Analytics – the truth is in there

Understanding consumers' behavior and their changing needs are key ingredients in good business decision-making. Valuable information in networks is underutilized, and this article describes Ericsson's approach to data analytics – how to create smart services from data.

ULF OLSSON AND JACO FOURIE

From a framework perspective, several levels of processing are required before bits become intellect. Figure 1 illustrates the relationships between data (bits), information (data in context), knowledge (interpreted information) and wisdom (the ability to make the right decisions). Here, Ericsson's holistic approach to cross-domain analytics, and how it allows many kinds of information to be interpreted using a common toolset, are discussed in practical terms. This implies that network information can be set into a user and service context, service behavior and acceptance can be related to network conditions, and consumer experience measurements can be related to all touch points.

Background

Information enables choice, relevance and accuracy in the decision-making process. Operators have been – and remain – the information gatekeepers of the telecom industry. To date, valuable information available in telecom networks has not been used to its full potential to make smart business decisions – but this situation is changing rapidly. That change is being driven by the need to shift focus from technology and services to user experience and relevance. As mobile broadband approaches global ubiquity, operators are spending more time on subscriber retention than acquisition, and this has resulted in them paying greater attention to the following stages of the customer lifecycle:

- target – enhance market awareness, proposition targeting and promotion relevance;
- buy – make the purchasing process smooth and enjoyable;
- set up – ensure efficient and seamless provisioning of the purchased service;
- use – guarantee that the service meets requirements for functionality and quality, and that customers are using the services available to them to the greatest possible extent;
- pay for – provide clear, convenient and easy-to-use payment facilities;
- support – enhance users' experience of customer care channels; and
- grow – include the flexibility to easily modify contract and service levels as customer needs evolve, or as a result of churn-management activities.

To provide effective management of the customer lifecycle, accurate and timely information is needed. Operators have an excellent opportunity to serve their customers better, but at the same time, they face issues created by the radically increasing amounts of data being generated by modern networks. Ericsson's focus is to address these issues, by handling the complexity and volume of data from the networks and other sources. In this way, correct and relevant information is provided for the processes used by operators to serve their customers with maximum efficiency.

It seems that operators are facing the classic combination of a challenge and an opportunity, both arising from the same basic phenomenon. The issue at hand is the surge in the amount of data generated by today's networks. Operators control an amazing amount of user data, demographic information and mobility patterns. Traffic to, from and among subscribers is yet another valuable source of information – moderated, of course, by opt-in, opt-out and other legislative and cultural limitations. The challenge: how to turn enormous numbers of bits into useful information? Traditionally, telephony systems have handled millions of call-detail records per day, resulting in the creation of customized management systems. The new traffic era, with different kinds of data traveling over airwaves through copper and through fiber, is generating massive amounts of data, stretching system capabilities to the limit. Ericsson's research into how to manage the data surge and push limits much further is outlined below, as well as how to widen the extent of available information to encompass all data handled by telcos.

From bits to insights

Analytics focuses on finding patterns and using complex pattern-matching techniques on new data to make sense of it. The sheer volume of information in networks makes it impossible for a person to process and absorb the multitude of events and data that is handled by a modern communication network. Consequently, the first level of analysis must be computer-processed. Intelligent programming looks for known patterns,

BOX A Terms and abbreviations

CEP	complex event processing
ENIQ	Ericsson Network IQ
KPI	key performance indicator
LTE	Long Term Evolution
OSS/BSS	Operations and Business Support Systems
ToD	time of day

with rules-based detection subsystems integrated with, or at least very close to, the network nodes generating and transferring data. The result of this process is filtered, aggregated and possibly reformatted data, which then becomes information – data with associated metadata. Data volumes resulting from this level of processing are more manageable, but they still tend to be associated with the network elements (or at least domains) where they originated.

