Deloitte.



Opportunities in Telecom Sector: Arising from Big Data

November 2015



Contents

Foreword	2
Message from Aegis Business School	3
1. Importance of Big Data in Telecom Industry	4
Benefits of Big Data for Telecom Value Chain	6
Challenges Faced by Telecom Operators	7
2. Key Technologies	9
Architecture	10
Key Technologies in Big Data	10
3. Scope of Big Data Implementation	15
Call Drop Analysis	17
Network Analytics	18
Churn Prediction	19
Customer Segmentation	21
Predictive Campaign	22
Location-Based Services	22
4. Use Cases from Telecom Industry	23
Customer Churn Predictive Model — Global Telecom Product and Service Provider	24
Marketing Mix Model — Large Telecommunication Company in the US	24
Big Data Implementation — Large Telecom Company in the US	25
Big Data Implementation — Large Online Retailer in Korea	25
About Deloitte Analytics	26
About Aegis School of Business, Data Science and Telecommunication	27
References	28

Foreword

Every day about 2.5 quintillion bytes of data is created.¹ This clearly indicates that we are in the era of Big Data. Big data has become a ubiquitous part of telecom industry because of the huge amount of data being generated every minute though connected world. The upgraded networks and the proliferation of smart devices has enabled the telecom operators to have access to a wealth of information about their customers' behavior, preferences, movement, etc. Not only human-to-human communication but also human-to-machine and machine-to-machine (M2M) communication generate huge amount of data which could be helpful for all industries including telecom. M2M communication is expected to overtake human generated data in the near future. According to Industry analysts, projections are of 32 billion devices generating 44 trillion GB of data by 2020.²

The telecom and technology companies have been using legacy analytics for years, however, the full potential is yet to be realized by leveraging on the huge amount of data that is generated every day from social networks, search engines, government portals, online businesses and other applications through real-time predictive analytics. The advanced analytics solutions can provide insights which can help in creating new business models and launch innovative products and services. By leveraging its own data and combining the data of different sources, telecom operators can gain deeper understanding of their customer interaction, product performance, and churn and thus can improve upon the customer experience and value addition. Telecom companies can use these insights to help other industries such as agriculture, healthcare, education to name a few.

With the growth of technology, big data will become crucial to understand customer, business and the industry itself. Hence, data analytics is becoming an integral part of every business.

In a study conducted by "Telecoms Intelligence", 47% of the operators had big data investments in place and 19% are expect to implement big data strategy at some point in 2015, with an additional 16% looking to implement big data in 2016 or beyond. Hyper-competition, slow revenue growth, and increasing network costs will compel telecom companies to deploy advanced analytics solutions. Increasing network costs will also encourage operators to adopt analytics solutions, which will allow them to improve data traffic management and optimize the utilization and performance of networks.

Although data analytics companies are experiencing 40% annual growth, there is a huge shortage of talent.³ Big data analytics need professional data scientists who can understand the technology of data analytics as well as the business objective of a telecom operator.

As telecom operators face high network and spectrum costs and intense competition, advanced analytics will help improving their profitability and gain a competitive advantage by enhancing customer experience and optimizing network usage. Articulating a clear value proposition in use of analytics will help justify and attract investments.



Hemant Joshi

Message from Aegis Business School

When the whole country is suffering with an epidemic of call drop, Govt. is clueless about the actual call drop rate and the exact reasons for the same. Current methods of call drop analysis and network analysis is not transparent and representative of actual status. We even don't know what the actual Current Call Drop Rate across the country or in different geography is. Prime Minister Shri Narendra Modi on 25th August this year has directed officials to initiate urgent steps to resolve the problem of call drops. And in the wake of public outcry and Prime Minister making comment TRAI has recommended to penalize operators for the same.

Absence of proper evidence based diagnostics is leading chaos. Various stakeholders TRAI, DOT, Operators and associations are not able to have meaningful dialogue and plan a concrete road-map to stop this menace.

While this debate is on Retail customer and business, continue to suffer due to call drop and poor quality of the network.

With big data technologies telcos can find out pattern and reasons of call drop in real time and send customers' apology text message and also refund money for dropped calls resulting in improved customer satisfaction and brand value.

We need to build a big data platform on which all the operators need to upload the sample data of CDR, network and billing data after masking subscribers details and this platform will give provide Dropped-call rate (DCR); patterns of Call Drop; reasons of call drop and pattern of billing of subscribers and Call Drop etc. This will bring some transparency in term of network performance and call drop pattern. Once TRAI has this pattern various stakeholders can discuss why, where and how the call drop is happening and take some proactive actions rather than just thinking of punitive actions against operators. At Aegis Big Data Product Factory we are developing a Call Drop Analysis platform to analyze call drop rate, patterns and reasons behind it.

Crystal ball of big data can help Telco's see a clear picture in real time as well as act on it too. Hope this white paper helps in addressing a few issues.

The white paper on Big Data in Telco is an attempt to put various use cases to develop action plan and formulate guidelines for managing ongoing challenges in telecommunication.



Bhupesh Daheria FOunder & CEO

Aegis School of Business, Data Science and Telecommunication



Indian telecom industry has crossed two decades post privatization of this sector. Since then a lot of innovation, consolidation, and maturation have happened in the industry, and today we have 12 major mobile telecom operators operating in the country. Presently, the total revenue of telecom operators is about ₹ 1.8 trillion with a burden of ₹ 2.5 trillion debt with a dwindling voice and SMS revenue. This is happening due to severe tariff competition in case of voice and declining SMS revenue due to the advent of new instant messaging applications, etc. Operators are facing disruptive technologies, rapidly changing business rules, intensified regulatory environment leading to eroding service margins.

In this scenario, the only stabilizing factor for telecom operators is revenue generated from data provisioning and driving value from this data. Telecom operators can use advanced analytics on customer and network data to generate a real-time view of customer preferences and network efficiency. This could empower them to make near real-time and fact-based decisions and hence enable a forward looking, focused, decisive, and action-oriented culture in the company.

