Journal of Communications Technology, Electronics and Computer Science, Issue 30, 2020
ISSN 2457-905X

Network Resource Management Drives Machine Learning: A Survey and Future Research Direction

Wasswa Shafik[1], S. Mojtaba Matinkhah[2], Mamman Nur Sanda[3]

Computer Engineering Department[12], Department of Physics[3]

Yazd University, Yazd Province, Iran[123]

Wasswashafik[1], mammannur10[3]{@gmail.com}, matinkhah@yazd.ac.ir[2]

Abstract— **Network resource management is one of the vibrant factors in the current dynamic technological computing paradigms. This reduces poor resource utilization; network resources include network devices, management networks, management systems, and management support organizations carrying out task planning, resource scheduling, and network managing among many more. It's now been observed within different networks that machine learning has been applied to help in carrying some network tasks. This paper surveys current approaches that have been done in managing resources with a deep focus on energy optimization, auto-scaling methods based on current machine learning approaches. We further present the proposed techniques that are identified by evolving this overview by investigating other state-of-the-art recommendations. This study provides a deep insight into proposed methods, loopholes that will aid researchers to design, model, or create novel frameworks that focus on the better options of resource management in contrast to the existing methods.**

Keywords— Resources Management; 5G Networks; SDN; NFV; Auto scaling Method; Machine Learning

I. INTRODUCTION

The 5G networking is anticipated in having increased user mobility, ultra-reliable communication with ultra-density, promising data rate of ultra-low latency to avail a diversity of services to the end-users, automatics, tactile Internets, virtual reality, and smart generations must be put in place to have better resource utilization. The 5G networks are anticipated further software configurable, agile with flexibility, scalable utilization of current emergent technologies like Network Slicing, Network Functions Virtualization (NFV), and Software-Defined Networking (SDN). These techniques require automatic management and dedicate each slice to a set of service types [1].

SDN and NFV technologies tend to contrast based on the functionality and resources allocation and utilization. For instance, the SDN extracts physical networking resources like routers and switches, transfers decision manufacture into a virtual network control plane. In this tactic, the control plane chooses where to send traffics, whereas the hardware endures to direct and handle the traffics in return as well.

The main aim of the NFV technologies is to virtualize all physical network resources and permits the network to grow with no addition of more peripherals. It's quite important to note that despite the fact both technologies (SDN and NFV) varies within flexible and dynamic networking architectures and tasking diverse roles in significant architectures with support [2].

Consequently, the SDN and NFV are subjected to seriously virtualization that enables network infrastructure and designs to extract software that can be implemented by fundamental software trans-versing hardware platforms and devices. At times, the SDN does execution over the NFV infrastructures, the SDN technology transfers onwards data packets from one network device to another device.

Further, still, the SDN's networking switch purposes during routing periods, policy applications, and definition run in a virtual machine anywhere on the network. Accordingly, the NFV offers elementary networking functionalities as well as the SDN technology panels and orchestrates resources for precise routines. More importantly, the SDN additionally permits configuration and conduct to be programmatically well-defined and altered [3].

The main SDN has a number of the ingredients the drive the positive functionalities among many they include the fast deliveries of directly programmable network controls, it is also agile and has properties of responsiveness where the SDN authorizes the administrators to regulate or manage network-wide traffics flow vigorously to encounter all the fluctuating requirements and demands, there is also network intelligence that is rationally centralized through the SDN controllers.

The SDN technologies provide programmable configuration, the SDN has possibilities of the allowing network managers can arrange, secure, regulate and tune network resources for better utilization by the use SDN programs that provide automatic applications, the SDN technology is vendor-neutral and standards-based since SDN is used open standards, and lastly on the list is that the SDN often restructures network design and processor operation through energy optimization [4].

The NFV technologies substitute the traditional network services that are often provided through devoted hardware with virtualized software with a great focus on the dynamic resource allocation and power management. This implies that the network services that are involved like a case for example load balancers, firewalls, routers, Wireless Area Network processing, and optimization devices at times are substituted by that software execution on virtual machines.

It's also vital to note that the Virtualized network functions are controlled by the hypervisor that the SDN accomplishes the NFV technology aids to save together capital expenditures and operating expenses. The network services mentioned above are used to necessitate specific, devoted hardware that can execute on standard commodity servers like x86 commodity hardware, and see the reduction of the costs during computation and so forth [5].

In this paper practically, we reveal the presence of the SDN and NFV technologies in different ways in infrastructures, architectures or frameworks, the proposed models or methods with an attempt for efficient network resource management through the use of machine learning techniques or other tactics identifying methods, its applicability and the loophole that are we have suggested alternative measures to minimize resource during networking, we explore the and present the research directions that in case followed, will provide solutions to the existing challenges since we have dug deeper on network resource auto-scaling method of the converged network.

Firstly, there comes the network resource management that sliced into five different isolated categories with miscellaneous resource management necessities where we present table 1 showing major resource management challenges. Then, different surveys on resource management are carried out on machine learning that includes reinforcement learning, learning Automaton among others, auto-scaling and so many others that the virtualized SDN controller allocates

Communicational resources to each slice to satisfy their minimum services and resources required.

The rest of this paper is structured as follows. In Section 2, we review provide main network resource management categories that include especially telecommunications, fog computing, could computing, Edge computing, Network service system, and Network production systems with associated challenges that are summarized in table 1. In Section 3, the analysis on the network resource management is provided in specific on energy efficiency based on the different learning domains including machine learning to mention but a few with a focus on energy management and its machine learning existing approaches or Methods, a summary of all other state-of-the-art machine learning methods.

In Section 4, We disclose more state-of-the-art resource management techniques in networking through the use of the common auto-scaling approach, a summarized table 2 is provided to illustrate all these methods identified in association to the methods, exploring persisting resource management issues to provide clear research direction based on field, challenges, and structures. We conclude in Section 5 and depicting our future research more.

II. NETWORK RESOURCE MANAGEMENT

In this section, we present several fields with a fundamental common approach to resource management, to begin with, telecommunication networks, fog networking, Cloud network computing, Edge network computing, Network service systems and lastly the Network production systems existing methods through clarifying these identified approached to their empowering machine learning in particular reinforcement learning and these approaches as discussed below:

A. Telecommunication Networks/ Tele-comm

Computing

A telecommunication network is referred to as a gathering of terminal nodes where links are connected to enable telecommunication between the terminals. These terminals are often connected to a node where the propagation of either an analog or digital information signal finished to a physical point-to-multipoint or point-to-point transmission mediums either optical fiber, wireless, or wired. All these need attention that resource management counts in telecommunication networks.

