International Journal of Science and Research Archive
eISSN: 2582-8185
Cross Ref DOI: 10.30574/ijsra
Journal homepage: https://ijsra.net/

(Review Article)

Check for updates

Case studies for binary decision diagram for 5G mobile networks optimization

Panagiotis Leliopoulos and Athanasios Drigas *

Net Media Lab, IIT NCSR Demokritos, Athens, Greece.

International Journal of Science and Research Archive, 2024, 11(02), 780–802

Publication history: Received on 19 February 2024; revised on 25 March 2024; accepted on 27 March 2024

Article DOI: https://doi.org/10.30574/ijsra.2024.11.2.0491

Abstract

In this paper, we study the binary decision diagrams and mobile network optimization. As we see about the binary decision diagrams, Lee and Akers first introduced the general concept of binary decision diagrams but later spread by Bryant. It turns out that, according to the European Interactive Digital Advertising Alliance (EDAA) order and reduction rules, binary decision diagrams are a standard form of representation. After that, we study the spanning trees and binary decision diagrams. An advantage of binary decision diagrams-based methods is that binary decision diagrams provide a concise and separate description of system success or failure. The effectiveness of binary decision diagram methods is measured by the size of them, which depends heavily on the selected series.

Keywords: ICT; 5G; Optimization; Case Studies; Binary Decision; Diagram; Mobile Networks

1 Introduction

The subject we concern about is the social network layer and HetNet resource management, where we give an example. A networking example must be essentially preemptive and is based on the fact that network nodes exploit user environment information. With the use of data analytics, resource needs that change from one location to another over a while become predictable. In addition to network data, behavioral and emotional analyzes from social networks and other sources need to be considered to predict where and how users can use the mobile network. We continue with the traffic loads and about users' behavior. We see some cases were analyzing the cell tower location data from hundreds of thousands of people; it is possible to quantify some of the most fundamental rules of human motion. Hence, a different type of social data collected by mobile service providers when adopting service plans (pricing) and other telecommunications products, which can be seen as the spread of social transmission. Also, location-based social network data contains three levels of information: a social level, a geographic level, and a time level.

Mobile network hotspots are significant. As we see, the identification process, for example, the size of the grid cell and the numbers or densities of users. Also, we see some real case studies about the number of hotspots and the scaling relationship between the number of access points to the population and the effect of grid size.

Also, we mention about subjects such as reactive to proactive networks where the proxy network serves users' requests ahead of time, the corresponding data is stored on the user's device, and when the request starts, the information is extracted directly from the cache instead of accessing the wireless network. Leveraging social networks where social networks carry out the vast majority of data traffic has played a crucial role in disseminating information via the Internet and will continue to shape how information is accessed.

Furthermore, we study the failure prediction, detection, recovery, and prevention, and we see the analysis between technologies is failing using big data. In this case, as we see, many incidents occur when the user is on the edge of a coverage area and moves to another, technologically different area. Also, we will see the self-healing in cellular

* Corresponding author: Athanasios Drigas

networks. Finally, the cell site failure prediction includes data from different data sources that are then integrated and then analyzed to know bandwidth trends.

Network monitoring and analysis of large-scale mobile telephony traffic describe large cellular networks with relatively high data rate connections and high requirements to be met. There are research studies were proposed a system to undertake this task, using Hadoop Map Reduce, HDFS, and HBase as an advanced distributed computing platform. Optimized bandwidth allocation for content delivery is mobile networks usually have a large number of users. Also, optimizing 5G networks with user behavior awareness is that big data can be used to analyze user mobility patterns, predict traffic trajectory, and pre-configure the network accordingly.

Finally, we study an extensive data-driven network operating system, where researchers propose a system that can maximize the efficiency of big data network operation. So, it can be accomplished by optimally allocating network resources to each access provider and user. The proposed system consists of two parts, the decision-making area, and the implementation area. We continue with the deployment of the cache server on mobile content delivery networks where it is essential to locate distributed cache servers on the content delivery networks as close to the end-user as possible to reduce response time and reduce shipping costs. We study the quality of experience modeling for network optimization, the quality of experience modeling to support network optimization and emotional techniques. Various services and applications using a set of quality-of-service parameters about quality of experience modeling and improving quality of service in cellular networks through self-regulating cells and self-optimized delivery. Cellular networks have a critical element on which the concept of mobility depends. In the end, we study emotional techniques, including user emotion, experience, expectation, and so on, while objective factors relate to both technical and non-technical aspects of services.

2 Binary Decision Diagrams

Binary decision diagrams are fundamental data structures that emerged in the 1980s as a means of effectively representing Boolean functions [1]. Methods for analyzing network reliability are usually based on Boolean algebra, error tree, diagram, and BDD. In the 1990s, a binary decision diagram (BDD) became increasingly popular for network analysis [2]. Lee and Akers first introduced the general concept of BDDs but later spread by Bryant, where it turns out that, according to the European Interactive Digital Advertising Alliance (EDAA) order and reduction rules, BDDs are a standard form of representation. The joint adoption of BDDs in Electronic Design Automation (EDA) is guided by efficient BDD handling algorithms and related software applications [3].

BDD now finds applications in many areas, such as logical synthesis, model control, verification, automated reasoning, accessibility analysis, and combination. Initial BDDs are applied, for example, to model control and logic synthesis. In this case, multiple terminal binary decision diagrams (MTBDDs) are used, for example, to calculate model probability properties and exploit zero-settlement binary decision diagrams (ZBDDs) to represent solid subsets and sparse arrays [1].

For some applications, it may be useful to combine features of different types of binary decision diagrams. The efficiency of in-flight space exploration is based on a reliable representation of the Boolean functions involved, which initially include a large number of binary variables set to 0 (appropriately represented by ZBDD) but, over time, expand some assignments to variables that become redundant (appropriately represented by BDD). So far, no solution combines the reduction rules for both BDD and ZBDD at the same time. Existing methods use either original BDDs or ZBDDs only and do not exploit these reduction rules [1].

3 Spanning Tress and BDDs

An advantage of BDD-based methods is that BDD provides a concise and separate description of system success/failure. Another advantage is that BDD can act as a platform to deal with various engineering problems such as common cause failure (CCF), imperfect peaks, terminal network reliability, etc. The effectiveness of BDD methods is measured by the size of the BDD, which depends heavily on the selected series of BDD variables. Therefore, a significant challenge facing BDD-based methods in determining the optimal ordering strategy and automated BDD generation for large systems. Although the literature shows that the BDD variable sequence improvement is complete NP, Friedman proposes an algorithm (time complexity 2 (3) n O n) to find the optimal sequence of variables. For small networks, existing order algorithms can usually generate a BDD short enough to calculate reliability. Some approaches may consider changes to network connectivity but rarely allow for multiphase features, such as different system configuration, variable failure rates, and network component repair activities, etcetera. [2].

The issue of BDDs is a mature and established research branch, which is applied to many cases. Since then, scientists have developed research projects in this area, seeking to improve BDD's ability to represent data or the efficiency of the algorithms that handle them. Some researchers are working on building a bridge between the minimum spanning tree (MST) and BDD. As address retrieval time depends on and is limited by the available memory access technology, they developed a different scheme based on the Bryant model and applied it to hardware that takes advantage of the processing speed of BDDs and hardware. With this format, they received experimental performance results of up to 229.3 million search results per second for the PacBell routing table with 6,822 prefixes, converted to 200 Gb / s output (with an average packet size of 1000 bits) [4]. Therefore, it is better to take into account traffic characteristics when developing a mobile network. Hence, to conduct an in-depth and detailed investigation of BDD optimization. [5].

3.1 Social Network Layer and HetNet Resource Management

A networking example must be essentially preemptive and is based on the fact that network nodes (i.e., base stations and mobile devices/smartphones) exploit user environment information. Also, anticipate the user requirements, and leverage their predictive capabilities to achieve significant resource savings. As a result, as part of the preventive networking example, network nodes monitor, learn and create user demand profiles to anticipate future requests, taking advantage of device capabilities and the vast amount of data available by intelligently leveraging traffic statistics and user environment information (e.g., file popularity, location, speed, and mobility patterns distributions). The example allows for better prediction when user content is requested with the number of resources required and which network sites must be pre-cached in the content [6].

Another trend we must mention is the online social networks (i.e., Facebook, Twitter, Digg) have become crucial in the distribution of user content [6]. With the use of data analytics, resource needs that change from one location to another over a while become predictable. In addition to network data, behavioral and emotional analyzes from social networks and other sources need to be considered to predict where and how users can use the mobile network. For example, when a social event, such as a marathon, occurs in a city, certain places such as the roads on the route can attract large crowds, resulting in potential congestion in those places during the event [5]. Thus, by exploiting the interdependence of people through users' social relationships and connections, future networks can learn correlation patterns in connected social and geographic data networks for better prediction and conclusions of user behavior [6]. Therefore, with this information provided by data analytics, operators can allocate more wireless network resources to the "hotshot" so that maximum traffic can be absorbed smoothly without sacrificing the QoE user [5].

Figure 1 shows a subtraction of the technological and spatial network layer overlapping with the social network layer. Since the content distribution of the nodes at the social network level is addressed through the nodes at the technological and spatial network level, the analysis of interactions between these two levels will bring further benefits to future networks [6].

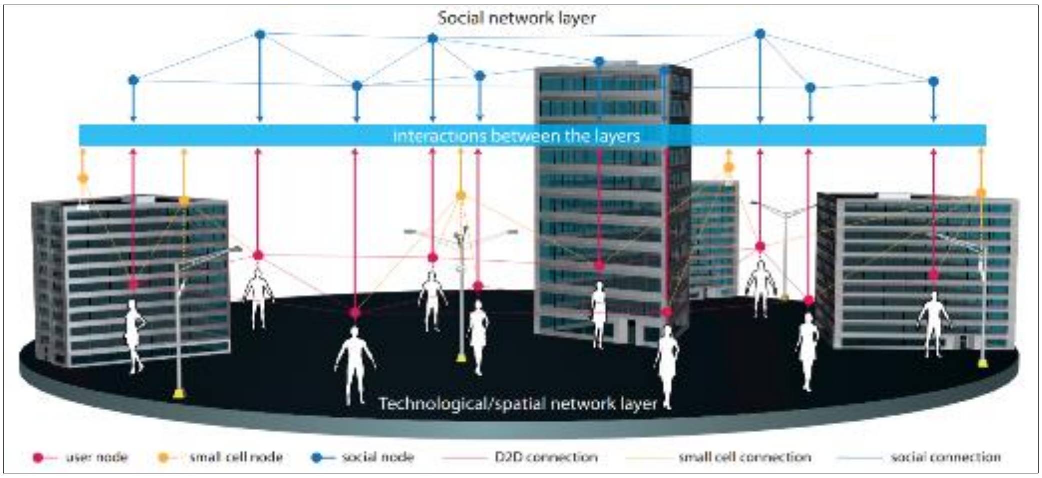

Figure 1 An illustration of an overlay of socially-interconnected and technological/spatial networks [6].

