A Device Performance and Data Analytics Concept for Smartphones' IoT Services and Machine-Type Communication in Cellular Networks

Kingsley A. Ogudo [1,*] , Dahj Muwawa Jean Nestor [2], Osamah Ibrahim Khalaf [3] and Hamed Daei Kasmaei [4]

[1] Department of Electrical & Electronics Engineering Faculty of Engineering and the Built Environment, University of Johannesburg, Johannesburg 0524, South Africa
[2] Department of Electrical & Mining Engineering, College of Science, Engineering and Technology (CSET) University of South Africa, Pretoria 0002, South Africa; dahjmuwawa@gmail.com
[3] College of Information Engineering, AI-Nahrain University, Baghdad 64074, Iraq; usama.ibrahem@coie-nahrain.edu.iq or usama81818@yahoo.com
[4] Department of Mathematics and Statistics, Faculty of Science, 3 Central Tehran Branch, Islamic Azad University, Tehran 00011, Iran; hamedelectroj@gmail.com or ham.daeikasmaei@iauctb.ac.ir
[*] Correspondence: kingsleyo@uj.ac.za

check for updates

Abstract: With the advancement of new technologies, the number of connected devices, the amount of data generated, and the need to build an intelligently connected network of things to improve and enrich the human ecosystem open new doors to modifications and adaptations of current cellular network infrastructures. While more focus is given to low power wide area (LPWA) applications and devices, a significant challenge is the definition of Internet of Things (IoT) use cases and the value generation of applications on already existing IoT devices. Smartphones and related devices are currently manufactured with a wide range of smart sensors such as accelerometers, video sensors, compasses, gyros, proximity sensors, fingerprint sensors, temperature sensors, and biometric sensors used for various purposes. Many of these sensors can be automatically expanded to monitor a user's daily activities (e.g., fitness workouts), locations, movements, and real-time body temperatures. Mobile network operators (MNOs) play a substantial role in providing IoT communications platforms, as they manage traffic flow in the network. In this paper, we discuss the global concept of IoT and machine-type communication (MTC), and we conduct device performance analytics based on data traffic collected from a cellular network. The experiment equips service providers with a model and framework to monitor device performance in a network.

Keywords: Internet of Things (IoT); device performance; low power wide area (LPWA); data analytics; smart sensors; predictive analytics; mobile network operators (MNOs)

1. Introduction

The expansion of data analytics and machine-to-machine communication (M2M) device performance is going to provide an inventive podium for mobile operators to innovate in the area of customer and service experience. The amount of connected low power devices is growing at a fast pace, as well as the amount of data generated by those devices which release a new height of services to improve the quality of life and enterprise productivity. Smartphones, smartwatches, and other wireless devices today are manufactured with a wide range of smart sensors to improve user experience in terms of security, health, location, and marketing applications. Sensors such as accelerometers, gyros, and fitness workout sensors are becoming a part of users' lives and are keeping them safe and

healthy. With the current architecture of the IoT system, mobile network operators (MNOs) are playing a significant role in the communication of IoT devices, providing both an adapted and efficient access layer. Overlooked as recent technology, IoT is becoming an essential part of the Internet [1], though it is a concept under continuous development [2]. Likewise, with constant interactions with diverse devices including sensors, actuators, cameras, and mobile smartphones, IoT constitutes a new source of revenue for mobile operators, provisioning and creating new faculties to implement IoT services. One of the fundamental elements of the IoT concept is M2M [3]. Several studies are being conducted, addressing different facets of the M2M on the current cellular infrastructure (2G/3G/LTE/LTE-A). Minor adjustments are expected from the legacy network infrastructures to handle the explosive amount of device data; to this end, an aggregation model of data traffic was proposed to benchmark the quality of services for smart devices, focusing on the basic data key performance indicators (KPIs), including the throughput and average response time of service applications [4]. M2M devices have different QoS requirements, depending on the application; for example, real-time monitoring may require less delay and a high traffic priority compared with standard monitoring [5]; many studies have been directed towards network provisioning [4], the management of a service's quality [6], and the application of predictive methods on network traffic [7]. In our case study, the performance is shifted towards devices. Figure 1 illustrates a simplified traffic model, in which traffic coming from multiple SIM-enabled devices are processed and transformed to provide analytics intelligence. The model provides a large partition between device, network, and service analytics. Mobile operators will play a significant role in connecting devices due to the existing transport infrastructure and the minimal modifications that will be needed to transfer a high load of traffic (with an indifference to building relatively new infrastructures).
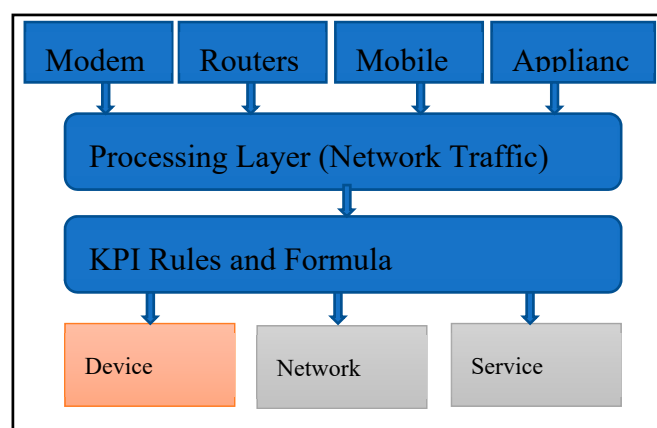

Figure 1. Scheme of a simplified device analytics model.

The introduction of 5G, network function virtualization, and cloud technologies are evident highlights of connected devices' prerequisites on networks [8].

The objective of this paper is to provide a practical model of network device performance, as a baseline for IoT and M2M systems. In the article, we use data analytics methodologies to explore devices' performance based on data application indicators; the experiment applies predictive analytics as well to classify device category, manufacturer, and type based on a defined set of rules.

2. Background

Many researchers have described the concepts of M2M and IoT [3,9]; while all definitions apply, the simplest way to outline IoT is the Internet of objects or the Internet of everything excluding humans. Devices such as smartphones, SIM-enabled sensors, smart meters, and IP cameras are all intelligently connected via the Internet to facilitate communications between systems, people, and devices, with the aim to improve a particular aspect of business or life [10]. Such devices use network infrastructures for

interconnections. Cellular networks with long-term evolution (LTE) and LTE-A (advanced) already support low power device interconnections; having more than 20 million connected devices entails an exponential growth of traffic in the network, which can result in high latency and network congestion.

