Resource optimisation in 5G and Internet-of-Things networking

Babatunde S Awoyemi[1*], Attahiru S Alfa[1,2] and Bodhaswar TJ Maharaj[1]

[*]Correspondence:
awoyemibabatunde@gmail.com
[1]Department of Electrical, Electronic and Computer Engineering, University of Pretoria, South Africa
Full list of author information is available at the end of the article

Abstract

Fifth generation (5G), the currently evolving communication standard, promises better performance in terms of capability, capacity, speed, latency, etc. than recent technologies such as WiMax, LTE and LTE-Advanced. Similarly, the internet-of-things (IoT), the newly developing internet computing paradigm, has the potential for providing seamless, efficient human-device and device-device communication and connectivity. Both 5G and IoT technologies are definite key players in achieving a smart, interconnected world. However, one great limitation is that the resources needed to drive 5G and IoT technologies are extremely limited. To address this challenge, efficient solution models that optimise the use of the scarce resources are required. In this paper, an investigation into the various optimisation approaches that are being explored for addressing resource problems in 5G and IoT is carried out. The solution approaches are categorised and strengths and weaknesses are revealed, while new and exciting research directions are discussed. One of the research areas identified, namely, the aspect of spectrum availability, is addressed. In addressing the spectrum scarcity problem of 5G and IoT, a solution model is developed whereby an allotted spectrum is employed by two networks simultaneously. The results obtained from the analysis show that with such arrangement, a marked improvement in resource usage and overall productivity of the 5G and IoT network is achievable.

Keywords: Fifth-generation; internet-of-things; resource optimisation; spectrum management; QoS provsioning

1 Introduction

Recent developments in communication and computer networking have brought about two new, exciting and highly promising paradigms - the fifth generation (5G) of wireless communication and the internet-of-things (IoT) computer networking and connectivity. A close look at these technologies reveals that 5G and IoT will be crucial and quite instrumental, both playing vital roles in the design and eventual realisation of a smart, interconnected, interdependent and highly productive world. It is becoming increasingly clear that most of the newly emerging and quickly evolving concepts of e-health, e-transport, e-banking, e-farming, e-security, etc. that are currently being designed and developed to drive the realisation of our smart cities will rely heavily on the successful roll-out of 5G and IoT technologies.

For 5G, the technical requirements of speed in the order of 1 Gbps and latency of less than 1 ms, as well as the promise of near-global coverage and always-on reliability have all come together in defining this new generational shift in wireless communication [1]. Also, 5G networks - in form, content and reach - are making laudable promises of providing applications that would have very high social and economic value and far-reaching impact [2]. Unequivocally then, 5G is distinguishing itself as one of the most invaluable tools that will help to achieve the much-talked-about, highly anticipated hyper-connected world [3]. It is exciting to note that since the concept and prospects of 5G emerged, research interests have been astonishingly high among all telecommunication stakeholders, especially in the academia and the industry.

Just as 5G is currently attracting global research interests among major stakeholders, IoT - the recently developed and still evolving internet connectivity paradigm - is attracting almost equal, if not more, attention and focus by various interest groups in the academia and industry in recent times [4]. Simply explained, IoT describes a novel and highly substantial internet reality for the near future - the possibility of simultaneously and seamlessly interconnecting several objects or 'things' (devices, machines, structures, buildings, gadgets etc.) through the internet to facilitate the provision of effective and efficient autonomous services, with as much less human intervention and/or participation as possible [5].

As promising as 5G and IoT are, a potential limitation to their prospects is the scarcity of resources for driving their operations. Resource allocation (RA), utilisation and optimisation thus constitute one of the most active research focus areas in 5G and IoT. In this paper, a comprehensive study on 5G and IoT is undertaken with keen interest in exploring the various approaches being employed in addressing their resource problems. Hence, in the paper, solution approaches are categorised, similarities and differences identified, strengths and weaknesses revealed, while new and exciting research directions for improving resource solutions for 5G and IoT networking are discussed. Furthermore, an attempt at addressing one of the identified research areas, namely, the aspect of spectrum availability for service provisioning, is provided. In seeking to address the spectrum scarcity problem for 5G and IoT, a solution model that makes simultaneous double usage of a spectrum space by two networks feasible is developed and analysed. The results obtained show that with such an arrangement, a marked improvement in resource usage and overall productivity of the 5G and IoT networks is achievable.

The key contributions of this paper are summarised:

- We established the various classes and/or categories of optimisation approaches that have been and/or are being developed and applied for solving RA problems in 5G and IoT networks.
- Considering the peculiarities of 5G and IoT, we identified open-ended problems and discussed areas where optimisation can be further explored in addressing RA problems for 5G and IoT networks.
- We investigated a possible solution to one of the identified problems, that is, the problem of spectrum scarcity for 5G and IoT networking.

The remainder of the paper is organised as follows: in Section II, a clear distinction between 5G and IoT is provided, but also the underlying basis for their joint consideration in this paper; Section III presents the key technologies that will drive 5G and IoT networks; Section IV discusses the concept of resource allocation in 5G and IoT; Section V establishes the general formulation of RA in 5G and IoT; Section VI provides examples of RA problems in 5G and IoT; in Section VII, the various optimisation approaches for solving RA problems in 5G and IoT are explored; Section VIII discusses the areas where optimisation can be applied further in driving 5G and IoT; Section IX provides contemporary research directions in RA for 5G and IoT; in Section X, one of the areas of research identified is critically examined and a possible solution model is developed and analysed and the concluding remarks are presented in Section XI.

2 Distinction between 5G and IoT

For the sake of clarity and to remove any form of ambiguity or disconnect, it is essential to distinguish between the two technologies of interest in this paper. On the one hand, 5G is predominantly the emerging communication standard that is currently being developed (in fact, some 5G 'trials' are already being experimentally deployed in some cities and countries across the globe) to improve significantly on the LTE/LTE-Advanced standards that are at present in use in major parts of the world. The would-be improvements in and promises of 5G are significant enough to necessitate the description of 5G as a generational shift in communication, even though at the moment, there are still ongoing debates and deliberations on the exactness in details of 5G. Nevertheless, massive progress with 5G is being made and clarity, compromises and consensus are fast being reached.

On the other hand, IoT has its foundation on, and is thus an extension of, the internet [6]. In differentiating the traditional internet from the newly developing IoT, the main distinguishing characteristic is that 'things' in IoT is a broader and more encompassing term. Unlike in the regular internet where the connecting devices over which inter-networking occurs are mostly computers, 'things' in IoT are not limited to traditional computers, but include all matters with which it is possible to connect, and/or such matters between which the exchange of information and communication is achievable. To put this differently, 'things' in IoT can be any uniquely identifiable fixed or mobile communicating object that is capable of collecting data, relaying information to other objects, processing relayed information collaboratively and taking action autonomously. The interconnection of these objects, coupled with the embedding of software that collect data and analyses results in a timely manner, makes it possible for IoT to provide intelligent services to human work and life in ways that are far beyond what the traditional internet currently accomplishes for us [7]. Services that IoT provides, such as connecting, communicating, reporting, directing, warning, operating and intercepting, would be seamless, and would also not require the active participation of humans [8].

From the above definitions of 5G and IoT, it is easy to pinpoint that even though these two technologies are in some ways characteristically unique and different, they do have a similar and/or agreeing interest or goal. The striking common interest in both technologies is the drive towards providing quick, efficient, reliable, affordable and accessible wireless communication and connectivity, both among humans and non-human elements, in pursuance of the realisation of the broader goal of creating an interconnected, highly functional smart world [9]. This linkage in interest and/or goal of 5G and IoT is pivotal and is one strong reason for studying them together in this paper.

Apart from the underlying interest that connects 5G and IoT, these two distinct technologies have some intriguing intersections that need to be pointed out. Essentially, while 5G is strictly a telecommunications paradigm, IoT is a much broader technology which, to a large extent, encapsulates major aspects of 5G. IoT can be viewed, in fact, as a progressive combination of the internet, broadband wireless mobile communication (of which 5G is its newest development), wireless sensor networks (WSNs), heterogeneous networks (HetNet) etc. to form a single but powerful interconnected communication network. This clearly shows that even though the goal and prospects of IoT are much broader than, and cannot be single-handedly realised by 5G technology, 5G will definitely be one of the greatest enablers of, and a key driver for the realisation of a viable and vibrant IoT networking [10]. This is because the promises of 5G (which include those of hyper-connecting pre-existing communication technologies such as 3G, 4G, WiMax, Wi-Fi, etc., providing higher coverage and availability, increasing network density in terms of cells and devices, improving data rates and latency etc.) are all critical and will play massive roles in providing a meaningful IoT realisation. On the other hand, because 5G will provide faster data transmission, it would simplify device management of connected devices, which means 5G would lead to significant growth in the IoT. Presently, we know that the latency that is achievable in current cellular networks limits many IoT applications. 5G will solve that, leading to increased efficiency and in effect, an higher potential of having more connected devices. Hence, IoT will greatly impact 5G just as 5G will be very helpful in the successful realisation of IoT. It is important to note that this interesting interplay between 5G and IoT technologies as currently established, as well as the striking similarity in their design goal, plus their unmistakable prospect to achieve the dream of building smart cities and a hyper-connected world, all make it imperative to study them together.

3 Key Technologies in 5G and IoT Networking

The key technologies that help define, describe and distinguish 5G and IoT networks are briefly summarised in this section. It is important to establish these technologies because they help give focus and direction on scope, implementation and further research in the field [11].

For 5G, the key technologies to drive its applications are not entirely new. Although 5G is considered a generational shift, it is being put forward as a development on pre-existing generations of wireless communication. Therefore, most aspects of 5G will have their technical and technological profiling as offshoots of past and present technologies [1, 12]. However, there are recent and currently evolving technologies that may not be very practicable for older generations or standards of communication but that will be key drivers for 5G. Among these evolving technologies, the ones that are most applicable to 5G are briefly discussed.

- **HetNet**: With HetNet, 5G (and other near future wireless communication paradigms) are developed to work in such a way that they can accommodate simultaneously two or more network configurations, standards, radio access technologies, architectures, transmission solutions, base stations, user demands, etc., in order to expand their mobile network capacity [13, 14]. For example, 5G wireless standards will develop on femtocells and/or picocells being able to better work alongside the more traditional macrocells, making it possible for 5G to achieve better productivity than current technologies.