For customers, this brings faster results, predictive power, and new depth to analytics in the following aspects:



Data-enhanced customer experience

Acquiring a deeper understanding of users and improving their experience at every touchpoint through high performance services, fast feedback and customized offerings

Data-driven efficiency

Taking advantage of actionable information available within the organization, combined with insights from the market, in order to work intelligently and reduce costs



Data driven growth

Establishing innovative offerings and generating new revenue streams sparked by big data

Big data promises to promote growth and increase efficiency and profitability across the entire telecom value chain.

Benefits of Big Data for Telecom Value Chain

Big data analytics brings in considerable value to decision making and provide more accurate and actionable insights which eventually help to build competitive advantages and a more efficient cost structure. In comparison to traditional data warehousing technologies, big data offers following advantages/ opportunities for telecom operators:

- 1. Prepare your networks for future demands: Big data helps businesses take advantage of the available information within their networks in order to make them robust, optimized, and scalable. It can help optimize routing and quality of service by analyzing network traffic in real time. Reviewing a network from a smartphone perspective helps reveal areas that require improvement. A revision of 3G-capable smartphone users running their devices on 2G could identify ways of making significant improvements based on intelligent analysis. Examination of user behavior can also play an important role in understanding how to better deliver media content, and thus directly impacting customer experience.
- 2. Understand customer experience: Big data, with its capabilities, makes it easier to understand your customers in detail right from network data and social media information, which further helps establish customer-centric KPIs which enables to understand user experience.

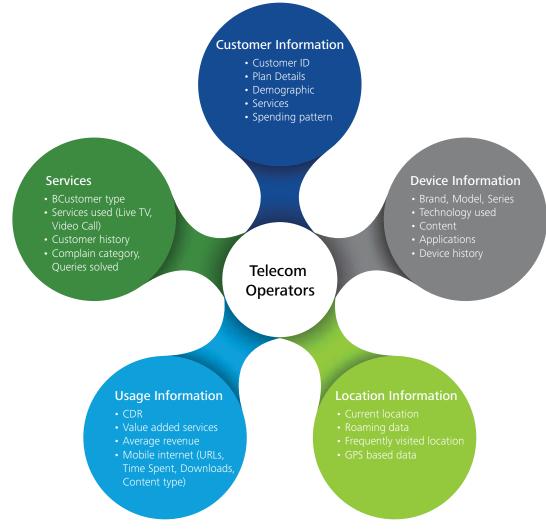
Insights can be deployed in customer call centers to better answer and solve concerns for the customer on the phone and allow them to flexibly and profitably modify subscriber calling plans immediately. They can be used to tailor marketing campaigns to individual customers using location-based and social networking technologies. They can also be used in Network Operations Centers (NOCs) to discover and handle any issues for bigger groups of users. Where relevant, information can be collected about the users' experience in real time. For example, call center staff could see if the customer on the phone has experienced problems at a particular location or while using a particular service. The result is an improved customer service, higher levels of customer satisfaction, and a decreased churn rate. Big data can also help analyze call data records in real time to identify fraudulent behavior immediately.

Entire telecom value chain can be benefited by leveraging the Big Data solutions and below listed are the proposed areas:



Challenges Faced by Telecom Operators

Big data has the potential to place communications services providers (CSPs) in a prime position to win the battle for customers and create new revenue streams. It provides them with a wealth of information about their customers' behaviors, preferences, and movements. Yet, many CSPs struggle to fully derive the greatest value from big data.



At a macro level, the big data analytics throws many challenges for CSPs because of its variety, velocity, and complexity.

- Variety: The social media networks, connected devices, government portals, call data records, billing information, etc., produce huge amount of data as shown in figure above. Most of the data coming from different sources is unstructured. The telecom players need to enrich their Call Data Records (CDR) with other information like location-based service, financial information, etc., to standardize the data for business intelligence platforms before the analysis can be done on it.
- Velocity: Every minute, Indians spend ₹ 1.85 million to shop online. Almost 100 hours of video is shared on YouTube every minute.⁴ The average time spent by a social media user in India is 2.5 hours a day.⁵ All this points to the fact that the data generation speed is tremendously high and to gain value from this data, it needs to be

processed in proper timeframe. This volume of data requires new real-time operational capabilities for various functions and that in turn demands increased data storage for compliance and potential future uses as well as new tools for mediating, managing, and archiving data within available time frames.

• **Complexity**: The user generated data is mostly unstructured and complex because of the lack of standard format to store data. The legacy network and storage devices do not have any specific format to store data which can be relevant for advanced analytics. The data varies with demographics, geography, life style, etc. Analytics may provide unwanted results if the data is not filtered properly.

At a micro level, telecom companies faces other challenges while adopting big data for advanced analytics.

01 0	 Determine a strategy how to leverage on the benefits of big data Define criteria for evaluating return on investments Determine business drivers and if big data can play a role in better insight
02 >	 Identifying and acquiring the skill sets required to understand and leverage big data Organizing business analysts team with strong knowledge of company ecosystem
03 •	 Flexibility of infrastructure to interact with various data formats and large volumes Cost and effort associated with scalability
04 >	 Identifying the best software and hardware solutions and determining the best overall infrastructure solution Transitioning from legacy systems to newer technology
05 つ	 Compromise of quality due to volume and variety of data Cost of maintaining all data quality dimensions: completeness, validity, integrity, consistency, timeliness, accuracy
06 O	 Re-evaluation of internal and external data policies and regulatory environment Privacy issues related to direct and indirect use of big data sources

Big data analytics need professional data scientists who can understand the technology of data analytics and marry it with the business objectives of telecom operator. While there is a shortage of skilled analytics professionals, small operators find it unaffordable to hire data scientists. According to analysts, analytics-based companies in India are expected to face a shortage of 200,000 data scientists.⁶

The infrastructure for data analytics needs high computational capabilities and storage space. It also needs flexibility to analyze different formats of data. Telecom operators tend to ignore these requirements as it is not a part of their core business and also to avoid more capital expenditures.