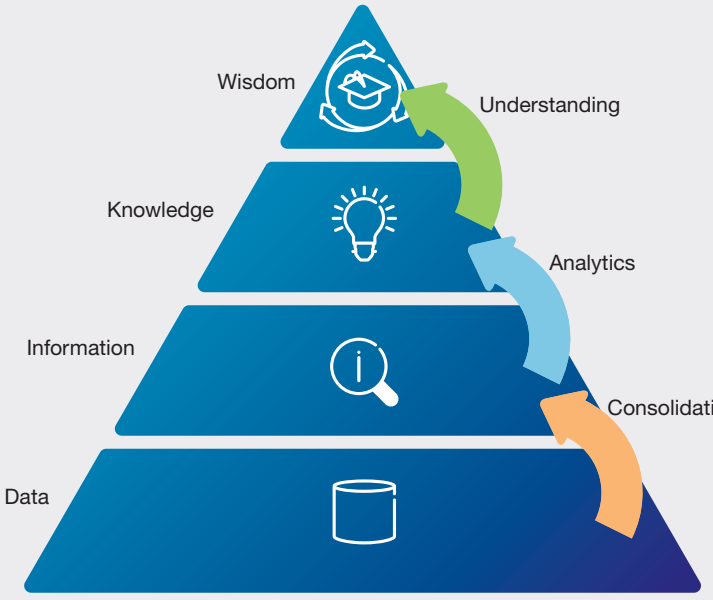
The next level of processing is mainly dedicated to detecting instances of different types of patterns. Here, the idea is to find patterns relating to network events in context – relating to who the user is, for example, which services they have accessed and what kind of device they are using. This process is, however, essentially about detecting predetermined patterns. At this point, the number of instances of a certain type of event can be displayed on an operator's dashboard showing where identified KPIs meet their targets.

This type of analytics can be compared to the production side of a mining operation: the whereabouts of the ore is known and its content is known, as is the technique used to extract it. It is not surprising that analytics is sometimes called data mining. However, there is another side to the mining process that has a similar parallel in the data field: namely prospecting.

Prospecting is the process of finding previously undiscovered wealth – the skills and tools of the prospector lead to the detection of buried treasure. The parallel in data-mining is usually referred to as ad hoc analysis, which involves twisting data around, associating and slicing it in various ways until the analyst finds previously unidentified patterns. Consequently, an interesting duality arises in smart business intelligence originating, on the one hand, from ad hoc analysis, and on the other, from production analytics. Ad hoc analysis requires flexibility, high performance and rapid recalculation of vast amounts of data using intuitive graphics tools. Production analytics drive the big boards in operation centers and feed business dashboards providing essential business intelligence reports. The challenge – which defines the architecture that needs to be built – is to encompass both aspects of analytics in a common, efficient and reliable infrastructure.

BOX B
Ericsson Network IQ provides service-aware and network-performance management with statistical trend analysis, by service or subscriber, and problem-management alarms for proactive problem indication.

FIGURE 1 From data to wisdom



Sources of information

In the telco world, traffic-handling nodes are naturally considered as primary sources of information, and they are indeed the source of most data. Increasingly, however, other types and sources of information are being used to create a holistic view of the user, including:

- user-related data such as name, demographics (gender and income bracket), account status, subscription history, and customer-reported issues; and
- service-related data such as frequently used services, average usage duration, from which device and from where.

The Ericsson portfolio includes products such as ENIQ Statistics and ENIQ Events, as well as the Analytics Suite set of tools that build on the comprehensive range of online charging products.

Architecture

Analytics today is different from traditional analytics in three ways:

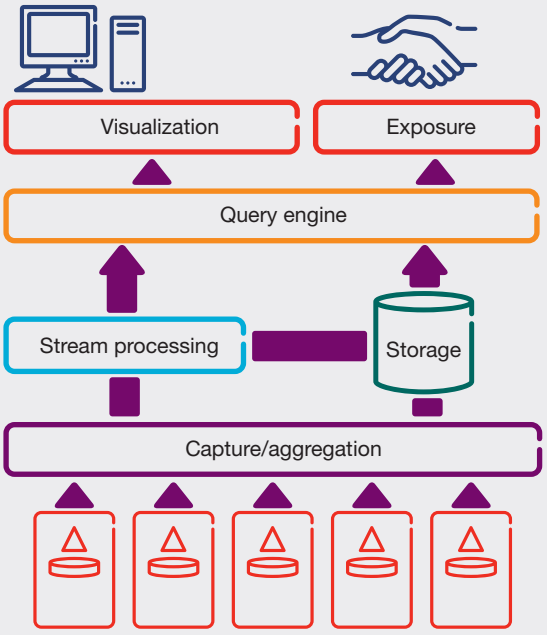
- reporting is no longer just about accounting and is instead moving toward understanding, implementing change, and benchmarking progress. In other words, reporting is moving from passive formatted reports to full-circle, active decision-making processes in which the periodic report is less important and the process of continuous change is the primary focus;
- there is a shift from a network-oriented approach – nodes and services – to users and user segments, with a focus on how users experience the services provided by the operator; and
- increasingly, focus is on finding new segments and innovative offerings that the operator can offer.

Figure 2 shows a high-level layer view of the new analytics architecture, illustrating how the flows from all information sources are unified to provide the above capabilities. Information sources exist in all domains of the operator's network – access, multi-access edge, standard services, custom services, transport, management – and possibly from business-to-business partners (indicated by the bottom row of icons in Figure 2).