Table 1.Network Resource Management Challenges and Approaches

Network Resource Management Entities	Network Resource Management Challenges	Network Resource Management Approaches	Reference
Telecommunication Networks	NFV elasticity	Auto-scaling apparatuses	[27]
	An IMS cloud elasticity execution	Load balancing Auto-scaling and	[28]
	LTE-A complexities, 5G mobilities	Automatic network management	[29]
Cloud networks, Fog Networks Edge Networks	Virtual machine allocation	Survey	[30]
	Resource optimization	Close studies on the resource challenges	[31]
	Resource management	Application of reinforcement learning, machine learning, extreme learning	[17]
Network service systems	Network staffing challenge	Operating difficulties	[32]
	Network staffing challenge	A simple linear programming	[33]
	Network staffing challenge	A dynamic programming	[34]
Network production systems	Emerged capacities Inventory management	Search (advanced) algorithms	[34]
		A dynamic programming	[35]

Standards have been put in place in the approach to proper network resource management like the emerging arena of the self-organizing networks through the introduction of the 3rd Generation Partnership Project (3GPP) detailed in [6]. To was further extended to create automatic scaling that is seen the NFVs though still it was challenged since the self-organizing networks had issues that were not sorted and cleared as dynamically the capacities of the network qualities and grade of service were compromised, the most known telecommunications resources that need management include computer networks, telephone network, and Internet as was it as detailed in [7-9].

B. Fog Networks / Fog Computing

Fog computing has attained popularity in a few years ago where resources or computation is done at the edge of the network, by this fact during the early studies carried out on this fog network paradigms, we illustrated some security and privacy resource issues that influence network resource management and proper execution of the networks. Within the same study, we provided basic definition, computing paradigms in networking that Jungle computing, cloud computing, Edge computing, provided the common fog network characteristics, and fog network taxonomy detailed in [10].

It was after the close focus on the transmission that is carried out during networking and revealed that several resources are over-utilized during the execution of the requests or networking running operations, this opened a gap that researchers need to bridge that is resource scheduling and manage resource pooling steams in approach to evaluating the execution delays [11]. More still, during the determination of the packet losses of networks in particular within wireless communication, the authors assessed the network packet losses where they stilled concluded factor of over resource management as well commutated in [12].

C. Cloud Networks / Cloud Computing

Notably, Cloud networks or cloud computing has existed in the number of years where network resources, computational processing are done within a cloud center. Here is quite clear in a condition where resources are pooled in once stream, chances of over utilizations are likely to rise in case the resource is not controlled, regulated or managed efficiently based on enabled data center network based on traffic demands of the computation [13].

It has been quite aired out for a couple of studies that one of the most persisting issues challenging cloud computing is resource management, it was clearly described in four main sorts including computation (compute), networking (network), powering (power), storing (storage) like it as detailed in [14, 15]. Alternative solutions were availed in approach to network resource management among which the solutions included reconfiguration of processing units, memories, determining the server numbers for only user-defined application to generate algorithmic fib numbers (Fibonacci number in the advanced algorithm) thus satisfying that the challenge is still open for more solution to manage the network resources [16, 17].

D. Edge Network / Edge Computing

There exists a thin line between the last two upper networks (Fog network and Cloud network) and Edge network computing. The concept of availing resource computation is not this much innovation since it has been around for many years due to the presence of content deliveries or distribution of networks (Distributed Networks). A Study carried about heterogeneous network issues management based on enhanced inter-cell interference coordination and the machine learning algorithms, have also opened another window for viewing resource management issues, a close study more still showed how best extreme learning can be used in managing the number of resources to gain better resource qualities and regulation [18, 19].

E. Network Service Systems / Network Systems

This is one of the core points of the network where the network resources almost poorly controlled, regulated, and management based on the dynamic checkouts and stochastic processing or execution of resource complexities.

A network service varies on an application executing at the network application layer and beyond, that affords different task including data processing, manipulation, storage, presentation, communication among other capabilities or capacities that are regularly executed through the use of client-server connectivity or peer-to-peer connectivity architectures based on application layer SDN [20].

Several service system challenges are raised for example on empowering social web of customer service entailing web of things like systems that are configured technologies and administrative networks in case the systems designed to deliver services that gratify the user desires requirements or requests, or ambitions of customers. It's important to note this Network Service system is often applied during service management, services marketing, service engineering, service design literature service operations as more commonly used in discussed during forecasts in buffering circuits or centers, determining the number of the agents to be included within a given time, staffing challenges and offloads of the network [21, 22].

F. Network Production Systems / Network Production Oriented Systems

Lastly on the list among the category as introduced are the network production systems. This is seen as commonly seen in the production system as the name insights. One of the systems that are used in identifying the challenge of network resource management where these are placed to produce products to satisfy consumer needs at a given time of production. To achieve this standard, some things are meant to hinted on through which resource tends to be managed poorly thus leading to backlog demands, in minimizing connectivity delay of the internet of things that allows some network devices to be connected with acceptable internet speeds, all these studies are done mainly done to minimize network resources here internet is inclusive [23].

The above earlier proposed solutions primarily depend on the controller-side detection and filtering embedded within the systems; hence leave the consumption of the control plane bandwidth resources and limited to quick response because of the switch-controller delay. The author offered INFAS, a system that runs on commodity servers installed near network devices, for protecting SDN against CPS as detailed [24]. More still, a distributed architecture in performing network analytics applications using machine learning techniques in the context of network operation and control of 5G networks was introduced [25, 26], as summarized table 1 as provided to illustrate Network resource management all the six discussed fields.

III. NETWORK ENERGY MANAGEMENT BASED ON MACHINE LEARNING APPROACHES

In this section, an analysis of the energy efficiency since almost every network computing paradigm including cloud, edge, fog, and mobile-edge among others are facing inefficiency in energy is presented, mainly revealing the main energy challenges basing on the machine language techniques in specific.

To begin with, Saha and Mitra [1] propose a systematic study of the inherent energy efficiency of several optimization algorithms for possible uses in FC. The time complexities of Multi-Application Provisioning along with Single Application Provisioning algorithms are compared. The results show a Randomized Algorithm and Approximation Scheme have potential efficiency towards optimization algorithms in Fog computing as both of them can ensure the Green computing nature.

Correspondingly, Ahvar et al. [2] introduce a taxonomy of different Cloud-related architectures ranging from fully centralized to completely distributed ones to evaluate and compare their energy consumption. The results show that as a result of not using an intra-data center network and large-size cooling systems, a completely distributed architecture consumes between less energy than fully centralized and partly distributed architectures respectively.

Some Fog-assisted IoT systems realized energy efficiency through working on local storage and local communication. For example, Nguyen Gia et al. [3] use FC at the edge of the network for remote monitoring of diabetic patients which offers local distributed storage. This system leverages the wearable TNs for energy efficiency. Even though the TNs is equipped with many types of sensors, it consumes minimum energy. Correspondingly, Tahir et al. [4] used Bluetooth sensors in three different modes periodic, sleep–awake, and continue to reduce energy consumption. The sensed data are transmitted to the relevant Fog and Cloud devices for further processing.

The simulation results show that the proposed technique reduces energy consumption. Similarly, other papers improve FNs by leveraging renewable energy sources. Jiang et al. [5] develop a model to study the energy sustainability of FNs powered by renewable energy sources. In their model leaky bucket used to shape and police traffic source for rate-based congestion control. The energy depletion probability and mean energy length calculated which is sensitive to certain traffic source characteristics in Fog networks.

On the contrary, for keeping energy efficiency, some authors design FNs with a power-saving goal instead of working on TNs. For example, Mukherjee et al. [6] propose the use of switches with active or idle mode according to the user's presence in its coverage for FNs. They compare the power consumption and delay of the proposed system to reduce power by approximately 89% and 94%, respectively, in comparison to the local Cloud server-based and remote Cloud server-based systems.