Transitioning the historical data from legacy system to new system is a challenge. Data quality is also a roadblock as different equipment provides data in different formats. Data could be inaccurate, and maintaining appropriate quality of data is a mammoth task for every company.

The governance and privacy issues are other major challenges as customers prefer not to share their personal data. Government policies and regulations restrict the operators for independent use of the data.

Key Technologies

Data architectures allow telecom companies to store new types of data, retain that data longer, and join diverse datasets together to derive new insights. The following reference architecture diagram represents an amalgam of those approaches that we see across our telecom clients.

Big data is where the data volume, acquisition velocity, or data representation limits the ability to perform effective analysis using traditional relational approaches or requires the use of significant horizontal scaling for efficient processing.

Data may be structured, unstructured or semi-structured, and is frequently and extensively inter-changed. Structured data only accounts for 20% of the big data stored in databases while the rest of 80% is unstructured data.⁷ Data from the internet — including data created by users, data exchanged in social networks, and data from physical sensing devices and the internet of things — is dynamic and unstructured.

Architecture

Data processing in typical big data processing system includes collection and preprocessing, storage, analysis, mining, and value application.

Data-source layer consists of data from enterprises, industry, the internet, and the internet of things. In the datacollection layer, the collected data is pre-processed. This pre-processing includes data cleanup and heterogeneous data processing.

In storage layer, structured, unstructured, and semi-structured data is stored and managed. In the data-processing layer, data is analyzed and mined so that users can analyze services, such as common telecommunications and internet services on the platform.

Key Technologies in Big Data

Big data technologies describe a new generation of technologies and architectures, designed to economically extract value from very large volume of a wide variety of data by enabling high-velocity capture, discovery, and/or analysis.

Many technologies in big data ecosystem have an open source origin. The popularity and viability of these open source tools have driven vendors to launch their versions of tools.⁸ For example, Hadoop framework in conjunction with additional software components such as R and range of NoSQL - Not only Structured Query Language tools such as Cassandra and Apache Hbase is the core of big data.

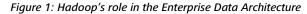
"Telecom providers have a treasure trove of captive data - customer data, CDR, call center interactions and tower logs. However, they have been unable to fully exploit the latent value. Disruptive technologies like - IoT, Big Data and Data Science - offer unprecedented opportunity for data-driven business optimization in key dimensions of cost efficiency, customer intimacy and revenue maximization." Prasad Shyam, VP and Global Head - BI & Analytics, IGATE

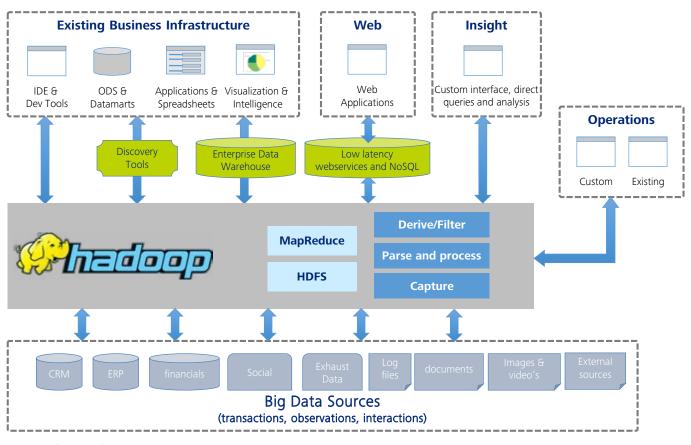
Hadoop

It is an open-source software framework for storing data and running applications on clusters of commodity hardware. It provides both distributed storage and computational capabilities. It is a distributed masterslave architecture that consists of the Hadoop Distributed File System (HDFS) for storage and MapReduce for computational capabilities.

Hadoop consists of following main components:

- MapReduce a distributed processing framework
- Hadoop Distributed File System (HDFS) distributed file system known
- Hadoop common Libraries and utilities used by other Hadoop modules
- Yarn Resource management framework for scheduling and handling resource requests from distributed applications





Source: Deloitte Analysis



Call Detail Records – CDRs

Apache Flume can ingest millions of CDRs per second into Hadoop, while Apache Storm processes those in real-time and identifies any troubling patterns.

- Apache Flume is a distributed, reliable, and available service for efficiently collecting, aggregating, and moving large amounts of streaming data into the Hadoop Distributed File System (HDFS)
- Apache Storm Real time data processing capabilities to enterprise Hadoop
- Hadoop facilitates long-term data retention for a root-cause analysis. This CDR analysis could be used to continuously improvise on call quality and to help in better marketing campaigns.

Fraud detection

The MapReduce Distribution for Hadoop can help build models that can flag anomalous phone calls that might indicate theft or hacking, both in business-to-business and business-to-consumer environments.

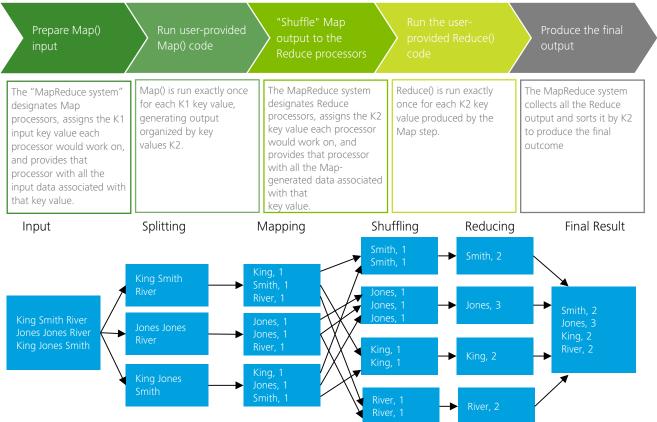
Recommendations - For new launches/ relevant communications

Hadoop Data Lake reduces sales friction and creates NPTB – Next Product To Buy competitive advantage similar to Amazon's advantages in e-commerce.

