Online ISSN 2424-6166. EKONOMIKA 2017 Vol. 96(2)
DOI: https://doi.org/10.15388/Ekon.2017.2.11004

THE CHALLENGES OF BIG DATA ANALYTICS IN THE MOBILE COMMUNICATIONS SECTOR

Kęstutis Saldžiūnas
BITĖ Group, Lithuania

Rimvydas Skyrius*
Vilnius university, Lithuania

Abstract. *The activities of the MNO (Mobile Network Operator) feature rapid development and business model innovations; one of their principal results is the communications infrastructure that is vital for economic growth. This dynamic and changing mode of operation (modus operandi) introduces high requirements for business decisions and overall informing to maintain competitiveness. One of the principal success factors in MNO activities is the application of contemporary information technologies, in particular technologies of business intelligence and analytics.*

The activities of MNO create large data volumes, leading to a significant potential discovery of insights from data. As a result, MNOs have been using analytical technologies to mine large data volumes for several decades, and experience accumulation started long before the term "big data" emerged in academia and business. The growing dynamics of activities drive the efficient use of analytical experience to boost competitive advantage. The goal of this paper is to define the most important features of the use of big data analytics in MNO business and any possible related challenges.

Keywords*: mobile network operators, analytical technologies, big data analytics.*

1. Introduction

The dynamic and competitive sector of mobile communications comprises an important part of the national economy in many countries, where in 2016, mobile services added 4.4% to gross national product (The Mobile Economy 2017).

Lately, the growth of MNO income has been experiencing controversial trends (Telco 2015: Five … 2015). For example, the turnover for Lithuanian MNOs has fallen 29% in the years 2009-2014. The general trend of the industry is two-sided: the income from phone calls, SMS and mobile data transmission is slowing down, while the use of mobile data experiences fast growth. This results in growing expenses for network support, the growth rate of which exceeds the income growth rate (Fig. No. 1).

* *Corresponding author:*
Faculty of Economics, Vilnius university, Sauletekio Ave. 9, LT-10222, Vilnius, Lithuania
E-mail: rimvydas.skyrius@ef.vu.lt

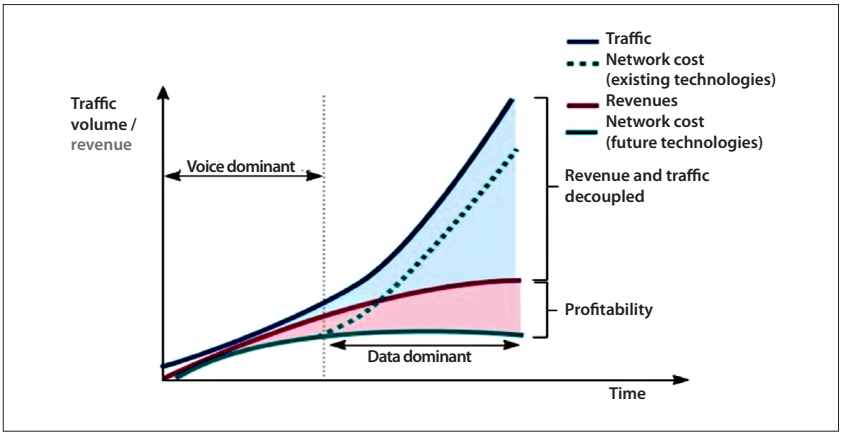

FIG. No. 1. **Divergence of MNO income and expense rates (Telco 2015: Five ... 2015)**

The reasons for the abovementioned growth are:

- An intense competition between MNOs;
- The proliferation of smart phones that pulls the use of mobile data;
- Spread of the internet, together with the growing space of digital services;
- The new, specialized services in communication and content, or the over-the-top services (OTT) are pushing out traditional MNO services – e.g., calls and messaging in Skype, Viber or WhatsApp applications, and create now needs for the customers.

The reasons for the growth of MNO costs are:

- The growing data flow puts extra strain on the current network and IT (Information technology) infrastructure, which has to be upgraded;
- A growing number and complexity of systems in use;
- The need to upgrade provided services or cooperate in reaction to the innovative services introduced by OTT on a permanent basis;
- The reduction of income from EU users on the road because of the market regulation taking place.