- **Cooperative diversity and relaying techniques**: Cooperative diversity proposes how an improved wireless channel conditioning can be realised for 5G. This is actualised by making the cooperating users (called nodes or relays) to form a 'virtual' multiple input, multiple output (MIMO) arrangement thereby providing the likelihood of diversity gains among spatially dispersed users [15]. The cooperating users use their antennas, as done in conventional MIMO systems, to assist each other in transmitting (or retransmitting) their data to a given destination user. Overall, a significant increase in reliability and capability of the 5G system is realised [16].
- **Massive MIMO and beamforming techniques**: With massive MIMO and improved beamforming capabilities, a large number of extra antennas are incorporated into the network to help focus energy into ever smaller regions of space, thereby providing huge improvements in both throughput and radiated energy efficiency [17]. Massive MIMO and beamforming are great enablers for the development of 5G networks, providing not just energy efficiency, but also improving security, robustness and efficient use of spectrum [18, 19],
- **Cognitive radio networks (CRN)**: The growing interest in CRN for 5G and other next-generation networks is because of its promise of proffering solution to the spectrum challenge by providing dynamic spectrum access and usage capabilities for such networks [20]. 5G will require large bandwidths and invariably, large spectrum spaces. Unfortunately, the spectrum is a limited resource. Scarcity in spectrum availability could potentially limit the productivity of 5G. By employing CRN in 5G, the spectrum challenge will be significantly curtailed [21, 22].
- **Cloud computing**: Cloud computing will be one of the most driving technologies for 5G. The massive amount of data that will be generated and required to be transmitted within the shortest time possible will demand that firstly, a large part of the data has to be stored in the cloud, and secondly, a sizeable part of the data processing must be carried out in the cloud as well [23]. Aspects of 5G such as vehicular autonomy will rely heavily on cloud computing [24, 25].
- **Non-orthogonal multiple access (NOMA)**: Recently developed multiple access technologies such as OFDMA and its variants have been successfully employed in the LTE/LTE-Advanced technologies. However, as promising as these technologies are, they may not be the best access technologies for 5G. The major reason for this is the fact that such technologies have the singular channel allocation limitation, meaning that a subchannel or subcarrier can only be utilised by at most one user in every time slot [26]. More recent works are proposing NOMA as a preferred access technology for 5G because it addresses the singular channel allocation challenge by making it possible for communication resources to be served concurrently to multiple users through non-orthogonal sharing techniques [27, 28].

It is necessary to state that each of the above-mentioned technologies for 5G, in itself, is broad and a substantial amount of research work on each one of them is currently ongoing. Very important work on 5G would actually be to develop designs that can link these technologies together accurately so that their strengths can be jointly annexed for achieving the goals and promises of 5G technology.

For IoT, the key technologies to drive it operations have been identified and classified in a number of ways [29]. However, it seems that the generally accepted classification of the driving technologies for IoT identifies four categories of technology, namely radio-frequency identification (RFID) technology, sensor technology, network communication technology and embedded system technology [30, 31]. These key technologies are briefly discussed.

- **Radio-frequency identification**: 'Things' in IoT must have a unique way in which they can be identified and addressed before they can be effectively connected. RFID is used to achieve this identification. RFID uses radio waves to identify and track items in real-time so as to get important information about their location and status, thus achieving the goal of providing identification for the various IoT objects [32].
- **Sensor technology**: Generally, sensors play a very important role in bridging the gap between the physical world and the information world. Advancement in sensor networks, especially in WSN, is one huge catalyst for driving connectivity among the many objects in IoT. IoT objects have sensors that perceive and collect data from their environment and generate information or raise awareness about the context. This makes it possible for the objects to monitor changes in their environment and to initiate appropriate responses [33, 34, 35]. Sensor technology is thus an important technology for IoT networking.
- **Network communication technology**: Sensors must have a way of relaying their sensed signals or communicating with one another or the network. Network communication technology of sensors consist of short-distance communication techniques as well as wide area network communication technologies. These common network communication technologies of sensors, such as Bluetooth, ZigBee, Wi-Fi, ultra wide band etc. are essential tools for effective IoT networking [36, 37, 38].
- **Embedded system technology**: Embedded system technology is the technology with which objects in IoT are made. It combines computer hardware and software, sensor technology, integrated circuit technology, electronic technology application etc. in producing these smart IoT objects [39]. The objects are made with the capacity to process information, self-configure and make decisions independently. Embedded system technology employs the latest advances in nanotechnology and miniaturisation to provide embedded intelligence in the objects themselves, thus making them extremely smart devices [40].

4 Resource Allocation in 5G and IoT Networking

As already established, both 5G and IoT are frontrunner candidates for driving emerging communication and connectivity networking. This implies that for both 5G and IoT, the expectations of fast, reliable and accessible services are very high. An immediate inference from this is that efficient service provisioning by these technologies would require that resources (such as spectrum, transmit power, bandwidth, data rate, time slot, electrical energy, memory, processor etc.) for driving their operations and achieving their expectations would be abundant and readily available. However, this is not always the case as, in reality, resources for driving wireless communication technologies are usually very limited and/or scarce [41]. It has been clearly stated that the limitation in resource availability is a major threat to new and daring telecommunication paradigms [42]. The reality of non-ubiquity in resource availability for accomplishing the goals of 5G and IoT is therefore an underlying problem of these technologies.

To overcome the limitation of resource scarcity, attempts at devising appropriate RA models/schemes for 5G and IoT are currently ongoing. RA in communication networks describe mechanisms for achieving the utmost productivity for the networks, thus overcoming the limitation in resource availability. The goal of RA in 5G and IoT is to efficiently coordinate the distribution and utilisation of the limited resources so as to achieve the overall best functioning for these networks. Approaches developed for addressing the limitations in resource availability and usability in 5G and IoT have been quite diverse and there does not seem to be a well-established, one-fits-all approach for solving these RA problems. In this paper, a thorough investigation of the various methods and/or approaches that have been developed and employed for addressing RA problems in 5G and IoT is carried out. The approaches identified are categorised and characterised. The strengths and weaknesses of the various approaches are pointed out and ideas for better solutions and practical implementations are then put forward.

5 Problem Formulation of RA in 5G and IoT Networks

RA problems in emerging technologies such as in 5G and IoT have been described as optimisation problems [43]. Optimisation is a well-developed analytical tool for solving a host of problems and has been used broadly in different fields of science, such as in mathematics, operations research, business finances and economics, engineering etc. In optimisation, there is usually an objective to be achieved, either maximising or minimising an entity or a number of entities, and this is always captured in the objective function. Then, there are limiting constraints that must be taken into consideration while seeking to achieve the objective. In solving optimisation problems, the constraints may not be violated, otherwise the solutions to such problems are unusable and unreliable, if they are ever obtained in the first place. The final component of all optimisation problems is the decision variables. The decision variables are the parameters to solve for in order to arrive at (optimal or near-optimal) solutions.

A general form of RA optimisation problem formulation for communication networks such as 5G and IoT has been developed in [43]. The general formulation gives a description of what the objective functions usually are, as well as the constraints and the decision variables, and how they are interconnected and/or interrelated. The formulation is briefly summarised as follows:

Let $\mathbf{p}$ and $\mathbf{q}$ be two vectors of dimensions a and b respectively. Also assume a set of positive integers $I = \{0, 1, 2, ...\}$. Then, assume we need to obtain the values of $\mathbf{p}$ and $\mathbf{q}$ for which a function $f(\mathbf{p}, \mathbf{q})$ is maximum, given that there are a set of constraints $g_i(\mathbf{p}, \mathbf{q}) \leq n_i$, $i = 1, 2, ..., r$, and that each variable is non-negative. The above formulation can be written mathematically as:

$$\max z = f(\mathbf{p}, \mathbf{q}) \tag{1}$$

subject to

$$g_i(\mathbf{p}, \mathbf{q}) \leq n_i, i = 1, 2, ..., r, \tag{2}$$

$$p_j \geq 0,\ j = 1, 2, ..., a, \tag{3}$$

$$q_k \in I,\ k = 1, 2, ..., b. \tag{4}$$

Equation (2) is more simply written as:

$$\mathbf{g}(\mathbf{p}, \mathbf{q}) \leq \mathbf{n},$$

where

$$\mathbf{g}(\mathbf{p}, \mathbf{q}) = \begin{bmatrix} g_1(\mathbf{p}, \mathbf{q}) \\ g_2(\mathbf{p}, \mathbf{q}) \\ \vdots \\ g_r(\mathbf{p}, \mathbf{q}) \end{bmatrix},$$

and $\mathbf{n} = [n_1, n_2, ..., n_r]^T$. If the problem was a minimisation problem, the function $z = f(\mathbf{p}, \mathbf{q})$ could be easily transformed to a form of maximisation function by simply negating the objective function, i.e., $\max w = -f(\mathbf{p}, \mathbf{q})$. From the general formulation given above, equation (1) is the objective function, equations (2)-(4) are the constraints, while p_j and q_k are the decisions variables. As an example, consider a 5G network with an amount of power at the base-station and a number of heterogeneous user devices to be serviced. In that case, equation (1) could be a maximisation of the total network capacity, vector $\mathbf{p}$ could be a set of transmission power for users, vector $\mathbf{q}$ could be subchannel allocation, which would usually take integer values of 0 or 1, and equation (2) could be the interference limit constraint or the power constraint.

6 Examples of RA Optimisation in 5G and IoT Networks

In this section, examples of RA problems in 5G and IoT, and the corresponding optimisation formulations developed for solving them are presented.

A mathematical formulation for RA in 5G using non-orthogonal multiple access for multi-user channel and power allocation was developed in [44]. In the model, a base-station services a set of K users over a bandwidth B divided into a set of $\mathbb{N}$ subchannels. Each subchannel $n \in \mathbb{N}$ can accommodate up to M users, where $M \leq K$ depends on receivers' design complexity and signal processing delay for successive interference cancellation. The goal is to determine which users are multiplexed to the subchannels, as well as to determine the optimal power allocation such that the sum-rate utility of the network is maximised. If $x_{k,n}$ is a binary representation of the channel allocation (1 when a user k has been allocated to a subchannel n and 0 otherwise) and $p_{k,n} \geq 0$ is the power allocated to user k on subchannel n, the optimisation problem is given as:

$$\max \sum_{k\in K}\sum_{n\in\mathbb{N}} R_{k,n} x_{k,n} \tag{5}$$

subject to

$$\sum_{k\in K}\sum_{n\in\mathbb{N}} p_{k,n} \leq P_{\max} \tag{6}$$

$$\sum_{n\in\mathbb{N}} p_{k,n} \leq P_k, \forall k \in K \tag{7}$$

$$\sum_{k\in K} p_{k,n} \leq M, \forall n \in \mathbb{N} \tag{8}$$

where $R_{k,n}$ is the data rate given by:

$$R_{k,n} = \frac{B}{\mathbb{N}} \log_2(1 + SINR_{k,n}), \forall k \in K, \forall n \in \mathbb{N}$$

and $SINR_{k,n}$ is the signal-to-interference-plus-noise ratio of user k on subchannel n. The objective function maximises the sum-rate utility of the system, the constraint in equation (6) restricts the total power consumed by all users over all subchannels to the maximum power available to the system $P_{\max}$, the constraint in equation (7) ensures that each user k can only use power within its assigned limit P_k, the constraint in equation (8) is to limit the maximum number of users multiplexed on each subchannel to no more than M. The formulation is both non-linear and non-convex because of the nature of $R_{k,n}$. Suboptimal solutions were obtained by using Lagrangian duality combined with dynamic programming.