MapReduce

MapReduce is a framework that allows massive scalability across hundreds or thousands of servers in a Hadoop cluster. It allows developers to write programs to process high volumes of unstructured data in parallel across a distributed cluster of processors or stand-alone computers.

Figure 2: MapReduce - parallel and distributed processing



Source: Deloitte

For Operators - Churn prediction

MapReduce can predict churners and non-churners after validation processes.⁹ It helps analyze all the datasets and is useful in data mining to identify unseen patterns which can be further used to analyze the customer behavior.

HIVE

Hive provides an SQL dialect called Hive Query Language (abbreviated HiveQL or HQL) for querying data stored in a Hadoop cluster. It is useful to move an existing data infrastructure to Hadoop when that infrastructure is based on traditional relational databases and the Structured Query Language (SQL).

Most suited for data warehouse applications, where relatively static data is analyzed, fast response times are not required, and when the data is not changing rapidly. Hive provides a familiar programming model for people who know SQL. It eliminates lots of boilerplate and sometimes tricky coding that would be required in MapReduce programming.¹⁰

NoSQL Databases

NoSQL database environment is a non-relational and largely distributed database system that enables rapid, ad-hoc organization and analysis of extremely highvolume, disparate data types.

NoSQL encompasses a wide variety of database technologies that are an "internet" solution for handling the rise in the volume of data stored, the frequency in which this data is accessed and performance needs.

For Operators - Storage of huge data sets and personalization

NoSQL is considered as primary data source to take product feeds from customers and maintain master catalog of all customer's products, names, categories, and brands. Further to store all the user generated content received from customer's website/various data centres. These datasets enable client to personalize customer's products that they might like, rather than just the single product they were on.

NoSQL systems have the feature to scale horizontally with very less constraints, so expanding the applications will be much easier as the developers need not worry about the storing Petabytes of data as well as latency issues that exists in RDBMS. It can usually scale across different physical servers easily without needing to know which server the data you are looking for is on.

Fraud detection

While having massive inflows of information, clients track reputations for users, compute spam scores for a certain piece of content, and all of this leads to a very high write loads. NoSQL databases will help clients to be resilient and perform replication assignment across different servers.

Internet of Things/Sensor Data

GPS/Sensor tracking data consists incremental data loads which needs faster processing and involves extensive data aggregations and mining in real time processing. NoSQL Databases help to have minimal impact on database or system architecture.

Content and Metadata Store – Applicable to Any Mobile App

For example, if you're building a content catalog, you may store tens of millions of different objects – unstructured content and metadata that may be hierarchical, sparse, free-form text, or varying length. Your data model needs to allow you to quickly add new attributes, without dealing with the complexity and time it takes to change a schema, to ensure faster and more efficient development cycles.

Hortonworks

Hortonworks Data Platform (HDP) is a massively scalable and 100% open source platform for storing, processing and analyzing large volumes of data. It is a key player in pushing Hadoop to an "enterprise-ready" state, which, in turn, will further boost adoption by organizations that are currently skeptical of the IT costs of running Hadoop clusters.

- Hortonworks Data Platform provides an open, stable, and highly extensible platform that makes it easier to integrate Apache Hadoop with existing data architectures and maximize the value of the data
- Next-generation MapReduce Architecture (also known as YARN) adds advancements in scalability, performance and high availability, decouples MapReduce from the resource management architecture and enables new application types to plug into Hadoop, including stream processing, graph processing, bulk synchronous processing and message passing interface (MPI)
- HDFS Federation, which allows Name nodes to act independently and without coordination with each other, HDFS Name Node High Availability improves data integrity and supports multiple failover options

For Operators - Data discovery

To explore new data types and large data sets that were previously too big to capture, store and process, business analysts use HDP extensively. This provides insights of clickstream, geo-location, sensor, server log, social, text and video data.

Enhanced customer experience

- Creating dynamic customer profiles
- Data Lake of customer 360-degree profile view



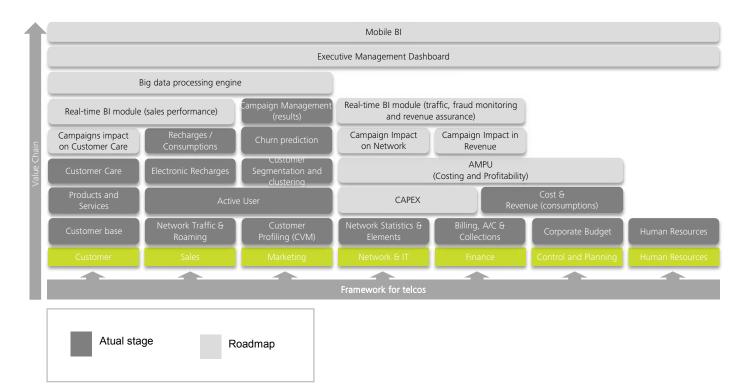
Scope of Big Data Implementation

Telecom industry has come a long way since it began its liberalization era and seen exponential growth in the last couple of years. Telecom companies have started using advance analytics to correlate, sanitize and extrapolate customer, and network data to generate a real-time view of customer preferences and network efficiency. They are collaborating with technology vendors to offer business customers analytics solutions along with core telecom services with an aim to increase revenue.

Following is the implementation roadmap that aims to create additional value for telecom operators to more actively support the financial and commercial strategies and to better understand customers through inclusion of social unstructured data using a big data engine.

