**Big Data Case Study**

**T-MOBILE**

**AIT 622**

**Spring 2019**



**BY**

**REDDIVARI SAI SARAN**

**G01142501**

**Instructor: Sriman Abhishek Yengaldas**

A picture containing clipart

Description generated with high confidence

**WHO**:

**Who they are, what do they do?**

T-Mobile telecommunication company which was founded in December 1999 by Deutsche Telekom, a German telecommunications company. It has headquarters located in Bonn, Germany. Over 44,000 employees are working under the T-Mobile telecommunication company [1]. T-Mobile is also available in different countries like Albania, Austria, Bosnia and Herzegovina, Croatia, Czech Republic, European Union, Greece, Hungary, Macedonia, Montenegro, Netherlands, Poland, Romania, Slovakia, United Kingdom (UK), and United States (USA) [1]. T-Mobile serves several products such as wireless PDAs, Cellular Telephones, Tablets, and PCs. T-Mobile platform can be accessible around the world through the mobile application and the websites.

**What is their role/purpose?**

The purpose of T-Mobile is to increase the customers by providing a good quality network, products, and services to the users which helps them improve their annual profits. According to the T-Mobile survey, almost 5.7 million customers were added in 2017, and the total revenue was 40.6 billion dollars [2]. The concept of providing a good quality of service is achieved by matching the demand, supply, and cost of the network or line. If the user is satisfied with all the requirements provided by the carrier (T-Mobile), there is an increase in the users or customers. By 2024, T-Mobile is planning to launch the 5G network that is six times faster than the 4G network, which provides the user richer and faster experiences such as mobile AR (Augmented Reality), HD videos, VR (Virtual Reality) and many other applications [3].

**NEED:**

**What is the domain (subject area) of the study? Why did they want/need to do a big data project?**

Telecommunication is one of the biggest markets in the world. Telecommunication company can share and maintain a network with the other network capacity companies. Wireless telecommunication provides data, telephone, signal transmission, internet, and other services to the users [4]. The biggest challenges for the telecom companies in big data is to understand the new product, improve the customer experience, reduce the service troll, implement the value-based network, reduce customer churn, forecast capacity more accurately [5]. Most of the telecom companies do not use the big data strategy. T-Mobile is the only company that makes use of the big data strategy where they manage to bring down the churn rate by 50% annually [6].

T-Mobile telecommunication company collects a massive amount of data every day through calls, messages, internet usage and so on. Almost 5.7 million customers use the T-Mobile network to receive the full amount of data; T-mobile decided to combine network data with multiple source system and database [2] [7]. T-Mobile in the USA has integrated big data to combine interaction data and customer transaction to predict customer desertion. T-Mobile has leveraged big data with the transaction data by using the CRM and the billing system. The company has implemented a big data solution with a well-known bank which can improve the support system for the customers even before the problem gets on the phone [7]. Big data is used as the platform to determine and understand the relationship between the customer. A huge amount of data is collected from different sources in different formats like unstructured, structured, and semi-structured type of data which is hard to process, clean and analyze the data.

The Big data solution combines different data sources such as internal and external data which provides the T-Mobile company with a practical approach to understand and analyze the information to identify the customer intent to purchase and customize the best deals and the offers to the customer.

**What questions did they want to answer?**

There are a few basic questions that they wanted to answer

* How to provide the best service, products, and network to the customers by using social media data?
* How to build a strong relationship with the customer?
* How to increase the number of customers yearly?

**Why was this a "big data" problem? (i.e., what are the “Vs.”?)**

The T-Mobile company needs to focus on maintaining the existing customers and gradually increase the interest of the customer to prefer the T-Mobile carrier rather than using the other carrier. The big data problem can be defined by using the 3 V’s.

Volume - Huge amount of data is collected from the different social media data such as calls, messages, internet, transaction details and from other sources. Almost three million records will be generated per hour. Herencia said that “The volume in the telecommunication doubles for every 40 months”. Since there was a huge amount of data generated it is so difficult to process and analyze the output [8].

Velocity - In telecom companies’ velocity is more important than the volume. There are billion or trillion amounts of data that is transmitted to and received from the user. Hence there should be quick access to the database to transmit the data. T-Mobile network performance data varies with the states, but it continually maintains 5.5-40 Mbps download speed, 3.0-20.7 Mbps upload speed, and 30-50 4G LTE Latency [8] [9].