2.1. Machine-Type Communication (M2M)

Mobile cellular networks have been deployed and optimized in recent years to accommodate human-to-human communication (H2H) traffic, which in most cases is characterized by a high demand on the downlink stream. However, with the application of IoT, H2H is being slowly overshadowed by M2M, which imposes high demands on the uplink traffic stream [11]. In machine-type communication, devices communicate between them or interacting with a remote server, with minimal human interaction. The use of smart electricity meters, intelligent vending machines, and intelligent traffic control devices are examples of machine-type communications. These devices are equipped with SIM cards that facilitate a unique identification in the network, using an IMSI (international mobile subscriber identity) and an IMEI (international mobile equipment identity). The upgrade of the current cellular infrastructure, to accommodate M2M, has been a topic of study for many researchers. Although 5G cellular support is viewed as favorable for IoT and M2M connections [5], research is highlighting the growth and challenges of M2M traffic on existing cellular networks [12].

Gohil et al. [13] studied the parallelism between M2M and H2H traffic by analyzing a week of data collected from a mobile operator in the USA. The study used vital parameters such as data volume on the uplink and the downlink. It was shown that M2M generates heavy uplink traffic. The nature of the M2M data stream was also reviewed in [14], which proved the complexity of M2M traffic patterns. David Aragao et al. [3] introduced a mechanism for congestion control in M2M, improving energy efficiency, evaluating the impact of M2M compared to H2H, and analyzing traffic priorities on different classes of devices. The research on M2M shows the preventive measures that mobile operators need to take to efficiently adjust to high, unbalanced, and possibly unstructured data traffic.

2.2. Mobile Networks Digitization and the 5G

Different network types can be used to connect IoT devices. Wireless networks such as traditional mesh networks, ad-hoc networks, satellite networks, sensors networks, Bluetooth, and so on ensure Internet connectivity. In personal area networks, IEEE 802.15.x provides a large number of application specifications [15]. Another communication standard implemented for the support of wireless communication is the extended Wi-Fi networks, using 900 MHz bands instead of the legacy 2.4 GHz Wi-Fi band and supporting the concept of IoT [16]. However, coverage and scalability are two parameters that limit the adoption of the mentioned technologies in IoT services and machine-type communications. Reliability, data security, and interference pose another threat due to the use of unlicensed frequency bands. Cellular mobile networks have already been supporting connectivity between different applications, going beyond the stream of just mobile phones, tablets, and computers to accommodate other applications such as connected buildings, smart meters, and vending machines. Mobile networks should guarantee security and reliability because they are strongly regulated. Current machine-type and IoT applications use LTE cellular technology because of its high-speed characteristics and reliability. As projected by many studies, the estimated number of connected devices by 2020 could be around 50 billion [17]. Other disciplines outside the telecommunications area are going to be major contributors to network traffic and load; for example, in the domain of healthcare, there is going to be a large number of devices connected for health and medical conditions in the smart city domain. This means the current traffic could grow exponentially; future demand in terms of network connectivity requirements is going to surpass what LTE systems can currently achieve effectively. Hence, a need for the MNO business model is to digitize their networks to support the future trends of IoT and MTC applications. More data capacity, higher throughputs, massive machine-type communication (mMTC), and ultra-low latency applications are the main drivers of the 5G network

system. The enhanced mobile broadband (eMBB), mMTC, and critical machine-type communication (uMTC) are the main group of use cases that will form the building blocks of 5G networks [18].

2.3. Real-Life Applications of IoT and M2M Communication

The connectivity of devices aims to improve human life. Let us look at a few real-life applications of the connected world. In terms of smart parking and intelligent parking reservation [19], the developed algorithm helps users to find free parking slots and estimate the distance to free parking spots. Smart CCTV monitoring systems, equipped with artificial intelligence, upload video footage to remote servers (data usage, throughput, and latency requirements). Smart power meters used daily regulate energy usage in houses; with the prepaid electricity model, the power meters register our daily usage and send the data usage to a remote server for monitoring [20]. The usage of smart heating systems to control and monitor appliance temperatures in homes, saving energy and efficiently managing costs [21], is another important application of IoT technology.

Smart refrigerators, smart lighting systems, health monitoring gadgets, and smart cars are all real-life illustrations of currently popular IoT applications. The devices linked to the mentioned applications upload a massive amount of data to remote central servers through communication networks.

3. Data Analytics and Machine Learning for Device Performance

Data analytics will be pivotal in performance tracking due to the nature of the M2M data stream and the projected increase in the number of connected devices. Various applications require different types of processing. Specific applications such as those in healthcare require a high transmission rate and lower delay compared to other applications—hence the need for flexible data processing and business case adaptation. MNOs will have to switch from network monitoring to service monitoring and device monitoring. In the area of IoT and MTC, the massive amount of data generated by devices would require intelligent analytics to represent the management of traffic [22]. On top of the analytics capability of smart device behavior, machine learning and predictive analytics are also required to exploit the knowledge that is somewhat hidden in the vast amount of data generated by those devices. Combining machine learning and big data presents a new challenge in terms of computational cost. High-performance servers and parallel computing create a need to ensure efficient processing and the retrieval of data. Data analytics is therefore crucial for both MNOs and IoT end service providers. The most commonly known big data challenges should be looked at and mitigated.

4. Problem of the Study

For many years, MNOs have focused continuously on QoS on the network level, so as to guarantee network quality as the main component of a good experience. Nonetheless, in recent years, there has been a paradigm shift to a service-oriented approach and toward customer experience management, where operators have focused on the services provided and on each client's experience. With the evolution of technology, the IoT, and MTC, the quality of device (QoD) will also be a fundamental concept of network monitoring.

Service and device performances as perceived by IoT service providers and MNOs are dissimilar. While the former provides the end-user service (on the application layer), the latter puts in place a transport mechanism for connectivity between end-user devices. Thus, similar to OTT (over the top) services such as WhatsApp, Skype, and Viber, the content of the IoT services can also be encrypted with end-to-end encryption processes for data security. In this case, MNOs do not have access to the content of the IoT packets but provide the communication channels.