Figure 4: Implementation framework – in Telecom Domain

"As we see the wave of DATA taking over the VOICE for telecom operators and the fact that every customer is leaving his footprint across various transactions, it is equally important to churn these insights and convert into meaningful consumption form, so that entire team takes decision cohesively. We are seeing great adoption of Analytics and Visualization in Telecom given the disparate sources of data, huge volume, decision cycles reduced and need to combine various functional data (Network and Finance) to get a consolidated view for decision making." Naveen Gattu, COO, Gramener



Source: Deloitte

Typically, telecom's big data engine comes with three types of resources, i.e., business supporting system (BSS), operation supporting system (OSS), and measurement report (MR).

- BSS data include billing, short message service record, call records, complaint records recharge history, and customers' demographic data. OSS manages communication network functions including network inventory, service provisioning, network configuration and fault management.
- OSS data include two categories, circuit switch (CS) and packet switch (PS).
- CS is related to the voice service supporting system.
 CS data reflects the voice service quality.
- PS is related to the mobile internet data service supporting system. PS data is also called mobile broadband (MBB) data, which are gathered by probes with deep packet inspection (DPI) technique. PS data describes customers' data usage behaviors such as mobile search queries, app usages, and streaming records.

Call Drop Analysis

In a highly competitive market, telecom companies while expanding broadband services focus on scaling up their network performance. A disruption or outage in network can lead to call drops and poor sound quality which harms the reputation of the telecom provider and can increase the attrition among its customers. Telecom companies should continuously monitor their networks for such disruptions and resolve root causes at the very early stages.

With call detail data collected at a rate of millions of records per second, employing such an analysis is quite challenging along with analyzing the exact reasons for call drops. Reasons for call drops are varied across:

- Network failures
- · Credit limits/ outstanding balance
- Handover issues

Dissatisfied customers may not always report on frequent call drops but would have a greater propensity to churn out in search of better services/coverage. To address these challenges, following approaches help:

- Analyze CDR data generated by customers, correlate with corresponding time interval network device logs and classify the reasons for call drops
- Drill down dashboards, which present a region-wise analysis on rate of call drops, reason for call drops and propose resolution to concerned telecom departments, is developed
- Enhanced visualizations, which assist in tracing location information based on frequency and revenue loss due to call drop reasons, are presented for analysis
- Subscriber calling pattern is considered for calculation of revenue loss per subscriber. Reports are generated on a Revenue Dashboard to indicate total revenue loss categorized per Base Station/Area/Division/City

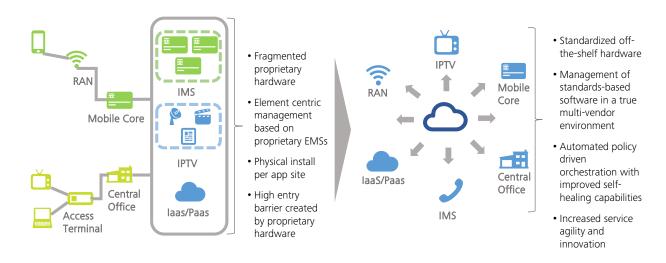
Apache Flume has the capability to ingest millions of call detail records into Hadoop. Apache Storm use this data to run pattern recognition algorithms which identify any troubling patterns in real-time.

Network Analytics

Enterprises across industries need high visibility into their networks from lowest granular to the edge to understand how it serves both internal and external customers. Transmission towers and their related connections form the spine of a telecommunications network. Failure of a transmission tower can cause service degradation. Replacement of equipment is usually more expensive than repair. There exists an optimal schedule for maintenance: neither too early, nor too late.

Telecommunication networks are migrating from traditional hardware and appliance-centric deployments to cloudbased deployments, with software as the critical component of all network functionality – NFV (Network Function Virtualization) or SDN (Software Defined Networking)8. Both of these aim at virtualizing network applications as well as the network connectivity.

Figure 5: Telecom Architecture shift towards a NFV/SDN Environment



Source: Deloitte

Functional areas in Transmissions/Network

- Address Activation and Provisioning: Activation and provisioning needs to be enabled via an intuitive portal that provides a simplified workflow, and pre-defined templates for standard activities such as service definition and composition, service activation, as well as service modification.
- **Change Management**: Software-based workflow should be implemented to acquire approvals for changes, and automatically bring about approved changes via the centralized orchestrator. Logically isolated test environments, built using SDN, can simulate multiple What-If scenarios and quantify impact of planned changes in a staging environment.
- **Inventory Management**: Auto discovery algorithms and version controlled archival systems can help in establishing a real-time topology view and inventory reporting system.
- Performance Optimization: Self-optimization capabilities need to be introduced in performance management modules which can optimize configuration based on current network performance e.g. scale up Virtual Machines (VM), add new Virtual Network Functions (VFN) instances for load balancing, configure new routes between VMs, etc.

Using network analytics, telecom companies can monitor real-time data from network, compute, and storage applications which examine data use to react more quickly to potential risks or failures. Big data analytics stores unstructured, streaming and sensor data from networks.

Big data analytics tools can virtualize their functions and process data on-demand enabling a-per-unit-of-informationprocessed models. This brings in revolutionary new paradigm for effective customer experience monitoring, quality of service tools and other monitoring or analysis solutions.

Hadoop stores unstructured, streaming, sensor data from the network. Telecom companies can derive optimal maintenance schedules by comparing real-time information with historical data. Machine learning algorithms can reduce both maintenance costs and service disruptions by fixing equipment before it breaks.

Churn Prediction

Retaining customers is one of the most critical challenges in the maturing mobile telecommunications service industry. Prediction of customers who are at risk of leaving a company is called as churn prediction in telecommunication. Acquiring a new customer is more expensive than retaining the old one.

With help of predictive models and machine learning algorithms, it is possible to accurately identify customers who are likely to lapse. Bringing together data collected on customer usage, complaints, transactions, social media, they can create factors which can identify customers at risk of moving out.

Industry trends show that annually there's an over 20-40% churn, especially in telecommunication industry.¹¹ The operator's success is determined by its ability to control and optimize processes, as well as identify growth opportunities and reach promptly to market changes.