Variety - Data can be received from different sources like photos, text files, messages, customer information, videos, social media updates, transaction details, and location details. Most of the data is in a structured and unstructured format, and this kind of format can be stored and harvested [10].

**CHALLENGES:**

Data can be collected from various data sources like Customer’s ID, services used by the customer, usage plan, demographic information of the customers, rating and feedback from the customers, social media, GPS, spending pattern, MSISDN (Mobile Station International Subscriber Directory Number), call recordings, transaction, information collected from cookies, IP address and Complaints [11] [12]. Since a huge amount of data is generated from various sources, the data needs to be analyzed and processed carefully to get the required output to satisfy or to improve the services of the customer.

Every information is accessed and collected by the T-Mobile telecommunication company except when the network is in roaming with the WIFI service provider, or another carrier they are going to receive the telephone number, call list, location, and some other information about the usage will be accessible to the carrier to enable that services. T-Mobile telecommunication company has the rights to access the information for different purposes such as [12]

* To verify identity and transactions.
* To confirm products and services.
* To provide a bill for product and services.
* To provide customer services and technical services.
* To protect T-Mobile rights, safety, property, and interest from the customers.
* To make a business decision for current and future services and product offerings.

T-Mobile uses a variety of electronic, procedural safeguard, and physical to protect information from unauthorized access. The data collected from the various sources should be secured because the data may include the customer’s password, address, email, photos, SSN (Social Security Number), and birth dates [12]. There was an incident in January 2006 the manager of internal network has accessed the customer's information which is evidence that he attempted to sell the data to others for identity theft [12]. So, the data need to be protected and encrypted to avoid hacking. T-Mobile receives a portion of 1.3 million warrantless law enforcement requests to the subscriber's information which may include phone numbers, location, and text messages which also needs to be secured. There are some federal laws that T-Mobile require to protect CPNI (Customer Proprietary Network Information) from unauthorized users.

T-Mobile telecommunication company collects a huge amount of data from various sources; the data may include the duplication of the information which may include the change in number, address, and phone numbers. Due to the duplication of the information, it is difficult to analyze the data and give feedback to the customer [12]. De-Duplication is the process which helps to detect the duplication information and helps to analyze the data to get the desired output. The collected data from a various source need to follow a proper format if the data does not follow the proper format then it will be more difficult to separate and analyze the data [9].

**Organization and Technical Challenges:**

Organization Challenges:

* Managing Cost: Every organization needs to manage costs and have the ability to change its business according to the market. By using reboot torestore software, T-Mobile provides a lower cost associated with the IT support and bring down the revenue loss associated with the systems [13].
* Infrastructure on data demands: The major problem faced by the organization is leveraging the right infrastructure for the best functionality. Customers use different devices for messaging, calls and internet access [13]. So the carrier should address the growing data needs. The organization should analyze the existing infrastructure and make changes wherever they want to expand the carrier capabilities and rise the data usage demand.
* Customer Complaints: The main deal for the organization is to satisfy the customer and provide solutions for their problem. To ensure the best service, the organization should have the communication technology, useful software tools, and computer system. The organization can also have the reboot solution to clear the malicious software and settings of the device used by the customer [13].

The organization's primary challenge is to stand out among its competitors in the market by offering better services, customer support and should build a brand for innovation. The organizations are challenged to provide better-enhanced performance, content and increased uptime. One of the technical challenges faced by the T-mobile is the network since they have the best frequency range of 1900MHz GSM PCS and AWS UMTS digital cellular network using 600MHZ, 700MHz, 850MHz, 1700MHz, and 2100MHz covering around the USA but still, there was an issue with the network and connection problem [12]. Around 48% of complaints are received regarding the network problem. T-Mobile company are working on the network issue and trying to provide a better network than the other carriers.

**STAKEHOLDERS:**

T-Mobile has recognized that it should adopt a stakeholder which orientation to not only to meet the needs of the various stakeholders. It also gains a significant advantage in the market and even to its competitors. It emphasizes the quality services and stakeholders has earned ethical conduct. In 2012 T-mobile was nominated for the world’s most ethical company for four consecutive years and also won 2011 Top 100 Military-Friendly Employers Award [14]. The people or the organization who invested in the T-Mobile stock are more interested in conducting the study and also the people who are interested in buying the shares from T-Mobile. If the services, support, and technology are improved, then more customers are going to join the T-Mobile then supply and demand will be more from the T-Mobile then the stock rate increases and more people will like to buy the shares.