By providing the transportation mechanism, MNOs can, through deep packet inspection (DPI), identify services by analyzing the patterns of traffic in the data packets, also known as the traffic signature. Figure 2 illustrates the communication concept, the traffic flow, and the data collection points in the network for this experimental study. The DPI engine allows the system to recognize IoT

and MTC services based on their signature (packets traffic pattern) and not their content, as well as other standard Internet services.
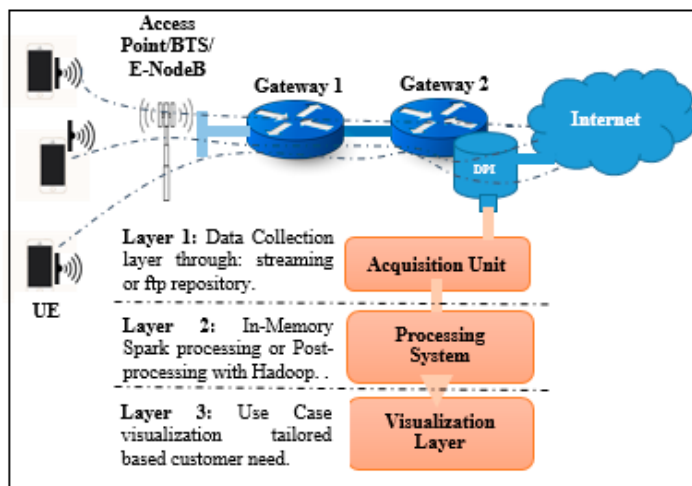

Figure 2. Data processing model for traffic capturing analytics.

The heterogeneity and the volume of M2M device data make IoT service data processing much more complicated—on the one hand, because of the complexity of data streams and, on the other, because of the amount of intelligence required to detect the traffic patterns in the network. Depending on the service application, M2M system processes events in either real time or non-real time [23]—hence the need for big data technologies and cloud computing.

5. Process Methodology

Figure 3 illustrates the methodology used in the study. Three significant steps were taken to complete the experiment. Data preprocessing was done during the experiment setup. The data processing step included data cleaning and understanding the variables and the definition of the use case model. The data output step embedded the designation of key performance indicators and user rules that should be applied to the analytics results. Finally, the predictive application step included short exploratory data analysis, data training, model fitting, and behavior prediction.
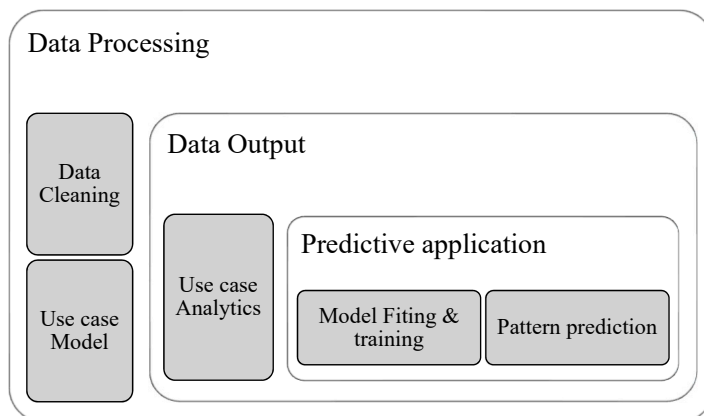

Figure 3. Process methodology.

In the experiment, we categorized devices by manufacturer, model, and type based on their performance, defined through data KPIs, to determine the impact of traffic in the network.

6. Experiment Setup

In this experiment, we collected five days of data from a mobile operator through port mirroring on one Gn interface link (on the GGSN) using Wireshark. The application runs on a Linux Ubuntu laptop with 16GB of RAM, an i7 processing Core, and 1 TB of storage. The sniffed packet was analyzed and filtered by a TAC (Type Allocation Code) device. The TAC device was correlated with the IMEI database to extract the device type, manufacturer, and model. Due to the privacy of the captured data, sensitive information was hashed. The hashed information included the TAC, IMEI, and model name if necessary.

7. Device Data Processing

The data processing was accomplished using SparkSQL and the R framework for machine learning and predictive analytics [24]. All unwanted fields were removed in the data cleaning process.

7.1. Understanding the Generated Data

The study concept used two sets of data. One was the user plane information collected from the operator network and the other was the mobile device IMEI database downloaded from kaggle [25]. Tables 1–3 describes the processed variables that we used in the experiment.

Table 1. Data field descriptions.

Parameter	Description
TAC	Type allocation code, eight digits of the IMEI, identifying the manufacturer and model
SUBS	Number of subscribers using that specific device.
APPRT	Application response time in milliseconds; An indication of how long the application takes to respond to the request
RTT_C	Round trip time to on the uplink, in milliseconds
RTT_S	Round trip time to on the Downlink, in milliseconds
TCP_DLP	TCP downlink packet in KB/sec
TCP_ULP	TCP uplink packet in KB/sec
VOLUMEDL	Downlink data volume in GB
VOLUMEUL	Uplink data volume in GB
RTX_DL	Downlink retransmitted data packet in GB
RTX_UL	Uplink retransmitted data packet in GB
DNSOK	Domain name server queries with successful termination
DNSERR	Domain name server queries with unsuccessful termination
Packet_Loss_DL	Packet loss in % on the downlink
Packet_Loss_UL	Packet loss in % on the uplink

**Table 2.** Calculated data field descriptions.

Calculated Parameters	Description
THT_DL (Throughput)	$8x\frac{VOLUMEDL\ (MBytes)}{ACTTIME\_DL\ (Sec)} = Mbits/s$
THT_UL (Throughput)	$8x\frac{VOLUMEUL\ (MBytes)}{ACTTIME\_UL\ (Sec)} = Mbits/s$
DER (DNS Error Rate)	$8x\frac{VOLUMEUL\ (MBytes)}{ACTTIME\_UL\ (Sec)} = Mbits/s$

Table 3. Enrichment information parameters.