Techniques such as decision trees, which enable long-term forecasting and early detection of customer's value loss and profiling, allow marketers to use variables to easily identify potential churners.

Figure 4: Predict customer churn



Source: Deloitte

Churn - Preventive approaches

Big data helps companies on the cutting edge of customer experience figure out why some customers leave and how to stop others before they do.

- Analyze social data and other public forms of data which could be used to understand customer's sentiment and needs before they become issues or problems that lead to churn.
- Use clickstream data to calculate the speed of downloaded data. This gives great proxy for understanding throughput speeds at the tower level, as well as handset speeds.
- Identify value proposition of individual customer loss and create targeted strategies.
- Analysis of large, heterogeneous, structured and unstructured volumes of data quickly and easily which enables to minimize acquisition costs and increase marketing efficiency.

Customer Segmentation

The process of segmenting the market or customer base into groups that behave similarly is known as customer segmentation. This approach is quite prevalent as telecom providers are growing to serve larger number of customers and is a critical component which enables telecom providers to make strategic decisions.¹² Some of the benefits that telecom companies can derive through segmentation are listed below:

- **Customer Value Segmentation** Identify loyal customers who have a high potential lifetime value, enabling targeted marketing and retention activities to reduce the churn rates.
- Create tailored products for customers More extensive customer segmentation allows to create products offering high utility for each segment depending on customer needs, thereby increasing customer satisfaction.

By implementing big data technologies telecom companies are able to bring together all the data they collect on customers and their usage history to run more effective customer segmentation which enables them to run a more targeted campaign. Combining multiple mass market and specialty compilers, membership databases, and thousands of unique customer files from publishers, direct merchants, seminars, associations, and others, creates a single platform to launch all direct marketing communication.

- **Identifying high-value and long-term customers**: Using integration of big data and additional attributes, it can contribute to reorder predictive models which helps in determining customers that are more likely to repeat purchases/buying patterns.
- Identify potential customers: Big data helps to identify new customer's base who can have high potential in the near future. This helps telecom companies to identify the target reach and cutting costs on non-associative client base.



Predictive Campaign

Today in this hyper-competitive environment, operators face major challenge in retaining the customers and in generating revenues from customers. Engaging with customers throughout their life-cycles has been key to be successful in the long run. Real-time analytics helps operators to analyze, correlate, and act on insights of data proactively and resolve customer churn and revenue losses.

Insights gleaned from real-time analysis of consumer data uncover purchasing patterns that are highly-personal, and drive responses that are ultra-timely. Customers get exactly what they want, when they want it. Businesses get real-time data that makes future offerings even more targeted. The result is loyal customers who know they are understood and are likely to stay put.

- Renewed focus on the individual: Marketing to masses, once backbone of marketing methodology kind of became distant memory. Consumers these days show clear preference which has content manufactured for the masses and they engage with content that has personal context. Digital technologies, fuelled by analytics and predictive intelligence, have finally enabled marketers to realize the one-to-one personalization.
- Shift to omni-directional: Connectivity revolution, driven by device proliferation and the internet of things, is to evolve their omni-channel strategies in a dozen new directions. Digital devices are multiplying and the leading marketers are staying one step ahead to capitalize on the customer opportunities.
- Event-driven communication: Consumers travel through today's world amid a series of data-driven and datagenerating events. Does the customer prefer a text message letting them know? Should the customer's plan be upgraded for free, based on preferences they selected? Aided by analytics, event-driven communications are more meaningful to both TMT companies and customers.

Applying big data analytics, telecom companies use applications to conduct high volume of campaigns tools which encompass data management, predictive analytics, campaign management and performance monitoring – to deal with volume of data that needs to be sifted through and to configure business rules and seamless integration with other operational systems.

Location-Based Services

Geographical location data related to mobile devices provide great insights of real-time information. These data sets can be utilized for various analytical services and representations. For example, a restaurant in suburb of Mumbai would value in knowing the population density within a specific geo radius during specific time intervals. The restaurant management can utilize this information to efficiently manage staff and adjust operating hours accordingly.

Organizations are increasingly turning to geospatial technology to harness the power of location information to assist with the toughest business challenges. Location-based positioning is driving new opportunities for telecom operators and enterprises to easily utilize infrastructure to support intelligent positioning services where Wi-Fi is in considerable usage.¹³

- · Cost savings through the consolidation, monitoring, and optimization
- Stimulating business by creating a digital strategy to access broadband and Wi-Fi services

Accessed information from device sensors -- GPS, Wi-Fi, Bluetooth, accelerometer records usage patterns and context in which it is being used. Once this data converts to a user profile, it can be utilized in many novel ways in consideration with usage context. A few of the areas can ascertain to identify following information:

- · Asserting real-time location data
- Geo-advertising, traffic, local search
- Location data from intelligent transportation environments Vehicular ad-hoc networks (VANET)¹⁴
- Data from indoor positioning systems RFIDs (radio-frequency ids) & Wi-Fi access points¹⁵

Use Cases from Telecom Industry

Customer Churn Predictive Model — Global Telecom Product and Service Provider

Description: A leading telecom products and global service provider wanted to identify customers who are at risk of not renewing their contracts and understand what factors influence their decision. The telecom company identified predictive modeling as key initiative for this analysis, furthermore they wanted to understand if it is possible to create strategic interventions which can influence customers in renewing their contract.

To perform this analysis Deloitte sourced the available data sources like:

- Customer demographic information
- Customer usage data
- Product information
- Reseller/distributor data

While bringing all these sources together, Deloitte has created additional variables to create generalized linear model technique to identify most influential factors which affect customer renewal rates.

Result: Deloitte helped in identifying the key factors such as contract duration, discount percentage, average prior contract length, distributor historical renewal rates influenced contract renewal rates significantly. Deloitte developed a visualization tool to help client identify which contracts are more likely to be renewed, based on the identified factors, intervention strategies are put in place to improve the probability of renewal. The client got benefited with saving millions of dollars in 2013.