4 Traffic Loads and Users Behavioral Case Studies

4.1 Behavioral Studies Using Mobile Phone Data

Mobile users often travel from one place to another throughout the city, e.g., they work in the central business district during the day and probably live in the suburbs at night. For these reasons, the circulation of each cell varies at different times of the day, which is called the "tidal effect." If resources are available in every cell with a fixed configuration, resource usage should be underestimated, and hotspot users may find it challenging to obtain good QoE during peak hours [5]. Recent data analysis by mobile service providers has led researchers to see an increased picture of human movement patterns. Nevertheless, some researchers are not agreed with these ideas as "universal laws of human motion." It is clear that by analyzing the cell tower location data from hundreds of thousands of people, it is possible to quantify some of the most fundamental human motion [7].

It is a different type of social data collected by mobile service providers when adopting service plans (pricing) and other telecommunications products, which can be seen as spreading social transmission. By analyzing the spread of these contagious social elements in a call graph, it may be possible to learn more about the social dynamics within a population [7]. On the other hand, many resources can be wasted idle in low traffic areas. Current and background data can be used by data analysis to predict traffic in high-density areas of networks. With the Cloud RAN architecture, the projected resource allocation to central core units can help accurately service the right place at the right time, i.e., to know when and where traffic peaks occur, causing minimal service disruptions [5].

The extensive use of mobile devices has dramatically enriched the urban user experience and promoted location-based social services in recent years. For example, Foursquare and Facebook have attracted billions of users' worldwide and generated vast site-based social networking data, offering opportunities to the researchers and challenges to explore mobile user behavior. Also, it is fundamental for designing more advanced site-based services, such as site-based marketing and disaster relief. Location-based social network data contains three levels of information: a social level, a geographic level, and a time level, as shown in Figure 2. The social level consists of social friendships, the geographic level displays users' historical check-in, and the time layer indicates time stamps for each check-in action [8].

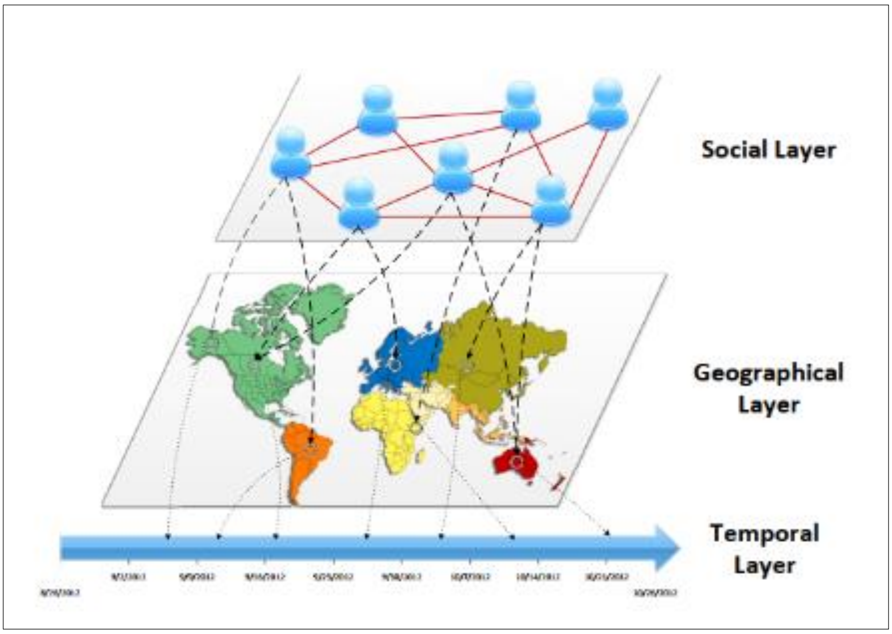

Figure 2 The Information Layout of location-based social networks [8].

Previous research has studied the social and spatial strata in location-based social networks (LBSNs) in socio-historical links, socio-spatial properties, geographical influence, and "geo-social" correlations, etcetera. Simultaneously, the time level in terms of temporal effects has rarely been studied to model user behavior on LBSNs. The time level on LBSNs is usually leveraged as an order indicator for a check-in connection to create location tracks, which has not been fully utilized. As observed in human movement, vital time cyclic patterns appear in terms of daylight hours and day of the week. Since human movement is observed as a stochastic process, the temporal characteristics created by a user's motion become very sparse in the long period of attributes. While the unobserved characteristics seriously affect the performance of the prediction. For example, a time pattern of visiting a restaurant at 10:00 a.m. and 12:00 p.m. could indicate the possible presence of the user in this restaurant at 11:00 am, although such a presence at 11:00 is. Hence, researchers can take advantage of these time circular patterns to model a user's time preferences. Social correlation shows that human movements are usually influenced by their social context, such as having lunch with friends (Figure 3)[8].

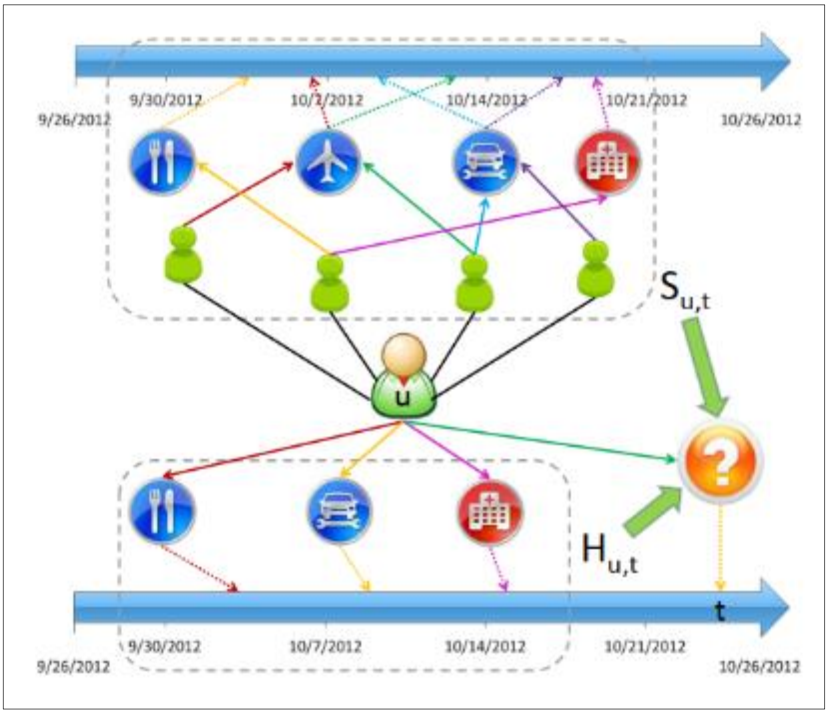

Figure 3 Mobile Behavior of User u at Time t [8].

Figure 3 shows a user you's mobile behavior at time t w.r.t. His/her personal check-in history Hu,t and friends' check-in history S u,t. Given the corresponding observations of Hu,t, and S u,t, the probability distribution over the check-in locations of u at time t is governed by the following formula: P(cu = l|t, Hu,t, S u,t), (1) where cu denotes the check-in location of user u. Various temporal information related to cyclic patterns can be implied by t (e.g., "2013-02-22 11:09:59 pm") to indicate a user's check-in state, such as a specific hour of the day (11:00 pm), a day of the week (Friday), or a month of the year (February), etc. [8].

In Figure 4, we can see the traffic loads of some typical applications, i.e., WeChat, news, and electronic music services, in three typical urban scenarios, i.e., business zones, restaurants, and accommodation at different times of the day for a week. All data is collected in a major city in northeast China. It is clear that traffic changes either within a day or between different days, i.e., time-varying characteristics. Also, the traffic loads in different zones are not the same, i.e., geographical variants. Also, the traffic of each application has its unique features, i.e., features related to the service [5].

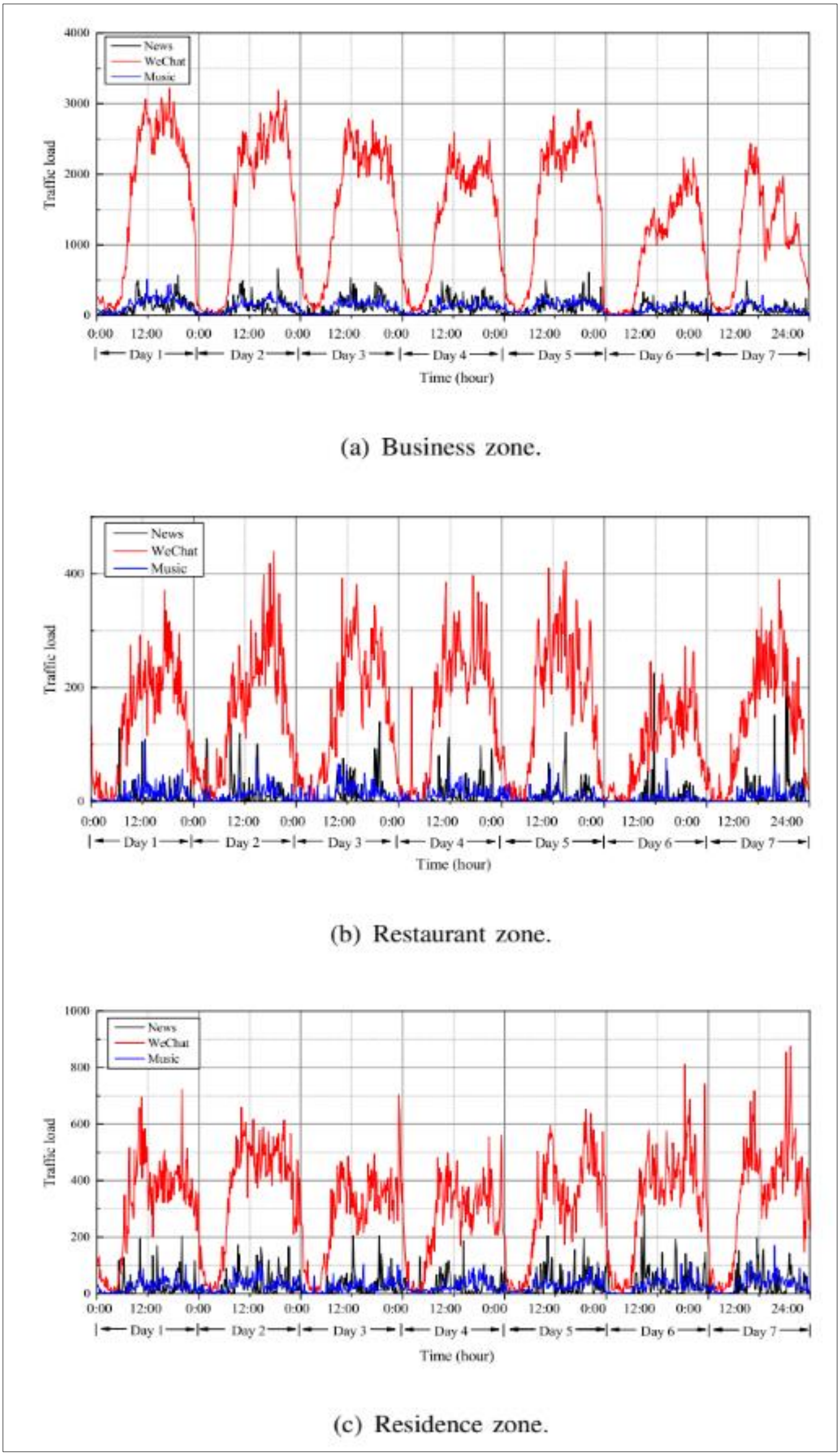

Figure 4 Measured traffic loads of several applications in typical urban scenarios [5].