Extra Parameters	Description
TAC	Type allocation code, eight digits of the IMEI, identifying the manufacturer and model
Manufacturer	Manufacturing company
Model	Model of the device as provided by the manufacturer
Class	Device class (smartphone, dongle, route, or any other class available in the file)
OS	Operating system of the device
LTE	Indication of whether the device supported an LTE network.

7.2. Use Case Model

In order to simplify the workflow on the proposed use case model, a hierarchical simulation was performed on the vertical axis, and performance was monitored on the horizontal axis as shown in Figure 4. A three-layer model was proposed on both the horizontal and vertical axes to manage devices based on specific services.
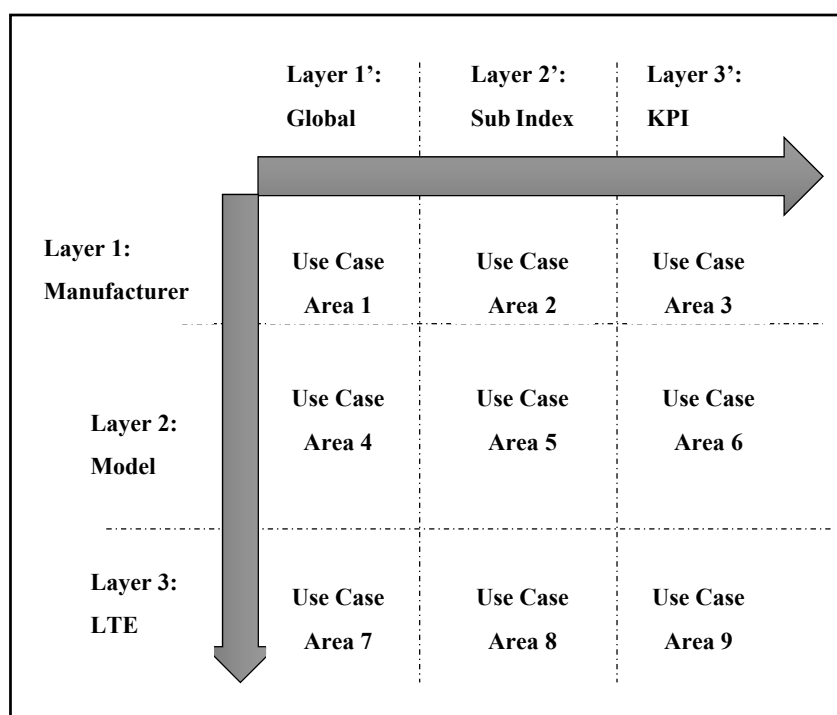
Figure 4. Three-layer horizontal and three-layer vertical analytics model.

Every use case is a result of the analytics done on the datasets taking into consideration the position of the use case compared to the horizontal and vertical axis. The definition of a use case is not standardized by any organization. However, in the proposed model for device analytics, we can see that, for three-layer horizontal and vertical axes, we deduced nine use cases.

Let V_L be the number of vertical layers and H_L the number of horizontal layers. The number of use cases that we can build from the system, U, is provided by a simple equation:

$$U = V_L {\cdot} H_L.$$

- Horizontal Layers: The horizontal layers aim to provide performance indicators. The concept of the global index (GI) is invoked to address different classes of users. Key parameter indicators are aggregated to form a key quality index.

- Vertical Layers: The vertical layers categorize the device and provide a drill-down mechanism from manufacturing to models. In this case, network operators can identify manufacturers or devices with poor performance.

8. Use Case Output

Based on the model proposed during the experiment, the device performance use cases are presented in this section along with the analytics. The experimented results are provided in the form of reports.

8.1. Use Case Layer 1: Performance Index per Manufacturer

Situated in position V_1H_1 of the analytics model, where V_1 is Vertical Axis Element 1, and H_1 is Horizontal Axis Element 1, Use Case 1 provides the key quality index per manufacturer in the network. Figure 5 outputs the results of the analysis. Only one service is considered in this case: video-streaming applications.
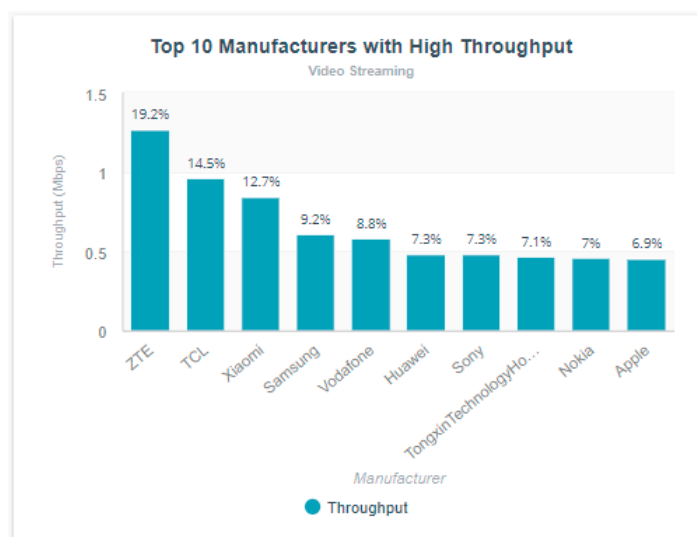

Figure 5. Use Case 1: Top 10 device manufacturers with high video throughput.

The above figure includes all the devices grouped by manufacturer and all the technologies included. The graph is influenced by the technology's coverage in the area where the network data is provided. The analytics shows that ZTE-based devices have the highest throughput when it comes to video streaming, representing 19.2% of all high-throughput devices. Popular manufacturers, including Samsung, Apple, Huawei, and Nokia appear in the report.