Capture and aggregation

This layer aggregates pieces of

FIGURE 2 Cross-domain analytics architecture



event information into a complete description of an event, enriching the data with relevant user and service information. Much of the existing business logic (after detecting and filtering out known patterns) can be executed here, significantly reducing the need for processing capacity higher up in the stack and thereby improving the cost-to-performance ratio.

Storage

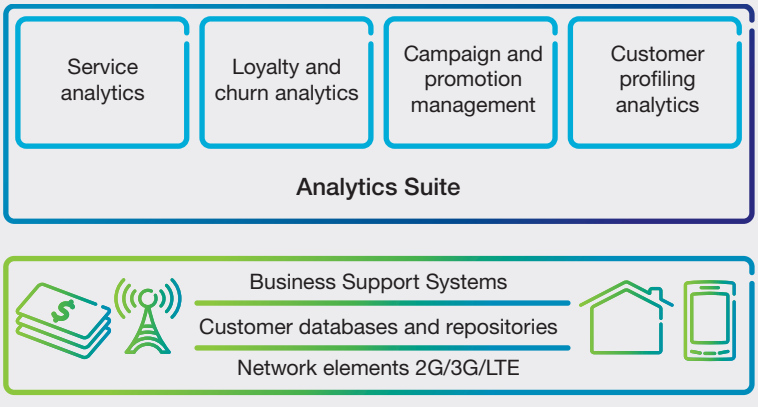
A substantial number of recent developments in analytics have taken place in the area of storage. The main focus of these developments has been on methodologies for handling, storing and manipulating data in order to keep costs down and performance up. Ericsson contributes its telecommunications experience to facilitate linear scaling and maximum flexibility for both traditional reporting and future analysis activities.

The cost of disk space can make up a substantial proportion of the total spending for an analytics system. Column databases, which radically increase query-execution performance and reduce storage space, are just one example of the technology improvements that contribute to reduced cost.

On a different track – sometimes referred to as part of the NoSQL trend – open-source initiatives such as the Hadoop framework and the accompanying Hadoop distributed file system are being used to explore other ways of addressing the issue of storage. Typically, these initiatives improve processing performance by enabling massive scale-out possibilities, where processing and storage load is shared across a large number of servers.

A combination of the two approaches seems to be the most beneficial solution: using the Hadoop approach in the lower levels where scale-out properties can be best exploited, and feeding into structured databases for the dashboard, report-generation and exploratory analysis work in the higher layers.

FIGURE 3 The layered approach to analytics



Stream processing

Real-time data-processing becomes possible as systems and data move from being costly, isolated and offline to being affordable, linked and online. This, in turn, leads to improved monitoring of services and faster identification of pattern changes. This enables immediate and accurate decision-making, which is key for modern telecommunications operators.

In particular, real-time processing can be used to merge streams and use time-based filtering logic to find cases where transactions take an unusually long time, for example. Complex event processing (CEP) systems, as they are often known, tend to be high-performance systems with limited intermediate storage, meaning that data needed on a long-term basis must be sent to storage. There is thus a need for communication between stream processing and data storage; the CEP engine usually needs reference information from the long-term storage area so that it can assess the events that occur.

For the storage and stream processing layers, Ericsson relies on qualified partners, as these technologies have applications in a wide range of industries and are outside the scope of Ericsson's core technologies.

Query engine

This level of the system contains the business logic – the rules and processing components that turn data sets and filtered events into summaries that can be presented on dashboards and tabulated in reports. The need for flexibility here is high: changing the parameters of a KPI, for example, should be controlled by the enterprise (in this case, the telco). In contrast, operators who want to know what is going on in their networks and with their customers, but who lack a budget for highly qualified analysts, need a vendor with a portfolio of ready-made analytics to provide clear answers to relevant questions.

This is one of the key benefits of the Ericsson Analytics Suite approach: it provides a comprehensive set of pre-integrated solutions to a number of significant operator issues such as churn prediction and campaign analytics. **Figure 3** shows the ever-expanding set of analytics applications supporting operators' current and future needs.

The query engine enables ad hoc queries to efficiently and intuitively support the search for new users through an analyst's ability to find undetected patterns. Top performance is essential in this area, as the analyst's ability to be creative and test ideas is reduced when processing takes several hours to complete.

In this layer, Ericsson's telco experience is a major differentiator, as domain knowledge is essential to be able to phrase the right questions and to know what information must be retrieved to answer those questions and how to organize queries for maximum effectiveness.