5 Influence of the spatial scale of aggregation

5.1 Hotspots

In hotspots, the identification process, such as the size of the grid cells in which they are collected, the numbers or densities of users, is another arbitrary parameter. Since researchers do not want to specify this value separately for each city, they consider that different sizes should be tested for each city and that it is reasonable to assume that this cell size can range from 500 meters to 2 km. The cell size should primarily be chosen based on a reasonable size for an urban hotspot. Below 500 m, it would be necessary to assemble consecutive hotspots. For example, for a=100 m (10^-2 km^2 cells), two consecutive hotspots could not be distinguished as quickly as two different pedestrian points.

On the contrary, a size of 2000 m can be considered as an upper limit for the same reasons. If two consecutive cells are classified as hotspots, it makes sense to recognize them as two separate neighborhoods. In the case of a=1000 m (1 km^2 cell), researchers considered that two neighboring hotspots are two different hotspots. The index values must be robust by changing the cell size [9].

Below (Figure 5), we can see some graphs. This graph shows eight different cities, with different sizes, the time evolution of the # hotspots / # cell ratio for two definitions of hotspots, and different grid cell sizes. The selected cities cover the full range of the population size distribution of 31 cities studied. It can be seen that the quality pattern remains the same regardless of the size of the grid for the pair (city, method) [9].

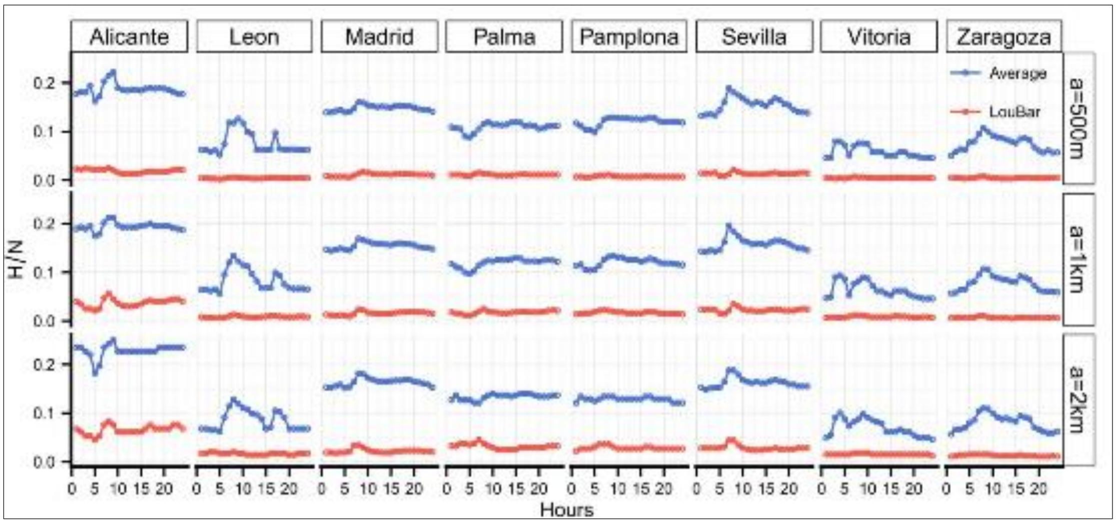

Figure 5 Time evolution of the ratio #hotspots/#cells for two hotspots [9].

5.2 Number Of Hotspots

Figure 6 shows the scaling relationship between the number of access points to the population and the effect of grid size. Here we see that the scaling effects and the value of the exponent are robust against a change (i) of the limit used to determine the access points and (ii) the size of the grid cells [87]. Scatter diagram and model adaptation line of the number of H points versus population size P for the 31 cities studied. The linear relationship in a log graph shows a power-law relationship between the two quantities, with an exponential value b, 1, which indicates that the number of activity centers in a city increases linearly with its population [9].

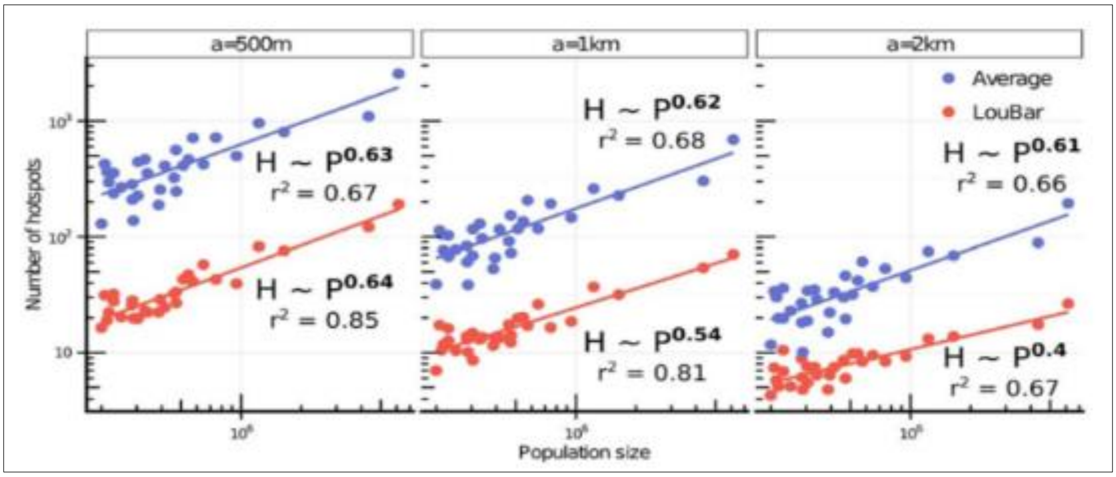

Figure 6 Scatter plot and model fit line of the number of hotspots H vs. the population size P for the 31 cities [9].

5.3 Stability of the Hotspots Hierarchy

Another exciting feature to check-in cities is the stability of their hotspots and the evolution of their relative importance in the city. Hence, the time of day is related to the evolution of the hierarchy of places in the city. For each city, researchers count the number of hourly bins during which each cell was a hotspot. As shown in Figure 7, the distribution of hotspots is "lifetime" (measured in several one-hour bins) for the eight largest cities in Spain. Each city has a number of its essential locations, those that constitute the center of the city. In addition to permanent hotspots, we can observe two other main groups, a set of intermediate hotspots (with a lifespan of half a day) and "intermittent" hotspots that exist only a few hours a day. Permanent hotspots are the most critical locations in the city in terms of population density. The results show that the city centers are stable both in space and time, regardless of size [9].

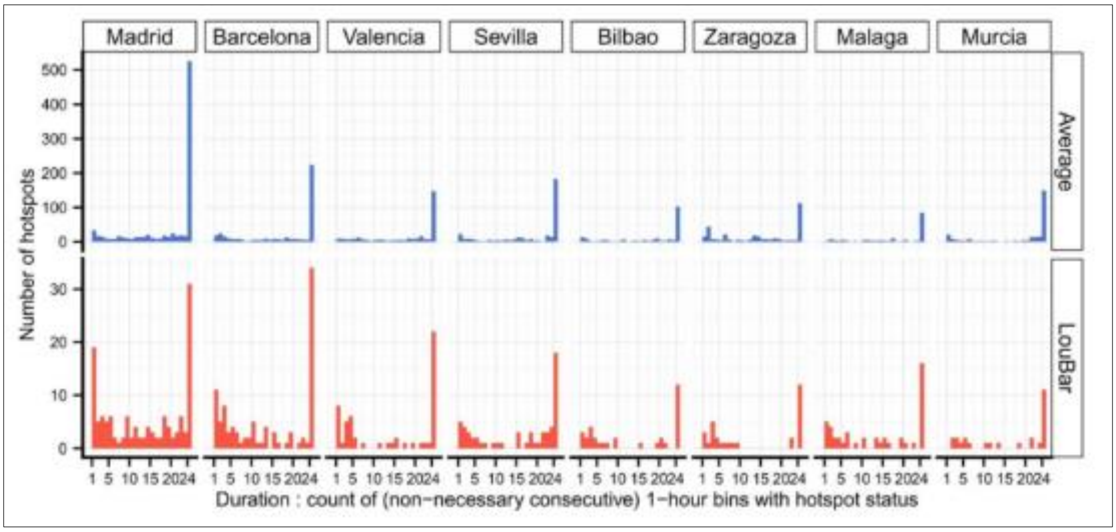

Figure 7 Histogram of lifetime duration of hotspots for eight cities and the two hotspots identification methods (top: 'Average' method and bottom: 'Loubar' method) [9].

5.4 Mobile Users' Regarding Mobility And Habits

As we can see from previous research studies, we summarize the below attributes for the individual and social networks, regarding mobile users' mobility and habits.

5.4.1 Socio-economic status through Calling Card

The vast majority of subscribers are in prepaid programs, which require the periodic purchase of scratch cards, a ubiquitous commodity immediately available in both urban and rural areas of the county [7].

5.4.2 Travel: Distances between Cellular towers

People in rural areas travel significantly more per month than people living in cities. One reason for this could simply be the small possible distances covered within the capital and the much longer distances within rural areas [7].

5.4.3 Frequency and volume

As previously thought, we can confirm that people living in urban areas tend to communicate almost 50% more than people living in rural areas [7].

5.4.4 Degree and average volume per degree

Previous studies have found that people in rural areas have relatively strong ties to fewer people, while people in urban areas tend to have more but weaker ties. People living in rural areas have a lower rate than those living in the capital. However, while these people in rural areas may have fewer ties, they have higher average social endurance. This effect validates and quantifies the qualitative theory about the role of the city in personal networks [7].

5.4.5 Decision Making

As the use of mobile network data reveals reservations about the analysis of human mobility, the measures proposed may be convenient when it comes to quantifying the bias they introduce. Such measures can be combined in many decision systems considering the proposed measures depending on the studies (Figure 8) [10].