Under such circumstances, it is natural that MNOs are balancing on the border of survival and are forced to constantly search for ways to improve service quality, business effectiveness and profitability. One of the directions to reduce costs and boost income is the reinforcement of competences in Big Data analytics. The preconditions for development of Big Data analytics, or BDA, are the following (Zalieckaitė, Skyrius 2012):

- Technology advances;
- The rapidly growing volumes of captured data in traditional information systems, especially those operating in businesses dominated by large numbers of small transactions;

- Newly formed data sources on the internet with data on website visits and activities in social networks, especially the growth of video content that, although structured in terms of formats and presentation, in the future will require a semantic analysis to understand its content;
- The need to record significantly more data on events related to service provision; such data would be used for cost accounting and quality assessment (e.g., location data);
- A faster introduction of new services.

2. Opportunities and Application Areas for Big Data Analytics

The data available for MNO has formidable utilization potential for ensuring the growth of MNO income, cost reduction and provision of quality services. Yet larger potential lies in creating a foundation for the expansion of strategies of services and network, where the analysis of customer data supports the discoveries of new customer needs, leading to new products or even new business models. This point is supported by Gartner (2012).

Banerjee (2013) provides a comparison of infrastructures required for traditional, or legacy analytics with those required for Big Data (Table No. 1):

TABLE No. 1. **A comparison of infrastructures required for legacy analytics and Big Data analytics**

Criterion	Legacy analytics	BD analytics
Storage cost	High	Low
Analytics	Offline	Real-time
Utilizing Hadoop	No	Yes
Data loading speed	Low	High
Data loading time	Long	Average 50%-60% faster
Data discovery	Minimal	Critical
Data variety	Structured	Unstructured
Volume	Gigabyte, terabyte	Petabyte, exabyte, zettabyte
Administration time	Long	Average 60% faster
Complex query response time	Hours, days	Minutes
Data compression technique	Not matured	Average 40%-60% more data compression
Support cost	High	Low
Hardware	Single powerful server	Many ordinary servers

Source: Banerjee (2013).

With this, Banerjee (2013) has split the areas of BDA application in MNO activities into four categories:

A. Real time analytics and decision-making

Network congestion control and optimization. Radio access network access is a limited resource and may experience congestion if user numbers grow. The aggregation of available information about clients, service use and location enables the prioritization of client flow, thus enabling to maintain the required quality of services.

Location-based information on customers and their activities. The availability of this information in real time enables event-driven marketing and reaction to client-related events in the right time and the right place, like sending an SMS about approaching a cost control limit or an attractive proposal of a new payment plan. Qiao et al. (2016) point out that available mobile internet data traffic contains huge potential in analyzing human mobility. Altogether, non-interactive services currently are hard to imagine.

Deep packet inspection. The volumes of data flow are analyzed together with its content. This data, when aggregated with client data, enables MNOs to inform clients about having reached a huge bill or proposing an attractive service package matching the usage pattern.

Network security and fraud prevention. The real time estimation of anomalies in client behavior may trigger the blocking of fraud calls, thus avoiding income loss due to unpaid debts.

Activity reports. An MNO has significant competitive advantage if it receives required activity data on an enterprise level – e.g., monitors the social networks for client reaction to promotions in real time and reacts appropriately.

B. "Smart" marketing

Optimized offers. The performed analysis enables the discovery of profitable and popular service packs and service support. It also supports proposal development on the basis of information on network usage, data volumes, loyalty points, events and rules. The cooperation with OTT providers also facilitates the development of innovative support modes, like developing offers for former users, thus boosting application and development of value-added services.

Cooperation with OTT providers. As Banerjee (2013) argues, the business model can work in two ways: an MNO may promote the use of OTT content partner applications where sponsored access is funded by such partners or include these applications as part of their data plans.

Customer retention. In addition to common segmentation, more customer retention tools are available:

- The detection of current or potential multi-SIM owners with appropriate action to follow;
- The detection of locations with a higher churn rate, with subsequent sending of appealing offers;
- The prevention of rotational churn to discourage customers from the illegal use of promotions that apply for new customers only;
- The social network analysis to define circles of friends and families, and to present offers through decision influencers.