In Figure 6, a similar performance is shown per manufacturer, now grouped by technology. In this case, the analyst or network operator can see the main contributor by network technology. LTE is currently the most recent deployed cellular technology in the market. For each manufacturer, the traffic distribution per technology is highlighted. The analysis reveals that most of the ZTE M2M traffic appears on 3G technology only. Samsung and Nokia have a percentage of traffic on 2G technology. Huawei and Apple devices benefit more from 4G/LTE with over 50% of video-related activities performed on 4G.
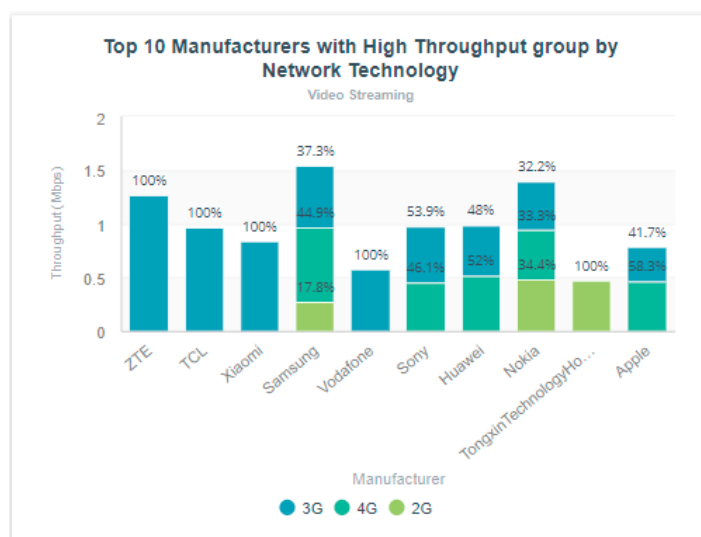
Figure 6. The top 10 device manufacturers with high video throughput, further grouped by technology.

8.2. Use Case Layer 2: Performance Index per Device Model

Layer 2 breaks down the analysis per device model complementing Layer 1 analysis. The analysis allows the user to pinpoint the device model performing well or performing poorly based on the defined KPIs. Table 4 presents the device model ranked by a descendent instantaneous throughput value. Three classes of devices appear in the analytics report: smartphones, dongles, and tablets.

Table 4. Performance index per device model based on throughput: Top 10 devices.

Model	Class	Time Delay (sec)	Packet_Loss (%)	Throughput (Mbps)
O***7	Smartphone	6.39	0.51	3.053
H***X	Smartphone	38.61	0.51	2.897
M***7	Dongle	0.5	0.5	2.248
LT***S	Smartphone	0.25	0.09	1.876
G***i	Smartphone	0.25	0.06	1.846
G***Plus	Smartphone	0.66	0.25	1.689
M***0	Dongle	3.29	1	1.655
G***Lite	Tablet	0.44	0.16	1.654
X***al	Smartphone	29.37	0.04	1.374
G***e	Smartphone	2.46	0.07	1.276

Table 5 presents the device model ranked by descendent response delay. It is not ideal for sensitive systems such as health monitoring or HD-video streaming devices.

Table 5. Worst 10 device models with high application delay.

Model	Class	Time Delay (sec)	Packet_Loss (%)	Throughput (Mbps)
E***1	Smartphone	42.37	0.15	0.446
Mobile***	Dongle	40.27	0.51	0.11
L***30HC	Tablet	23.71	1	0.307
G***3	Smartphone	18.77	0.39	0.44
STL100-1	Smartphone	12.71	0.18	0.404
X***al	Smartphone	10.34	0.22	0.856
X800	Feature Phone	9.78	0.18	0.052
G***e	Smartphone	9.72	0.5	0.653
A***5	Feature Phone	9.21	0.62	0.636
E3***1	Dongle	6.44	0.07	0.181

8.3. Use Case Layer 3: Performance Index per Technology

Different network technologies provide different network experiences. It is, therefore, essential to perform device analytics based on technology and distributed by the used operating systems.

In Figure 7, the analytics result is based on the throughput. We can see that, for 4G technology, Android-based devices present high throughput in video streaming applications, followed by iOS devices. However, any combination of KPIs can be used.
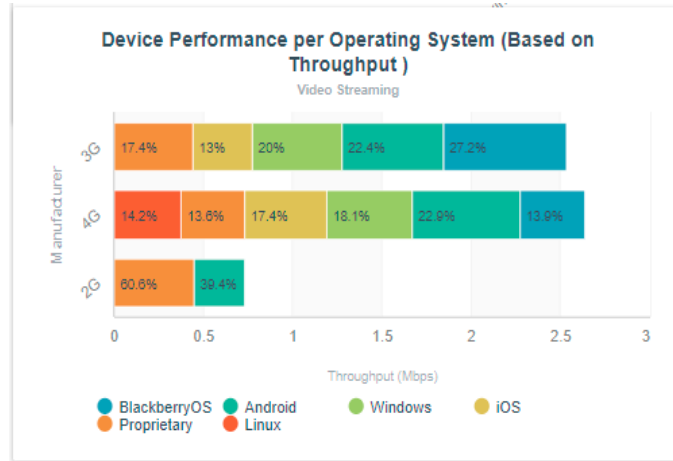

Figure 7. The Layer 3 use case result.

In Figure 8, the analytics is based on packet loss to determine which technology is susceptible to packet loss. High packet loss is registered on the uplink on both technologies. In the graph, we can see that there is a high rate of packet loss on the uplink than on the downlink. M2M applications use more uplink stream than downlink—hence the need to optimize uplink transmission.
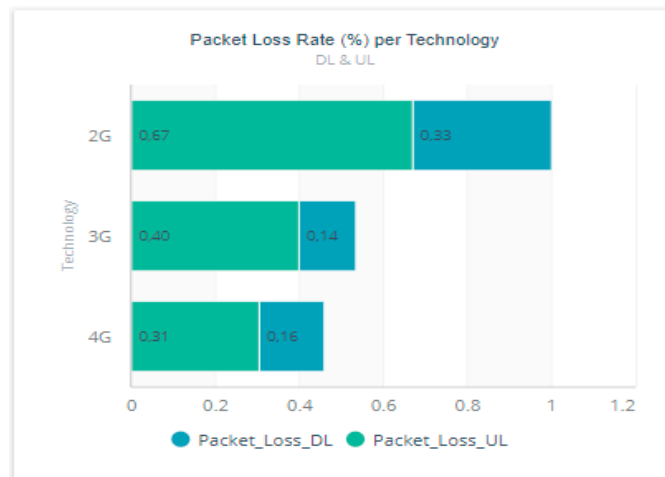

Figure 8. The Layer 3 use case result based on packet loss.

Another aspect of device analytics is to help the operators in expanding the competitive force into the market. The adoption of specific or individual devices is presented in Figure 9, which displays the number of subscribers possessing devices of a certain manufacturer. The top manufacturer retains many more subscribers than any other.