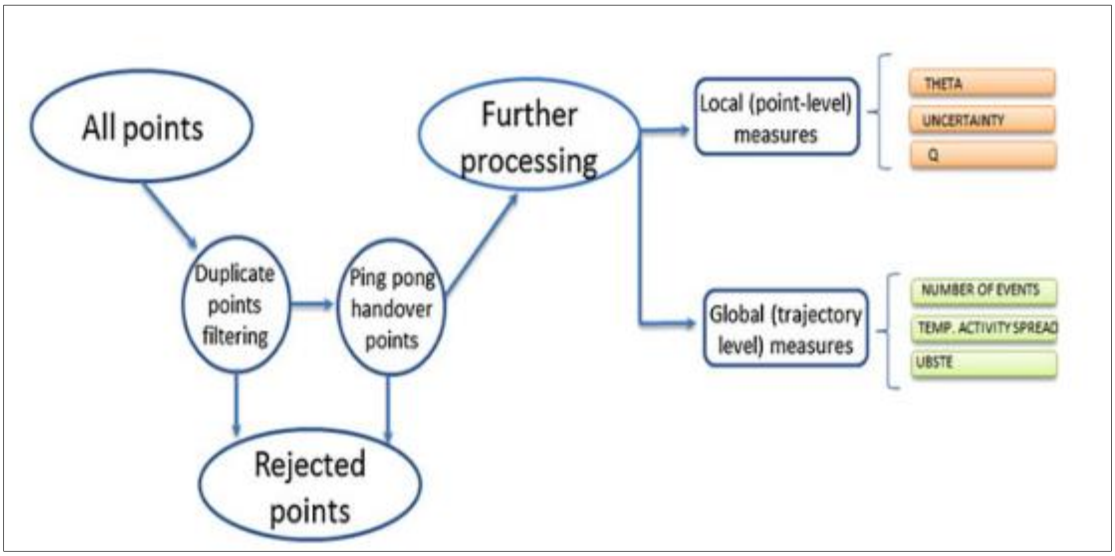

Figure 8 Decision system for points filtering and user's selection [10].

6 From Reactive To Proactive Networks

The preventive approach uses the existing heterogeneous cellular network. Also, it includes the design of radiofrequency resource management forecasting techniques to maximize the efficiency of 5G networks. Such predictability can be utilized to minimize the maximum load on cellular networks, preventing pre-temporary storage of the desired information to selected users before requesting it. Utilizing the powerful processing capabilities and large memory storage of smartphones allows network operators to serve predictable peak requests during off-peak hours proactively. When the proxy network serves users' requests ahead of time, the corresponding data is stored on the user's device, and when the request starts, the information is extracted directly from the cache instead of accessing the wireless network [6].

To this end, new machine learning techniques need to be developed to find optimal compensation between predictions that result in the recovery of content that users ultimately do not request and requests that are not predicted on time. Analyzing user traffic and temporarily storing content locally on small base stations (SBS) and the end-user can significantly reduce traffic bounce, especially when networks are flooded with similar content requests. Therefore, the goal is to predict and infer future events intelligently. Hence, it is a complex problem exacerbated by the large sample of data caused by large and sparse data. Indeed, data failure is a significant challenge, as it may not always be possible for a single user to collect enough data to predict its patterns accurately [6].

6.1 Leveraging Social Networks

The vast majority of data traffic is carried out by social networks, which have played a key role in disseminating information via the Internet, and will continue to shape how information is accessed. For example, geotagging data in social networking applications can help operators track where people generate mobile data traffic for better small cell development. A byproduct of this helps operators with other aspects of network design, such as small cell transfer, managing multilevel interference (knowing which cell the user will connect to next), power management, and greener networks serving only users when you are close to a microcell [6].

6.2 Network Capacity and Coverage

PMs should be aware of their long-term installation goals in terms of network capacity, coverage, number, and location of base stations. New resource allocation strategies are also needed to meet different traffic requirements across the coverage area. To achieve these goals, PMCs monitor the quality of network service (QoS) through smartphone driving tests. The measurement results are collected from selected smartphones, or specific test mobile phones in their networks analyzed with specialized software. However, this is not cost-effective due to excessive time and human resources and is also inaccurate due to limited test samples [5].

Hence, in recent years, researchers in telecommunications networks have begun to look at the big data in the design toolbox. Featuring hundreds of tuned parameters, the wireless network design updated by big data analytics received the most attention. However, other types of networks also received increasing attention. The massive amount of data that can be collected from the networks, together with the distributed modern high-performance computing platforms, can lead to a new economical design space (e.g., reduction of the total property cost by using a dynamic virtual network topology adaptation) in comparison to classical approaches (i.e., static virtual network topologies) [11].

6.3 Network Design and Big Data

The uses of Big Data analytics can offer a new way of tackling these problems. Network analysis includes monitoring, real-time and user data analysis, mobile networks, and service providers. There are several stages that BDD approaches can help Providers to develop and operate their networks more efficiently [5]. Those new standards promise to transform networks from waterproof data tubes to frame-oriented networks. The Role of Big Data in Mobile Network Design is vital for the design of cellular networks. Also, we can see below the following classifications [11]:

- Failure related: This includes error tolerance (i.e., detection and correction), predictions, and prevention techniques that use big detailed data in cellular networks.
- Network monitoring: This shows how extensive large-scale data analysis is as a large-scale tool for monitoring data traffic on cellular networks.
- Cache related: Explores how much data can be used for content delivery, cache placement and distribution, location caching, and preventive caching.
- Network optimization: Big data can be involved in various issues, such as predictable wireless resource allocation, interference avoidance, network optimization in the light of experience quality (QoE), and flexible network design in the light of consumption [11].

A Big data-driven Framework for Mobile Data Network Optimization When talking about optimizing a mobile network, it is essential to gather as much information as possible [11]. Extensive networks and their users create a wealth of data, for which the use of sizeable analytical data is vital to the analysis of the colossal amount. This framework includes several stages, starting with big data collection, storage management, execution of detailed data, and the last stage of the procedure is network optimization. Three case studies were used to show that the proposed framework could be used for mobile network optimization [11].

Resource Management in HetNets Mobile Network Providers (MNOs) may use big data to provide real-time and historical analysis between users, mobile networks, and service providers. MNOs can benefit from BDD approaches to the operation and development of their network, which can be done in several stages [11]. Below we can see them.

6.4 Network Design

Due to incomplete statistics, evolving node B (eNB) sites are not optimally optimized. Hence, it can be addressed if sufficient information (user and network) is provided for analysis. Big data analytics can help MNOs make better decisions about deploying eNB in mobile networking. The researchers suggested using the network and anonymous user data (e.g., dynamic location information, and other service features) [11].

6.5 The 5G User Network Focus

The Network Design improving the user experience, providing higher data rates, and reducing latency are considered a vital goal of a 5G system. In most traditional development cases, eNB websites are not optimized due to a lack of sufficient statistics. By monitoring mobile devices, their detailed activities can be recorded to provide real-time information about where, when, and what information about mobile users on the network. A viable solution is to use both the network and anonymous user data, including dynamic location information and various other service features [5]. Researchers expect the following to be critical elements in designing a user-centric 5G access network [11].

6.6 Personalize Local Content Delivery

The access network needs to evolve from being an agnostic user and service, merely acting as a blind pipe connecting the user to the central network, to become user-centric. The researchers suggested several steps to achieve content delivery and are as follows [11]:

- Obtaining traffic and user information: Traffic characteristics (application type, server address, and port number, etcetera.) are gathered through packet analysis and analyzed using a clustering algorithm (e.g., k-media) to execute traffic markup (news, sports, etcetera.) [11].
- Analyze and predict user requirements: This step is achieved with a large data analysis algorithm (an algorithm such as collaborative filtering is recommended), which uses the traffic features mentioned above, and therefore suggests content based on the user's interests [11].
- Local caching and content management: Famous content is provided in the form of local copies. Content that may be of interest to users should be cached locally after downloading from the application server [11].
- Content Delivery: When the user initiates a request, the system will check if the content is stored locally, ready to be sent directly. High data analytics can provide content recommendations, so it will check if it is stored locally and push it directly to the user [11].

7 Provide a flexible network and development of functions

The researchers note that processing the characteristics of both the local user and the service using extensive analytical data can significantly contribute to the rigid development of networks and functions in the following ways [11]:

7.1 Flexible network development

As 5G will support different low costs (AP) access points, such as AP coverage and AP hot spot, the use of extensive analytics data can be useful in predicting traffic characteristics, thus creating the basis for achieving dynamic network development for APs [11].

7.2 Flexible application functionality

Application functionality is the analysis and the prediction of peripheral requirements of users and services can be performed using extensive analysis data [11]. Therefore, the bulk, speed, and variety of data must be processed with advanced analysis techniques, turning the data into practical knowledge. To better understand traffic trends, it is imperative to analyze the respective content and events [5]. Given the knowledge gained from large data sets, PMCs can make wise decisions about where and how to deploy computer networks in a cellular network. It also allows them to anticipate traffic trends and prepare plans for future investments [5].

8 Resource Allocation

Resource requirements vary depending on the density and usage patterns of mobile network subscribers. Using the cloud RAN architecture, the right place at the right time can be served through predictable resource allocation, so minimal downtime can be achieved [11]. Hence, HetNets with small cells can be used to coordinate interpolation between macro and small cells. Schemes such as enhanced intercellular interpolation coordination (eICIC) in LTE-Advanced effectively allow resource allocation between intervening cells and improved load balancing between cells in HetNets [11].

EICIC allows macrocells to evolve at node B (MeNB) and neighboring small eNBs (SeNBs) to transmit data in isolated subframes, thus preventing MeNB interference to SeNB. For the implementation of eICIC, a particular type of subframe called Almost Blank Subframe (ABS) has been identified that carries minimal and more basic control information. It is worth noting that ABS subframes are transmitted at reduced power and that the network administrator can control the configuration of this subframe. Many factors contribute to determining the macro-cell ratio to small cell ABS. Such as traffic load in a specific area, type of service, and so on. The best ABS ratio varies dynamically, and this is since the interferences between cells change over time for the factors [11].

In a BDD system, optimization of radio resource allocation can be achieved by using network analytics. The development of BDD optimization features in MeNB would allow them to collect and analyze big raw data derived from eNB (e.g., service attributes and traffic attributes) in real-time, thus allowing a quick response [11].

9 Intervention Coordination

Within a small cell HetNet, tuning interferences between macro and small cells must be performed in the time domain instead of the frequency domain, e.g., the enhanced Inter-Cell Tuning Interference (eICIC) LTE-Advanced mode [5]. A new enhanced ICIC (eICIC) scheme for multilevel networks has developed for LTE-Advanced, which provides time-resource allocation between network tiers for better performance. The idea of eICIC is specifically designed to handle potential downlink interference problems in multilevel networks with the simultaneous development of macros, micro, and HeNBs. Hence, good macroeconomic coverage and performance over a wide area are still essential. So one of the causes is how to develop and manage these small base stations without compromising performance [12].

Such a scheme allows efficient allocation of resources between intervening cells and improves load balancing between cells in HetNet. The basic principle behind eICIC is that it allows a macro of ENBs (menBs) and its adjacent small ENBs (SeNBs) to transmit data in separate subframes, that is, to be kept rectangular in the time domain, mostly avoiding interference from menB to SeNBs. Thus, when communicating with cellular acne UEs, SeNBs use subframes orthogonal to their macrocell adjacent, thus avoiding possible interference from the MeNB. Meanwhile, SeNBs can transmit to UE modules in their cell in any subframe regardless of whether MeNB is currently transmitting data [5].

For eICIC purposes, a particular type of subframe is defined, i.e., the Almost Blank Subframe (ABS), which carries no data but only minimal control information, e.g., the reference signal, the mandatory system information, and so on. Thus, there is no interference in the data signals, while the interference caused by the control signals can also be mitigated. In an LTE system, a radio frame consists of ten subframes. Each can be used either as an average subframe or as ABS by the eNB, except for subframes 0 and 5. The decision on how to configure the ABS subframes is made by the network operator [5].

