computing," *Sensors*, vol. 22, no. 13, p. 4969, Jun. 2022, doi: 10.3390/s22134969.
[20] S. Yadav and S. Nanivadekar, "Hybrid optimization assisted green power allocation model for QoS-driven energy-efficiency in 5G networks," *Cybern. Syst.*, pp. 1–16, Feb. 2023, doi: 10.1080/01969722.2023.2175147.
[21] J. L. Gallardo, M. A. Ahmed, and N. Jara, "Clustering algorithm-based network planning for advanced metering infrastructure in smart grid," *IEEE Access*, vol. 9, pp. 48992–49006, 2021, doi: 10.1109/ACCESS.2021.3068752.
[22] T. Shiobara, P. Palensky, and H. Nishi, "Effective metering data aggregation for smart grid communication infrastructure," in *Proc. 41st Annu. Conf. IEEE Ind. Electron. Soc.*, Nov. 2015, pp. 002136–002141, doi: 10.1109/IECON.2015.7392417.
[23] H. Bao and L. Chen, "A lightweight privacy-preserving scheme with data integrity for smart grid communications," *Concurrency Comput., Pract. Exper.*, vol. 28, no. 4, pp. 1094–1110, Mar. 2016, doi: 10.1002/cpe.3527.
[24] M. Wan, W. Shang, P. Zeng, and J. Zhao, "SRDA: A secure routing and data aggregation approach for wireless smart meter," *J. Commun.*, vol. 11, pp. 108–117, Jan. 2016, doi: 10.12720/jcm.11.1.108-117.
[25] T.-W. Sung, Y. Xu, X. Hu, C.-Y. Lee, and Q. Fang, "Optimizing data aggregation point location with grid-based model for smart grids," *J. Intell. Fuzzy Syst.*, vol. 42, no. 4, pp. 3189–3201, Mar. 2022, doi: 10.3233/JIFS-210881.
[26] G. Wang, Y. Zhao, J. Huang, and R. M. Winter, "On the data aggregation point placement in smart meter networks," in *Proc. 26th Int. Conf. Comput. Commun. Netw. (ICCCN)*, Jul. 2017, pp. 1–6, doi: 10.1109/ICCCN.2017.8038499.
[27] U. Das and V. Namboodiri, "A quality-aware multi-level data aggregation approach to manage smart grid AMI traffic," *IEEE Trans. Parallel Distrib. Syst.*, vol. 30, no. 2, pp. 245–256, Feb. 2019, doi: 10.1109/TPDS.2018.2865937.
[28] J.-F. Huang, G.-H. Zhang, and S.-Y. Hsieh, "Real-time energy data compression strategy for reducing data traffic based on smart grid AMI networks," *J. Supercomput.*, vol. 77, no. 9, pp. 10097–10116, Sep. 2021, doi: 10.1007/s11227-020-03557-8.
[29] X. Tian, Q. Song, and F. Tian, "Multidimensional data aggregation scheme for smart grid with differential privacy," *Int. J. Netw. Secur.*, vol. 20, no. 6, pp. 1137–1148, 2018, doi: 10.6633/IJNS.201811.
[30] O. Y. Al-Jarrah, Y. Al-Hammadi, P. D. Yoo, and S. Muhaidat, "Multi-layered clustering for power consumption profiling in smart grids," *IEEE Access*, vol. 5, pp. 18459–18468, 2017, doi: 10.1109/ACCESS.2017.2712258.
[31] S. Sabino and A. Grilo, "Routing for efficient alarm aggregation in smart grids: A genetic algorithm approach," *Proc. Comput. Sci.*, vol. 130, pp. 164–171, Jan. 2018, doi: 10.1016/j.procs.2018.04.026.
[32] I. Popović, A. Rakić, and I. D. Petruševski, "Multi-agent real-time advanced metering infrastructure based on fog computing," *Energies*, vol. 15, no. 1, p. 373, Jan. 2022, doi: 10.3390/en15010373.
[33] R. C. Diovu and J. T. Agee, "Overview of data aggregation schemes in a smart grid AMI network," *J. Commun.*, vol. 14, pp. 787–801, Sep. 2019, doi: 10.12720/jcm.14.9.787-801.
[34] E. Inga, R. Hincapié, and S. Céspedes, "Capacitated multicommodity flow problem for heterogeneous smart electricity metering communications using column generation," *Energies*, vol. 13, no. 1, p. 97, Dec. 2019, doi: 10.3390/en13010097.