Finally, Optimizing the Cloud layer is also promising. For example, Alharbi et al. [7] study the offloading of virtual machine services from the Cloud to the Fog considering the impact of the virtual machine workload and the proximity of FNs to TNs with various data rates. The result shows that the optimum placement of virtual machines significantly decreases the total power consumption. To conclude the different solutions discussed above, Liu et al. [8] compare TN computing with FN computing.

The authors formulated two multi-objective optimization problems to minimize the required energy and time, where the time assignments and the transmit power are jointly optimized. For the local computing mode, the closed-form expression of the optimal time assignment for energy harvesting solved by a convex optimization problem, and the effects of the scaling factor between the minimal required energy and time on the optimal time assignment are analyzed. For the FC model, closed-form and semi-closed-form expressions of the optimal transmit power and time assignment for offloading by adopting the Lagrangian dual method is derived using the Karush–Kuhn–Tucker (KKT) conditions and Lambert W Function. Simulation results show that, when the sensor device has poor computing capacity or when it is far away from the FN, the FC model is the better choice; otherwise, the local computing is preferred to achieve better performance.

One of the proposed solutions for conflict resolution between energy consumption and delay minimization is leveraging a broadcast system of the sensed data for all FNs and decisions based on the fading channels between TNs and FNs. In this way, the system will become relaxed of firstly determining which FN to associate with. For instance, Abkenar et al. [9] proposed a new protocol, where the TNs immediately broadcast their data with a certain transmission rate to potentially all FNs. Upon receiving the data, FNs determine whether to process the data locally or forward to the Cloud center, depending on whether they are overloaded or not. The fading channels between TNs and FNs allow a mathematical analysis of the energy consumption tradeoff between TNs and FNs by varying the broadcasting rate of the TNs. Multicast routing with multiple constraints of QoS has been proved to be an NP-complete problem.

Nevertheless, Kadhim and Seno [10] propose an energy-efficient multicast routing algorithm in Vehicular networks based on Software Defined Networks and FC using a partitioning technique to reduce the time complexity and overhead. The algorithm classifies the multicast requests based on their application type and deadline constraint and schedules them along with their priority. The results reveal the reduction of multicast energy consumption while meeting the critical multicast sessions deadline. Mukherjee et al. [11] propose a mathematical model for the delay and energy reduction in the 5G Femtolet-based Fog network. The results of the simulation show that the proposed model reduces the energy consumption and delay than the FC-based existing IoT paradigm.

More still, Fang et al. [12] propose an energy-efficient algorithm in fog-assisted industrial wireless sensor network which uses a hierarchy process to achieve the trade-off between security, transmission performance, and energy consumption. The proposed trade-off can effectively select the secure and robust relay node. Simulation results show the algorithm not only prevents the appearance of network holes and defends bad-mouthing attacks but also help the balance of the network load.

Within the close parameters, a multi-sensor geo-Fog proposed by Mishra et al. [13] with minimal delay in which the sensor and geospatial data processing and analysis take place inside the FNs. The shortest path to the victim region is determined using an intelligent K* heuristic search algorithm. The simulation results show that the reduction of both energy consumption and the average delay at a significant pace. Vales et al. [14]

propose a data replication algorithm based on an original edge service managed by an adaptive distance for node clustering. The adaptive distance is evaluated by distance from consumers to the data storage location, spatiotemporal data popularity, and the autonomy of each battery-powered node.

The results indicated that the algorithm offers more responsive data access to consumers, reduces core traffic, and depletes fairly the available battery energy of edge nodes. Similarly, Cen et al. [15] propose an edge content caching strategy for FNs by mining out mobile network behavior information using a pre-mapping containing content preference information and geographical influence by an efficient non-uniformed accelerated matrix completion algorithm. Simulation results verified by theoretical analysis show the energy efficiency achieved by the proposed algorithm.

Toor et al. [16] through iFogSim simulator and formal verification show having adaptive performance and energy-aware scheme significantly reduce energy consumption in the power saver mode. Besides, the proposed scheme also establishes its effectiveness concerning the QoS metrics. Scarpiniti et al. [17] develop VirtFogSim which is a MATLAB-supported software toolbox that allows the dynamic joint optimization and tracking of the energy and delay performance of Mobile-Fog-Cloud systems for the execution of applications described by general Directed Application Graphs (DAGs). To check both the actual optimization and scalability capabilities of the VirtFogSim toolbox, some experimental setups featuring different use cases and operational environments are simulated, and their performances are compared.

Comparably, Djemai et al., 2019 [18] propose a services placement strategy with energy consumption and applications delay minimization. The effectiveness of the proposal is proved in the iFogSim environment. Rahbari and Nickray [19] propose a scheduling algorithm for allocating resources of modules in FC. The algorithm simulated in iFogsim as a standard simulator for FC. The results show that energy consumption, execution cost, and sensor lifetime are better than those of the first-come-first-served (FCFS), concurrent, and delay-priority algorithms.

Isa et al. [20] propose a Mixed Integer Linear Programming model to optimize the number and locations of the primary and backup processing servers with the aim of the energy consumption of both the processing and networking equipment are minimized. The results show that geographical constraints increase energy consumption and increasing the number of processing servers that can be served at each candidate FNs can reduce the energy consumption of networking equipment besides reducing the rate of energy increase of networking equipment due to increasing level of demand.

Tinini et al. [21] devise an integer linear programming (ILP) optimization and graph-based heuristics to choose when to activate FNs and how to determine wavelengths on a time-and-wavelength division multiplexing passive optical network to support the fronthaul. The results show that Cloud-Fog RAN in future 5G networks can consume less energy as well as latency. Dong et al. [22] propose a cooperative FC system to process offloading workload on the entire Fog layer by data forwarding. A joint optimization problem of QoS and energy in an integrated FC process with fairness, by proving its convexity, is solved by a fairness cooperation algorithm. Results show that the algorithm can effectively reduce the time overhead and the energy consumption compared to the distributed optimization algorithm.

Vu et al. [23] proposed a joint energy and latency optimization scheme for upstream IoT offloading services with task assignments in 5G Fog radio access networks. Simulation results show that the scheme outperforms current methods regarding energy consumption and load balancing while maintaining IoT service satisfaction. Li et al. [24] study the computation offloading problem of the coexistence and cooperation between FC and Cloud computing by jointly optimizing the offloading decisions, the allocation of computation resources, and transmit power. They propose an energy-efficient computation offloading and resource allocation algorithm to minimize the system cost. The simulation results show the proposed algorithm can effectively decrease the system cost.

Ma et al. [25] propose a mixed-integer linear programming (MILP) model to optimize the power consumption of the vehicular Fog which comprises FNs connected to the central data center through an optical infrastructure. Taking into account different number of task assignments problems (called software matching problem), the results verified a significant power saving. Kim et al. [26] propose an energy optimal offloading scheme based on a probabilistic priority model of Cloud tasks over FC networks.

The optimization problem jointly minimizes the task popularity-based energy consumption of TNs and the FNs. Simulation results show that the scheme offers better performance compared to the conventional offloading scheme, which operates exactly within the determined delay bound. Dinh et al. [27] formulate the energy efficiency maximization problem, considering the local user clustering constraint specific to Fog-RANs to face radio resource and interference management as well as beamforming.