Another practical framework for RA formulation in 5G is provided in [45]. In the work, energy efficiency for 5G networks is optimised through appropriate RA solutions. Two different energy efficiency problems are considered and the optimisation is performed with respect to subchannel assignment and transmit power allocation. Using the same notations as defined in the previous example, the bit/Joule energy efficiency of the kth user η_k is given as:

$$\eta_k = \frac{\sum\limits_{n\in\mathbb{N}} \log_2(1 + SINR_{k,n})}{p_{c,k} + \sum\limits_{n\in\mathbb{N}} p_{k,n}} \tag{9}$$

where $p_{c,k}$ is the circuit power used to operate the transmitter of user k. Given that w_k is the weight attached to user k, the weighted minimum of the energy efficiencies η is defined as:

$$\eta = \min_{k=1,\ldots,K} w_k \eta_k. \tag{10}$$

The energy efficiency of the entire system is measured using the global energy efficiency metric ψ given as:

$$\psi = \frac{\sum_{k \in K} \sum_{n \in \mathbb{N}} \log_2(1 + SINR_{k,n})}{p_c + \sum_{k \in K} \sum_{n \in \mathbb{N}} p_{k,n}} \tag{11}$$

where p_c gives the total circuit power dissipated in the system. ψ can be seen as the benefit-cost ratio of the system, defined as the ratio between the sum achievable rate and the total consumed power. On the other hand, maximising the (weighted) minimum of the energy efficiencies makes it possible to achieve a more fair RA policy. The minimum energy efficiency maximisation is given as:

$$\max_{\{p_{k,n} \geq 0\}_{k,n}} \quad \eta \tag{12}$$

subject to

$$\sum_{n \in \mathbb{N}} p_{k,n} \leq \bar{P}_k, \forall k \tag{13}$$

$$\sum_{n \in \mathbb{N}} \log_2(1 + SINR_{k,n}) \geq \theta_k \forall k \tag{14}$$

while the global energy efficiency maximisation problem is given as:

$$\max_{\{p_{k,n} \geq 0\}_{k,n}} \quad \psi \tag{15}$$

subject to

$$\sum_{n \in \mathbb{N}} p_{k,n} \leq \bar{P}_k, \forall k \tag{16}$$

$$\sum_{n \in \mathbb{N}} \log_2(1 + SINR_{k,n}) \geq \theta_k \forall k \tag{17}$$

where $\bar{P}_k$ and θ_k are the maximum power and minimum achievable rate of user k respectively. Both these problems were solved using fractional programming. To achieve computationally manageable solutions, the tool of sequential convex programming was leveraged, in which case the idea was to solve a sequence of easier problems whose solution converges to a local solution of the original problem.

An example of RA optimisation in IoT is found in [46]. In the model developed and investigated, there are a number of machine-type devices using machine-to-machine (M2M) communication in an LTE-Advanced system. The RA goal is to maximise the random access efficiency while ensuring that the random access delay requirement of the various machine-type devices is met. The random access efficiency R_{eff} shows how efficiently the devices can be accessed (in terms of number of devices successfully accessed) with the resources provided in each communication frame. If a device is contending for random access and there are L random access opportunities and C devices in total, then R_{eff} is written as:

$$R_{eff} = C \times \left(\frac{1}{L}\right)\left(1 - \frac{1}{L}\right)^{C-1}. \tag{18}$$

If Γ is the set of available random access opportunities, $E[D]$ is the random access delay of devices and D_{req} is the minimum delay requirement, the random access efficiency optimisation problem is formulated as:

$$\max R_{eff} = C \times \left(\frac{1}{L}\right)\left(1 - \frac{1}{L}\right)^{C-1} \tag{19}$$

subject to

$$L \in \Gamma \tag{20}$$

$$E[D] \leq D_{req}. \tag{21}$$

R_{eff} is maximised when $L = C$ for any given value of C.

7 Classification of Optimisation Solutions for RA in 5G and IoT Networks

There are a good number of optimisation techniques that have been established and well-documented for solving RA problems [43]. However, the peculiarities and scope of 5G and IoT make the direct application of conventional optimisation techniques improbable, therefore making it imperative to develop specific, efficient optimisation techniques for solving RA problems in 5G an IoT networking. Interestingly, optimisation has been successfully adopted for addressing some of the most complex emerging communication problems, such as in heterogeneous CRN [20]. It can be argued that with the necessary modification and/or adaptation, optimisation can similarly be used to address RA problems in 5G and IoT. This argument is corroborated by the ingenuity provided in several works on RA optimisation solutions for 5G and IoT that are studied in this section. What is also interesting in the study carried out in this section is the different optimisation approaches that are being adapted and employed for addressing RA problems in 5G and IoT. In cases where modifications have been required, peculiarities of 5G and IoT that necessitated such modifications/adaptations and/or were considered in achieving them are highlighted. The various approaches and/or techniques of optimisation that have been used either directly or by some form of adaptation to address RA problems in 5G and IoT are categorised for easy comprehension and comparison.

7.1 Self-Optimisation

Many authors have argued that because of the complexity and broadness of 5G and IoT, the most appropriate optimisation technique adaptable for their RA problems is self-optimisation. With self-optimisation, each device (or user) is enabled to optimise its own resources to provide the needed services and enhance the overall benefit of the larger network [47]. Generally for IoT, self-optimisation helps shift the focus from design and deployment of a single element or a few elements operating autonomously to a large complex ecosystem of a network of autonomous elements. Utility functions and organic computing are often used in distributed autonomic computing systems to achieve self-optimisation. Examples of self-optimisation applications in RA for 5G and IoT are discussed.

In [48], the author established that achieving the desired quality of experience in 5G will require that devices be enabled to self-optimise. The work then proposed the use of neural networks as an efficient technique for adaptive estimation and self-optimisation of the quality perceived by users or user devices in 5G scenarios.

Self-optimisation has been investigated as a solution to one of the major challenges of RA in IoT - the problem of integrating and managing the various large-scale HetNet that combine to make up the IoT. The authors in [35] presented a self-optimising sensor network management system that has the ability to automatically configure itself, while at the same time conserving energy. The developed management system was built to be quite flexible and robust enough to endure severe structural changes. Developing such management systems is critical for effective IoT networking and being able to solve the ensuing RA problems through self-optimisation techniques is an important positive step towards the realisation of the benefits of IoT.

The authors in [49] developed an intelligent system for controlling traffic lights at intersections to help solve the problem of traffic congestion. The problem was studied and solved as a self-optimisation problem in IoT. The system was developed as an IoT system with the monitoring sensors being viewed as the connected things. An RA problem was formulated where the design goal was to optimise the crossroad traffic light time parameters according to the number of vehicles in the traffic, while the objective was that of minimising the time cars spent waiting in intersections. The system also proposed a solution for managing the special situations encountered when a rapid intervention vehicle (such as an ambulance, fire department or police vehicle) crosses an intersection.

In [50], the authors first argued that the static single-application arrangement of WSN would be a serious limitation to its scalability in IoT applications. The authors then developed a self-optimisation platform for finding optimal mapping between applications and resources in the WSN paradigm for IoT. The goal was to propose an architecture for efficiently adopting WSN in IoT when the demands and scale of the network are changing. In their proposition, WSN would be seen and employed as not just a system for sensing tasks but also a platform for dynamically assigning resources. When nodes in WSN are developed as resource providers, they can be used to leverage operational capacity across multiple IoT platforms. Hence, their topology shifts from just node availability to resource utility.

There has been some criticism of the use of self-optimisation in IoT as well. For instance, in [51], the authors argued that IoT, compared to other traditional fields of autonomic computing, would be characterised by an open-ended and highly dynamic ecosystem with variable workloads and resource availability. These characteristics would make it difficult to implement self-awareness and self-optimisation capabilities. Another challenge of self-optimisation is the ability to change a system's behaviour to achieve a desired functionality while still maintaining a balance between quality of service (QoS) provisioning and resource usage.

7.2 Classical Optimisation

Another approach to optimisation that has been extensively used for developing and solving RA problems in 5G and IoT is classical optimisation. In classical optimisation, RA problems in 5G and IoT are developed and solved using well-established optimisation techniques such as linear programming and convex optimisation. The developed RA problems somehow fit almost perfectly into one or more of these classical optimisation techniques so that they are easily solved with as little modification as possible. A number of works that have employed classical optimisation in addressing their RA problems in 5G and IoT are discussed.

Authors in [44] proposed that for 5G networks, the non-orthogonal multiple access scheme would be a preferred multi-user access scheme to the more conventional orthogonal frequency division multiple access (OFDMA) scheme employed in LTE/LTE-Advanced. On that basis, the authors established joint power and channel allocation for 5G as an non-deterministic polynomial-time (NP)-hard optimisation problem. The developed RA problem was solved using the classical optimisation approach of Lagrangian duality and dynamic programming.

In [52], the authors set up an experiment in which a number of sensor nodes in an IoT scenario were deployed to record temperature data. By employing a linear programming model, the collected data was used to predict the temperatures of other sensor nodes accurately. The temperature predictions made it possible to reduce overall node sampling rate for the network significantly. Furthermore, an optimal number of nodes to be deployed was arrived at, thus reducing the amount of energy the sensors consumed.

Authors in [33] addressed the challenge of ensuring the soundness in the quality and quantity of data collected from the sensors in IoT networks through classical optimisation. The challenge was addressed by developing an independent regional connectivity model that guaranteed global connectivity with satisfactory quality of data service. The optimisation model developed had as its decision variables the sensing radius and communication range of the different types of sensors. The overall objective was to minimise the number of sensors used for monitoring a given region, while the major constraint was that of providing a guarantee to a certain degree for the sensing coverage and connectivity region at all times.

7.3 Multi-Objective Optimisation

In multi-objective optimisation, the problem usually presents a number of objectives to be achieved simultaneously. Sometimes, these objectives are rather conflicting, making the resulting RA problem complex and difficult to solve. For instance, the objectives of an RA problem in 5G could be to minimise the time taken in transmitting an amount of data, to minimise the amount of power consumed in transmitting these given data while at the same time maximising the total data transmitted over a period of time. These objectives are somewhat conflicting, as it would be problematic to minimise power and maximise capacity simultaneously. Multi-objective optimisation techniques are used for addressing those kinds of problems.

The authors in [53] explained that because of the coupling in the various objectives of 5G, such as minimising energy and power consumption, maximising average user data rates and throughput of network, the objectives cannot be treated separately.

Hence, achieving the high expectations of 5G requires the development of efficient network systems that can handle all the conflicting 5G objectives simultaneously and arrive at optimal solutions. This implies that a design framework that can handle multiple objectives and can support the search for the best attainable operating point has to be used. A case study of employing massive MIMO systems for downlink in a 5G network was studied. Three conflicting objectives - high average user rates, high average area rates, and high energy efficiency - were simultaneously optimised using multi-objective optimisation.

In [51] the authors proposed a methodology for automating the efficient deployment of IoT applications in the presence of multiple optimisation objectives and variable operational circumstances. The developed model was based on an off-line exploration phase that collects relevant profiling information for optimisation before the actual deployment, and a runtime phase that autonomously adapts the deployment and configuration towards changing operational circumstances. The goal was to autonomously optimise the QoS and other qualitative attributes of IoT applications. Furthermore, the trade-off between coverage and QoS when different configurations are deployed was also investigated.

The authors in [34] developed a multi-objective optimisation model that brings IoT into manufacturing and helps to achieve higher and better productivity. In the developed problem, manufacturing things were embedded with sensors and RFID so they could communicate with each other. The optimisation technique employed developed a real-time data-driven optimisation model, which characterised the loT-based manufacturing execution system succinctly. The objective was to sense the changes in some relevant control parameters and to quickly adjust these parameters based on the sensed real-time manufacturing information, thus eliminating possible disturbances in the production process.