However, determining an appropriate macro-cell ABS ratio to micro-cells depends on many factors, e.g., the types of services, the traffic load in the given area, etc. As is well known, service behaviors in small cells vary over time. Also, the traffic patterns of the individual services are changing. Thus, intercellular interference does not remain constant. Therefore, the optimal ABS ratio essentially changes dynamically over time [5].

In a BDD system, network analyzes can be used to optimize the allocation of radio resources. Resource allocation can adapt to both environmental changes and traffic changes based on information obtained from data analysis. Hence, to enable a fast response, some BDD optimization functions can be deployed in MeNB so that they can collect and analyze early big data from eNB, e.g., service and traffic characteristics. As a result, the performance of each shell and user can be optimized. It can be done by regularly processing raw data for statistics and automatically detecting traffic fluctuations to predict optimized ICIC parameters, such as the ABS ratio [5].

Besides, a global optimization process can jointly optimize the location and traffic requirements of users of multiple computers. For example, a particular SeNB can be turned off to prevent interference with its nearby SeNB, which may have higher traffic due to its higher signal-to-noise ratio (SINR). Also, reducing energy consumption may be another optimization goal to consider [5].

10 Failure Prediction, Detection, Recovery, and Prevention

10.1 Analysis Between Technologies Failed Using Big Data

Many of these events occur when the user is on the edge of a coverage area and moves to another, technologically different area, e.g., moving from a 3G (BS) base station to a 2G BS. Standard solutions to address such deficiencies are either conducting motion tests or performing network simulation. The proposed solution uses extensive data analytics (Hadoop platform) to analyze the messages of the Basic Application System (BSSAP) exchanged between the nodes of the Basic Subsystem (BSS) and the Mobile Switching Center (MSC). Another comparison was made with the Performance Index (KPI) approach, and the results were in favor of the proposed approach [11].

10.2 Intelligent LTE Network Optimization Based On Data

Using all-signaling data and user interface and wireless network data, combined with Self-Organized Network (SON) technologies, full-scale automatic network optimization could be performed. The authors developed a smart cellular network optimization platform based on signaling data. This system includes three main stages [11]:

- Defining network performance indicators by exporting XDR keywords: External Data Representation (XDR) contains the necessary signaling information (e.g., causes of process failures and signaling types). The status of a complete signaling process can also be recognized by the XDR (e.g., success or failure of the installation and release of the signaling). These indicators can be searched from many dimensions and levels (e.g., user level, cell, and grid-level) [11].
- Troubleshooting: Service installation rate, delivery success rate, and dropout rate are among the network signal level states that can be reflected by XDR-based network performance indicators. Network hardware with unsatisfactory performance indicators can be further analyzed, and this can be done by further excavating the initial signaling of the respective indicators [11].
- Providing best practice solutions: Recognized and solved problems can offer an optimization experience. As a result, a diversity of network issues can be verified. For example, when a cell has a low delivery success rate, as defined by the relevant indicators, the ratio is proposed to be the low delivery success rate. The solution would be to tune the overlapping coverage areas formed between the source and target cells and the parameters (e.g., shifting the decision threshold and initiating delivery) [11].

11 Anomaly Detection in Cellular Networks

11.1 Self-Healing In Cellular Networks

An abnormal and disturbing service can be detected by examining the Call Detail Record (CDR) of users in a specific area. CDRs are generated during a call and include caller IDs, call duration, caller location, and cell ID where the call was initiated or received. The process is done first, is the CDR data collected from the network nodes, and stored on a proxy server. The second phase begins with distributing CDRs collected in the relevant departments (e.g., data warehouses and billing services). The Hadoop platform processes massive data sets that require distributed processing in groups of computers. The process must be carried out within a reasonable period, not to degrade the quality of the services provided. Three cases of use are suggested by researchers for a process of self-healing in cellular networks [11]:

1- Data Reduction: The operation and maintenance database (O&M) can be used for troubleshooting purposes. However, the size of the database is relatively large as it contains data related to both standard and degraded intervals, which makes it challenging to edit. The authors proposed this parallel process independently by analyzing each BS separately. They chose the degraded interval detection algorithm (the degraded interval is a time when the BS behavior is degraded), and these intervals were detected by comparing the BS KPIs with a certain threshold. This algorithm was paralleled with its application as a map function. A field was added to identify each BS, and all fields are added by a reduction function [11].

2- Sleep cell detection: Cell interruption or sleep cells are a common problem in cellular networks. Users are directed to adjacent cells instead of the nearest and most optimal cell. According to the algorithm described in, sleep cells can be detected using adjacent BS measurements, thus calculating sleep cell disruption. The detection process is based on the resource output period (ROP), where each BS generates configuration management (CM), fail management (FM), and performance management (PM) data every 15 minutes. If the number of deliveries suddenly drops to zero, and a malfunction is indicated by the cell performance (PI) indicators, it is considered a sleeper cell. Researchers suggest the use of the above algorithm based on the principle of essential data. Propose to divide the ground into compartments that are the maximum distance between neighbors, where each BS within the partitioned area is tested sequentially by the presence of the algorithm, and this is made by examining the data of its neighbors. This method was compared with other approaches (e.g., lack of KPI and KPI), and most of the simulation was detected inoperative [11].

3- KPI Correlation-based diagnosis: Map-Reduce was used to implement this algorithm in parallel, the correlation process and the creation of a PI list sorted by correlation was implemented as a map and reduce the functions, respectively [11].

11.2 Cell Site Failure Prediction

Abrupt discontinuation of services can have serious consequences, and that is why it is essential to keep communication equipment, such as cellular locations, in good working order. The challenge for the researchers is to analyze the user's bandwidth at the cellular level. Due to the size and variety of the data collected, it is necessary to use extensive analytical data to process them. Data from different data sources are then integrated and then analyzed to know bandwidth trends [11].

12 Network Monitoring

12.1 Monitoring and analysis of large-scale mobile telephony traffic

Large cellular networks have pretty high data rate connections and high requirements to be met. However, with the ever-expanding data rates, data volumes, and requirements for detailed analysis, this approach seems to have limited scalability. Researchers, therefore, proposed a system to undertake this task, using Hadoop Map Reduce, HDFS, and HBase (a distributed storage system that manages the storage of structured data and stores them in key or value pairs) as an advanced distributed computing platform. They took advantage of its ability to handle large volumes of data while operating on commodities. The proposed system was developed on the central side of a commercial cellular network and handled 4.2 TB of data per day via 123 Gbps links with low cost and high performance [11].

A good example is the China Unicom, China's largest WCDMA 3G mobile phone company with 250 million subscribers in 2012, has introduced an industrial ecosystem. The researchers highlighted it as an ecosystem-based on telecommunications operators based on a large data platform. The aforementioned big data platform has been developed to retrieve and analyze data generated by mobile internet users. To optimize storage, enhance performance, and speed up database transactions, the authors proposed a platform that uses HDFS for distributed storage. The cluster had 188 nodes used for data storage, performing statistical data analysis, and management nodes. 186 M.S., Computer Networks Compared to the Oracle database, it is noted that the system achieved four times lower import rate. The query rate was also compared to an Oracle database, and HBase performed better, taking into account the impact of the record size [11].

13 Cache and content delivery

13.1 Optimized Bandwidth Allocation For Content Delivery

Mobile networks usually have many users, and with the growth of Internet-based applications, it has become necessary to allocate the required bandwidth to meet users' expectations and ensure a competitive level of service quality. Cellular networks can provide an Internet connection to their users at any time. In terms of the base station, the impact of forwarding the same video content to multiple users on the same base station is enormous. The idea is based on the sharing of the base station's wireless channel by a group of users who want to download the content. Thus, saving the base station resources, providing a better data rate for cluster users, and providing an opportunity for non-cluster users to take advantage of the stored resources (bandwidth).

It should be noted that the assembled users can receive the contents from the cluster head using short-range communication techniques such as Wi-Fi Direct and Device to Device (D2D). Two conditions must be met before creating a cluster of users. First, users who request the same content are the ones who form the cluster. Second, users must be or will be within walking distance of each other. For this reason, the authors suggested the use of extensive analytical data to determine the proximity of users and to group users in the cluster(s). A cluster head is then selected from nearby users, and the process is repeated among the base station users until there is either a cluster of users or a free (cluster-free) user. The simulation was performed for a base station network, and the results showed faster content delivery and improved user-level performance [11].

13.2 Tracking And Caching Popular Data

The number of social network users is vast. However, important and vital events attract much attention, and, as a result, much content is shared on these networks. When a particular video or event goes viral, this sharing will ultimately burden the network, as the requested content will have to travel with the network on its way to the servers. So researchers suggest monitoring popular websites and social media, analyzing data, determining if there is a growing interest in specific content. Also, by which age group and temporary storage of popular data for a specific database station. The result would be content temporarily stored for users to be faster and without the network burden [11].

13.3 Proactive Caching In 5G Networks

Cache-enabled base stations can serve mobile subscribers by predicting the most strategic content and storing it in their cache. An approach proposed by the authors used high data analysis and machine learning to develop a preventive caching mechanism, predicting the distribution of content popularity on 5G mobile networks. After collecting raw data, i.e., user traffic, the massive data platform (Hadoop) has the task of anticipating user requirements, exporting useful information such as Location Area Code (LAC), Hypertext Transfer Protocol (HTTP) -Uniform Resource Identifier (URI), Tunnel Endpoint Identifier (TEID) -DATA and TEID request for control and data levels. Experimental testing of this work at 16 base stations, as part of a functional cellular network, resulted in 100% satisfaction of requests and 98% backhaul unloading [11].

14 Optimizing The Resource Allocation In LTE-A/5G Networks

To improve the user experience, the authors proposed an approach that uses user and network data, such as modulation and log files, alarms, and database entries/updates. This approach is based on extensive analytical data for the processing of the data, as mentioned earlier. The primary goal is to provide an optimal solution to the problem of radio resources to RAN users and to have a minimum delay time between the resource request and its assignment. The goal is done by identifying the user and network behavior, which is a task that fits well with big data. The proposed framework includes three stages [11]:

First stage: This process is performed in the eNB system; it processes the data from the cellular and primary network side. Binary values are obtained by comparing cell-level KPIs with the corresponding threshold values, thus keeping the binary grid up to date. This process is repeated at regular intervals [11].

Second stage: Repeating the same steps as in the first step. However, this process is performed on subscriber-level data to obtain a subscriber KPI and maintain a binary table [11].

Third stage: This is enabled when a user initiates a resource allocation request. A binary pattern is created based on user requirements. This pattern is delivered later in step 2 for updating the binary table (if required) and integrating the new values into the order that represents the requested bandwidth. To determine which PRBs suit the user, the fuzzy binary pattern algorithm was used for this purpose [11].

14.1 Framework Development For Big Data Empowered By SON For 5G

Researchers proposed a framework called Big Data Empowered SON (BSON) for 5G cellular networks. According to the authors, what makes BSON different from SON are three main features [11]:

- Having complete information about the status of the current network.
  • Having the ability to predict user behavior.
  • Having the ability to connect between network response and network parameters [11].