Marketing Mix Model — Large Telecommunication Company in the US

Description: Leading telecommunication firm in the US wanted to analyze effectiveness of marketing investments pursued. It wanted to understand further if there were any channels/medium which were more effective in certain regions or customer segments. In addition, optimize their marketing spends across channels to drive maximum ROI.

Deloitte made use of data on the weekly marketing spends of the client, external data on how many people consumed the content in each channel and also the competitor marketing activity data. Bringing together all the data, a panel dataset was created at weekly level which was used for further analysis. Several new variables were created with the marketing spend variables by using lag and ad-stock transformations. These variables were then used in the analysis to build a mixture model which predicts the effects of the client's and competitor marketing activities on the number of additional contracts that were signed.

Result: During this analysis, marketing activities have helped in attaining new contracts, while retaining existing contracts.TV and Print were identified being most effective channels to bring in new business and identified promotional activities which were effective in driving sales. Results from mixture models and running various optimization scenarios, Deloitte helped the client in optimizing their marketing spends which led to a 27% increasing in annual revenues.

Big Data Implementation — Large Telecom Company in the US

Description: With telecom provider capturing increasing volumes of data each day, traditional databases are proving to be expensive and slow. A large telecommunication client was looking for an alternative low-cost solution that could help in handling the large volumes of data without compromising on the performance.

Deloitte worked with the client to identify jobs which didn't require real-time processing and off loaded the large volumes of data related to these tasks to a more cost optimized Hadoop cluster solution. Deloitte architected the solution using Big Data technology enabled by Apache Hadoop to implement an extract, transform and load (ETL) process that could handle the volumes of usage data. Usage data was sourced and ingested using SQOOP and loaded into HDFS. Transformation, aggregations and cleansing were performed using PIG, HIVE and MapReduce depending on the complexity of the operations. Results were transferred back to EDW and other data-marts using either SQOOP or HIVE JDBC connection.

Result: Deloitte demonstrated to the client that Hadoop was a viable alternative to Teradata to process large volumes of Telecom usage data. With this implementation, client was able to reduce client was able to reduce operational costs from \$200K per TB to \$2K per TB.

Big Data Implementation — Large Online Retailer in Korea

Description: Online retailers are capturing more and more data on their customers and their browsing patterns. The client started its big data team to capture browsing data of its customers and wanted to identify how it can leverage this data to create additional value through introduction of data mining solution and analysis of data modeling.

Deloitte leveraged the available purchase data and click information data for the analysis and suggested additional information retrieval for future analysis. Identified associations between products and used it to improve current recommendation engine. Further RFM (Recency, Frequency, and Monetary) analysis were carried out to identify customer value and were applied across all models. In addition, Deloitte developed loyalty models to identify the most loyal customers which helped client in targeted marketing.

Result: Deloitte developed visualizations which gave insights into how the associations between products change with user-context. In addition, improvised to improve click-through rates of their recommendation engine by 76% which resulted in 2% increase in revenues.

About Deloitte Analytics

Deloitte pairs its vast industry experience with a unique, unbiased approach to analytics, turning everyday information into useful and actionable insights for decision makers. Combined with traditional management tools, Predictive Project Analytics gives Deloitte clients the ability to track, predict, and understand the real-time status and risk of any project during all stages of the project lifecycle.

Deloitte Analytics generated approximately USD 2.4 billion in global BA service revenue in 2014, representing 14% growth year over year, with 11,000 professionals. Industry coverage includes banking and insurance, consumer products and retail, life sciences, media and entertainment, and manufacturing.

Our global practice ensures to deliver implementation of a Big Data Analytics solution provides a unique set of capabilities to achieve customer's key business, technical and operational challenges. Having strong focus on innovation, breaks through labs to meet clients demands and Highly Immersive Visual Environment (HIVE) labs, as well as a breadth of analytics accelerators, including analytics value maps, target operating models and rapid solution prototyping. All is available through its global network of 21 Global Delivery Centers and 25 Deloitte Greenhouses.

Contact:

Tel: +91 20 6624 4600 Email: inideas-tmt@deloitte.com Website: www.deloitte.com/in

Acknowledgements:

Santosh Anoo	Monojit Mazumdar
Senior Director	Senior Director
Anu Peisker	Kalyani Dacha
Manager	Specialist Master
Hari Mahendra Singh	Abhishek Kuppili
Senior Consultant	Consultant

About Aegis School of Business, Data Science and Telecommunication

The old models of producing general managers and functional specialists are no longer adequate in today's fast changing environment. Organizations need leaders who are able to harness the power of innovation and lead them through a rebirth - they need 'Renaissance Leaders'. "How do we create a Renaissance Leader? Are they born or are they made?"

We founded Aegis to answer these questions. Founded in the year 2002 in Indore (India) with support from Bharti Airtel to develop cross functional telecom leaders. Aegis is the first institution in India to offer one year Masters in Telecommunication Management (MPTM) which is commonly known as MS in Telecom Management and MBA in Telecom Management. In the following years, Aegis established itself as a leading school in telecom management and shifted its campus to Mumbai, the economic nerve center of India. The DNA of Aegis lies in the entrepreneurial spirit of initiating unique specialized programs backed by industry support in terms of knowledge and content and is effectively reflected in all its endeavors. Aegis current delivers programs to executives over 25 countries around the world using mUniversity platform. Aegis has currently academic alliance with Telecom Ecole de Management based in France.