7.4 Heuristics and Meta-heuristics

Heuristics and meta-heuristics have both been widely applied in addressing RA problems in 5G and IoT. Heuristics do not necessarily employ analytical or scientific derivations but only apply logical reasoning for solving particular problems in 5G and IoT. Meta-heuristics are employed for problems that have the possibility of obtaining more than one local ‘optimal’ solution to a given RA problem. Though heuristics and meta-heuristics, in most cases, only proffer suboptimal solutions, such solutions are usually obtained in a much reduced time frame and with much less computational complexity, even for large networks. For these reasons, both heuristics and meta-heuristics have been applied extensively for addressing RA problems. Examples of the use of heuristics and meta-heuristics in solving RA problems in 5G and IoT are discussed.

The authors in [54] modified a real-coded genetic algorithm to optimise both the configuration of base-stations and the amount of energy consumed in a 5G network. In the developed model, users were located in several dense areas while the goal was to provide services to as many users as possible, using the lowest number of base-stations and lowest transmit power by locating those base-stations in optimal positions. The real-coded genetic algorithm used had to be modified because the decision variables (transmit power and location) were continuous values.

The modification was achieved by introducing a base-station crossover rate to shuffle less and by using small standard deviation values.

In [55], a genetic algorithm heuristic was developed and employed to achieve optimal placement and power allocation of RFID devices in an IoT environment. The network topology design developed produced a flexible deployment layout for the RFID readers that optimised both the location and power level of the readers. The authors in [32] introduced a genetic annealing algorithm that could achieve even better RFID topology design for placement of the RFID devices. The results they obtained by the use of the heuristic were shown to outperform existing placement designs.

Objects in IoT need to make their resources available in ways that are quite flexible. In [56], the authors proposed a distributed optimisation protocol based on a consensus algorithm to solve the problem of RA and management in IoT heterogeneous networks. In the proposed protocol, an IoT scenario was created where nodes involved in the same IoT tasks need to adjust their task frequency and buffer occupancy. The work showed that, using the proposed protocol, the network can converge to a solution where resources are homogeneously allocated among clusters of nodes, and overall, network productivity is significantly improved.

Table 1 contains summary of the optimisation solution approaches for RA in 5G and IoT, as already discussed.

8 Areas of Application of Optimisation in RA for 5G and IoT Networking

In this section, a summary of the key areas where RA optimisation is being and/or can be applied in 5G and IoT networks is provided.

1 **Optimising heterogeneity for 5G and IoT networking**: The 5G and IoT networks are going to be heterogeneous in nature. Several aspects of heterogeneity have to be covered: heterogeneous networks, heterogeneous devices and/or objects, heterogeneous user conditions, heterogeneous service requirements etc. Optimisation can be employed for addressing the heterogeneity problem in 5G and IoT, such as allocating an appropriate amount of power or bandwidth to different categories of users, based on their demand or priority profile. For instance, the work in [20] gives an idea of how the various heterogeneous considerations can be factored into one main design and/or formulation and appropriate RA models can be developed to solve the resulting optimisation problems. Both optimal and near-optimal solutions can be achieved by studying the problems' structure to discover clues that can be employed in solving them. Also, heuristics can be developed to solve the resulting RA problems, as achieved in [15].

2 **Application in WSNs**: WSN is an indispensable technology for the successful development and deployment of IoT. In the not too distant future, global 5G and IoT penetration will result in the proliferation of the number of sensor nodes deployed in various locations for collecting 5G and IoT information (the number of sensors required would probably be in their billions or zillions!). Very importantly therefore, sensors that are energy-saving must be developed, simply because the sensor nodes are usually battery-powered and thus have

limited lifespans [33]. Optimisation is important in developing energy-saving sensor models of WSN for IoT application.

3 **Application in RFID**: There are at least four technical areas where optimisation is applicable in RFID technology for IoT, as identified by [57], namely in solving the problem of optimal placement of the RFID readers, in solving the problem of load balancing for the RFID readers, in solving the problem of data allocation for the RFID tags and finally, in solving the problem of collision during tag reading, especially when there are too many tags for a single reader to identify.

4 **Application in embedded system**: Embedded systems design has an intrinsic difficulty - the need to handle a lot of data (for example big images) in a very short time (for example in the case of moving images) [58, 59]. Efficient memory management optimisation models have been and are still being developed to address this and other related problems of embedded system in IoT networking.

5 **Addressing delay problems**: Several kinds of data traffic are generated and transmitted in typical 5G and IoT environments. Moreover, most 5G and IoT services have real-time and reliability expectations. Long delays and/or packet losses in the course of transmitting data can have a significant, undesirable impact on the overall service performance. Unfortunately, network links in 5G and IoT would often be limited in the amount of bandwidth over which transmission could occur, while servers in the processing layer too might also have finite capacities. This is usually the case in practical IoT networks such as in M2M communication. In such cases, machine-type devices seeking to gain access to the network may at times experience delay, loss of packets and/or even connectivity. Therefore, optimising congestion control and queue management is imperative in IoT, if the performance requirements of service provisioning are to be met. The authors in [46] proposed an RA and access control mechanism that addressed network performance degradation caused by these types of occurrences.

6 **Economic aspect of 5G and IoT**: Optimisation has been extensively used in solving economic problems, especially in the social sciences. Interestingly, both 5G and IoT will require sound economic considerations in their design and deployment. For instance, [57] mentioned that the economic efficiency enhancement in RFID applications has to be properly studied. This also applies to the economic implications of changing from the traditional macrocells to microcells and/or femtocells in HetNet for 5G application. Optimisation can be used in analysing these economic considerations and arriving at best decisions on the minimum cost of deployment, time frame for return on investment, pricing etc.

Other aspects of application of optimisation in RA for 5G and IoT that are worth mentioning are increasing the performance of IoT by deploying heavyweight application components on faster hardware, reducing the amount of communication and network latencies between distributed components, and optimising the overall energy consumption of the application components on the different application platforms.

Table 2 gives a summary of the various areas of application of optimisation in RA for 5G and IoT networks.

9 Contemporary Research Directions and Focus Areas

The following areas of RA optimisation in 5G and IoT networking are identified as open-ended problems that still require active and adequate research/investigation, if the prospects of 5G and IoT are to be fully realised:

1 **Spectrum availability for service provisioning**: Radio-frequency spectrum is a necessary resource for effective 5G and IoT networking. The spectrum, as already established, is a scarce/limited resource. The limitation in spectrum availability poses a big challenge to 5G and IoT functionality. New attempts at improving spectrum availability and usability, such as currently being developed in CRN, are imperatives for successful 5G and IoT networking. Similarly, investigating the most appropriate optimisation techniques that can help in improving spectrum allocation for 5G and IoT networks is an important area of research.

2 **Improving sensing capabilities of 'things'**: For IoT especially, efficient deployment will rely heavily on the capability of its 'things' to sense accurately and timeously for quick decisions to be made. Poor sensing will greatly hamper the development of IoT. Hence, a major research focus is that of designing and providing very reliable sensing mechanisms for connected objects in IoT. Importantly, more investigations of optimisation techniques that can be employed to significantly improve the time, resources, number of sensors etc. required for an effective IoT realisation are necessary.

3 **HetNet integration, coordination and management**: IoT can be viewed as an example of HetNet, not only from the perspective of the computation capability of the various objects within the network, but also from the point that networks and communication technologies used for interconnecting their objects and the services they offer are going to be heterogeneous in nature [60]. Both 5G and IoT will run on a number of HetNet designs and applications, as well as a deep reliance on cloud computing, to achieve their ends. However, HetNet in itself has been undergoing a great deal of research and development in recent times. Establishing the right linkages for macrocells, microcells, picocells and femtocells, combining wireless communication standards such as Wi-Fi, 3G, 4G, LTE-Advanced etc. with other technologies like fibre optics and relay networks in generating swift, smooth and seamless communication is what HetNet is seeking to accomplish. The integration of HetNet and cloud computing into both 5G and IoT is therefore crucial. So is the coordination and management of heterogeneity and cloud computing in 5G and IoT networking. These aspects of 5G and IoT are still open-ended research areas. Optimisation is being employed actively in accomplishing many HetNet designs and strategies for 5G and IoT networking.

4 **Optimising system functionality that incorporates scalability and robustness requirements**: Efficient deployment of application components in 5G and IoT networks has been said to be multi-objective optimisation problems. As mentioned earlier, the optimisation objectives in those problems, in

most cases, conflict with one another (for instance, performance versus energy consumption). In such cases, obtaining single solutions that simultaneously optimise the objective can be a herculean task, and resource trade-off has to be made. Authors have had to employ Pareto optimisation in addressing those kinds of problems. However, because IoT is essentially an open-ended ecosystem of heterogeneous resources, this makes the crisp definition of Pareto-optimal solutions difficult because of an incomplete view of the external factors and uncertain circumstances that might influence the optimality of such solutions [51]. Hence, arriving at optimisation solution methods that, alongside Pareto optimisation, can adequately address the kinds and peculiarities of RA problems for 5G and IoT is still an open research topic. Suggestions such as the one in [51] where Pareto optimisation was combined with reinforcement learning to yield acceptable solutions are very good developments.

5 **Improving system security**: Security is of paramount importance in the deployment of both 5G and IoT. Arguably the most demanding of concerns and/or requirements for the widespread realisation of many of the IoT visions is its security [61]. There are a number of threat implications of an expanding IoT, as described in [62]. Hence, it is imperative that the IoT technologies such as the RFID and sensors be adequately protected from malicious attacks. In [31], security concerns of the RFID and the sensor networks are highlighted. The possible RFID security threats identified include replication attack, channel blocking attack, forgery attack, impersonation attack and tampering attack. There are security issues too at the transmission or transport layer. Of serious concern is the possibility of cross heterogeneous networks attack. The application layer also poses security issues, such as being able to select the same database content according to different access, providing user privacy information protection, solving the leakage of information tracking problem, taking the computer forensics, destroying the computer data, protecting electronic products and software intellectual property etc. Optimisation can be employed in strengthening security with the IoT. Optimisation can be useful in providing models that can help in taking quick and decisive action that need to be taken to manage and protect the RFID label identities, security and privacy. Strengthening the transport layers' cross-domain authentication and cross-network authentication would also require employing some optimisation in the network.

6 **Addressing the problems of large communication overheads and efficient data management**: The authors in [63] argued that the traditional approach of migrating raw data to centralised points for data storage and analysis was likely to incur debilitating communication and energy costs, which could affect the environment negatively in the future. Developing models to optimise communication overhead and storage mechanisms are therefore necessary, as these two usually have the most significant impact on the amount of energy the network consumes. Optimisation tools are therefore very relevant in achieving this goal.