**REQUIREMENTS:**

**Hardware Required:**

* HP Itanium rx8640
* 144 GB of RAM
* Operating System: HP-UX v.11.31
* 24 Itanium 2 9100 processors running at 1.6 GHz

**Software Required:**

* SAS v.9.2
* SAS CLA v.2.2
* Tableau
* SQL Injection
* Oracle WebLogic Server

Every day a huge amount of data is collected from the various data source. To store that volume of data, T-Mobile telecommunication company required a vast physical infrastructure. So, T-Mobile has adopted IBM corporation Netezza data warehouse where it can store 15 billion rows of data per day on 30 racks of real gear. Even T-Mobile has used Hadoop 2.0 which is an open-source software to store the data and to analyze [15].

**RESULTS:**

To completely assess the data, T-Mobile combined huge subscriber and network data among multiple databases and source systems. They used several tools and technologies to store, analyze and visualize data. Also, T-Mobile made use of different data zones to achieve their business objectives. Customer data zone had a complete view of each customer that helps in attacking customer dissatisfaction. Products and services zone drive innovation based on when and how different types of services and products were used. Customer experience zone tracks information about the channels that interact with customers and use this data to optimize the service levels. Business operations zone defines the best areas for optimal performance based on all the billing, accounting, finance and risk management information. These physical zones place physical data storages in a virtualized environment, which help T-Mobile identify the complex systems, differences in data definitions, duplicate or incompatible data and prevents it [6].

T-Mobile makes use of tribal customer model to tackle the churn rate. The tribal customer model is based on the intensity of influence of a person or a group on others through the social network. Based on the level of influence, Customer Lifetime Value is calculated, and most valuable customers are determined. The churn expectancy of a customer is also based on billing analysis, drop call analysis and sentiment analysis. These analyses are combined into an integrated single-view for customer care and retail store associates to provide customer-specific offers [2]. This approach caused a drop in monthly leaving customers from 100,000 customers leaving in the first quarter of 2011 to 50,000 customers in the second quarter of 2011. Since then, T-Mobile focusses on retaining its loyal, high lifetime customer value subscribers as well as upgrading its customers to high-quality products, leading to high customer satisfaction and additional revenue.

This project helped T-Mobile understand the importance of analyzing data that they collect, identify useful patterns from that and modify their plans. Because of this project, T-Mobile gained insights into the factors affecting the churn rate of the customers and have provided offers to customers based on their profiles [22].

**CRITIQUE:**

T-Mobile can improve its customer service in numerous ways. Firstly, it needs to pay attention to the number of complaints filed by customers periodically and analyze the pattern in the complaints based on the type of complaint, sometimes a complaint is filed by the same customer and in total, geographic location of the customer filing the complaint. An effective solution should be provided based on these analyses. Then, the company should try to build trust factor and confidence in the customers by improving the quality of their service. Also, T-Mobile bought a low-band spectrum of 600MHz for covering long distances, penetrating walls, etc., in more than 900 cities and towns in 32 states of the USA. However, only a few Samsung, LG, OnePlus and Motorola phones can use this new spectrum. Also, T-Mobile takes a few years to deploy the spectrum across the US. To, improve the service and thereby it's turned over, T-Mobile must take necessary actions to deploy the spectrum at the earliest [22].

**Explanation/Defined Terms:**

1. MSISDN- MSISDN is defined as Mobile Station International Subscriber Directory Number which is a unique number to identify the subscription of the customer in a GSM or UMTS mobile network [16].
2. CPNI- CPNI is defined as Customer Proprietary Network Information who collects the data from the telecommunication companies about the customer’s telephone calls [17].
3. Hadoop 2.0- Apache Hadoop is an open-source software which is used to solve the problem, store data, process big data by using MapReduce program model and computation. Hadoop 2.0 is the second iteration of the Hadoop framework [18].
4. UMTS- UMTS is defined as Universal Mobile Telecommunication System which is the third-generation cellular system for the network-based GMS standard [19].
5. SAS- SAS is a software which is used for predictive analytics, business analytics, multivariate analysis, data management, and advanced analytics [20].
6. SQL Injection- SQL Injection is a code injection technique which might destroy the database and it is the most common web hacking technique [21].