Visualization

Flawless analytics results are of little use if they cannot be converted into human understanding and corresponding action. Effective tools to display bar charts, trend indicators, cluster diagrams or any other kind of visual representation are absolutely vital to unlock the real potential of knowledge hidden deep in the data.

This is a rapidly evolving area, where high-performance, flexible tools are required. Users should be able to adjust their display to suit the task at hand; futile attempts at accurate guesswork from an engineer's desk far from the operational realities is unlikely to provide the solutions needed in the flux of day-to-day network operations.

Exposure

Finally, analytics results can also be used in other systems, made available as database tables or possibly as online services. User profiles and call-success data are simple examples of information that can be made available to a customer care workstation in real time.

Aggregated mobility data that is made available for traffic monitoring systems to assess traffic congestion is a somewhat more complex usage scenario. In this layer, telco expertise and experience are vital in order to interpret and combine data correctly.

Use cases

By using the data in the OSS/BSS domain efficiently, an operator can improve the relevance and efficiency of decision-making in many disciplines, ranging from product management, marketing and sales to service management and customer care.

The following use case shows how product management and customer care areas can use OSS/BSS data to build customer profiles that provide more efficient and relevant customer care, maximizing customer satisfaction, targeting and retention.

Customer profiling is centered on a customer profile card, which is a multi-dimensional view of a user that can be utilized for several different business processes.

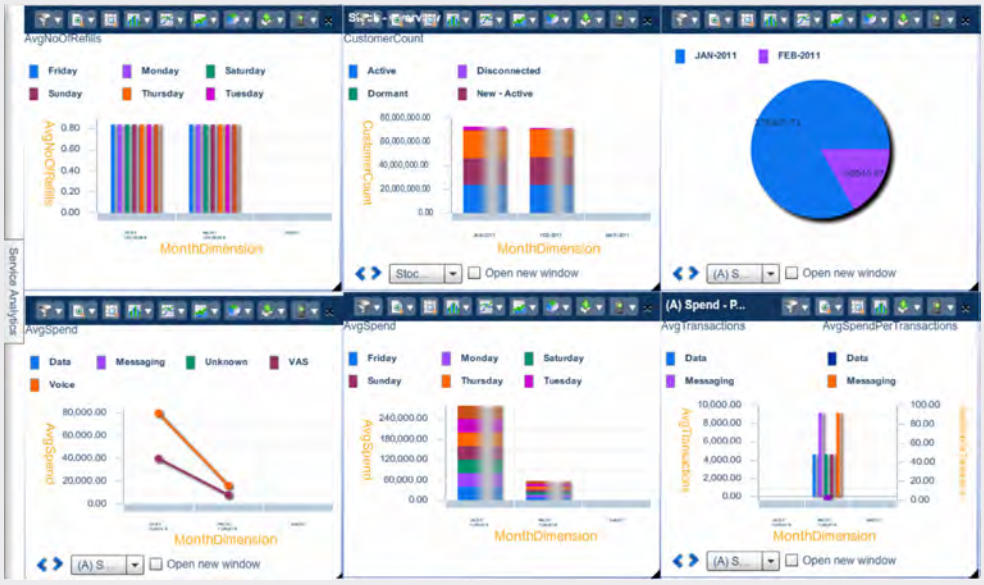
The type of information held in a profile card includes:

- demographic data;
- behavioral characteristics/segments;
- device(s) details;
- broadband profile;
- social-network score;
- churn score; and
- customer experience for each product.

The above categories can be further broken down by category-specific attributes. Some examples of attributes for the broadband profile are:

- browsing profile – such as preferred browser and average session duration;
- news profile – including frequency and most frequently used news services;
- download profile – such as device

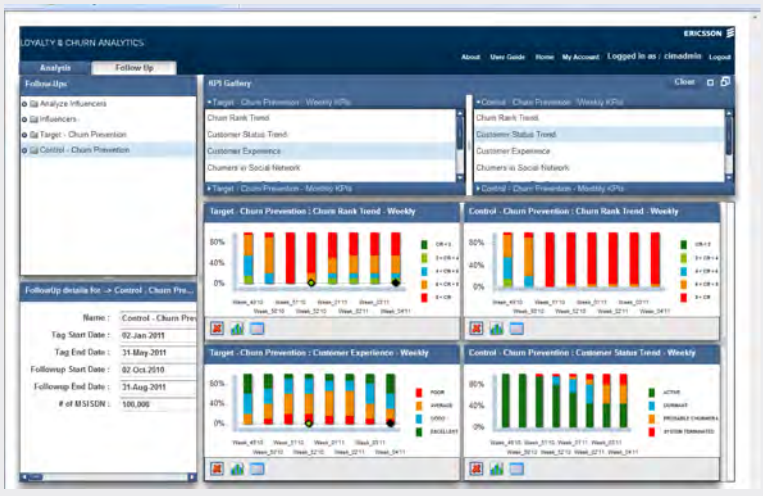
FIGURE 4 **A typical analytics dashboard**