C. Operational efficiency

Proactive customer service. The creation and application of error correction knowledge base by constant monitoring of service use and incident forecasting.

Intelligent network planning. The planning system may perform modeling of service demand, estimate network loads and predict resource exhaustion.

Cell site optimization. By automatically managing cell interaction, 4G networks are intended to become self-optimizing in managing power consumption, balancing traffic and minimizing interference. This function would significantly increase its value if connected with contextual user data.

Mobile network load reduction. By joining network information with available customer information, intelligent decisions on subscriber offloading to Wi-Fi network could be made.

D. Innovative business models

Demand satisfaction. By encouraging clients to express their needs, and knowing their precise location, an MNO may match the needs with the offerings of business entities in the nearby areas.

Payments. By cooperating with business partners, such as retail chains, an MNO may offer an infrastructure for electronic payments. Transaction data thus becomes available and enables generating more revenue by customization and personalization.

Provision of data access and network API to third parties. By leveraging access to their network management, APIs and the available data, an MNO may enable the creation of new innovative services by themselves or by external developers.

E. Enhanced customer experience

Profiling and segmentation. Real-time capture of data on customer and service use that enables customer segmentation by usage, interests, location, social class, churn propensity etc.

Analysis of Web page usage. Real-time capture of data on visited pages, leading to an analysis of navigation scenarios and browsing results.

All the named types of BDA application may be grouped into two groups: creation of additional revenue sources and enhanced effectiveness of activity. The majority of industry respondents believe it is more important for operators to harness the power of Big Data to drive new revenues streams externally than to drive efficiencies internally. By 2016, almost every operator to which Big Data is relevant have embarked upon their strategy with a view to bringing greater advantages in customer retention, segmentation and targeting as well as network planning and optimization.

3. Challenges of BDA Use

3.1. Challenges in Analysis

Data is useless without a multistep process of their analysis, during which most of created value comes out of practical application of discovered insights (Labrinidis and Jagadish 2012). Analysis starts with data extraction and accumulation. A significant part of data is redundant, so the data is filtered, deduplicated and compressed. In real-time analytics, data are analyzed at the moment of capture, thus avoiding their record.

The collected data are structured, semistructured or unstructured. The latter are unsuitable for formal analysis until converted into structured data; they have to be cleansed, transformed, integrated and aggregated. An important point in reducing data to manageable volumes is to avoid loss of valuable information. For example, data on failed calls has been ignored for a long time due to the lack of legal obligations to accumulate this data for the MNOs in European Union.

The analysis process should be automatic to avoid broken rules and mistakes. However, in practice, analysis is often performed using Excel spreadsheets. This is a flexible and easily accessible analysis mode. However, it is also quite time consuming, and this problem is solved by assigning more employees to achieve the required rate of analysis.

The results of analysis are largely influenced by the match between the analysis task and the logical model of a database. Therefore, this model has to be developed by database experts in cooperation with domain experts. Another direction would be to use analytical technologies independently of a database logical model.

The accumulated data are not sufficiently reliable because of errors, duplication and missing data. Data are cleansed by applying known models of data validation or error recognition. The fact that data are redundant and contain significant “noise” is an advantage, because it enables multiple modeling and correlation analysis that eliminates separate deviations and reveals hidden models and relationships. The use of data validity cross-checks also makes it simpler to compensate missing data or correct mistakes.

The problem of BD analysis is the missing integration between database systems and analytical tools performing non-SQL processing: data mining and statistical analysis.

This stage takes longer because the data have to be exported, statistically processed, and the results have to be imported back into traditional relational database (Agrawal et al. 2012).

A common mistake in implementing BDA is to assign most of attention to data collection, organization and transformation into data (Nicholson 2013) Thus, the focus over the most important stage – insight development – is lost, and the produced information is not used for insights on how to improve customer services or upgrade business processes. This stage requires the presence of a person with specific data analysis skills, and if this is not specified in the project, data collection is pointless. For instance, customized proposals may have a reverse effect – allow the customer to act rationally, thus reducing overall revenue.