The suggested framework contains functional and functional blocks and includes the following steps [11]:

- Data collection: A data set is collected from all sources of information in the network (e.g., subscriber, cell, and network kernel levels) [11].

- Data Transformation: This involves converting big data into the correct data [11].

This process has many steps, starting with:

  - a: Classification of data according to key business and business objectives (OBO), such as accessibility, sustainability, integrity, mobility, and business intelligence [11].
  - b: Consolidation/diffusion stage and the result of this stage are more significant KPIs obtained by consolidating multiple performance indicators (PIs) [11].
  - c: According to the KPI impact on each business objectives, KPIs are ranked [11].
  - d: Filtration occurs at KPIs, affecting the OBO less than a predetermined limit [11].
  - e: Correlate for each KPI and find the network parameter (NP) that affects it [11].
  - f: Order the relevant NP for each KPI depending on the correlation power [11].
  - g: Cross-correlation of each NP with finding a vector that quantifies its relation to each KPI [11].
- 3- Model: Learn from the correct data obtained in step 2 that will contribute to developing a network behavior model [11].
- 4- SON engine run: New NPs are identified, and new KPIs are identified using the SON engine in the model [11].
- 5- Validation: If a new NP can be evaluated with expert knowledge or previous pilot experience, proceed with the changes. If the simulated behavior corresponds to KPIs, proceed with the new NPs [11].
- 6- Retraining / Improving: If the validation in step 5 was not successful, give feedback to the shift concept block, which will update the behavior model [11].

14.2 Optimization Of 5G Networks With User Behavior Awareness

Wireless resource optimization must be done according to user requirements and 5G services. According to the authors, big data can analyze user mobility patterns, predict traffic trajectory, and pre-configure the network accordingly. APs log each user's historical access point (AP) list. In this way, the data can be uploaded to a central unit for editing or a target AP in the event AP service change. Hence, using a large data analysis algorithm, the collected data is analyzed to predict traffic trajectories [11].

---

15 Extensive Data-Driven Network Operating System.

Researchers proposed a system that can maximize the efficiency of big data network operation. So, it can be accomplished by optimally allocating network resources to each AP and user. The proposed system consists of two parts [11]:

15.1 1- Decision-Making Area

The Decision-making area is responsible for collecting and managing the user, network upgrades, configuration, service, and terminal status information. This domain leverages high data analytics to provide the necessary configuration needed to initialize the network. For this domain to work correctly, it must understand the full picture of user and service requirements and the distribution of functionality on the network [11].

15.2 2- Implementation area

The implementation area is mainly for reporting the status of the user, the terminal and the network, the dynamic development, and the configuration of the network. Depending on the requirements after obtaining the personalized configuration, this domain can use the dynamic AP functions and configuration to create multi-connection terminals [11].

15.3 Deployment Of Cache Server On Mobile CDN

Cellular networks are currently experiencing an explosive increase in data traffic [5]. With the development of network applications and the increase of the network population, content delivery has become a central issue for current networks. The client-server structure is the conventional solution for delivering content to end-users, where the proxy server can be placed between the client and the server to store some popular content temporarily. Content Delivery Networks (CDNs) have been supported for many years, with copies of the original content being temporarily stored on a group of geographically distributed content servers [13].

MNOs considered CDN as an effective delivery method for popular content such as blockbuster movies [5]. The optimization of ICIC parameters (e.g., ABS ratio) can be achieved by periodically processing primary data to obtain statistics and automatically detect traffic fluctuations. Deploying a cache server on a mobile CDN Popular content (e.g., movies) can be delivered over a CDN, a method that is considered sufficient by many MNOs. Distributed cache servers need to be close to users to achieve fast response and reduce delivery costs. Due to the unique features of RAN, it was the primary interest of the authors [11]. The primary purpose of having their CDNs is to reduce operating costs while providing adequate support to their core businesses. It is essential to locate distributed cache servers on the CDN as close to the end-user as possible to reduce response time and reduce shipping costs, e.g., a distributed cache server that works with a central cache server a hierarchical CDN [5].

As content servers are placed close to the clients, the request can be satisfied by providing copies to the content server to the clients, thus reducing the response time. Peer to peer (P2P) networks have also been implemented for content delivery, where P2P users can communicate directly with each other for content sharing without having to rely on a central server. Compared to conventional solutions, in P2P, each node can act as a client to request content, and also the server to deliver content. Compared to the above CDN and P2P, the main difference is that the delivery of content to mobile social networks (MSN) is based on social characteristics, while the others are not. Content delivery to MSN has its inherent characteristics, as follows [13].

- Content is delivered based on social behaviors, including social ties, communities, and so on. For example, if there is a social connection between two mobile users in the same community, the content may be distributed [13].
- The amount of mobile content is enormous and continues to grow [13].

Researchers report that every MSN user can not only request content frequently but also create content to share with other mobile users who have the same interest. Therefore, because content delivery to MSN has the above different properties compared to other existing networks, there are new challenges in managing MSN. On the other hand, the cooperation of controllers and switches is required to reduce the control traffic and data traffic caused by delivering large volumes of content [13].

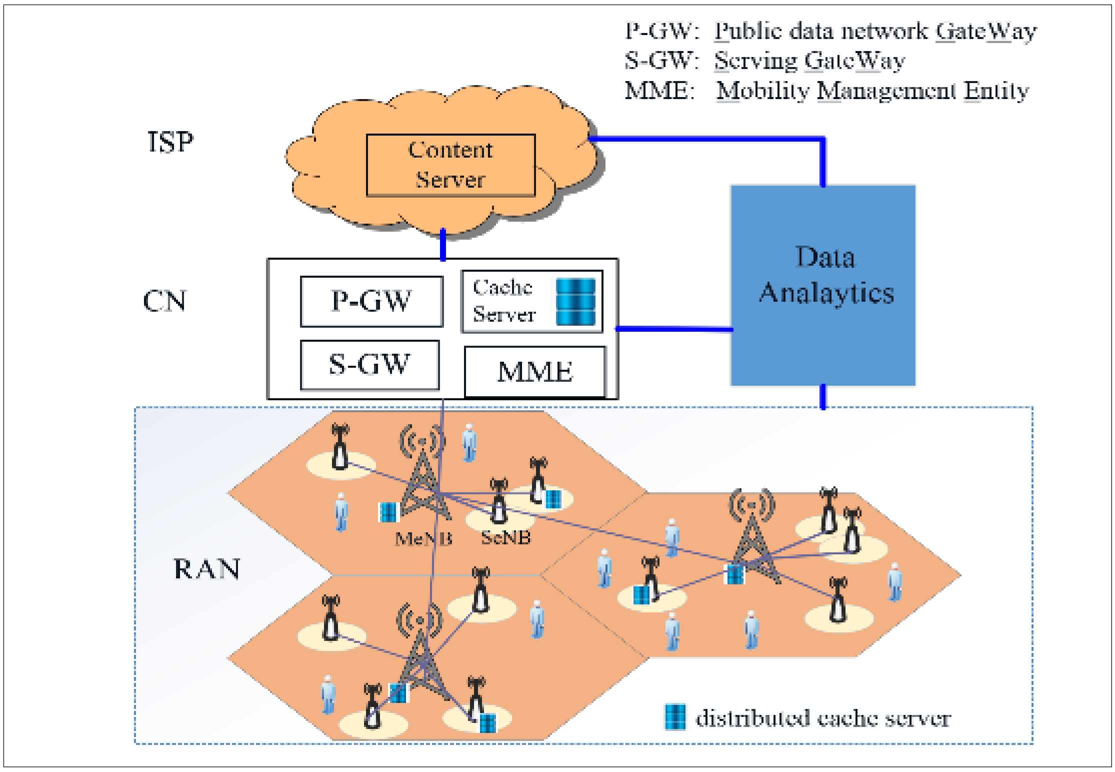

Figure 9 Illustration of the BDD cache server deployment [5].

However, the cache access rate to the distributed cache server may be lower than that of the central cache server. Sometimes the distributed cache server still needs mobile data traffic to bypass the relevant cache server via the backhaul link in improper placement. Therefore, it is vital to select the optimal location for the caching servers on the hierarchical CDN. In this section, only RAN is in our primary interest, as it has unique features compared to fixed CDNs [5].

As shown in Figure 9, it may be useful to place the distributed cache with the MNO wireless access network, allowing more efficient distribution of content on the edge of the network. Thanks to the hierarchical structure of the heterogeneous network with small cells, the MeNB cell site is another good location for the distributed cache server, as it is usually located in the center of the local network. Due to the backhaul capability among the eNBs or the eNBs to the CN, which is expected to be significantly enhanced in 5G networks, there is minimal concern about traffic congestion and transmission delays. Therefore, it is not necessary to deploy all MeNBs with a single distributed cache server. Additionally, a distributed cache server can be deployed in parallel with a SeNB if needed. The cost of storage and streaming equipment, features, and traffic load in a given area are crucial factors determining a cache server [5].

15.4 Metrics For Computational Effort

Algorithms are usually judged by the complexity of their space and time. So, the top BDD nodes and several BDD functions to quantify the space and time effort required for the supervisor's composition. Performing a supervisor composition twice with the same input and algorithm configuration will give the same result. Top BDD Nodes Due to the space explosion state, the complexity of the space is a limiting factor when applying the supervisor composition. During the supervisor configuration, the number of BDD nodes used to describe the system generally varies. Space effort can be measured by the maximum number of BDD nodes used in the synthesis [14].

Since reduced sorted BDDs are used, which are minimal representations, the peak BDD nodes used is the minimum number of BDD nodes required to represent the categories to solve the synthesis problem. In CIF, BDD nodes are stored in a hash table. Each new node is allocated to an entry in the hash table. Once the hash table has reached a specific fill rate, the garbage collection is used to release the most used entries. We only count the used BDD nodes, i.e., the hash table entries containing relevant information about the BDDs in use [14].

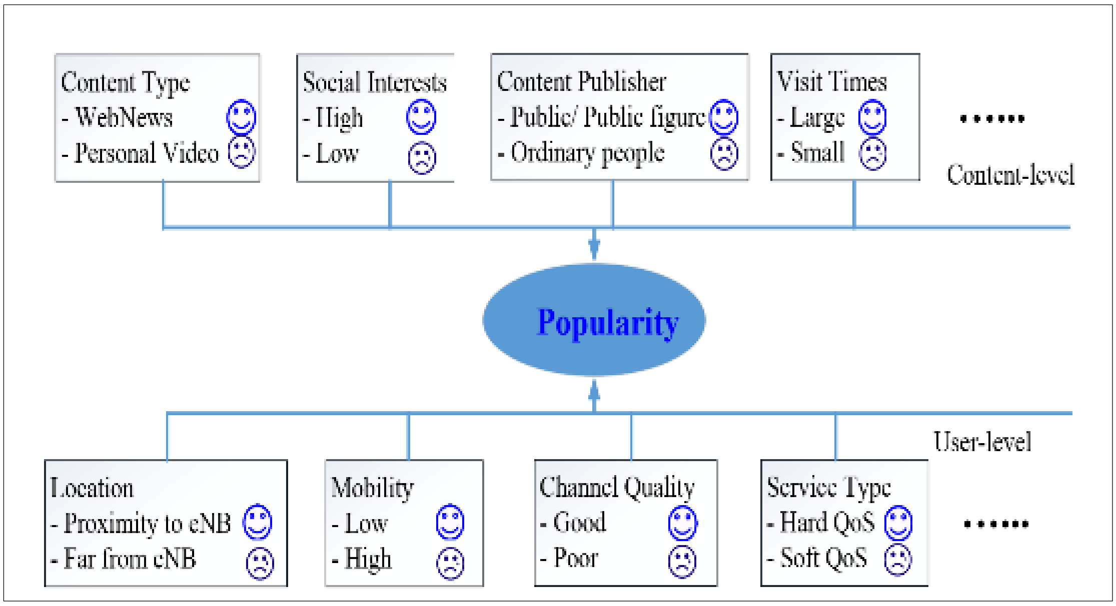

Figure 10 Different factors influencing content popularity in the mobile CDN [5].