The problem revealed to be a non-convex optimization problem and need some level heuristics: Splitting the energy-efficient user scheduling into two parts: an equivalent sum-rate maximization problem, then, greedily activating the most energy-efficient FNs. Local beamforming is done with fixed user scheduling to meet the requirement of the low computational complexity of FNs. Simulation results show that the algorithm not only optimizes energy efficiency but also provides sufficient user rates and fairness. By the same token, Chen et al. [28] formulate the joint optimization of the communication resources and computation resources in FC to maximize the overall system energy efficiency, in which multiple FNs and end-users are taken into account which is a mixed-integer non-linear problem.

The leveraged heuristic here is first converting the problem to a non-probability form and then efficiently solving the problem based on the Dinkelbach algorithm and Lagrange duality approach. Besides, an imperfect channel state information is used to tackle the expanding scale of the problem. The simulation results reveal the usefulness of the proposed scheme, especially when the scales are huge. Similarly, Liu et al. [8] formulate two multi-objective optimization problems to minimize the required energy and time, where the time assignments and the transmit power are jointly optimized. The convex optimization problems solved by using the Lagrangian dual method and the Karush–Kuhn–Tucker (KKT) conditions and Lambert W Function. Simulation results show that the proposals achieve an operative performance.

Rahbari and Nickray [19] propose a greedy knapsack-based low-latency and energy-efficient scheduling algorithm for allocating resources of modules in FC. It is one of the simplest integer-optimization problems which can be expressed algebraically. It gives rise to a linear problem in which a linear expression is to be maximized or minimized in non- negative integers, subject to a linear inequality constraint. The results show that the energy consumption is improved. Another solution proposed for persuading the FNs for cooperation to increase the capability of the Fog layer and is using mechanism design which is a well-known method in game theory. One of the disciplines of such mechanism design is motivating the FNs through monetary profit.

In the illustration of the above, Ridhawi et al. [29] design the shared profit gains for the cooperating Fogs which offered guaranteed QoS to users to not only maximize their satisfaction but also increase the Fog profit gains. Khalid and Javaid [30] propose a coordination based energy management system as-a-service on Fog for the better management of energy. A social network problem is formulated as a game theory-based coalition method between the produced and the required energy. The surplus energy (pay-off) will be distributed among the energy-deficient players using Shapley value that unevenly distributes the power according to the demand. The experimental results show that extra power is saved and distributed among energy-deficient consumers. In the same way, Mebrek et al. [31] propose a Nash Equilibrium problem (NEP), which allows the trade-off between consumed energy by the system and QoS. The simulation experiments and comparison with two benchmarks show the efficiency of the proposal.

Machine learning is another proposed methodology for energy efficiency in FC. Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it to learn for themselves. Like, Ridhawi et al. [29] build a learning mechanism that relies on online and offline simulation results to build guaranteed workflows for new service requests. Shukla et al. [32] propose a cluster-level task manager that employs Reinforcement Learning to identify optimal configurations at the server- and cluster-levels. A load-balancing scheduler assign services based on the learned configurations to co-optimize latency as well as energy.

This system achieves up to 28% energy saving compared to best-case cluster scheduling techniques on a 4-server Fog cluster. In the same way, Mebrek et al. [31] propose a joint optimization of resource allocation and workload dispatching over a Fog-Cloud system. To break the curse of large-scale systems, a Reinforcement Learning-based algorithm is proposed to allow users to learn the optimal policy without having a priori knowledge of the dynamic statistics of the system. The simulation experiments and comparison with two benchmarks show the efficiency of the proposal.

Resource allocation and task scheduling are the NP-hard problems in FC. Each application includes several modules that require resources to run. Some other groups of researchers are working to find an optimized solution to the mathematical formulation of the problem. In this way, they can formulate the problem of how to guarantee the level of delay in response to the computation needs of TNs. To this aim, the key tool is the optimization of the offloading ratio. For example, Chen et al.,[33] develop an energy-efficient computation offloading scheme for industrial IoTs in an FC mode. The proposed algorithm for offload optimization is accelerated gradient with a fast convergence speed tested by some numerical simulations. Sun et al. [34] also propose an FC architecture that the energy and time-efficient computation offloading and resource allocation is formulated into the energy and time cost minimization problem. The proposed algorithm to solve the problem, improving the energy consumption and completion time of application requests.

On the other hand, meta-heuristic is used for solving computationally hard tasks such as the solution of NP-complete problems, for which no known algorithm can compute an exact solution. Such algorithms provide a sufficiently good solution to an optimization problem, especially with incomplete or imperfect information or limited computation capacity. The reason behind their prevalence is that metaheuristics may make few assumptions about the optimization problem being solved, and so they may be usable for a variety of problems. Compared to optimization algorithms and iterative methods, metaheuristics do not guarantee that a globally optimal solution can be found on some class of problems which should be applied with prudence. Such metaheuristics include simulated annealing, tabu search, iterated local search, variable neighborhood search, etc. These metaheuristics can both be classified as local search-based or global search metaheuristics.

Ridhawi et al. [29] obtained a service level agreement of the users employing a modified tabu-based search mechanism that uses previous solutions when selecting new optimal choices. Performance evaluation results demonstrate significant gains in terms of service delivery success rate, QoS, reduced power consumption for Fog and Cloud data centers, and increased Fog profits. In the same way, Tajiki et al. [35] propose a near-optimal heuristic algorithm for solving energy-aware service function chain rerouting which although the problem itself is very hard, the solution is in polynomial time. Their simulation results show that the heuristic algorithm is appropriate for large-scale networks.

Likewise, Usman et al. [36] propose a decision model that considers the resource constraints of CPU, Memory, and Storage. Using the integrated resource allocation model will improve resource utilization as well as the reduction in energy consumption by a datacenter. Similarly, Butt et al. [37] propose two meta-heuristic algorithms: Genetic Algorithm (GA) and Binary Particle Swarm Optimization (BPSO) to balance the set of requests on servers of Cloud and Fog. The proposed algorithms are compared with existing algorithms to measure efficiency. The proposed GA and implemented BPSO are used to minimize the response time.

Comparably, Djemai et al., 2019 [18] propose services placement in a Fog architecture. The placement strategy taking into account the system's energy consumption and applications delay violations minimization with a Discrete Particle Swarm Optimization algorithm. Simulation is done in iFogSim and the results compared with other heuristics such as Binary Particle Swarm Optimization (BPSO), Dicothomous Module Mapping (DCT), CloudOnly, IoTFogOnly, IoTCloud (IC) and FogCloud (FC) placement approaches.

Establishing a data collection path model is another suggested solution for minimizing the energy on FNs. As an example, Wang et al. [38] design both an energy-efficient as well as trustworthy protocol. The mobile data collection path with the largest utility value is created, which can avoid visiting needless sensors and collecting unreliable data which results in outperforming traditional data collection methods in both energy and delay. Likewise, Tajiki et al. [35] also face with maintaining an acceptable level of network path survivability and a fair allocation of the resource between different requests in the event of faults or failures. Energy consumption constraints applied for optimization if the real-time traffic flows pass through

Cloud switches to jointly decrease the network side-effects of flow rerouting and energy consumption of the FNs.