# References

|  |  |
| --- | --- |
| [1] | wikipedia, "T-Mobile," [Online]. Available: https://en.wikipedia.org/wiki/T-Mobile. [Accessed 12 March 2019]. |
| [2] | T-Mobile, "2017 Annual Report Changing Wireless for good," T-Mobile, [Online]. Available: https://explore.t-mobile.com/2017-annual-report. [Accessed 12 March 2019]. |
| [3] | T-Mobile, "5G for everyone, everywhere.," T-Mobile, [Online]. Available: https://www.t-mobile.com/5g. [Accessed 12 March 2019]. |
| [4] | CollegeGrad, "Telecommunications Industry," CollegeGrad, [Online]. Available: https://collegegrad.com/industries/telecommunications. [Accessed 12 March 2019]. |
| [5] | Datameer, Top Six Telecommunications Big Data Use Cases, New York: Datameer , 2016. |
| [6] | DATAFLOQ, "T-Mobile USA Cuts Downs Churn Rate By 50% With Big Data," DATAFLOQ, [Online]. Available: https://datafloq.com/read/t-mobile-usa-cuts-downs-churn-rate-with-big-data/512. [Accessed 12 March 2019]. |
| [7] | P. Consulting, "Case Studies," Propel Consulting, [Online]. Available: http://www.propelconsulting.us/Case%20Studies.html. [Accessed 12 March 2019]. |
| [8] | BBVA, "The five V’s of big data," BBVA, [Online]. Available: https://www.bbva.com/en/five-vs-big-data/. [Accessed 12 March 2019]. |
| [9] | T-Mobile, "T-Mobile Network Performance Data," T-Mobile, [Online]. Available: https://www.t-mobile.com/coverage/network-performance-data. [Accessed 12 March 2019]. |
| [10] | X. R. Maintenance, "The V's of Big Data: Velocity, Volume, Value, Variety, and Veracity," XSi Rethink Maintenance, [Online]. Available: https://www.xsnet.com/blog/bid/205405/the-v-s-of-big-data-velocity-volume-value-variety-and-veracity. [Accessed 12 March 2019]. |
| [11] | Exastax, "How Big Data and Analytics Can Drive Profitability for Telecom Operators," Medium, [Online]. Available: https://medium.com/@exastax/how-big-data-and-analytics-can-drive-profitability-for-telecom-operators-b47827bfad7c. |
| [12] | T-Mobile, "T-Mobile Privacy Statement," T-Mobile , [Online]. Available: https://www.t-mobile.com/responsibility/privacy/privacy-policy. [Accessed 27 March 2019]. |
| [13] | Faronics, "IT CHALLENGES FACED BY TELECOMMUNICATIONS PROVIDERS," Faronics, [Online]. Available: https://www.faronics.com/news/blog/it-challenges-faced-by-telecommunications-providers. [Accessed 27 March 2019]. |
| [14] | D. F. E. Initiative, "T-Mobile’s Strong Ethical Culture," [Online]. Available: https://danielsethics.mgt.unm.edu/pdf/t-mobile.pdf. [Accessed 27 March 2019]. |
| [15] | LightReading, "Big Data Saves T-Mobile Big Bucks," LightReading, [Online]. Available: https://www.lightreading.com/analytics/big-data/big-data-saves-t-mobile-big-bucks/d/d-id/709611. [Accessed 27 March 2019]. |
| [16] | Webopodia, "MSISDN," [Online]. Available: https://www.webopedia.com/TERM/M/MSISDN.html. |
| [17] | TechTarget, "Customer proprietary network information," [Online]. Available: https://searchnetworking.techtarget.com/definition/CPNI. |
| [18] | TechTarget, "Hadoop 2," [Online]. Available: https://searchdatamanagement.techtarget.com/definition/Hadoop-2. |
| [19] | TechTarget, "UMTS," [Online]. Available: https://searchmobilecomputing.techtarget.com/definition/UMTS. |
| [20] | CBR, "SAS," [Online]. Available: https://www.cbronline.com/what-is/what-is-sas-4935497/. |
| [21] | W3School, "SQL Injection," [Online]. Available: https://www.w3schools.com/sql/sql\_injection.asp. |
| [22] | J. BRODKIN, "New T-Mobile upgrade may boost your coverage—if you have the right phone," ars TECHNICA, [Online]. Available: https://arstechnica.com/information-technology/2018/06/new-t-mobile-upgrade-may-boost-your-coverage-if-you-have-the-right-phone/. [Accessed 3 April 2019]. |