Functions from applying this algorithm in the JavaBDD2 library are reused to measure the BDD nodes used. Top used BDD nodes are a reproducible measurement. Running a supervisor configuration twice with the same input delivers the same peak used by BDD nodes. [14]. After collecting data related to all relevant factors in the coverage area for a long time, cluster analysis can be used as a feasible method in data analysis to help MNOs develop cache servers in RAN [5].

• Pre-processing: Not all raw data is suitable for analysis, so some data must be deleted. For example, incomplete and redundant data must be filtered. Therefore, the main features can be selected from the rest of the data, e.g., traffic load, type of service, backhaul use, and recovery delay [5].

• Grouping: Each eNB position is considered as a group point. Each property specifies the value of a point dimension. Thus, there are many points, i.e., eNB Positions represented by a multidimensional vector. According to the principle of this grouping, all points can be divided into two non-overlapping groups, i.e., one is for cache server development while the other is not [5].

Analysis capabilities are integrated into the hierarchical CDN using collective information data architecture. Each cache server has a tracking agent to collect log information. This tracker sends log status information to the data analysis function block, determining when / which content to outsource and where to place the copies [5].

Highly popular content is more likely to be placed on cache servers to improve the cache access ratio. As shown in Figure 10, popularity usually depends not only on the content itself but also on the users. Also, user mobility can often cause cached content to change, resulting in inadequate content storage. Therefore, the data analytics feature must analyze data related to content and users to accurately determine or predict the popularity of content [5].

16 QoE Modeling for Network Optimization

Various services and applications using a set of QoS parameters (e.g., packet loss, latency, and jitter). However, management can be more effective when end-users perceive the quality, i.e., QoE is taken as an optimization target instead of QoS. To this end, automatic and accurate real-time QoE estimation is the first step. Data analytics can help model and monitor QoE in a different heterogeneous environment, which is essential for optimizing the global network [5].

In this direction, it is expected that there will be improved backhaul capability in 5G networks, and this will result in minimizing the concerns related to the delay and traffic load of backhaul transmissions. Optimal cache server placement depends on some factors, such as the characteristics and traffic load in a given area and the cost of storing and streaming equipment. According to help MNOs where to deploy caching servers, data analysis methods can be considered as a viable solution [11].

16.1 QoE Modeling To Support Network Optimization

Researchers believed that managing services and applications needed more than just relying on QoS parameters. Instead, they suggested that quality (i.e., QoE), as perceived by end-users, be considered an optimization goal. In addition to technical factors, non-technical factors (e.g., users' feelings, habits, and expectations, etcetera.) can affect QoE. A profile for each specific user that includes the non-technical factors would help evaluate QoE. Analytical data can determine what QoE affects users' devices and network services and resources. The next step is for network optimization functions to respond to the problem and select the optimal action accordingly [11].

Improving QoS in cellular networks through self-regulating cells and self-optimized delivery Cellular networks have a critical element on which the concept of mobility depends. This item is the delivery success rate, which ensures the continuity of the call while the user moves from one cell to another. Operators' efforts ensure that each cell has a list of manually configured and optimized adjacent cells. Therefore, it is vital to note the high probability that these cells will not adapt when a quick response is required due to a sudden network change. The researchers presented two methods that used the big data to introduce a self-formulated and self-optimized delivery process, the first correlated with the newly introduced cells, while the second dealt with the existing cells. A processor runs after the collection period and aims to check the files to see if they are marked as new cells (where Self-Configuration Analytics starts) or not (where Self-Optimization Analytics starts) [11].

As shown in Figure 11, the data required to estimate QoE comes from both the network and the users. In addition to the technical factors, various non-technical factors can affect QoE results, e.g., device type, user emotions, habit, expectation, etc. Thus, in QoE evaluation, it is useful to create an individual profile for each user, which is a user model that represents the preferences, habits, and interests of users. A user usually does not want to spend much time answering questions to create a profile model. Alternatively, a user profile can be constructed and monitored using detailed data with indirect information collected by a profile collector installed on mobile devices [5].

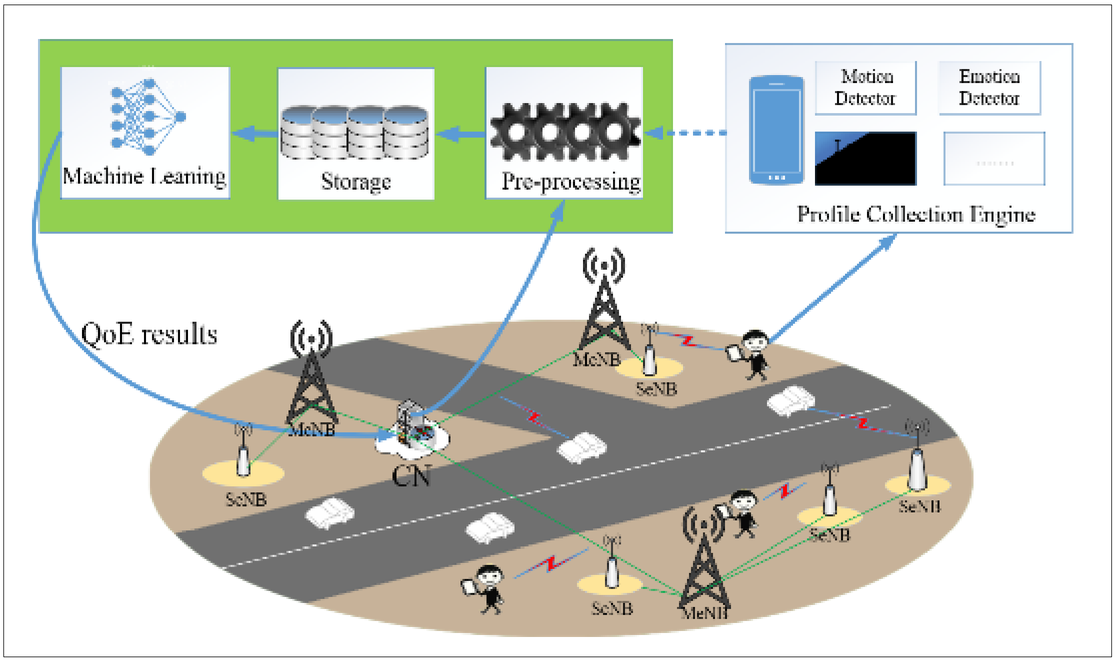

Figure 11 Illustration of QoE modeling driven by Big Data [5].

17 Emotional Techniques

User activities are monitored and compared to identify similarities and differences [5]. For example, emotional expression is vital as an essential part of our daily lives. Some researchers point out that one conveys information about current needs through the expression of emotions. According to communicate the information to others, the emotions expressed are observable in a person's face, postures, and verbal communication. Mobile-mediated communication does not always have less emotional bandwidth than face-to-face communication [15]. Hence, the motion detector output to the profile picker may include (but is not limited to) the number of clicks and scrolling on the screen. Emotional user behavior can be extracted from a detected user behavior by emotional computational techniques [5].

Meanwhile, network data, including QoS parameters, is collected through measurement and signaling on the network. All data is stored in a database for further processing [5]. Hence, it can be useful for counting the experience quality (QoE), defined as the successive overall quality of an application or service, and is commonly used to capture the accurate perception of users. In general, the factors influencing QoE can be classified into subjective and objective categories. Subjective factors include user emotion, experience, expectation, and so on, while objective factors relate to both technical and non-technical aspects of services. For example, video services are ultimately monitored by human observers; the observer's subjective opinion is the best QoE indicator of video services. Therefore, subjective tests are the most relevant QoE measurements because the results are obtained directly from humans. However, subjective quality assessment requires significant human resources and time. Due to the limitations of regular subjective measures, researchers rely on simple objective measures for video services [16].

A machine learning mechanism is then used to determine the relationship between influencing factors and QoE through artificial intelligence. Machine learning techniques make it possible to make more accurate decisions over time, even when data sets are incomplete or new situations arise [5]. These methods are mainly based on specific objective measurements, such as maximum signal-to-noise ratio (PSNR), structural similarity index (SSIM) measurement, video quality measurement (VQM), and so on [16].

Next, another approach, called pseudo-subjective quality assessment (PSQA), was proposed because of its combined advantages of both subjective and objective methods. For some standard approaches, such as the neural network model used in the pseudo-subjective quality evaluation (PSQA) evaluation method, the QoE model must be trained before being used to evaluate performance. Analyzing large data sets leads to information about the actual user experience, which may need to integrate social data [5]. The quality assessment model is first trained from the results collected from subjective approaches, which reflect the effects of subjective factors. In convergence, this model can be used to predict the quality of video services without any human involvement (Figure 12) [16].

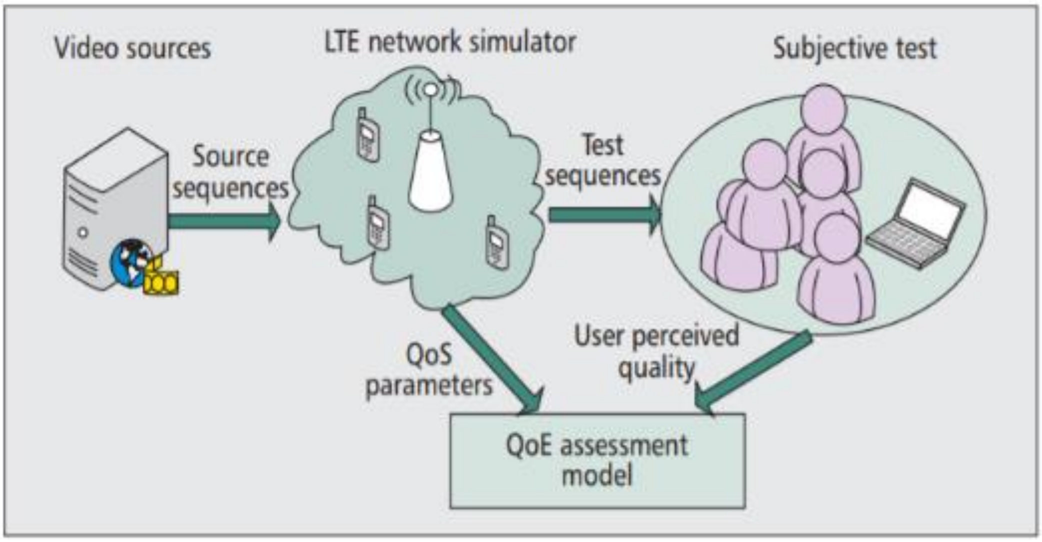

Figure 12 Framework of the video assessment method [16].