Recent research on 5G and IoT networking has been seeking to address and solve the above-mentioned problems. Although this paper, being a survey paper, has

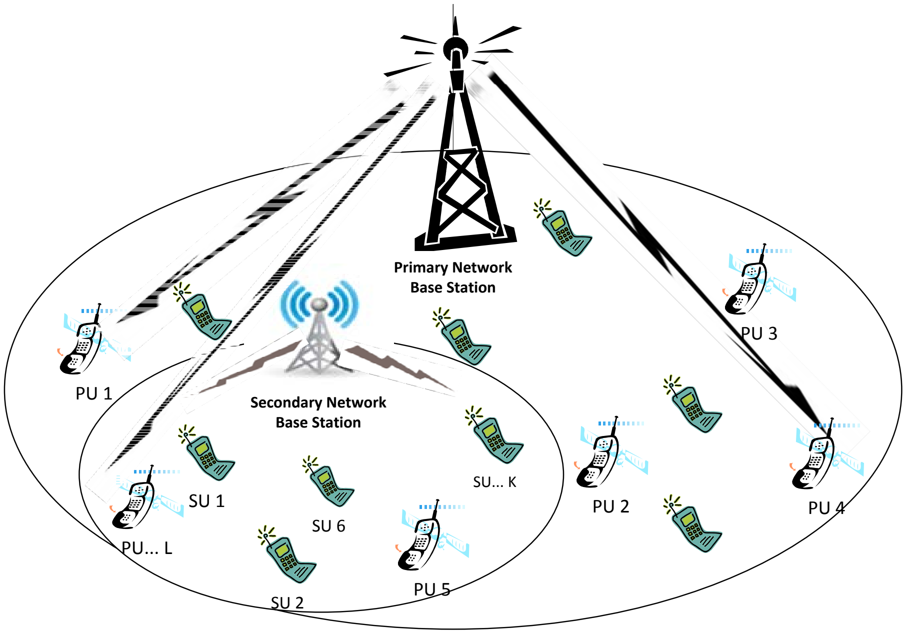

Figure 1 System model for primary-secondary networking in 5G and IoT

focussed on reviewing the progress made so far and opportunities still unfolding with RA optimisation for 5G and IoT, the next section of the paper is our suggestion and/or an attempt at what we understand can be a very promising way of addressing one of the problems identified, namely the problem of spectrum scarcity for effective 5G and IoT functionality. We seek to investigate an appropriate solution model which we believe can be very effective in solving the spectrum scarcity problem for 5G and IoT networking.

10 Improving Spectrum Availability for Service Provisioning in 5G and IoT

In this section, we attempt to develop and analyse a spectrum allocation model that can effectively address the problem of spectrum scarcity for 5G and IoT networking. The main concept in the RA design is for the 5G and/or IoT network to be enabled to simultaneously 'double' use its allotted spectrum space by two networks so as to improve its overall productivity. Even though the concept of spectrum co-use is not entirely new, its practical application to 5G and IoT networks is yet to be fully developed. We therefore attempt to solve the spectrum scarcity problem for 5G and IoT by investigating a RA optimisation solution that uses spectrum double usage as its basis for improving the scarcity problem.

10.1 System Model

The system model is shown in Fig. 1. A similar model has been developed for CRNs in [42] but the model is equally good for 5G and IoT. In the model, a primary network is made to operate alongside a secondary network over the same spectrum space in an underlay arrangement. There are a number of practical scenarios that can fit this type of arrangement. For instance, in a 5G network, the primary network could be a macrocell while the secondary network could be a femtocell or picocell designed to work alongside simultaneously the macrocell to improve coverage or the QoS experienced by the users/clients. For an IoT example, the primary network could be centrally controlled M2M communication for a parking system in a smart city, while the secondary network could be D2D communication in specific places such as malls or banking halls within that smart city. A third scenario where this kind of arrangement is possible is in underlay CRN developed for 5G and/or IoT application. In each of the above-mentioned scenarios of possible 5G/IoT set-up, the expectation is that of an improvement in the spectrum availability and usage, as well as a significant improvement in overall service provisioning for all users in the system.

In the model presented in Fig. 1, the primary network (made up of primary users (PUs)) is designed to work simultaneously with the secondary network (made up of secondary users (SUs)) over the allotted spectrum space. The SUs have access to the entire spectrum but must transmit within a predefined interference level that is permissible to the PUs. The system is heterogeneous in nature. Both the primary and the secondary networks are capable of operating using different configurations of modulation schemes, power levels, interference etc. Channel heterogeneity is incorporated by the use of an OFDMA platform, which makes it possible for different slices of the frequency band to be used by different users at the same time. User heterogeneity is integrated in that the users are classified and serviced based on some predetermined criteria.

10.2 Analysis of Model

In the analysis provided in this section, it is assumed that the primary network's use of spectrum (and eventual service provisioning) is guaranteed and the PUs are unaffected in any way by the secondary network, since the permissible interference constraint is not to be violated by the SUs' transmission. Hence, attention is dedicated to analysing the secondary network's optimal usage of the spectrum (and other resources) given its power limitations and other important constraints by which it is bound. It is imperative to note, however, that the overall spectrum usage for the developed model is in reality the combination of both primary networks' usage (which is not factored into the analysis) and the optimal co-use of the spectrum by the secondary network (analysis of which is presented below).

In the model, there are K heterogeneous SUs, L PUs and N subchannels within the coverage region of the system. The primary network is controlled by a primary network base station while the secondary network is controlled by a secondary network base station (SNBS). The K heterogeneous SUs are divided into two categories. The categories are differentiated as K_1: SUs with minimum rate guarantee and K_2: SUs with best effort service. The corresponding sets of these two categories of SUs are denoted as κ_A and κ_B, respectively. Category one SUs are given higher priority and their minimum rate demands are met first. The remaining resources are thereafter shared among the category two SUs based on their proportional rate constraint.

The SNBS selects the subchannels for each SU and relays this decision to the selected SU through a separate control channel. It is assumed that the communication between the SU and the SNBS over the control channel is error-free. All subchannels are also modelled to be in slow fading. The data rate c for each subchannel is dependent on the modulation scheme assigned to that subchannel. For this work, the modulation schemes considered are BPSK, 4-QAM, 16-QAM and 64-QAM, which transmit $c = 1, 2, 4$ and 6 bits per OFDMA symbol respectively. To achieve a given bit error rate (BER) ρ value at the receiver, the minimum amount of power $P_r(c, \rho)$ required over any given subchannel is also dependent on the modulation scheme employed. For the modulation schemes considered in this work, the minimum power for BPSK modulation is given as $P_{BPSK}(1, \rho) = N_\phi[erfc^{-1}(2\rho)]^2$ while for the M-ary QAM, the minimum power is given as $P_{M-QAM}(c, \rho) = \frac{2(2^c-1)N_\phi}{3}[erfc^{-1}(\frac{c\rho\sqrt{2^c}}{2(\sqrt{2^c}-1)})]^2$ where $erfc(x) = (\frac{1}{\sqrt{2\pi}})\int_x^\infty e^{\frac{-t^2}{2}}dt$ is the complementary error function, $\pi = (22/7)$, and N_ϕ is the single-sided noise power spectral density, which is assumed to be the same for all subchannels.

For a given value of ρ, as the number of bits assigned to a subchannel increases, the transmit power increases as well, albeit non-linearly. The subchannel power gain matrix between the SNBS and the SUs is given as $H^s \in R^{K\times N}$. $H^s_{k,n}$ therefore denotes the power gain between the SNBS and the kth SU at the nth subchannel. The power required to transmit $c_{k,n}$ bits over the nth subchannel to the kth SU with a BER threshold ρ is given as

$$P_{k,n}(c_{k,n}, \rho) = \frac{P(c_{k,n}, \rho)}{H^s_{k,n}}. \tag{22}$$

The power gain matrix between the SNBS and the PUs is given by $H^p \in R^{L\times N}$. $H^p_{l,n}$ therefore denotes the subchannel power gain between the SNBS and the lth PU at the nth subchannel.

Let R_k be the minimum data rate that must be assigned to the kth SU in κ_A and γ_k be the predetermined value of the normalised proportional fairness factor for each SU in κ_B. Also let data rate R_i indicate the rate for the element i in κ_B. Let $\Phi_n = \Sigma_{k=1}^K P_{k,n}$ be the total power of the nth subchannel with $P_{k,n}$ being the transmit power of the kth SU over the nth subchannel. Let $H^p_{l,n}$ be the magnitude of the interference channel gain between the lth PU and the SNBS over the nth subchannel. Let ε_l be the threshold interference power to the lth PU from all the SUs and let $P_{\max}$ be the maximum transmit power of the SNBS. The RA optimisation problem for the developed 5G/IoT network is formulated thus:

$$\begin{gathered}\max z = \sum_{n=1}^{N}\left(\sum_{k=1}^{K_1} w_1 c_{k,n} + \sum_{k=1}^{K_2} w_2 c_{k,n}\right); \\ c_{k,n} \in \{0, 1, 2, 4, 6\}\end{gathered} \tag{23}$$

subject to

$$\sum_{n=1}^{N} c_{k,n} \geq R_k;\ \forall k \in \kappa_A \tag{24}$$

$$\frac{R_k}{\sum\limits_{i \in \kappa_B} R_i} = \gamma_k;\ \forall k \in \kappa_B \tag{25}$$

$$\sum_{n=1}^{N} \sum_{k=1}^{K} P_{k,n} \leq P_{\max} \tag{26}$$

$$\sum_{n=1}^{N} \Phi_n H_{l,n}^{p} \leq \varepsilon_l;\ l = 1, 2, ..., L \tag{27}$$

$$c_{k,n} = 0\ if\ c_{k',n} \neq 0,\ \forall k' \neq k;\ k = 1, 2, ..., K. \tag{28}$$

The objective function (23) gives the total data rate achievable by all the SUs in the secondary network. Constraint (24) shows that the minimum data rate for each of the category one SUs must be met. In constraint (25), the proportional fairness factor γ_k is used to determine how much of the remaining capacity is assigned to each SU in category two. Constraint (26) explains that the total transmit power of all the SUs cannot be greater than the maximum transmit power of the SNBS. Constraint (27) shows that the interference from all the SUs to each PU must not be greater than the value of the permissible interference to the PUs. Constraint (28) is the mutually exclusive constraint, meaning that no single subchannel can be assigned to two or more SUs at the same time. In other words, subchannel n can no longer be assigned to a user k if it has been assigned to any other user k' that is not k. It is important to note that the constraint in equation (28) is only applicable if the 5G/IoT network has been developed using an orthogonal multiple access technology such as the OFDMA, as being considered in this analysis. However, if an NOMA technology has been modelled, the constraint would be different and has to be modified. In NOMA networks, more than one user can be multiplexed on a given subchannel. In such a case, if M $(1 \leq M \leq K)$ is the number of users that can be assigned to subchannel n, then only users $(K - M)$ cannot be assigned to subchannel n. In such a case, the constraint in equation (28) would be more accurately given as:

$$c_{k,n} = 0\ \forall k \in (K - M);\ k = 1, 2, ..., K.$$

The above NOMA consideration of equation (28) would definitely give an even better performance for the RA problem than in the OFDMA scenario, as the restriction of a single user being assigned to one subchannel is relaxed. The cost, however, is that the signalling overhead for NOMA is higher. We retain the OFDMA consideration for the analysis carried out in this paper as a kind of 'worst case' scenario for an 5G/IoT application.