[35] E. Inga, S. Cespedes, R. Hincapie, and C. A. Cardenas, "Scalable route map for advanced metering infrastructure based on optimal routing of wireless heterogeneous networks," *IEEE Wireless Commun.*, vol. 24, no. 2, pp. 26–33, Apr. 2017, doi: 10.1109/MWC.2017.1600255.
[36] E. Inga, J. Inga, and R. Hincapié, "Maximizing resource efficiency in wireless networks through virtualization and opportunistic channel allocation," *Sensors*, vol. 23, no. 8, p. 3949, 2023, doi: 10.3390/s23083949.
[37] G. Wang, Y. Zhao, J. Huang, Q. Duan, and J. Li, "A K-means-based network partition algorithm for controller placement in software defined network," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6, doi: 10.1109/ICC.2016.7511441.
[38] W. A. Jabbar, S. Annathurai, T. A. A. Rahim, and M. F. M. Fauzi, "Smart energy meter based on a long-range wide-area network for a stand-alone photovoltaic system," *Exp. Syst. Appl.*, vol. 197, Jul. 2022, Art. no. 116703, doi: 10.1016/j.eswa.2022.116703.
[39] S.-V. Oprea and A. Bâra, "An edge-fog-cloud computing architecture for IoT and smart metering data," *Peer-Peer Netw. Appl.*, vol. 16, no. 2, pp. 818–845, Mar. 2023, doi: 10.1007/s12083-022-01436-y.
[40] S. Khalid, M. Sarfraz, V. Singh, and A. Allahloh, "A novel model for IoT-based meter using ATmega328P microcontroller and Google cloud store," in *Advances in Electromechanical Technologies*. 2021, pp. 639–648, doi: 10.1007/978-981-15-5463-6_57.
[41] D. B. Avancini, J. J. P. C. Rodrigues, R. A. L. Rabêlo, A. K. Das, S. Kozlov, and P. Solic, "A new IoT-based smart energy meter for smart grids," *Int. J. Energy Res.*, vol. 45, no. 1, pp. 189–202, Jan. 2021, doi: 10.1002/er.5177.
[42] H. Harada, K. Mizutani, J. Fujiwara, K. Mochizuki, K. Obata, and R. Okumura, "IEEE 802.15. 4G based Wi-SUN communication systems," *IEICE Trans. Commun.*, vol. 100, no. 7, pp. 1032–1043, 2017, doi: 10.1587/transcom.2016SCI0002.
[43] A. Yusri and M. I. Nashiruddin, "LORAWAN Internet of Things network planning for smart metering services," in *Proc. 8th Int. Conf. Inf. Commun. Technol. (ICoICT)*, Jun. 2020, pp. 1–6, doi: 10.1109/icoict49345.2020.9166455.
[44] L. Zhang, S. Yu, X. Ding, and X. Wang, "Research on IoT RESTful web service asynchronous composition based on BPEL," in *Proc. 6th Int. Conf. Intell. Human-Machine Syst. Cybern.*, vol. 1, Aug. 2014, pp. 62–65, doi: 10.1109/IHMSC.2014.23.
[45] H. Gibet Tani and C. El Amrani, "Cloud computing CPU allocation and scheduling algorithms using CloudSim simulator," *Int. J. Electr. Comput. Eng. (IJECE)*, vol. 6, no. 4, p. 1866, Aug. 2016, doi: 10.11591/ijece.v6i4.10144.
[46] S. Wang, A. Zhou, C.-H. Hsu, X. Xiao, and F. Yang, "Provision of data-intensive services through energy- and QoS-aware virtual machine placement in national cloud data centers," *IEEE Trans. Emerg. Topics Comput.*, vol. 4, no. 2, pp. 290–300, Apr. 2016, doi: 10.1109/TETC.2015.2508383.
[47] S. Mehmi, H. K. Verma, and A. L. Sangal, "Simulation modeling of cloud computing for smart grid using CloudSim," *J. Electr. Syst. Inf. Technol.*, vol. 4, no. 1, pp. 159–172, May 2017, doi: 10.1016/j.jesit.2016.10.004.
[48] Microsoft. (2023). *Microsoft SQL Server 2019 Express*. Accessed: Mar. 5, 2023. [Online]. Available: https://www.microsoft.com/en-us/download/details.aspx?id=101064