Samsung has more subscribers as shown in Figure 9, followed by Huawei, Nokia, and Apple, respectively. Adoption by device type is illustrated in Figure 10. The most frequently used device type is deduced from the dataset. The most frequently used communicating devices are smartphones, followed by tablets, dongles, and legacy 2G-capable devices.
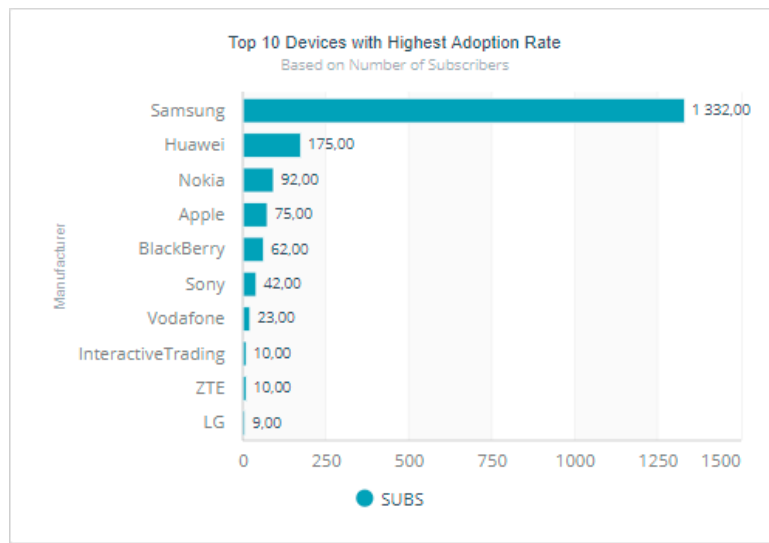
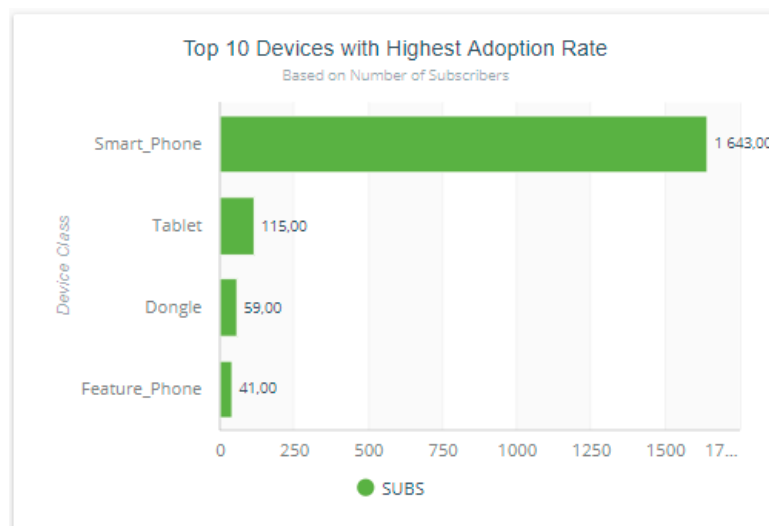**Figure 9.** The Layer 3 use case based on device adoption.

Figure 10. The Layer 3 use case based on device class adoption.

9. Predictive Analytics for Device Behavior

Machine learning and predictive analytics are becoming common practice in most of the analytics businesses. That is to say, the need to see into the future is crucial. We ran two predictive algorithms, a support vector machine (SVM) and C5.0, on the dataset to determine device models likely to experience performance problems. Predictive algorithms, even with low accuracy, can provide value for businesses.

For this study, we consider, for a good video streaming experience, the following: a 1 Mbps minimum download, an acceptable application delay of 0.5 s, a packet round trip time (RTT) of less than 0.5 s, a packet loss of less than 1%, and a domain name server (DNS) resolution failure rate of less than 2%. The predictive rule is as follows:

if (throughput > 1 Mbps & APPRT<0.5 sec & RTT<0.5 sec & packet_loss < 1% & DER<2%)

```
{pred_goodPer = 1; //Set the pred_goodPer to 1 if the condition is met.}

else {pred_goodPer = 0; //Set the pred_goodPer to 0 if the condition is not met}
```

Before applying the rule, we divide the data into training and testing sets. The rule was applied to the training set and evaluated on the testing set. Seventy percent of the data was used for the training set and 30% was used for the testing set. Prediction learning was applied to 3G and 4G technologies, as the data performance of 2G devices is vastly limited by default.

Concise exploratory data analysis, as illustrated in Figure 11, shows the number of well versus poorly performing devices in the dataset. The rule can be modified based on the tester's definition of acceptable video quality specifications.
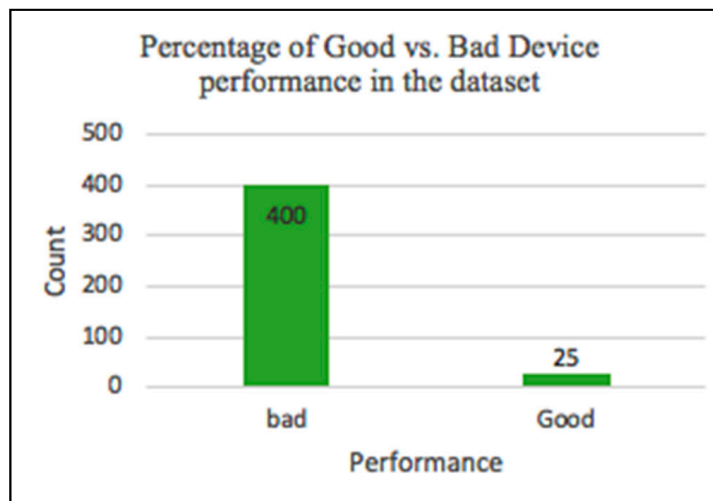

Figure 11. Percentage of well and poorly performing 3G4G devices in the dataset.