BD analysis is of little value if the end users cannot understand the analysis and interpret its results. The analyst has to have tools to verify the reliability of produced results:

- An accessibility of assumptions;
- A repeatability of analysis steps;
- An accessibility of data sources (preservation of metadata).

The last stage – the visualization of produced results – is of crucial importance, because it largely supports insight interpretation and use in making decisions.

3.2. Data Management Challenges

The MNO companies have several decades of experience in managing large data sets. However, the principal problems still remain:

- *Data silos.* Data silos are a typical feature of MNOs. Their existence is conditioned by the organizational structure and possibly large number of systems in use with their own data collections. However, complete data encompassing the whole organization are the most valuable, so data has to be integrated, including unstructured and non-standard data. As noted by He et al. (2016), in an MNO activities there is a considerable controversy in data placement: while charging and billing systems are traditionally centralized and allow company-wide analytics, large volumes of other important data, such as device data, cell site data, network data, back office data etc. are scattered throughout the organization.
- *Data centralization.* A centralized repository provides for cost saving, common standards and security, while a decentralized repository is more flexible and leaves more freedom for analysis, but contradictions regarding redundant and controversial data may emerge (Skyrius 2013).
- *Data volume.* Traditional data management systems experience problems while handling large data volumes (Jacobs 2009). Meanwhile, data volumes keep grow-

ing and exceeding the capacity of traditional data storage, processing and analytical technologies. Akyildiz et al. (2016) suggest that an alternative to move big data analytics to the cloud offers advantages on cost, efficiency and scalability, if the current issues on cloud connectivity and security are reduced to acceptable levels.

- *Data processing velocity.* The notion of BDA encompasses not only large data volumes, but also the capacity to process them in acceptable time. Usually the business accepts the results being 1-2 days late.
- *Data privacy and security.* Data on private persons and their service use may be collected and stored in conformance with legal acts in power. For instance, in Lithuania, MNOs are obliged to keep detail service data for no longer than 6 months, and only for issuing bills for services provided or solving disputes, as well as providing information to the authorities eligible. Musolesi (2014) points out that, regardless of available regulations, breaches of privacy will most likely occur, and suggests introducing stricter checks for privacy control.

3.3. Challenges related to organizational changes

The success of BDA development largely belongs upon successfully coping with organizational challenges in the following factor groups:

1. Management support

Value perception. A rather likely source of resistance – traditionally an MNO has a developed data warehouse that satisfies most of the internal information needs, and there is a history of attempts to use products bearing the BDA label with subsequent disappointment because of the failure to achieve expected results. An estimate of the level of management's motivation to create value from using advanced technologies would be helpful in implementing BDA (Skyrius 2013), if there are signs of disappointment and distrust in technology after the inflated wave of user expectations.

System management. BDA is based on open source platform, and as such may be considered an advantage because of significantly lower costs when compared to traditional vendor solutions. On the other hand, management may be opposing the open source technologies, because they require a professional internal team, and development is possible only with partial or full usage of internal development resources.

Establishment of new jobs. The fact that, apart from systems and data, additional staff will be needed, is often neglected. New or current employees will be required to possess specific data analysis skills: analytic experts, statisticians, data scientists. For instance, data warehouse engineers traditionally have skills handling data sets in transactional environments; however, they will have to reorient their skills for working with immutable data sets. Obviously, it should be expected that similar tasks in BDA, compared to traditional analytics, will take longer, at least in the beginning.

2. Enterprise processes

Cooperation. Apart from creating of the analytics team, it is no less important to create a cross-domain team cooperation to spread intelligence culture and data-supported decision skills in the company. One of the possible approaches is to have an analyst in all subject domains (Galbraith 2014). Failure to solve communication issues will preserve barriers for information sharing, and the number of separate domain silos will only grow (Banerjee 2013).

Decision making process. For MNO, the development of analytical skills impacts the decision-making process and speeds it up. Decisions may be made faster and in lower management layers because of data-driven approaches instead of experience and gut feeling. Such changes may be opposed by the management, because for them, information "democracy" reduces the asymmetry of the informing level that has previously given them advantage. Another important aspect is whether a MNO is ready to act in real time, establishing a real-time decision function in the business unit structure (Galbraith 2014).