BOX C

Profile cards are updated continuously as users consume operator services. If a user's behavior changes due to a change in income, for example, the profile card will change accordingly. Consequently, a profile card supplies the operator with accurate user information, improving the relevance of operator interactions with the user. This builds loyalty and gives users a greater sense of intimacy with their service provider.

FIGURE 5 Effect of anti-churn campaign to targeted set of users compared with churn score for a control set of users to whom the campaign was not offered.



- updates, content including music or movies, volume and ToD; and
- shopping pattern – such as shopping services visited.

By examining the values held in a subscriber's profile card, customer care representatives can, for example, gain an understanding of how users experience the services they subscribe to, and evaluate how valuable a specific customer is to the operator and whether a user is at risk of churning due to poor service experience.

Using profile cards enables targeted campaigns and advertising based on the profile card scores. For example, a specific user profile might indicate that the user is a strong influencer in their social network and is interested in tennis. Based on this information, a special advertisement that includes tickets to a local tennis event can be sent to that user. This user is more likely to spread this information further within their social network than a user with a low social-network score.

Operators can use the churn score to target potential churners with specific retention campaigns to try to reduce the risk of these users churning. Analytics Suite enables operators to identify potential churners based on:

- experience – the user's satisfaction level based on their experience of the operator's services;
- social-network score – taking into consideration the churn score of influential people in the user's network (indicating whether they have a high churn score or have churned recently);
- account status, including prepaid balance information, postpaid bill data and pattern tracking information;
- refill and bill-payment data highlighting changes in patterns, indicating an increased churn risk; and
- changes in usage patterns.

This data is used to set churn scores that can be used by product management, marketing and sales to effectively target retention activities and campaigns to specific users. Churn scores are updated continuously to provide simple benchmarks for monitoring the effectiveness of user-retention activities. By simply comparing the before and after values, operators can estimate the effectiveness of each activity.

Challenges

The first – and possibly most obvious – challenge is raw performance. Today's technology provides enough capacity to handle both current data volumes and volumes predicted for the immediate future. However, to meet the volume demands envisioned in the long term, an additional level of cutting-edge technology is needed. This is one of the drivers for Ericsson's significant investment in data-management research.

However, the capacity increase can't be allowed to interfere with other vital characteristics. Thus, the second challenge is how to reconcile capacity with the flexibility and convenience that the next generation of telco analysts will require. Ericsson's solution to this is to ensure that logic and presentation can be easily created and modified using flexible and powerful design tools.

Finally, the third challenge will be how to balance the technological possibilities with the ethical and legal limitations imposed by business practice and emerging legislation. Monitoring and contributing to EU and other legislation is of prime importance, as these laws will help to outline the requirements for consumer integrity protection functionality such as opt-in mechanisms. Additionally, Ericsson conducts consumer research to ensure that the systems can be adapted to fit not just formal requirements, but also to meet consumers' expectations on how, when and for what purposes their data should be used.

Conclusion

Ericsson already has the technology and products in place to solve a wide range of telco analytics challenges, combining the capabilities of network, service and user analytics. The aim is for the company to evolve these capabilities continuously, in dialog with customers to provide the best, most cost-effective and most adaptable telco analytics solution on the market.

Ulf Olsson

has a background in software architecture for distributed military Command and Control systems. He joined Ericsson in 1996, working mainly with architecture issues concerning packet-based systems, such as packet PDC, GPRS, UMTS and the CDMA2000 packet core network. He is now senior expert with Business Unit Multimedia, Systems Management group. He focuses on data management, business support systems, IMS and developer-oriented issues. He holds an M.Sc. in engineering physics from the Royal Institute of Technology (KTH), Stockholm.

Jaco Fourie

began his career in data communication, specializing in the X-Series Standards. He joined MTN South Africa in 1994 with focus on value-added services and charging. In 1999, he joined Ericsson, working with product management for charging systems. He is now senior expert and head of Business Development and Strategy with Business Unit Multimedia, Solution Area Revenue Management, where he focuses on the future evolution of the Ericsson BSS portfolio to suit the demands the Networked Society will put on next-generation BSS.