In this way, the energy consumption and the reliability of the selected paths are optimized, while the QoS constraints are met and the network congestion is minimized. Similarly, Oma et al. [39] propose a model where disconnected nodes in a fault-tolerant tree-based FC are partitioned into groups, and FNs in each group are connected to a different candidate node to reduce the energy consumption of each new parent node. Selecting multiple candidate parents for recovery algorithm for each disconnected node results in the reduction of the electric energy consumption of each new parent. In like manner, Rafi et al. [40] combine Low-Energy Adaptive Clustering Hierarchy (LEACH) protocol with Dijkstra's Algorithm under a Cloud environment, which optimizes the energy consumption based on shortest path selection. The algorithm includes load balancing by picking an appropriate cluster head (CH) node among its alternates by calculating its traffic situation with the sink/FN or Cloud. The simulation results show the algorithm improves the network efficiency as well as the durability of the whole network.

Until now, another solution is the energy-balancing scheduling algorithm which is the different approach for optimizing energy efficiency in FC. Luo et al. [41] propose a task scheduling algorithm for energy balancing in a multi-Cloud to multi-Fog architecture. The algorithm is based on the transmission energy consumption of TNs and uses a dynamic threshold strategy to schedule requests in real-time, thereby guaranteeing the energy balancing of TNs without increasing the transmission delay. Experimental results illustrate the proposed algorithm reduces service latency, improves FN efficiency, and prolong TNs life cycle through energy balancing. Akbar and Jamalipour [42] propose a new energy-aware algorithm for energy balancing which comprises of optimization models to reduce transmission power and transmission rate of TNs and delay of the network and topology between each TN and FN is defined to find the best FN for serving the TN, while the energy balancing among all FNs is guaranteed. Simulation results show the proposed algorithm can yield efficient energy balancing among all FNs.

Still more, proactive caching can contribute to solving an energy minimization problem. Xing et al. [43] use proactive caching and correlated task arrival patterns to formulate a long-term weighted-sum energy minimization problem with a three-slot correlation to jointly optimize computation offloading policies and caching decisions subject to rigorous per-slot deadline constraints. Because the proposed mixed-integer problem is again NP-hard, they used semi-definite relaxation as a theoretical upper bound and based on the offline solution, a sliding-window based online algorithm under arbitrarily distributed prediction error was proposed., the Effectiveness of computation caching as well the sliding-window algorithm is verified by numerical examples by comparison with several benchmarks.

Saraswat et al. [44] assume that an Edge, Fog, and Cloud layers compute the periodic task within the specific deadline. Using the total delay and energy consumption, the fractions of the task that are computed at each layer are derived to reduce energy consumption. The impact of the data size, network topologies, deadline, and characteristics of the sensors on the energy consumption, delay, and accuracy of the system is studied. Adhikari and Gianey [45] propose an offloading strategy that uses the Firefly algorithm for finding an optimal computing device based on two QoS metrics such as energy consumption and computational time. The main objectives of this strategy are to minimize the computational time and the energy consumption of the Fog applications with minimum delay. Over comparisons, computational time, energy consumption, CO2 emission, and temperature emission of the Firefly algorithm is proved operative.

Anzanpour et al. [46] propose dynamic multi-goal management for the energy dissipation of IoT-based wearable systems. The health status of the user of a wearable device, the continuity of monitoring, and the accuracy of collected data are parameters considered in the goal hierarchy to select a proper system management policy at run-time to achieve the most significant goal at a given time. The evaluation of the developed prototype shows that the proposed system can reduce power consumption and prevent data loss due to battery shortage.

IV: RESOURCE MANAGEMENT TECHNIQUES BASED ON AUTO-SCALING

In section, we briefly explain avail some of the proposed different Auto-scaling approaches that avoid the wastage of unnecessary Resource utilization and dealing with variable traffic patterns, providing more flexibility than once managed well, proper management of network resources can be achieved. When it comes to energy efficiency in line with resource allocation is one of the vigorous goalmouths in contemporary diurnal distributed dynamic computing since it reduces costs. Researchers often endeavor to create ideal arrangements for an asset the executives in the computing including Fog, Cloud, Edge, Mobile Edge, among others that have several parts, like, the task planning virtual machine position, remaining task at hand combination.

Table 2. A Summarized Resource Management Techniques Based on Auto- Scaling

Proposed Approaches	**Problems / Challenges Associated**	**Research loophole/ Future Research**	**Reference**
Self-adaptive making approach	Decision making	Investigate how the latency of scaling actions. Decision-making process Extending the work for managing energy in cloud	[82]
Dynamic resource scaling framework	Resource Wastage Fast data handling	Investigate efficient strategies for migrating the system from the current resource configuration to the new one recommended by DRS	[83,90]
Proactive virtual resource management framework	Traffic increment Decrements of resources Latency	Scaling of applications like databases Examine framework in the real large-scale cloud environment	[84,89]
Auto-scaling Performance Measurement Tool (APMT)	Poor manageability Synchronization challenges Performance issues	To acquire highly performance recipes to use auto-scaling solutions and policies and combinations	[85,88]
Cloud resource management frame-work	Quality of Services Grade of Service Performance Resource Wastage	Predictive auto-scaling techniques to provide quantitative resource scaling recommendation	[86,87]

There are several Auto-scaling methods used to Auto-scaling of Resources management among which they include Self-Adaptive Trade-off Decision Making for Auto-scaling Cloud-Based Services, it dynamically and adaptively adjust its behaviors to (a) adjust and determine the trade-offs decisions at runtime, and (b) extract the decisions that produce well-compromised trade-offs regarding all related objectives.

Meta-heuristic Algorithms (MAs) in the cloud computing, design Multi-Objective Ant Colony Optimization (MOACO) to search the optimal before near-optimal trade-offs decisions for cloud auto-scaling, the evaluation of decisions' overall quality for all objectives, the MOACO are encouraged to explore more information approximately the trade-offs surface that set aside computational efforts.

Other methods do include Auto-Scaling for Real-Time Stream Analytics (DRS), Multilayered Cloud Applications Auto-scaling Performance Estimation (APMT), A Proactive Virtual Resource Management Framework in Cloud (PRMRAP), Cloud Resource Management with Turnaround Time-Driven Auto-Scaling (TTDA).

The Self-adaptive decision assembly tactic or approached discussed in [82], discloses the effectiveness of style over the others approaches, provides better quality of trade-offs, and significantly smaller desecration of the necessities and optimization of the quality of services. The authors in [83] depicted that DRS is proficient in spotting sub-optimal resource allocation and making rapid and effective resource allocation and intelligence to handle fast data arrival and delivery of effect on time.

This approach of PRMRAP [84] shows that the faster and cost-effectiveness since PRMRAP considers VM's vertical resizing and horizontal resizing both. Further, APMT approached well detailed in the [85] revealed that the Auto-scaling performance measurement problem has been resolved and it is faster in construct to the early mentioned approaches. Lastly on the list method showed guarantee constant service quality and the proposed schedule-based mechanism is a better choice when workload variation can be forecast and also avoid wastage of network resources [86].