To fill this gap, researchers focus on critically evaluating video streaming over an LTE network, taking into account both the objective parameters of the network and the subjective factors of the users. A reliable QoE evaluation method is essential for wireless operators in designing effective radio resource management programs to meet the expected future demand for video streaming services at satisfactory QoE levels. Therefore, the researchers report a large-scale subjective evaluation campaign to create a database to evaluate the QoE of wireless video services using an LTE wireless simulator, which creates attenuated video sequences for different values the controlled QoS parameters. Accordingly, human observers' opinion scores are collected and averaged to generate the MOS as a QoE measurement for the given attenuated test sequence [16].

A new QoE evaluation method based on a two-step structure is also proposed by researchers to achieve good accuracy with achievable implementation complexity. Compared to the existing PSQA scheme, to avoid any local minimums and improve the achievable accuracy, particle optimization (PSO) is required to post-process Neural Networks (NN) weights when necessary. Samples that include QoS parameters and QoE measurements in the database can be used for training. Also, for the performance characterization of the proposed method "outside" the educational ensemble. Finally, the accuracy improvements with the proposed QoE evaluation method are characterized by our numerical results to demonstrate the advantages of the proposed PSO with the help of NN weight after processing [16].

Hence, data analytics can determine what EMPs need to know, which affects QoE quality across all devices, services, and network resources. Then network optimization features can instantly identify the cause of the problems and select the best action accordingly. In general, the ultimate goal of network optimization is to maximize QoE for users with the appropriate resource allocation while minimizing infrastructure costs through data analysis [5].

18 Conclusion

In this paper we study about the Binary Decision Diagram and the mobile network optimization. As we see about the binary decision diagrams, Lee and Akers first introduced the general concept of BDDs but later spread by Bryant, where it turns out that, according to the European Interactive Digital Advertising Alliance (EDAA) order and reduction rules, BDDs are a standard form of representation. After that, we study the spanning trees and BDDs. An advantage of BDD-based methods is that BDD provides a concise and separate description of system success/failure. The effectiveness of BDD methods is measured by the size of the BDD, which depends heavily on the selected series of BDD variables.

The next subject we concern about is the social network layer and HetNet resource management. A networking example must be essentially preemptive and is based on the fact that network nodes (i.e., base stations and mobile devices/smartphones) exploit user environment information. As a result, as part of the preventive networking example, network nodes monitor, learn and create user demand profiles to anticipate future requests, taking advantage of device capabilities and the vast amount of data available by intelligently leveraging traffic statistics and user environment information. We continue with the traffic loads and about users' behavior. If we can analyze a cell tower location data from hundreds of thousands of people, it is possible to quantify some of the most fundamental rules of human motion. In this case, a different type of social data collected by mobile service providers when adopting service plans (pricing). Other telecommunications products can be seen as the spread of social transmission. Also, location-based social network data contains three levels of information: a social level, a geographic level, and a time level.

Mobile network hotspots are essential, and as we see, the identification process, such as the size of the grid cells (in which they are collected) and the numbers or densities of users, are other parameters. For example, for a=100 m, two consecutive hotspots could not be distinguished as quickly as two different pedestrian points. On the contrary, a size of 2000 m can be considered as an upper limit for the same reasons. Hence, the scaling relationship between the numbers of access points to the population and the effect of grid size.

Also, we cover areas such as reactive to proactive networks where the proxy network serves users' requests ahead of time. The corresponding data is stored on the user's device, and when the request starts, the information is extracted directly from the cache instead of accessing the wireless network. Leveraging social networks where the vast majority of data traffic is carried out by social networks have played a vital role in disseminating information via the Internet and shaping how information is accessed.

Network capacity and coverage where PMs should be aware of their long-term installation goals in network capacity, coverage, number, and location of base stations. Network design and big data where big data in mobile network design are vital aspects for the design of cellular networks. Network design is due to incomplete statistics; evolving node B (eNB) sites are not optimally optimized. In this case, researchers suggested the use of network and anonymous user data. The 5G user network focus is the network design improving the user experience, providing higher data rates, and reducing latency are considered a vital goal of a 5G system.

Furthermore, we study the failure prediction, detection, recovery, and prevention, and we see the analysis between technologies that failed using big data. In this case, as we see, many incidents occur when the user is on the edge of a coverage area and moves to another, technologically different area, e.g., moving from a 3G (BS) base station to a 2G BS. Also, we see the self-healing in cellular networks. If a field was added to identify each BS, and a reduction function adds all fields. Finally, we see the cell site failure prediction includes data from different data sources that are then integrated and analyzed according to the bandwidth trends.

The network monitoring and analysis of large-scale mobile telephony traffic describe the large cellular networks, with relatively high data rate connections and high requirements to be met. Researchers, therefore, proposed a system to undertake this task, using Hadoop Map Reduce, HDFS, and HBase as an advanced distributed computing platform. Optimized bandwidth allocation for content delivery is mobile networks usually have a large number of users. With the growth of Internet-based applications, it has become necessary to allocate the required bandwidth to meet users' expectations and ensure a competitive level of service quality. Also, optimizing 5G networks with user behavior awareness is big data used to analyze user mobility patterns, predict traffic trajectory, and pre-configure the network accordingly. Hence, using a large data analysis algorithm, the collected data is analyzed to predict traffic trajectories.

Finally, we focus on an extensive data-driven network operating system. Researchers propose a system that can maximize the efficiency of big data network operation. So, it can be accomplished by optimally allocating network resources to each AP and user. The proposed system consists of two parts, the decision-making area, and the implementation area. We continue with the deployment of cache server on mobile CDN where It is essential to locate distributed cache servers on the CDN as close to the end-user as possible to reduce response time and reduce shipping costs, e.g., a distributed cache server that works with a central cache server on a hierarchical CDN. We study the QoE modeling for network optimization, the QoE modeling to support network optimization and emotional techniques. Various services and applications using a set of QoS parameters about QoE modeling and Improving QoS in cellular networks through self-regulating cells and self-optimized delivery. Cellular networks have a critical element on which the concept of mobility depends. In the end, emotional techniques including user emotion, experience, expectation, and so on, while objective factors relate to both technical and non-technical aspects of services.

Compliance with ethical standards

Acknowledgments

The Authors would like to thank the SPECIALIZATION IN ICTs AND SPECIAL EDUCATION: PSYCHOPEDAGOGY OF INCLUSION Postgraduate studies Team, for their support.

Disclosure of conflict of interest

The Authors proclaim no conflict of interest.

References

[1] T. van Dijk, R. Wille, and R. Meolic, 'Tagged BDDs: combining reduction rules from different decision diagram types', in 2017 Formal Methods in Computer Aided Design (FMCAD), IEEE, 2017, pp. 108–115.

[2] J.-M. Lu, F. Innal, X.-Y. Wu, Y. Liu, and M. A. Lundteigen, 'Two-terminal reliability analysis for multi-phase communication networks', Eksploat. Niezawodn., vol. 18, no. 3, 2016.

[3] L. Amarú, P.-E. Gaillardon, and G. De Micheli, 'An efficient manipulation package for biconditional binary decision diagrams', in 2014 Design, Automation & Test in Europe Conference & Exhibition (DATE), IEEE, 2014, pp. 1–6.

[4] C. Yanez-Marquez, I. Lopez-Yanez, O. Camacho-Nieto, and A. J. Arguelles-Cruz, 'BDD-based algorithm for the minimum spanning tree in wireless ad-hoc network routing', IEEE Lat. Am. Trans., vol. 11, no. 1, pp. 600–601, 2013.

[5] K. Zheng, Z. Yang, K. Zhang, P. Chatzimisios, K. Yang, and W. Xiang, 'Big Data Driven Optimization for Mobile Networks towards 5G', Nov. 2015. doi: 10.13140/RG.2.1.2389.1923.

[6] E. Bastug, M. Bennis, and M. Debbah, 'Living on the edge: The role of proactive caching in 5G wireless networks', IEEE Commun. Mag., vol. 52, no. 8, pp. 82–89, 2014.

[7] N. Eagle, Y.-A. de Montjoye, and L. M. Bettencourt, 'Community computing: Comparisons between rural and urban societies using mobile phone data', in 2009 international conference on computational science and engineering, IEEE, 2009, pp. 144–150.

[8] H. Gao, J. Tang, X. Hu, and H. Liu, 'Modeling temporal effects of human mobile behavior on location-based social networks', in Proceedings of the 22nd ACM international conference on Information & Knowledge Management, ACM, 2013, pp. 1673–1678.

[9] T. Louail et al., 'From mobile phone data to the spatial structure of cities', Sci. Rep., vol. 4, p. 5276, 2014.

[10] C. Iovan, A.-M. Olteanu-Raimond, T. Couronné, and Z. Smoreda, 'Moving and calling: mobile phone data quality measurements and spatiotemporal uncertainty in human mobility studies', in Geographic information science at the heart of Europe, Springer, 2013, pp. 247–265.

[11] M. S. Hadi, A. Q. Lawey, T. E. El-Gorashi, and J. M. Elmirghani, 'Big data analytics for wireless and wired network design: A survey', Comput. Netw., vol. 132, pp. 180–199, 2018.

[12] K. Pedersen, Y. Wang, S. Strzyz, and F. Frederiksen, 'Enhanced inter-cell interference coordination in co-channel multi-layer LTE-advanced networks', Wirel. Commun. IEEE, vol. 20, pp. 120–127, Jun. 2013, doi: 10.1109/MWC.2013.6549291.

[13] Z. Su, Q. Xu, H. Zhu, and Y. Wang, 'A Novel Design for Content Delivery over Software Defined Mobile Social Networks', IEEE Netw., vol. 29, pp. 62–67, Jul. 2015, doi: 10.1109/MNET.2015.7166192.

[14] S. Thuijsman, D. Hendriks, R. Theunissen, M. Reniers, and R. Schiffelers, 'Computational effort of bdd-based supervisor synthesis of extended finite automata', in 2019 IEEE 15th International Conference on Automation Science and Engineering (CASE), IEEE, 2019, pp. 486–493.

[15] B. Zhu and H. Li, 'Designing finger movement on mobile phone touch screen for rich emotional expression', 2014 Asia-Pac. Signal Inf. Process. Assoc. Annu. Summit Conf. APSIPA 2014, Feb. 2015, doi: 10.1109/APSIPA.2014.7041634.

[16] K. Zheng, X. Zhang, Q. Zheng, W. Xiang, and L. Hanzo, 'Quality-of-experience assessment and its application to video services in LTE networks', IEEE Wirel. Commun., vol. 22, no. 1, pp. 70–78, 2015.