From [15], it has been shown that constraint (25) can be equivalently rewritten as:

$$R_1 : R_2 : \ldots : R_{K_2} = \tilde{\gamma}_1 : \tilde{\gamma}_2 : \ldots : \tilde{\gamma}_{K_2}\ \forall k \in \kappa_B, \tag{29}$$

where$\tilde{\gamma}_k$ represents the product of γ_k and $\sum_{i \in \kappa_B} R_i$. The developed problem of RA in 5G and IoT networks is non-linear because of the power constraint. However, after carrying out some reformulation (similar to that carried out in [42]), the RA problem for 5G and IoT networking is reformulated as an integer linear programming (ILP) problem. The reformulation process is not carried out in this paper because of space limitations. The resulting formulation is given as:

$$z^* = \max_x [(w_1 \odot \boldsymbol{b}_1)^T \boldsymbol{x}_1 + (w_2 \odot \boldsymbol{b}_2)^T \boldsymbol{x}_2] \tag{30}$$

subject to

$$\boldsymbol{B}_i \boldsymbol{x}_1 \geq R_k;\ \forall k \in \kappa_A \tag{31}$$

$$\boldsymbol{B}_j \boldsymbol{x}_2 = \tilde{\gamma_k};\ \forall k \in \kappa_B \tag{32}$$

$$\boldsymbol{p}^T \boldsymbol{x} \leq P_{\max} \tag{33}$$

$$\boldsymbol{H}^p[\boldsymbol{A}(\boldsymbol{p} \odot \boldsymbol{x})] \leq \boldsymbol{\varepsilon}_l \tag{34}$$

$$\boldsymbol{0}_N \leq \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{1}_N \tag{35}$$

$$\boldsymbol{x}_{1,2} \in \{0,1\},\ w_1,\ w_2 \in \mathbb{R}^+. \tag{36}$$

The reformulated ILP problem in equations (30) - (36) is solved using the branch-and-bound approach of classical optimisation to obtain optimal solutions. For large 5G/IoT networks, heuristics can be developed to reduce computational complexity but usually at a cost of compromising optimality, as only suboptimal solutions can be obtained with the use of heuristics.

10.3 Results and Discussion

In obtaining the results presented in this section, the developed model is simulated using a combination of MATLAB and YALMIP simulators. An OFDMA system with $N = 64$ subchannels is designed, with the fading being random multipath. The number of PUs, $L = 4$. Category one SUs, $K_1 = 2$ have their minimum data rate requirements as 64 bits/user, while category two SUs, $K_2 = 2$ have the remaining data rate proportionally distributed among them, with a fairness factor equal for all users. The BER requirement is set to $\rho = 0.01$ for all SUs. The interference to the SUs caused by PUs' transmission is considered as noise at the SUs and its power spectral density is given as 0.01 mW/subchannel. All simulations are generated using 100 random channel pairs H^s and H^p. To determine and compare the QoS performance, results of the average individual user data rate, throughput and outage probability of the heterogeneous network are presented and discussed. In the results presented in Fig.2 - Fig.5, a weight of unity has been assumed for both SU categories.

The average user data rate achieved for each category of user over a varying interference power to the PUs is shown in Fig. 2. The maximum acceptable interference power to each of the PUs, ε_l was varied between $20 - 30 dBm$ with the available SNBS power set at 12 dBm, and then later increased to 30 dBm. First, we note

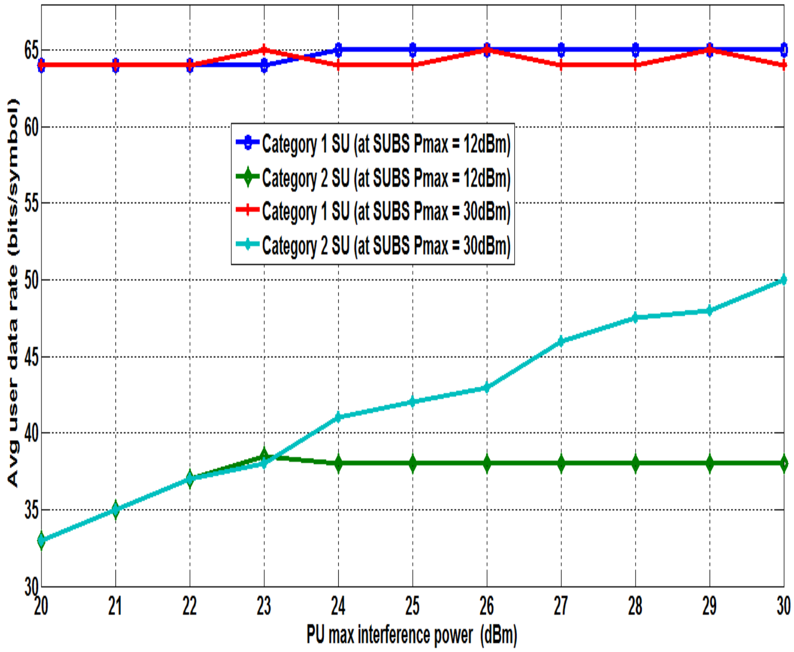

Figure 2 Average data rate of user (bits per OFDMA symbol) versus maximum interference power to PUs at different SNBS power for the categories of SUs.

that below 20 dBm interference, the problem becomes infeasible. Also, it can be seen that when the problem is feasible, the minimum data rate requirement for category one SUs is achieved at all points. Furthermore, the figure also shows that the algorithm achieves a similar result to about 22 dBm of maximum interference power. Beyond this limit, the average rate for users in both categories begins to stabilise when the SNBS maximum power is at 12 dBm. However, the average rate for users in category two SUs continues to increase when the SNBS maximum power is increased to 30 dBm. The reason for this is that with higher power at the SNBS, the average data rate of the users is greatly improved if all the other constraints are unchanged. It is also very significant to observe that the algorithm would rather increase the average rate of the category of SUs with best effort rate demand when it has such an opportunity than it would the category of SUs with a minimum rate demand.

In Fig. 3, the throughput or total data rate of the system against varying values of interference power to the PUs is presented. The maximum interference power to PUs was also varied between $20dBm$ and 30 dBm for similar values of SNBS power (12 dBm and 30 dBm). The result clearly shows that the 5G/IoT network will generally achieve better QoS in terms of throughput as the interference power to the PUs is relaxed (i.e., when it assumes a higher value). Also, it can be seen that for a higher SNBS power, the throughput keeps improving, unlike its lower SNBS power counterpart where the throughput quickly stabilises even with an increasing interference limit.

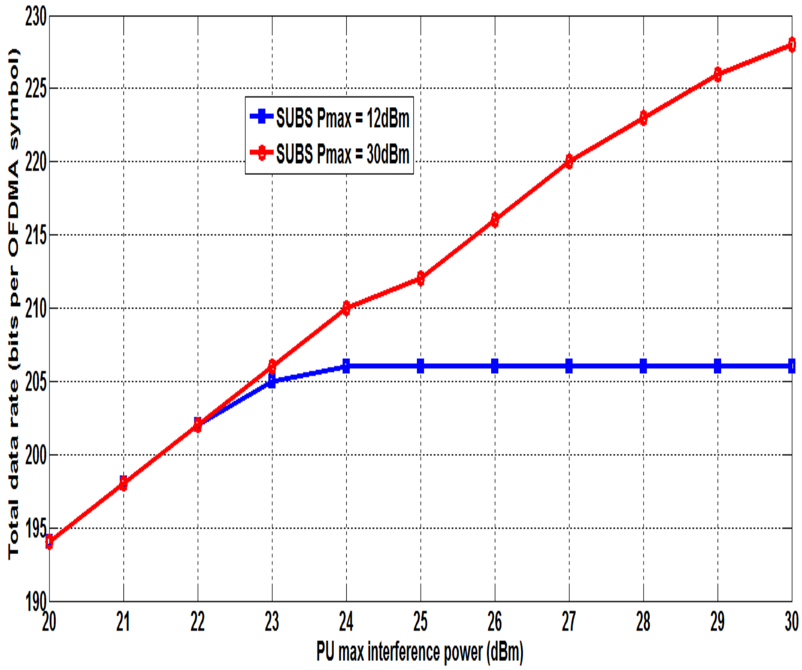

Figure 3 Total data rate (bits per OFDMA symbol) versus maximum interference power to PUs for different SNBS power.

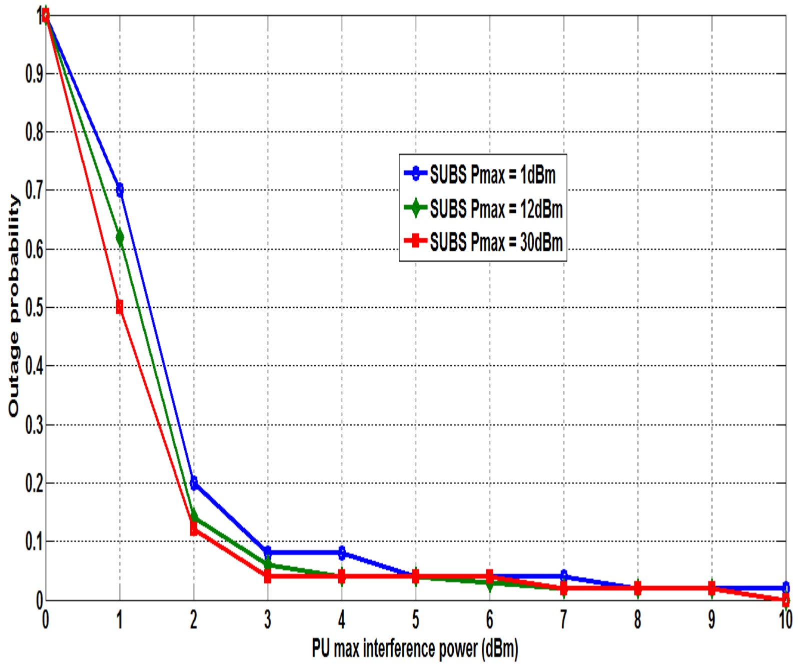

Figure 4 Outage probability versus maximum interference power to PUs for different SNBS power.

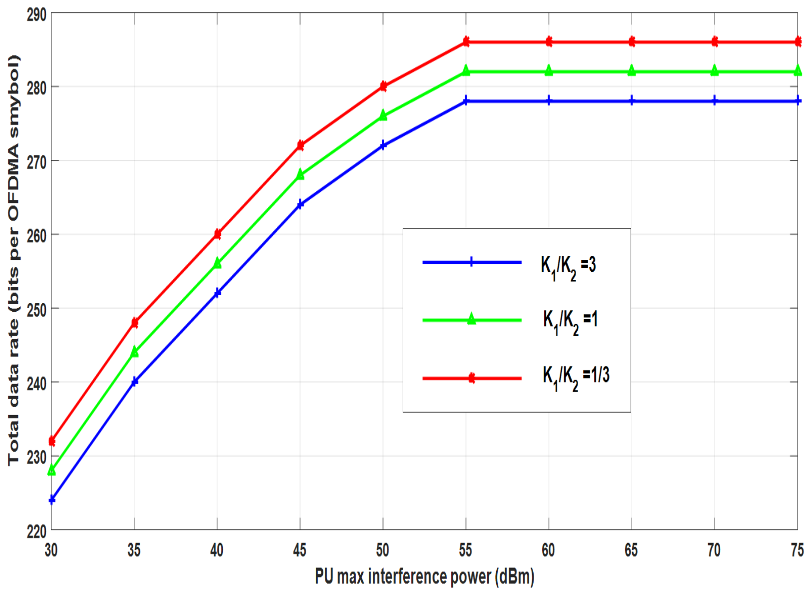

Figure 5 Total data rate against the maximum interference to PUs for different possible combinations of categories of SUs. Maximum transmit power is set at 30 dBm

The outage probability is the probability that the formulated problem will be infeasible given the current constraints and conditions. In Fig. 4, the outage probability over a varying amount of interference power to the PUs is shown for different values of SNBS power. From the figure, it can be depicted that the outage probability decreases with an increasing interference power limit to the PUs. It can also be easily observed that the outage probability generally improves (by achieving lower values) with an increasing SNBS power ($P_{\max}$). This implies that for a given value of interference power to PUs, the outage probability would be better at a higher SNBS power than it would at a lower SNBS transmit power.