In general, the resistance to BDA initiatives depends upon whether the new analytical skills strengthen or reduce the importance of available competences (Jay 2014). It is recommended to establish a position of Chief Digital Officer to cope with expected resistance in the case when available competences lose their importance.

Standards. In a highly dynamic environment of product development that is taking place in MNOs, the entire focus is given to product development, and common standards ensuring service usage data collection and integrity may be violated.

3. Employee support

Employee resistance is equally possible for the same reasons as the management's. The BDA implementation strategy should be planned accordingly: to be ready to withstand strong criticism of data warehouse and be ready to perform work in accordance with common standards. The implemented solutions should create real and visible value by addressing known existing problems and being financially viable.

The retraining of employees should be motivated by contemporary state-of-the-art technologies used in BDA, as well as by the growing job market demand for such professionals.

4. BDA implementation strategy

Two strategies are possible: top-down and bottom-up. According to Acker et al. (2013), traditionally, the most used strategy is the top-down type, which defines the business problem at first, and only then the search for required data is initiated. The weaknesses of the first strategy are the lack of accessible data and weak understanding of the data collection process. The application of this strategy most often fails to achieve significant

results, except for the cases where an MNO has excellent data analysis competencies. To avoid these problems when there's no experience with BDA, the second strategy is recommended, where the available data are analyzed first, and only then a problem to be solved is defined.

In more detail, the second strategy, according to the EMC report on Big data opportunities (EMC 2011), includes the following stages:

1. Control of data hygiene. A starting point would be the reorganization of available data stores and their preparation for the new type of analysis. The overhaul of existing data stores enables the retirement of outdated systems and the renewal of selected stores to maintain. An important function, as pointed out by He et al. (2016), is data preprocessing for data from geolocation sensors, video cameras, network logs; this data should be normalized, integrated with other related data and tagged wherever possible.
2. Staff training. After having evaluated the available BDA skills for IT and business staff, a plan should be developed for skill transfer or development.
3. Data value estimation. Business and IT individuals should be brought together to:
   - Define the most valuable data pools;
   - Evaluate data stores to be analyzed and consider ways for their expansions and development;
   - Pinpoint unstructured data sets and prioritize the ones to be converted into acceptable the format;
   - Define the most valuable data use scenarios with the shortest deployment time;
   - Define potential areas of business growth and related analysis questions; focus on answers to the iterative sessions of these questions.
4. Development of new dimensions for data sets. The development of analytical competencies will lead to interest in new methods for data combinations and uses. This activity would include the collection of new types of data, the discovery of new data sources and a data set combination to create new value and insights.
5. Evaluation of data analysis consequences. An early familiarization with legal requirements for data portability, security and compliance to ensure that discovered insights may be applied without creating legal issues.
6. Use of insights. The discovered insights are only as valuable as they are actionable, enabling decisions or action. Use of BDA at MNO creates opportunities to discover insights faster and more often, provided that the organization is agile enough to act fast.

To justify initial BDA expectations, a pilot project would be a reasonable step to start, where a dedicated and autonomous team with mixed competences collects and analyzes data. In this way, it is highly likable that unexpected opportunities will emerge and be

used to solve real-life business problems. This approach to begin the project allows the evaluation of the real benefits of BDA, an accumulation of expertise in working with BDA and understanding of long-term requirements to achieve the expected targets.

Conclusions

Big data is a widely known and massively promoted term, covering complex and thorough analytical and organizational work. MNOs are constantly watching the dynamics of market and technologies, aiming at efficient activities and investing into innovations. Due to the availability of huge volumes of data, MNOs have considerable experience in handling this data and gradually developing a data management strategy. One of the cornerstones of such strategy is big data analytics. With the support of BDA, MNOs may use available resources more efficiently, implement innovative services, leverage insights on customer behavior, altogether improving the velocity of enterprise processes.

Regardless of opportunities and benefits enabled by BDA, its implementation creates numerous challenges, starting with technical ones, related to data volume and quality, and continuing with organizational ones, related to changes in decision making and overall organization culture. The largest obstacle for a decision to invest into BDA technology is for short term sufficient capacity of available systems to cope with growing data volumes and satisfy information needs. The expected investment costs are driven up by the need to attract top-qualified employees, although this doesn't mean that investment into hardware will not be required.