V. CONCLUSION

This paper mainly surveyed and discussed network resource allocation and resource management efficiency aimed at the future computing setups in approach to better resource utilization and management. We presented some entities that are involved to manage the network resources and their support in organizations carrying out their tasks, it has been concluded that machine learning has been applied in the number of network tasks.

We compressively surveyed all current studies on resource allocation, management, and focused approach to energy efficiency. A comparative order of the proposed techniques that are identified through evolving overviews by investigating other state-of-the-art recommendations is further presented. This study provided a deep insight into proposed methods, loopholes that will aid researchers to design, model, or create novel frameworks that focus on the better options of resource management in contrast to the existing methods.

Acknowledgments

The authors acknowledge the Department of Computer Engineering, Electrical Engineering, for their extended assistance in editing the paper to this better quality.

REFERENCE

[1] J. Tang, B. Shim, T. Q. Quek, "Service Multiplexing and Revenue Maximization in Sliced C-RAN Incorporated With URLLC and Multicast eMBB," IEEE J. Sel. Areas Commun., vol. 37, no. 4, pp. 881-895, Apr. 2019.

[2] H. Hantouti, N. Benamar, T. Taleb and A. Laghrissi, "Traffic Steering for Service Function Chaining," IEEE Commun. Surveys Tuts., vol. 21, no. 1, pp. 487-507, 2019.

[3] X. Cheng, Y. Wu, G. Min, and A. Y. Zomaya, "Network Function Virtualization in Dynamic Networks: A Stochastic Perspective," IEEE J. Sel. Areas Commun., vol. 36, no. 10, pp. 2218-2232, Oct. 2018.

[4] V. Nguyen, A. Brunstrom, K. Grinnemo, and J. Taheri, "SDN/NFV-Based Mobile Packet Core Network Architectures: A Survey," IEEE Commun. Surveys Tuts., vol. 19, no. 3, pp. 1567-1602, 2017.

[5] R. Urgaonkar, U. C. Kozat, K. Igarashi, and M. J. Neely, ―Dynamic resource allocation and power management in virtualized data centers, ‖ in Network Operations and Management Symposium (NOMS), 2010 IEEE. IEEE, 2010, pp. 479–486.

[6] Hämäläinen, S., Sanneck, H., Sartori, C.: LTE self-organizing networks (SON): Network Management Automation for Operational Efficiency. Wiley, Chichester (2012)

[7] Pearl, J.: Causality: Models, Reasoning and Inference, 2nd edn. Cambridge University Press, New York (2009)

[8] M. Balmakhtar and A. Rajagopal, Virtual network function (VNF) resource management in a software-defined network (SDN). 2019.

[9] M. Mosahebfard, J. Vardakas, K. Ramantas, and C. Verikoukis, "SDN/NFV-based network resource management for converged optical-wireless network architectures."

[10] S. Mostafavi and W. Shafik, "Fog Computing Architectures, Privacy and Security Solutions," J. Commun. Technol. Electron. Comput. Sci., vol. 24, pp. 1–14, 2019.

[11]W. Shafik and S. M. Matinkhah, "Admitting New Requests in Fog Networks According to Erlang B Distribution," in 2019 27th Iranian Conference on Electrical Engineering (ICEE), 2019, pp. 2016–2021.

[12]W. Shafik and S. M. Matinkhah, "How to use Erlang B to determine the blocking probability of packet loss in wireless communication," in presented at the 13th Symposium on Advances in Science & Technology, 2018.

[13]M. Paliwal and D. Shrimankar, "Effective Resource Management in SDN enabled Data Center Network based on Traffic Demand," IEEE Access, 2019.

[14]Parikh, S., Patel, N., Prajapati, H.: Resource management in cloud computing: classification and taxonomy. CoRR (2017)

[15]Jennings, B., Stadler, R.: Resource management in clouds: survey and research challenges.J. Netw. Syst. Manage. 23, 567–619 (2015)

[16] Li, H., Venugopal, S.: Using RL for controlling an elastic web application hosting platform.In: International Conference on Automatic Computing, pp. 205–208 (2011)

[17] Rao, J., Bu, X., Xu, C.-Z., Wang, K.: A distributed self-learning approach for elastic provisioning of virtualized cloud resources. In: 19th Annual IEEE International Symposium on Modelling, Analysis, and Simulation of Computer and Telecommunication Systems, pp. 45–54 (2011)

[18] W. Shafik, M. Matinkhah, M. Asadi, Z. Ahmadi, and Z. Hadiyan, "A Study on Internet of Things Performance Evaluation," J. Commun. Technol. Electron. Comput. Sci., vol. 28, pp. 1–19, 2020.

[19]A. M. Sattar, Ö. F. Ertuğrul, B. Gharabaghi, E. A. McBean, and J. Cao, "Extreme learning machine model for water network management," Neural Comput. Appl., vol. 31, no. 1, pp. 157–169, 2019.

[20]S. Yan et al., “Field Trial of Machine-Learning-Assisted and SDN-Based Optical Network Management,” in Optical Fiber Communication Conference, 2019, pp. M2E–1.

[21] Gans, N., Koole, G., Mandelbaum, A.: Telephone call centers: tutorial, review, and research prospects. Manuf. Serv. Oper. Manage. 5(2), 79–141 (2003)

[22] W. Shafik and S. M. Matinkhah, “Privacy Issues in Social Web of Things,” in 2019 5th International Conference on Web Research (ICWR), 2019, pp. 208–214.

[23] A. Yousefpour, G. Ishigaki, and J. P. Jue, “Fog computing: Towards minimizing delay in the internet of things,” in 2017 IEEE international conference on edge computing (EDGE), 2017, pp. 17–24.

[24] T. Li, H. Salah, X. Ding, T. Strufe, F. H. Fitzek, and S. Santini, “INFAS: In-Network Flow mAnagement Scheme for SDN Control Plane Protection,” in 2019 IFIP/IEEE Symposium on Integrated Network and Service Management (IM), 2019, pp. 367–373.

[25]C. Hernández-Chulde and C. Cervelló-Pastor, “Intelligent Optimization and Machine Learning for 5G Network Control and Management,” in International Conference on Practical Applications of Agents and Multi-Agent Systems, 2019, pp. 339–342.

[26] A. K. Singh, V. Dravid, and Y. K. Bae, “Priority based resource management in a network functions virtualization (nfv) environment,” Jun-2019.