Fig. 5 shows the total data rate against the maximum permissible interference to the PUs when the various categories of heterogeneous users are also combined differently. The maximum transmit power at the SNBS is fixed at 30 dBm. The results show that as the permissible interference to PUs increase, the total data rate also increases until it achieves a maximum possible value. The reason for this is that at a larger amount of permissible interference, the SUs transmit at a higher rate, thus achieving better overall capacity. The total data rate does not increase indefinitely too because, at some point, other constraints come into play. The results further show that the more category two SUs in the network (in comparison to category one SUs), the better the overall throughput of the system. The reason for this is that it is easier to satisfy category two SUs because of the flexibility in their demand compared to the category one SUs whose rate expectations are quite static and probably high.

11 Conclusion

5G and IoT technologies, the respective near-future communication standard and emerging computing paradigm, both have great prospects to help create a smart, interconnected and highly functional world. Achieving their promised provision of autonomous, reliable and excellent ubiquitous wireless communication, computing and internet services requires that their often limited and/or scarce resources be judiciously administered for efficiency and effectiveness. Hence, adequate models that address their RA and management problems need to be investigated, and many such investigations are currently being carried out. In this paper, optimisation approaches that are being employed in addressing RA problems in 5G and IoT networks were examined. The approaches were classified based on their outstanding characteristics. Furthermore, strengths and weaknesses of the solution approaches were discussed. Importantly, areas that require deeper research were identified and suggestions and/or directions to improve current RA optimisation solutions for 5G and IoT networks were drawn up. One of the identified problems, i.e., the problem of spectrum scarcity, was studied in depth and a practicable solution model for improving spectrum availability in RA for 5G and IoT was developed and analysed.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This research was supported by the Advanced Sensor Networks SARChi Chair program, co-hosted by University of Pretoria (UP) and Council for Scientific and Industrial Research (CSIR), through the National Research Foundation (NRF) of South Africa.

Author details

[1]Department of Electrical, Electronic and Computer Engineering, University of Pretoria, South Africa. [2]Department of Electrical and Computer Engineering, University of Manitoba, Canada AND Department of Electrical, Electronic and Computer Engineering, University of Pretoria , South Africa.

References

1. Luo, F.-L., Zhang, C.: 5G Standard Development: Technology and Roadmap, p. 616. Wiley-IEEE Press, ??? (2016). http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7572796
2. Gohil, A., Modi, H., Patel, S.K.: 5G technology of mobile communication: A survey. In: 2013 International Conference on Intelligent Systems and Signal Processing (ISSP), pp. 288–292 (2013). doi:10.1109/ISSP.2013.6526920
3. Kumar, S., Gupta, G., Singh, K.R.: 5G: Revolution of future communication technology. In: 2015 International Conference on Green Computing and Internet of Things (ICGCIoT), pp. 143–147 (2015). doi:10.1109/ICGCIoT.2015.7380446
4. Xiaoli, X., Yunbo, Z., Guoxin, W.: Design of intelligent internet of things for equipment maintenance. In: Intelligent Computation Technology and Automation (ICICTA), 2011 International Conference On, vol. 2, pp. 509–511 (2011). doi:10.1109/ICICTA.2011.412
5. Yu, Z., Tie-ning, W.: Research on the visualization of equipment support based on the technology of internet of things. In: Instrumentation, Measurement, Computer, Communication and Control (IMCCC), 2012 Second International Conference On, pp. 1352–1357 (2012). doi:10.1109/IMCCC.2012.318
6. Singh, D., Tripathi, G., Jara, A.J.: A survey of internet-of-things: Future vision, architecture, challenges and services. In: Internet of Things (WF-IoT), 2014 IEEE World Forum On, pp. 287–292 (2014). doi:10.1109/WF-IoT.2014.6803174
7. Pereira, P.P., Eliasson, J., Kyusakov, R., Delsing, J., Raayatinezhad, A., Johansson, M.: Enabling cloud connectivity for mobile internet of things applications. In: Service Oriented System Engineering (SOSE), 2013 IEEE 7th International Symposium On, pp. 518–526 (2013). doi:10.1109/SOSE.2013.33
8. Zheng, L., Chen, S., Xiang, S., Hu, Y.: Research of architecture and application of internet of things for smart grid. In: Computer Science Service System (CSSS), 2012 International Conference On, pp. 938–941 (2012). doi:10.1109/CSSS.2012.238
9. Burange, A.W., Misalkar, H.D.: Review of internet of things in development of smart cities with data management amp; privacy. In: Computer Engineering and Applications (ICACEA), 2015 International Conference on Advances In, pp. 189–195 (2015). doi:10.1109/ICACEA.2015.7164693
10. Al-Falahy, N., Alani, O.Y.: Technologies for 5G networks: Challenges and opportunities. IT Professional **19**(1), 12–20 (2017). doi:10.1109/MITP.2017.9
11. Ma, Z., Shang, X., Fu, X., Luo, F.: The architecture and key technologies of internet of things in logistics. In: Cyberspace Technology (CCT 2013), International Conference On, pp. 464–468 (2013). doi:10.1049/cp.2013.2173

12. Meng, X., Li, J., Zhou, D., Yang, D.: 5G technology requirements and related test environments for evaluation. China Communications **13**(Supplement 2), 42–51 (2016). doi:10.1109/CC.2016.7833459
13. Munir, H., Hassan, S.A., Pervaiz, H., Ni, Q., Musavian, L.: User association in 5G heterogeneous networks exploiting multi-slope path loss model. In: 2017 2nd Workshop on Recent Trends in Telecommunications Research (RTTR), pp. 1–5 (2017). doi:10.1109/RTTR.2017.7887875
14. Wang, M., Chen, J., Aryafar, E., Chiang, M.: A survey of client-controlled hetnets for 5G. IEEE Access **5**, 2842–2854 (2017). doi:10.1109/ACCESS.2016.2624755
15. Awoyemi, B.S., Maharaj, B.T., Alfa, A.S.: Resource allocation in heterogeneous cooperative cognitive radio networks. Int. J. Commun. Syst., 1–14 (2016). doi:10.1002/dac.3247
16. Afana, A., Erdogan, E., Ikki, S.: Quadrature spatial modulation for cooperative MIMO 5G wireless networks. In: 2016 IEEE Globecom Workshops (GC Wkshps), pp. 1–5 (2016). doi:10.1109/GLOCOMW.2016.7849011
17. Larsson, E.G., Edfors, O., Tufvesson, F., Marzetta, T.L.: Massive mimo for next generation wireless systems. IEEE Communications Magazine **52**(2), 186–195 (2014). doi:10.1109/MCOM.2014.6736761
18. Liu, G., Hou, X., Jin, J., Wang, F., Wang, Q., Hao, Y., Huang, Y., Wang, X., Xiao, X., Deng, A.: 3D-MIMO with massive antennas paves the way to 5G enhanced mobile broadband: From system design to field trials. IEEE Journal on Selected Areas in Communications **PP**(99), 1–1 (2017). doi:10.1109/JSAC.2017.2687998
19. Vook, F.W., Ghosh, A., Thomas, T.A.: MIMO and beamforming solutions for 5G technology. In: 2014 IEEE MTT-S International Microwave Symposium (IMS2014), pp. 1–4 (2014). doi:10.1109/MWSYM.2014.6848613
20. Awoyemi, B.S., Maharaj, B.T., Alfa, A.S.: Optimal resource allocation solutions for heterogeneous cognitive radio networks. Digital Communications and Networks, 1–11 (2016). doi:10.1016/j.dcan.2016.11.003
21. Zhang, Z., Zhang, W., Zeadally, S., Wang, Y., Liu, Y.: Cognitive radio spectrum sensing framework based on multi-agent arc hitecture for 5G networks. IEEE Wireless Communications **22**(6), 34–39 (2015). doi:10.1109/MWC.2015.7368822
22. Hong, X., Wang, J., Wang, C.X., Shi, J.: Cognitive radio in 5G: a perspective on energy-spectral efficiency trade-off. IEEE Communications Magazine **52**(7), 46–53 (2014). doi:10.1109/MCOM.2014.6852082
23. Wubben, D., Rost, P., Bartelt, J.S., Lalam, M., Savin, V., Gorgoglione, M., Dekorsy, A., Fettweis, G.: Benefits and impact of cloud computing on 5G signal processing: Flexible centralization through cloud-RAN. IEEE Signal Processing Magazine **31**(6), 35–44 (2014). doi:10.1109/MSP.2014.2334952
24. Falchetti, A., Azurdia-Meza, C., Cespedes, S.: Vehicular cloud computing in the dawn of 5G. In: 2015 CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON), pp. 301–305 (2015). doi:10.1109/Chilecon.2015.7400392
25. Tao, M., Ota, K., Dong, M.: Foud: Integrating fog and cloud for 5G-enabled V2G networks. IEEE Network **31**(2), 8–13 (2017). doi:10.1109/MNET.2017.1600213NM
26. Lei, L., Yuan, D., Ho, C.K., Sun, S.: Power and channel allocation for non-orthogonal multiple access in 5g systems: Tractability and computation. IEEE Transactions on Wireless Communications **15**(12), 8580–8594 (2016). doi:10.1109/TWC.2016.2616310
27. Ding, Z., Peng, M., Poor, H.V.: Cooperative non-orthogonal multiple access in 5G systems. IEEE Communications Letters **19**(8), 1462–1465 (2015). doi:10.1109/LCOMM.2015.2441064
28. Liu, X., Liu, Y., Wang, X., Lin, H.: Highly efficient 3d resource allocation techniques in 5g for noma enabled massive mimo and relaying systems. IEEE Journal on Selected Areas in Communications **PP**(99), 1–1 (2017). doi:10.1109/JSAC.2017.2726378
29. Khan, R., Khan, S.U., Zaheer, R., Khan, S.: Future internet: The internet of things architecture, possible applications and key challenges. In: Frontiers of Information Technology (FIT), 2012 10th International Conference On, pp. 257–260 (2012). doi:10.1109/FIT.2012.53
30. Tan, L., Wang, N.: Future internet: The internet of things. In: Advanced Computer Theory and Engineering (ICACTE), 2010 3rd International Conference On, vol. 5, pp. 5–3765380 (2010). doi:10.1109/ICACTE.2010.5579543
31. Xingmei, X., Jing, Z., He, W.: Research on the basic characteristics, the key technologies, the network architecture and security problems of the internet of things. In: Computer Science and Network Technology (ICCSNT), 2013 3rd International Conference On, pp. 825–828 (2013). doi:10.1109/ICCSNT.2013.6967233
32. Chuang, P.-J., Tsai, W.-T.: On RFID network topology design for internet of things. In: Future Generation Communication and Networking (FGCN), 2014 8th International Conference On, pp. 25–28 (2014). doi:10.1109/FGCN.2014.13
33. Dou, R., Nan, G.: Optimizing sensor network coverage and regional connectivity in industrial IoT systems. Systems Journal, IEEE **PP**(99), 1–10 (2015). doi:10.1109/JSYST.2015.2443045
34. Zhang, Y., Sun, S.: Real-time data driven monitoring and optimization method for IoT-based sensible production process. In: Networking, Sensing and Control (ICNSC), 2013 10th IEEE International Conference On, pp. 486–490 (2013). doi:10.1109/ICNSC.2013.6548787
35. Xinhua, Z., Hong, L.: A self-reconfigurable sensor network constructon research in the paradigm of internet of things. In: Computer Science Service System (CSSS), 2012 International Conference On, pp. 311–314 (2012). doi:10.1109/CSSS.2012.85
36. Bontu, C.S., Periyalwar, S., Pecen, M.: Wireless wide-area networks for internet of things: An air interface protocol for IoT and a simultaneous access channel for uplink IoT communication. IEEE Vehicular Technology Magazine **9**(1), 54–63 (2014). doi:10.1109/MVT.2013.2295068
37. Tian, L.: Applications and prospects of internet of things based on mobile communication network. In: Proceedings of 2012 2nd International Conference on Computer Science and Network Technology, pp. 1167–1170 (2012). doi:10.1109/ICCSNT.2012.6526132
38. Hsieh, H.C., Lai, C.H.: Internet of things architecture based on integrated PLC and 3G communication networks. In: 2011 IEEE 17th International Conference on Parallel and Distributed Systems, pp. 853–856 (2011). doi:10.1109/ICPADS.2011.73
39. Lee, J., Stanley, M., Spanias, A., Tepedelenlioglu, C.: Integrating machine learning in embedded sensor systems for internet-of-things applications. In: 2016 IEEE International Symposium on Signal Processing and