The most favorable strategy for MNOs to achieve success with BDA is to execute small pilot projects in cross-functional teams with available data and looking for insights with instant practical use in decision-making. Fast success would reinforce management trust in technology and provide ground for upgrading to larger-scale technologies for big data analytics in the future.

REFERENCES

Acker O., Blockus A., Pötscher F. (2013) Benefits from big data & a new approach for telecom industry. Accessed on April 2, 2016 at: <http://www.strategyand.pwc.com/reports/benefiting-big-data

Agrawal D., Bernstein P., Bertino E., Davidson S., Dayal U., Franklin M., Gehrke J., Haas L., Halevy A., Han J., Jagadish H. V., Labrinidis A., Madden S., Papakonstantinou Y., Patel J.M., Ramakrishnan R., Ross K., Shahabi C., Suciu D., Vaithyanathan S., Widom J. (2012). Challenges and Opportunities with Big Data – A community white paper developed by leading researchers across the United States. Accessed on September 21, 2017 at http://cra.org/ccc/docs/init/bigdatawhitepaper.pdf

Akyildiz I.F., Nie S., Lin S.-C., Chandrasekaran M. (2016) 5G Roadmap: 10 key enabling technologies. *Computer Networks*, Vol.106, p.17-48.

Banerjee A. (2013) Big Data & Advanced Analytics in Telecom: A Multi-Billion-Dollar Revenue Opportunity. *Heavy Reading*, December 2013.

EMC report (2011). Big data: big opportunities to create business value. Accessed on April 2, 2016 at: <http://www.emc.com/microsites/cio/articles/big-data-big-opportunities/LCIA-BigData-Opportunities-Value.pdf>

Galbraith, J.R. (2014). Organization design challenges resulting from big data. *Journal of Organization Design*, Vol.3, No.1, p.2-13.

Gartner (2012). Market Trends: Big Data Opportunities in Vertical Industries. Accessed on March 26, 2016 at https://www.gartner.com/doc/2077415/market-trends-big-data-opportunities

He Y., Yu F.R., Zhao N., Yin H., Yao H., Qiu R.C. (2016) Big Data Analytics in Mobile Cellular networks. *IEEE Access Special Section on Theoretical Foundations for Big Data Applications: Challenges and Opportunities*, 1985-1996.

Jacobs, A. (2009). "The Pathologies of Big Data." *Communications of the ACM*, Vol.52, No.8, p. 1-12.

Labrinidis A., H. V. Jagadish, 2012. Challenges and Opportunities with Big Data. Accessed on April 2, 2016 at: <http://vldb.org/pvldb/vol5/p2032_alexandroslabrinidis_vldb2012.pdf>

Musolesi M. (2014) Big Mobile Data Mining: Good or Evil? *IEEE Internet Computing*, Vol.18, No.1, p.78-81.

Nicholson S.R. (2013) The Critical element of successful Big Data strategy. Accessed on April 2, 2016 at <http://www.socmedsean.com/comic-the-critical-element-of-a-successful-big-data-strategy>

Qiao Y., Chang Y., Yang J., Liu J., Kato N. (2016) A Mobility Analytical Framework for Big Mobile Data in Densely Populated Area. *IEEE transactions on Vehicular Technology*, Vol.66, No.2, p.1443-1455.

Skyrius R. (2013). *Verslo informacija: poreikiai ir jų tenkinimo keliai*. Vilnius: Vilniaus universitetas, 271 p.

Telco 2015: Five telling years, four future scenarios. IBM. Accessed on April 2, 2016 at https://www-05.ibm.com/cz/gbs/study/pdf/GBE03304USEN.PDF

The Mobile Economy 2017 (2017). Accessed on October 6, 2017 at: https://www.gsma.com/mobileeconomy/

Zalieckaitė L., Skyrius R. (2012). Analitinės programinės įrangos diegimo į mokymo procesą tyrimas. *Informacijos mokslai. Mokslo darbai.* Vilniaus universitetas. 2012, Nr. 61, 144-155.