Compared to the SVM algorithm that provided an accuracy of 99.7% on the training set as shown in Figure 12, the C50 algorithm provided an accuracy of 94% on the training set as shown in Figure 13. The two algorithms were then run on the testing set to evaluate the models.

```
> print(svmfit)
Support Vector Machines with Radial Basis Function Kernel

298 samples
 20 predictor
  2 classes: 'bad', 'Good'

No pre-processing
Resampling: Cross-Validated (10 fold)
Summary of sample sizes: 268, 268, 268, 268, 268, 268, ...
Resampling results across tuning parameters:

  C     Accuracy   Kappa
  0.25  0.9397701  0
  0.50  0.9397701  0
  1.00  0.9397701  0

Tuning parameter 'sigma' was held constant at a value of 1.635051e-06
Accuracy was used to select the optimal model using the largest value.
The final values used for the model were sigma = 1.635051e-06 and C = 0.25.
```

Figure 12. Training result for the support vector machine (SVM) algorithm.

```
> print(cartfit)
C5.0

298 samples
 20 predictor
  2 classes: 'bad', 'Good'

No pre-processing
Resampling: Cross-Validated (10 fold)
Summary of sample sizes: 268, 268, 268, 268, 268, 269, ...
Resampling results across tuning parameters:

  model  winnow  trials  Accuracy   Kappa
  rules  FALSE    1       0.9966667  0.9782609
  rules  FALSE   10       0.9966667  0.9782609
  rules  FALSE   20       0.9966667  0.9782609
  rules   TRUE    1       0.9966667  0.9782609|
  rules   TRUE   10       0.9966667  0.9782609
  rules   TRUE   20       0.9966667  0.9782609
  tree   FALSE    1       0.9966667  0.9782609
  tree   FALSE   10       0.9966667  0.9782609
  tree   FALSE   20       0.9966667  0.9782609
  tree    TRUE    1       0.9966667  0.9782609
  tree    TRUE   10       0.9966667  0.9782609
  tree    TRUE   20       0.9966667  0.9782609

Accuracy was used to select the optimal model using the largest value.
The final values used for the model were trials = 1, model = rules and winnow = TR
UE.
```

Figure 13. Training Result for the C50 algorithm.

Figures 14–16 respectively show the performance of the machine learning algorithms, C50 algorithm and SVM algorithm on the testing dataset. We can conclude that using the C50 algorithm, because of its excellent performance on the testing set, will allow us to predict poorly performing devices in the future.

```
Models: C50, SVM
Number of resamples: 10

Accuracy
         Min.     1st Qu.   Median    Mean      3rd Qu.   Max.      NA's
C50 0.9666667 1.0000000 1.0000000 0.9966667 1.0000000 1.0000000    0
SVM 0.9333333 0.9333333 0.9333333 0.9397701 0.9333333 0.9655172    0

Kappa
         Min. 1st Qu. Median    Mean 3rd Qu. Max. NA's
C50 0.7826087       1      1 0.9782609       1    1    0
SVM 0.0000000       0      0 0.0000000       0    0    0
```

Figure 14. Machine learning (ML) algorithm performance comparison on the training set.

```
> confusionMatrix(predictionc50, devicetesting$performance)
Confusion Matrix and Statistics

          Reference
Prediction bad Good
      bad  119    0
      Good   1    7

              Accuracy : 0.9921
                95% CI : (0.9569, 0.9998)
   No Information Rate : 0.9449
   P-Value [Acc > NIR] : 0.006276

                 Kappa : 0.9292
 Mcnemar's| Test P-Value : 1.000000

           Sensitivity : 0.9917
           Specificity : 1.0000
        Pos Pred Value : 1.0000
        Neg Pred Value : 0.8750
            Prevalence : 0.9449
        Detection Rate : 0.9370
  Detection Prevalence : 0.9370
     Balanced Accuracy : 0.9958

      'Positive' Class : bad
```

Figure 15. C50 algorithm performance on the device testing dataset.

```
> confusionMatrix(predictionsvm, devicetesting$performance)
Confusion Matrix and Statistics

          Reference
Prediction bad Good
      bad  120    7
      Good   0    0

              Accuracy : 0.9449
                95% CI : (0.8897, 0.9776)
   No Information Rate : 0.9449
   P-Value [Acc > NIR] : 0.59875

                 Kappa : 0
 Mcnemar's Test P-Value : 0.02334

           Sensitivity : 1.0000
           Specificity : 0.0000
        Pos Pred Value : 0.9449
        Neg Pred Value :    NaN
            Prevalence : 0.9449
        Detection Rate : 0.9449
  Detection Prevalence : 1.0000
     Balanced Accuracy : 0.5000

      'Positive' Class : bad
```

Figure 16. SVM algorithm performance on the device testing data set.

10. Conclusions

This paper introduces the concept of IoT and MTC with a focus on devices and associates data analytics with machine learning to showcase the importance of device monitoring. In line with this

idea, a conceptual model of a device monitoring system is presented through data analytics use cases and predictive applications. The C50 model presented a better performance on both the training and the testing datasets. The matricidal system model provides a foundation for use cases that can be used to monitor device behavior. Latency, packet loss, throughput, and DNS performance were considered as performance indicators in the experiment, and represent the foundation for the current KPI monitoring that, among the many other network requirements, needs to be carefully monitored.

11. Future Studies

IoT and MTC are technologies that will overwhelm the market in the future. They are emerging technologies, and many open holes are being identified and researched by scientists. As connected devices increase, more challenges open up in the area of data security. Data protection and security are two topics that are being looked at by many researchers [26]. Many kinds of studies will address the challenges of the current network infrastructure and the migration to 5G, providing smooth communication for end devices so as to support high-definition video streaming at a Gbps level [27].