Information Technology (ISSPIT), pp. 290–294 (2016). doi:10.1109/ISSPIT.2016.7886051
40. Kovácsházy, T., Wacha, G., Dabóczi, T., Erdős, C., Szarvas, A.: System architecture for internet of things with the extensive use of embedded virtualization. In: 2013 IEEE 4th International Conference on Cognitive Infocommunications (CogInfoCom), pp. 549–554 (2013). doi:10.1109/CogInfoCom.2013.6719308
41. Awoyemi, B.S., Maharaj, B.T., Alfa, A.S.: Resource allocation for heterogeneous cognitive radio networks. In: Proc. IEEE WCNC, pp. 1759–1763 (2015). doi:10.1109/WCNC.2015.7127734
42. Awoyemi, B.S., Maharaj, B.T., Alfa, A.S.: QoS provisioning in heterogeneous cognitive radio networks through dynamic resource allocation. In: Proc. IEEE AFRICON, pp. 1–6 (2015). doi:10.1109/AFRCON.2015.7331941
43. Awoyemi, B.S., Maharaj, B.T., Alfa, A.S.: Solving resource allocation problems in cognitive radio networks: A survey. EURASIP J. Wireless Commun. Netw. **2016**(1), 176 (2016). doi:10.1186/s13638-016-0673-6
44. Lei, L., Yuan, D., Ho, C.K., Sun, S.: Joint optimization of power and channel allocation with non-orthogonal multiple access for 5G cellular systems. In: 2015 IEEE Global Communications Conference (GLOBECOM), pp. 1–6 (2015). doi:10.1109/GLOCOM.2015.7417761
45. Zappone, A., Sanguinetti, L., Bacci, G., Jorswieck, E., Debbah, M.: A framework for energy-efficient design of 5G technologies. In: 2015 IEEE International Conference on Communications (ICC), pp. 1845–1850 (2015). doi:10.1109/ICC.2015.7248593
46. Oh, C., Hwang, D., Lee, T.: Joint access control and resource allocation for concurrent and massive access of M2M devices. Wireless Communications, IEEE Transactions on **PP**(99), 1–1 (2015). doi:10.1109/TWC.2015.2417873
47. Sánchez-González, J., Sallent, O., Pérez-Romero, J., Agustí, R.: A comparison of different optimisation search methodologies for self-optimisation in wireless cellular networks. In: 6th Joint IFIP Wireless and Mobile Networking Conference (WMNC), pp. 1–8 (2013). doi:10.1109/WMNC.2013.6549012
48. Pierucci, L.: The quality of experience perspective toward 5G technology. IEEE Wireless Communications **22**(4), 10–16 (2015). doi:10.1109/MWC.2015.7224722
49. Turcu, C.E., Gaitan, V.G., Turcu, C.O.: An internet of things-based distributed intelligent system with self-optimization for controlling traffic-light intersections. In: Applied and Theoretical Electricity (ICATE), 2012 International Conference On, pp. 1–5 (2012). doi:10.1109/ICATE.2012.6403461
50. Oteafy, S.M.A., Hassanein, H.S.: Towards a global IoT: Resource re-utilization in WSNs. In: Computing, Networking and Communications (ICNC), 2012 International Conference On, pp. 617–622 (2012). doi:10.1109/ICCNC.2012.6167496
51. Ramakrishnan, A.K., Naqvi, N.Z., Bhatti, Z.W., Preuveneers, D., Berbers, Y.: Learning deployment trade-offs for self-optimization of internet of things applications. In: Proceedings of the 10th International Conference on Autonomic Computing (ICAC 13), pp. 213–224. USENIX, San Jose, CA (2013). https://www.usenix.org/conference/icac13/technical-sessions/presentation/ramakrishnan
52. Hu, L., Zhang, Z., Wang, F., Zhao, K.: Optimization of the deployment of temperature nodes based on linear programing in the internet of things. Tsinghua Science and Technology **18**(3), 250–258 (2013). doi:10.1109/TST.2013.6522584
53. Bjornson, E., Jorswieck, E., Debbah, M., Ottersten, B.: Multi-objective signal processing optimization: The way to balance conflicting metrics in 5G systems. arXiv, 1–10 (2016). doi:arXiv:1406.2871v2 [cs.IT]
54. Sachan, R., Choi, T.J., Ahn, C.W.: A genetic algorithm with location intelligence method for energy optimization in 5G wireless networks. Discrete Dynamics in Nature and Society **2016**, 1–10 (2016). doi:10.1155/2016/5348203
55. Botero, O., Chaouchi, H.: RFID network topology design based on genetic algorithms. In: RFID-Technologies and Applications (RFID-TA), 2011 IEEE International Conference On, pp. 300–305 (2011). doi:10.1109/RFID-TA.2011.6068653
56. Colistra, G., Pilloni, V., Atzori, L.: Task allocation in group of nodes in the IoT: A consensus approach. In: Communications (ICC), 2014 IEEE International Conference On, pp. 3848–3853 (2014). doi:10.1109/ICC.2014.6883921
57. Wang, D., Ip, W.H.: Review on modeling and optimization problems about RFID technology and applications. In: Control and Decision Conference (CCDC), 2013 25th Chinese, pp. 1258–1263 (2013). doi:10.1109/CCDC.2013.6561118
58. Salem, K.H., Kieffer, Y., Mancini, S.: Memory management in embedded vision systems: Optimization problems and solution methods. In: 2016 Conference on Design and Architectures for Signal and Image Processing (DASIP), pp. 200–207 (2016). doi:10.1109/DASIP.2016.7853820
59. Salem, K.H., Kieffer, Y., Mancini, S.: Formulation and practical solution for the optimization of memory accesses in embedded vision systems. In: 2016 Federated Conference on Computer Science and Information Systems (FedCSIS), pp. 609–617 (2016)
60. Athreya, A.P., Tague, P.: Network self-organization in the internet of things. In: Internet-of-Things Networking and Control (IoT-NC), 2013 IEEE International Workshop Of, pp. 25–33 (2013). doi:10.1109/IoT-NC.2013.6694050
61. Xu, T., Wendt, J.B., Potkonjak, M.: Security of IoT systems: Design challenges and opportunities. In: Computer-Aided Design (ICCAD), 2014 IEEE/ACM International Conference On, pp. 417–423 (2014). doi:10.1109/ICCAD.2014.7001385
62. Covington, M.J., Carskadden, R.: Threat implications of the internet of things. In: Cyber Conflict (CyCon), 2013 5th International Conference On, pp. 1–12 (2013)
63. Ali, N.A., Abu-Elkheir, M.: Data management for the internet of things: Green directions. In: Globecom Workshops (GC Wkshps), 2012 IEEE, pp. 386–390 (2012). doi:10.1109/GLOCOMW.2012.6477602

Tables

Table 1 Summary of optimisation approaches for RA 5G and IoT networking.

S/N	Optimisation approaches	Features	Drawbacks
1.	Self-optimisation	Each object optimises its own resources, employs ideas of utility functions and organic computing.	Difficult to implement owing to complexity of 5G and IoT networks; ability of individual users to maintain the right balance between resource usage and rapid change in system behaviour can be difficult to achieve.
2.	Classical optimisation e.g. linear programming, convex programming, etc.	Gives optimal solutions, solutions can act as bounds for solutions obtained using other approaches.	Most problems do not fit into any classical optimisation model and proving convexity of problems can be a herculean task.
3.	Multi-objective optimisation e.g. cooperative and non-cooperative game, Nash bargaining, etc.	Good with problems that have multiple objectives, uses ideas from game theory.	Solution models can be complex and analytical modelling of solutions are usually difficult to achieve.
4.	Heuristics and meta-heuristics e.g. greedy algorithms, genetic algorithms, etc.	Quick solutions, good with large problems, very practicable.	Mostly suboptimal solutions; solutions are problem-specific and non-transferable.

Table 2 Summary of areas of application of optimisation in RA for 5G and IoT networking.

S/N	Areas of application	Reference examples	Features
1.	Optimising heterogeneity for 5G and IoT networking	[41], [42], [20]	5G and IoT must incorporate various aspects of heterogeneity, such as heterogeneous networks, heterogeneous devices and/or objects, heterogeneous user conditions, heterogeneous service requirements etc. Optimisation can be employed for addressing the heterogeneity problem in 5G and IoT, such as power or bandwidth allocation to different categories of users.
2.	Application in WSNs	[33]	High number of sensor nodes to be deployed in various locations for collecting 5G and IoT information. Energy saving, time saving, the number of nodes to be deployed, etc. can be addressed through RA optimisation.
3.	Application in RFID	[57]	Optimisation is important in addressing the problem of optimal placement of the RFID readers, in addressing the problem of load balancing for the RFID readers, in addressing the problem of data allocation for the RFID tags and in addressing the problem of collision during tag reading, especially when there are too many tags for a single reader to identify.
4.	Application in embedded system	[58], [59]	Efficient memory management optimisation models are being developed to address the need to handle a lot of data (for example big images) in a very short time.
5.	Addressing delay problems	[46]	Time delay, loss of packets and/or even connectivity in 5G and IoT due to limitations in the amount of bandwidth over which transmission occurs and finite capacities. Optimising congestion control and queue management is an imperative.
6.	Economic aspect of 5G and IoT	[57]	Both 5G and IoT will require sound economic considerations in their design and deployment, e.g. the economic implications of changing from the traditional macrocells to small cells for 5G and IoT application. Optimisation to be used in analysing economic considerations and arriving at best decisions for minimum cost of deployment, time frame for return on investment, pricing etc.