Millstones:

- In 2002, Aegis launched Masters in Telecommunication Management (MPTM) with Bharti Airtel's support.
- · In 2005, Aegis launched campus in Mumbai.
- In 2010, Aegis started Aegis Graham Bell Awards for innovation in Telecom, Internet, Media and Edutainment (T.I.M.E) and Social, Mobility, Analytics and Cloud (SMAC) for developing an ecosystem for fostering innovation in India. This award in organized with support of Cellular Operators Association of India (COAI); Convergence India; Deloitte and Telecom Centre of Excellence (TCOE).
- In 2011 Aegis launched mMBA, an MBA program on tab.
- In 2013, Aegis launched mUniversity.mobi MOOC platform with the Education market place
- In 2014, Aegis partnered with Telecom Sector Skill Council, which is part of NSDC, Govt. of India initiative and trained over 3000 candidates for MP Government, Vodafone, Ericsson, Serco, First Source, D.Y. Patil University and Bharati Vidyapeeth University on various vocational courses like mobile repair, optical fiber splicing, customer care and in-store promoter.
- In 2014, Joined and with CETTM, MTNL
- In 2015, Aegis in association with IBM has launched Post Graduate Program (PGP) in Business Analytics & Big Data, India's first holistic Data Science Program and PGP in Cloud Computing. It's a joint certificate of IBM & Aegis and delivered by IBM experts and Aegis faculties.
- Launched Big Data Product Factory

Contact:

Aegis School of Business, Data Science and Telecommunication

Phone: +91 22 2570 2815

Mahesh, Sector 15 Central Business District (CBD), Belapur, New Mumbai, Maharashtra, 400614

CETTM, MTNL, Heranadani Gardens, Powani, Mumbai

www.aegis.edu.in

www.mUniversity.mobi/Aegis

Email: info@aegis.edu.in

References

- 1 What is big data? See: https://www-01.ibm.com/software/data/bigdata/what-is-big-data.html
- 2 Digital Cosmos to Include 32 Billion Devices Generating 44 Trillion GB of Data by 2020. See: http://www industrytap.com/digital-cosmos-include-32-billion-devices-generating-44-trillion-gb-data-2020/28791
- 3 Search on for the big data professionals. See: http://www.thehindu.com/news/cities/Hyderabad/search-onfor-the- big-data-professionals/article6155919.ece
- 4 28 Must See Social Media Statistics. See: http://www.socialmediatoday.com/ content/28-must-see-social-media-statistics
- 5 Digital, Social & Mobile in India in 2015. See: http://wearesocial.net/tag/india/
- 6 Search on for the big data professionals. See: http://www.thehindu.com/news/cities/Hyderabad/search-on-forthe-big-data-professionals/article6155919.ece
- 7 Read structure and unstructured data. See: http://www.snia.org/sites/default/files/p17_SNIA_ BigDataCommittee_v2revised.pdf
- 8 Gigamon Announces Vision of Network Functions Virtualization for Tools in Service Provider Infrastructure. See: https://www.gigamon.com/company/newsroom/press-releases/gigamon-announces-vision-network-functionsvirtualization-tools#sthash.xSVtN35k.OlnLrpNj.dpuf
- 9 Hadoop framework. See: http://www.sas.com/en_us/insights/big-data/hadoop.html
- 10 Map Reduce : http://creately.com/diagram/example/h21wfdxq2/MapReduce
- 11 Operationalizing SDN and NFV Networks. See: https://search.deloitteresources.com/home?k=SDN&app=DNET&r =dttindustry%3D%22%C7%82%C7%8255533a546563686e6f6c6f67792c204d6564696120616e642054656c 65636f6d6d756e69636174696f6e737c54656c65636f6d6d756e69636174696f6e73%22 Predict churn in telecommunication domain. See: https://www.zu.de/daily-wAssets/pdf/2014/MA-Thesis_ Weinheimer.pdf
- 12 Process of segmenting. See: http://www2.deloitte.com/content/dam/Deloitte/uk/Documents/technologymedia-telecommunications/deloitte-uk-a%20people-based-telecom-business%20-%20Breathing-new-life-intosegmentation-strategies.pdf
- 13 The Forrester Wave™: Advanced Data Visualization (ADV) Platforms, Q3 2012
- 14 Vehicular ad-hoc networks. See: http://www.csie.ntpu.edu.tw/~yschen/mypapers/TS-2010-1-2.pdf
- 15 The Impact of Licensed Shared Use of Spectrum

28



Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee ("DTTL"), its network of member firms, and their related entities. DTTL and each of its member firms are legally separate and independent entities. DTTL (also referred to as "Deloitte Global") does not provide services to clients. Please see www.deloitte.com/about for a more detailed description of DTTL and its member firms.

This material is prepared by Deloitte Touche Tohmatsu India LLP (DTTILLP). This material (including any information contained in it) is intended to provide general information on a particular subject(s) and is not an exhaustive treatment of such subject(s) or a substitute to obtaining professional services or advice. This material may contain information sourced from publicly available information or other third party sources. DTTILLP does not independently verify any such sources and is not responsible for any loss whatsoever caused due to reliance placed on information sourced from such sources. None of DTTILLP, Deloitte Touche Tohmatsu Limited, its member firms, or their related entities (collectively, the "Deloitte Network") is, by means of this material, rendering any kind of investment, legal or other professional advice or services. You should seek specific advice of the relevant professional(s) for these kind of services. This material or information is not intended to be relied upon as the sole basis for any decision which may affect you or your business. Before making any decision or taking any action that might affect your personal finances or business, you should consult a qualified professional adviser.

No entity in the Deloitte Network shall be responsible for any loss whatsoever sustained by any person or entity by reason of access to, use of or reliance on, this material. By using this material or any information contained in it, the user accepts this entire notice and terms of use.

©2015 Deloitte Touche Tohmatsu India LLP. Member of Deloitte Touche Tohmatsu Limited

Deloitte Touche Tohmatsu India Private Limited (U74140MH199 5PTC093339), a private company limited by shares, was converted into Deloitte Touche Tohmatsu India LLP, a limited liability partnership (LLP Identification No. AAE 8458), with effect from October 1, 2015.