References

1. Zanella, A.; Bui, N.; Castellani, A.; Vangelista, L.; Zorzi, M. Internet of Things for Smart Cities. *IEEE Internet Things J.* **2014**, *1*, 22–32. [CrossRef]
2. Atzori, L.; Antonio, I.; Giacomo, M. The Internet of things: A survey. *Comput. Netw.* **2010**, *54*, 2787–2805. [CrossRef]
3. Vieira, D.; De Castro, M.F.; Aragao, D. A mechanism to control the congestion in machine-to-machine communication in LTE-A networks. In Proceedings of the 2017 IEEE Symposium on Computers and Communications (ISCC), Heraklion, Greece, 3–6 July 2017; pp. 794–797.
4. Dighriri, M.; Alfoudi, A.S.D.; Lee, G.M.; Baker, T. Data Traffic Model in Machine to Machine Communications over 5G Network Slicing. In Proceedings of the 2016 9th International Conference on Developments in eSystems Engineering (DeSE), Liverpool, UK, 31 August–2 September 2016; pp. 239–244.
5. Lien, S.-Y.; Chen, K.-C. Massive Access Management for QoS Guarantees in 3GPP Machine-to-Machine Communications. *IEEE Commun. Lett.* **2011**, *15*, 311–313. [CrossRef]
6. Ogudo, K.A.; Nestor, D.M.J. Modeling of an Efficient Low Cost, Tree Based Data Service Quality Management for Mobile Operators Using in-Memory Big Data Processing and Business Intelligence use Cases. In Proceedings of the 2018 International Conference on Advances in Big Data, Computing and Data Communication Systems (icABCD), Durban, South Africa, 6–7 August 2018; pp. 1–8.
7. Nestor, D.M.J.; Ogudo, K.A. Practical Implementation of Machine Learning and Predictive Analytics in Cellular Network Transactions in Real Time. In Proceedings of the 2018 International Conference on Advances in Big Data, Computing and Data Communication Systems (icABCD), Durban, South Africa, 6–7 August 2018; pp. 1–10.

8. Tesanovic, M.; Nekovee, M. mmWave-Based Mobile Access for 5G: Key Challenges and Projected Standards and Regulatory Roadmap. In Proceedings of the 2015 IEEE Global Communications Conference, San Diego, CA, USA, 6–10 December 2015; pp. 1–6.

9. Loper, M.L.; Swenson, B. Machine to Machine Trust in Smart Cities. In Proceedings of the 2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS), Atlanta, GA, USA, 5–8 June 2017; pp. 1887–1889.

10. Kortuem, G.; Kawsar, F.; Sundramoorthy, V.; Fitton, D. Smart objects as building blocks for the internet of things. *Int. Comput.* **2011**, *14*, 44–51. [CrossRef]

11. Shariatmadari, H.; Ratasuk, R.; Iraji, S.; Laya, A.; Taleb, T.; Jäntti, R.; Ghosh, A. MACHINE-TYPE COMMUNICATIONS: CURRENT STATUS AND FUTURE PERSPECTIVES TOWARD 5G SYSTEMS. *IEEE Commun. Mag.* **2015**, *53*, 10–17. [CrossRef]

12. Chen, K.-C.; Lien, S.-Y. Machine-to-machine communications: Technologies and challenges. *Ad Hoc Networks* **2014**, *18*, 3–23. [CrossRef]

13. Gohil, A.; Modi, H.; Patel, S.K. 5G technology of mobile communication: A survey. In Proceedings of the 2013 International Conference on Intelligent Systems and Signal Processing (ISSP), Gujarat, India, 1–2 March 2013; pp. 288–292.

14. Casale, G.; Zhang, E.Z.; Smirni, E. Trace data characterization and fitting for Markov modeling. *Perform. Eval.* **2010**, *67*, 61–79. [CrossRef]

15. Zheng, K.; Ou, S.; Alonso-Zárate, J.; Dohler, M.; Liu, F.; Zhu, H. Challenges of massive access in highly dense LTE-advanced networks with machine-to-machine communications. *IEEE Wirel. Commun.* **2014**, *21*, 12–18. [CrossRef]

16. Sun, W.; Lee, O.; Shin, Y.; Kim, S.; Yang, C.; Kim, H.; Choi, S. Wi-Fi could be much more. *IEEE Commun. Mag.* **2014**, *52*, 22–29. [CrossRef]

17. Ericsson, "More than 50 Billion Connected Devices," February 2011. Available online: http://www.ericsson.com/res/docs/whitepapers/wp-50-billions.pdf. (accessed on 1 April 2018).

18. Nikaein, N.; Ksentini, A. Toward Enforcing Network Slicing on RAN: Flexibility and Resources Abstraction. *IEEE Commun. Mag.* **2017**, *55*, 102–108.

19. Aydin, I.; Karakose, M.; Karakose, E. A navigation and reservation based smart parking platform using genetic optimization for smart cities. In Proceedings of the 2017 5th International Istanbul Smart Grid and Cities Congress and Fair (ICSG), Istanbul, Turkey, 19–21 April 2017; pp. 120–124.

20. Metering, A.S.; Visalatchi, S.; Sandeep, K.K. Smart energy metering and power theft control using arduino & GSM. In Proceedings of the 2017 2nd International Conference for Convergence in Technology (I2CT), Mumbai, India, 7–9 April 2017.

21. Jo, H.C.; Kim, S.; Joo, S.K. Smart heating and air conditioning scheduling method incorporating customer convenience for home energy management system. *IEEE Trans. Consum. Electron.* **2013**, *59*, 316–322. [CrossRef]

22. Desarkar, A.; Das, A.; Bhatt, C.; Dey, N.; Ashour, A.S. Big-Data Analytics, Machine Learning Algorithms and Scalable/Parallel/Distributed Algorithms. *Int. Things Big Data Technol. Gener. Healthc.* **2017**, *23*, 159–197.

23. Akbar, A.; Carrez, F.; Moessner, K.; Sancho, J.; Rico, J. Context-aware stream processing for distributed IoT applications. In Proceedings of the 2015 IEEE 2nd World Forum on Internet of Things (WF-IoT), Milan, Italy, 14–16 December 2015; pp. 663–668.

24. Team, R. Core. R: A Language and Environment for Statistical Computing. Available online: http://www.R-project.org. (accessed on 1 April 2018).

25. Richard Nagyfi, Type Allocation Code for Mobile and Wireless Devices. Available online: https://www.kaggle.com/sedthh/typeallocationtable/downloads/typeallocationtable.zip/1 (accessed on 1 April 2018).

26. Weber, R.H. Internet of Things—New security and privacy challenges. *Comput. Law Secur. Rev.* **2010**, *26*, 23–30. [CrossRef]

27. Mu, M.; Broadbent, M.; Farshad, A.; Hart, N.; Hutchison, D.; Ni, Q.; Race, N. A scalable user fairness model for adaptive video streaming over SDN-assisted future networks. *IEEE J. Sel. Areas Commun.* **2016**, *34*, 2168–2184. [CrossRef]