[27] Network Functions Virtualisation – Update White Paper. ETSI (2013)

[28] Evolving Mobile Core to Being Cloud Native. Cisco White Paper (2017)

[29] Sesia, S., Toufik, I., Baker, M.: LTE - The UMTS Long Term Evolution: From Theory to Practice, 2nd edn. Wiley, Chichester (2011)

[30] Mann, Z.A.: Allocation of virtual machines in cloud data centers - a survey of problem models and optimization algorithms. ACM Comput. Surv. 48(1), 11 (2015)

[31] Manvi, S.S., Shyam, G.K.: Resource management for Infrastructure as a Service (IaaS) in cloud computing: a survey. J. Netw. Comput. Appl. 41, 424–440 (2014)

[32] Gans, N., Koole, G., Mandelbaum, A.: Telephone call centers: tutorial, review, and research prospects. Manuf. Serv. Oper. Manage. 5(2), 79–141 (2003)

[33] Defraeye, M., Van Nieuwenhuyse, I.: Staffing and scheduling under nonstationary demand for service: a literature review. Omega 58, 4–25 (2016)

[34] Liu, Y., Watt, W.: Stabilizing customer abandonment in many-server queues with timevarying arrivals. Oper. Res. 60(6), 1551–1564 (2012)

[35] Buyukkaramikli, N.C., van Ooijen, H.P., Bertrand, J.W.: Integrating inventory control and capacity management at a maintenance service provider. Ann. Oper. Res. 231(1), 185–206 (2015)

[36] S. Saha and A. Mitra, “Towards Exploration of Green Computing in Energy Efficient Optimized Algorithm for Uses in Fog Computing,” in ICICCT 2019 – System Reliability, Quality Control, Safety, Maintenance and Management, 2020, pp. 628–636.

[37] E. Ahvar, A. Orgerie, and A. Lébre, “Estimating Energy Consumption of Cloud, Fog and Edge Computing Infrastructures,” IEEE Trans. Sustain. Comput., pp. 1–1, 2019.

[38]T. Nguyen Gia et al., “Energy efficient fog-assisted IoT system for monitoring diabetic patients with cardiovascular disease,” Future Gener. Comput. Syst., vol. 93, pp. 198–211, Apr. 2019.

[39]S.M. Matinkhah and W. Shafik, “Broadcast Communication Analysis for 5G Media Radio Access Networks,” in 16th Conference on Broadcast and Exhibition.

[40] J. Jiang et al., “Sustainability Analysis for Fog Nodes with Renewable Energy Supplies,” IEEE Internet Things J., vol. 6, no. 4, pp. 6725–6735, Aug. 2019.

[41] A. Mukherjee, D. De, and R. Buyya, “E2R-F2N: Energy-efficient retailing using a femtolet-based fog network,” Softw. Pract. Exp., vol. 49, no. 3, pp. 498–523, 2019.

[42] H. A. Alharbi, T. E. H. Elgorashi, and J. M. H. Elmirghani, “Energy Efficient Virtual Machine Services Placement in Cloud-Fog Architecture,” ArXiv190410224 Cs, Apr. 2019.

[43] Y. Liu, K. Xiong, Y. Zhang, L. Zhou, F. Lin, and T. Liu, “Multi-Objective Optimization of Fog Computing Assisted Wireless Powered Networks: Joint Energy and Time Minimization,” Electronics, vol. 8, no. 2, p. 137, Feb. 2019.

[44] F. S. Abkenar, Y. Zeng, and A. Jamalipour, “Energy Consumption Tradeoff for Association-Free Fog-IoT,” in ICC 2019 - 2019 IEEE International Conference on Communications (ICC), 2019, pp. 1–6.

[45] A. J. Kadhim and S. A. H. Seno, “Energy-efficient multicast routing protocol based on SDN and fog computing for vehicular networks,” Ad Hoc Netw., vol. 84, pp. 68–81, Mar. 2019.

[46] A. Mukherjee, P. Deb, D. De, and R. Buyya, “IoT-F2N: An energy-efficient architectural model for IoT using Femtolet-based fog network,” J. Supercomput., Jun. 2019.

[47] W. Fang, W. Zhang, W. Chen, Y. Liu, and C. Tang, “TME2R: Trust Management-Based Energy Efficient Routing Scheme in Fog-Assisted Industrial Wireless Sensor Network,” in 5G for Future Wireless Networks, 2019, pp. 155–173.

[48] M. Mishra, S. K. Roy, A. Mukherjee, D. De, S. K. Ghosh, and R. Buyya, “An energy-aware multi-sensor geo-fog paradigm for mission critical applications,” J. Ambient Intell. Humaniz. Comput., Sep. 2019.

[49] R. Vales, J. Moura, and R. Marinheiro, "Energy-aware and adaptive fog storage mechanism with data replication ruled by spatio-temporal content popularity," J. Netw. Comput. Appl., vol. 135, pp. 84–96, Jun. 2019.

[50] Y. Cen, Y. Cen, K. Wang, and J. Li, "Energy-Efficient Nonuniform Content Edge Pre-Caching to Improve Quality of Service in Fog Radio Access Networks," Sensors, vol. 19, no. 6, p. 1422, Jan. 2019.

[51] A. Toor et al., "Energy and performance aware fog computing: A case of DVFS and green renewable energy," Future Gener. Comput. Syst., vol. 101, pp. 1112–1121, Dec. 2019.

[52] M. Scarpiniti, E. Baccarelli, and A. Momenzadeh, "VirtFogSim: A Parallel Toolbox for Dynamic Energy-Delay Performance Testing and Optimization of 5G Mobile-Fog-Cloud Virtualized Platforms," Appl. Sci., vol. 9, no. 6, p. 1160, Jan. 2019.

[53] S. M. Matinkhah, W. Shafik, and M. Ghasemzadeh, "Emerging Artificial Intelligence Application: Reinforcement Learning Issues on Current Internet of Things," in 2019 16th international Conference in information knowledge and Technology (ikt2019), p. 2019.

[54] D. Rahbari and M. Nickray, "Low-latency and energy-efficient scheduling in fog-based IoT applications," Turk. J. Electr. Eng. Comput. Sci., vol. 27, no. 2, pp. 1406–1427, Mar. 2019.

[55] I. S. M. Isa, M. O. I. Musa, T. E. H. El-Gorashi, and J. M. H. Elmirghani, "Energy Efficient and Resilient Infrastructure for Fog Computing Health Monitoring Applications," ArXiv190401732 Cs, Apr. 2019.

[56] W. Shafik, S. M. Matinkhah, and M. Ghasemazade, "Fog-Mobile Edge Performance Evaluation and Analysis on Internet of Things," J. Adv. Res. Mob. Comput., vol. 1, no. 3.

[57] Y. Dong, S. Guo, J. Liu, and Y. Yang, "Energy-Efficient Fair Cooperation Fog Computing in Mobile Edge Networks for Smart City," IEEE Internet Things J., pp. 1–1, 2019.

[58] D.-N. Vu et al., "Joint energy and latency optimization for upstream IoT offloading services in fog radio access networks," Trans. Emerg. Telecommun. Technol., vol. 30, no. 4, p. e3497, 2019.

[59] Q. Li, J. Zhao, Y. Gong, and Q. Zhang, "Energy-efficient computation offloading and resource allocation in fog computing for Internet of Everything," China Commun., vol. 16, no. 3, pp. 32–41, Mar. 2019.

[60] R. Ma, A. A. Alahmadi, T. E. H. El-Gorashi, and J. M. H. Elmirghani, "Energy Efficient Software Matching in Vehicular Fog," ArXiv190402592 Cs, Apr. 2019.

[61] J. Kim, T. Ha, W. Yoo, and J. Chung, "Task Popularity-Based Energy Minimized Computation Offloading for Fog Computing Wireless Networks," IEEE Wirel. Commun. Lett., vol. 8, no. 4, pp. 1200–1203, Aug. 2019.

[62] T. H. L. Dinh, M. Kaneko, and L. Boukhatem, "Energy-Efficient User Association and Beamforming for 5G Fog Radio Access Networks," in 2019 16th IEEE Annual Consumer Communications Networking Conference (CCNC), 2019, pp. 1–6.

[63] X. Chen, Y. Zhou, B. He, and L. Lv, "Energy-efficiency fog computing resource allocation in cyber physical internet of things systems," IET Commun., vol. 13, no. 13, pp. 2003–2011, May 2019.

[64] I. A. Ridhawi, M. Aloqaily, Y. Kotb, Y. Jararweh, and T. Baker, "A Profitable and Energy-Efficient Cooperative Fog Solution for IoT Services," IEEE Trans. Ind. Inform., pp. 1–1, 2019.

[65] A. Khalid and N. Javaid, "Coalition based game theoretic energy management system of a building as-service-over fog," Sustain. Cities Soc., vol. 48, p. 101509, Jul. 2019.

[65] A. Mebrek, M. Esseghir, and L. Merghem-Boulahia, "Energy-Efficient Solution Based on Reinforcement Learning Approach in Fog Networks," in 2019 15th International Wireless Communications Mobile Computing Conference (IWCMC), 2019, pp. 2019–2024.

[66] S. Shukla, D. Ghosal, K. Wu, A. Sim, and M. Farrens, "Co-optimizing Latency and Energy for IoT services using HMP servers in Fog Clusters," in 2019 Fourth International Conference on Fog and Mobile Edge Computing (FMEC), 2019, pp. 121–128.

[67]S. Chen, Y. Zheng, K. Wang, and W. Lu, "Delay Guaranteed Energy-Efficient Computation Offloading for Industrial IoT in Fog Computing," in ICC 2019 - 2019 IEEE International Conference on Communications (ICC), 2019, pp. 1–6.

[68]H. Sun, H. Yu, G. Fan, and L. Chen, "Energy and time efficient task offloading and resource allocation on the generic IoT-fog-cloud architecture," Peer--Peer Netw. Appl., Jul. 2019.

[69] M. M. Tajiki, M. Shojafar, B. Akbari, S. Salsano, M. Conti, and M. Singhal, "Joint failure recovery, fault prevention, and energy-efficient resource management for real-time SFC in fog-supported SDN," Comput. Netw., vol. 162, p. 106850, Oct. 2019.

[70]M. J. Usman et al., "Integrated Resource Allocation Model for Cloud and Fog Computing: Toward Energy-Efficient Infrastructure as a Service (IaaS)," in Advances on Computational Intelligence in Energy: The Applications of Nature-Inspired Metaheuristic Algorithms in Energy, T. Herawan, H. Chiroma, and J. H. Abawajy, Eds. Cham: Springer International Publishing, 2019, pp. 125–145.

[71] W. Shafik, S. M. Matinkhah, and M. Ghasemzadeh, "Internet of Things-Based Energy Management, Challenges, and Solutions in Smart Cities," J. Commun. Technol. Electron. Comput. Sci., vol. 27, pp. 1–11, 2020.

[72] T. Wang, L. Qiu, G. Xu, A. K. Sangaiah, and A. Liu, "Energy-efficient and Trustworthy Data Collection Protocol Based on Mobile Fog Computing in Internet of Things," IEEE Trans. Ind. Inform., pp. 1–1, 2019.
[73] R. Oma, S. Nakamura, D. Duolikun, T. Enokido, and M. Takizawa, "Energy-Efficient Recovery Algorithm in the Fault-Tolerant Tree-Based Fog Computing (FTBFC) Model," in Advanced Information Networking and Applications, 2020, pp. 132–143.
[74] A. Rafi, Adeel-ur-Rehman, G. Ali, and J. Akram, "Efficient Energy Utilization in Fog Computing based Wireless Sensor Networks," in 2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), 2019, pp. 1–5.
[75]J. Luo et al., "Container-based fog computing architecture and energy-balancing scheduling algorithm for energy IoT," Future Gener. Comput. Syst., vol. 97, pp. 50–60, Aug. 2019.
[76] F. S. Abkenar and A. Jamalipour, "EBA: Energy Balancing Algorithm for Fog-IoT Networks," IEEE Internet Things J., vol. 6, no. 4, pp. 6843–6849, Aug. 2019.
[77] H. Xing, J. Cui, Y. Deng, and A. Nallanathan, "Energy-Efficient Proactive Caching for Fog Computing with Correlated Task Arrivals," in 2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), 2019, pp. 1–5.
[78]S. Saraswat, H. P. Gupta, T. Dutta, and S. K. Das, "Energy Efficient Data Forwarding Scheme in Fog Based Ubiquitous System with Deadline Constraints," IEEE Trans. Netw. Serv. Manag., pp. 1–1, 2019.
[79] M. Adhikari and H. Gianey, "Energy efficient offloading strategy in fog-cloud environment for IoT applications," Internet Things, vol. 6, p. 100053, Jun. 2019.
[80] A. Anzanpour, H. Rashid, A. M. Rahmani, A. Jantsch, N. Dutt, and P. Liljeberg, "Energy-efficient and Reliable Wearable Internet-of-Things through Fog-Assisted.
[81]S. M. Matinkhah and W. Shafik, "Smart Grid Empowered By 5G Technology," in 2019 Smart Grid Conference (SGC), 2019, pp. 1–6.
[82] Tao Chen"Self-Adaptive Trade-off Decision Making for Auto-scaling Cloud-Based Services" IEEE TRANSACTIONS ON SERVICES COMPUTING, VOL. 10, NO. 4, JULY/AUGUST 2017.
[83] Tom Z. J. Fu, Jianbing Ding, Richard T. B. Ma, Marianne Winslett, Yin Yang, and Zhenjie Zhang"DRS: Auto-Scaling for Real-Time Stream Analytics" IEEE/ACM TRANSACTIONS ON NETWORKING, VOL. 25, NO. 6, DECEMBER 2017.
[84]Qizhi Zhang, Haopeng Chen, Zhida Yin"PRMRAP: A Proactive Virtual Resource Management Framework in Cloud" 2017 IEEE 1st International Conference on Edge Computing.
[85] Anshul Jindal, Vladimir Podolskiy and Michael Gerndt"Multi-layered Cloud Applications Auto-scaling Performance Estimation" 2017 IEEE 7th International Symposium on Cloud and Service Computing.
[86] W. Shafik and S. A. Mostafavi, "Knowledge Engineering on Internet of Things through Reinforcement Learning," Int. J. Comput. Appl., vol. 975, p. 8887.
[87]Amazon Web Services, Inc. AWS CloudFormation_Con_guration Management & Cloud Orchestration. [Online]. Available: https://aws.amazon com/cloudformation/
[88] J. Dejun, G. Pierre, and C.-H. Chi, ``EC2 performance analysis for resource provisioning of service-oriented applications," in Proc. Int. Conf. Service- Oriented Comput., Berlin, Germany, 2009, pp. 197_207.
[89] H. Andrade, B. Gedik, K.-L. Wu, and P. S. Yu, "Processing high data rate streams in system S," J. Parallel Distrib. Comput., vol. 71, no. 2, pp. 145–156, Feb. 2011. system S," J. Parallel Distrib. Comput., vol. 71, no. 2, pp. 145–156, Feb. 2011.