Pipeline
Technology for Service Providers.
www.pipelinepub.com Volume 11, Issue 3

The Unstructured Data Structure

By Nancee Ruzicka

In a 2013 report, FICO cautions “80% of big data is unstructured.” If that’s the case, then why are we spending so much time and energy on structured data? The answer, it seems, is that the data currently most valuable to service providers is structured. Service providers collect data from the network for important tasks like billing, revenue assurance, policy management, service assurance, and fulfillment. For the majority of service provider operations, structured data captured from customer transactions and network elements reveals everything they need to know. Structured data is continuously collected and saved, so data bases and warehouses are readily available to be searched and analyzed to uncover trends and spot opportunities. In fact, service providers are required by law to capture and store customer transaction data and they often maintain many years’ worth of data for each customer.

While revenue assurance and fraud management rely almost exclusively on transaction data and network factors to analyze user behavior, find lost revenue, and determine the sources of fraudulent activity, accurately analyzing that same data to determine marketing behavior and evaluate the customer experience requires the inclusion of unstructured data. Traditional solution providers, such as cVidya, that use customer transaction and network data for fraud management and revenue assurance, are now expanding into the marketing analytics space. This new line of business builds upon existing solutions; however, these vendors also recognize that access to the unstructured data relating to customer behavior and attitudes is critical to service providers as they define CEM (Customer Experience Management) strategies to reduce churn and increase up sell. Other vendors, such as AsiaInfo, are providing "contextual awareness" to integrate both static and dynamic data sources to provide service providers with valuable customer insights.

Incorporating unstructured data is producing results. Several service providers report success in identifying customers likely to churn by using advanced analytics and integration with unstructured data. One of these service provider identified 75% of the customers that were likely to churn by targeting only a specific 10% of its total customers as indicated by the analysis. That segment is now the focus of offers and support designed to improve their unique customer experience.

Complex events require complex solutions

One of the most difficult challenges service providers face is data volume. Service providers collect massive amounts of data and that in and of itself can be one of the biggest challenges facing them today. Vodafone India, for example, collects 32 billion records per day and the largest service providers regularly collect 40 million records *per second!*

Much of the data that is readily available to service providers is structured. Data about billions of transactions, millions of customers, thousands of devices and hundreds of products are all captured or calculated in a consistent manner. However, what service

providers are missing, and what is becoming valuable to marketing and product management functions, is the unstructured data regarding perception, problems, and desires being expressed by customers during calls to customer care, online during chat sessions, or in online and social networking communities like Facebook and Twitter. There are also additional clues that can be found in videos and promoter index scoring that could be valuable to a service provider if properly decoded and analyzed.

Processing billions of user transactions, correlating and distributing that data to multiple OSS and BSS solutions, ensuring partner settlements, eliminating fraud, and improving the customer experience are all drivers for implementing big data management solutions like Hadoop. But there's more to it than data storage and computing power. Managing the volume of events, processing customer transaction records, and correlating unstructured data requires complex event processing that collects and analyzes every scrap of data available via network elements, probes, and partner elements.

Traditionally, the only data link between the customer and the network has been the data record (xDR) used for billing. However there is much more customer and network data that, when applied to correlated unstructured data, can be used to identify quality of service (QoS) issues, understand churn propensity, track products, support sales efforts, make customer care more responsive, and improve the overall customer experience. Analytics-enriched complex event processing, predictive modeling using multi-dimensional and non-linear analysis that captures millions of values from dozens of sources, delivers customer insight benefits beyond what is possible given existing standalone mediation systems. Complex event processing is the key to correlating disparate structured and unstructured data sources in order to identify opportunities for generating new revenue, recovering revenue that is currently being lost, and improving the overall customer experience.

Vendors of analytics tools and complex event processing platforms are increasingly acquiring and partnering with providers of tools that analyze high volumes of unstructured data and deliver a user-friendly presentation of results. Those applications, when integrated with the tools that analyze structured data collected by all service providers, create a powerful platform for processing, analyzing and presenting a more complete picture of customer experience and behavior.

So what?

Increased mobility, LTE, connected devices, cloud applications, and video are driving data volumes ever higher with no end in sight. Correlating unstructured data with structured customer usage records and network events, processing those events, and analyzing all of these behaviors in real time creates a daunting task that can only be accomplished using sophisticated event processing, analytics, and automation. Solutions like Hadoop, while valuable, address only the volume of transactions and processing demands.

When using analytics to understand customer behavior and buying patterns, it is possible to monitor and even understand the activities of any given customer; but it is another problem entirely to take targeted action. One example is delivering a partner offer based on customer location. For instance, when location and customer preference data indicates that a customer is near a coffee shop, numerous OSS/BSS solutions must be tasked in order to present an offer to the customer and then provision, activate, and set up billing for the offer in near real time.

And once the transaction is completed, the offer must be turned off and settlement completed with the partners. There are numerous tasks that must be completed before, during, and after the transaction to ensure success. A glitch in any one creates problems, costs money, and diminishes the experience.

Until the necessary OSS/BSS solutions are well integrated and operating processes are automated to the extent that each can be quickly configured to create and provision a new offer that is dependent on location or any other dynamic variable, service providers will not be comfortable offering a full menu of digital services.

Another complaint that service providers have when it comes to analytics is the quality of the data available for analysis. With data stored in hundreds of OSS/BSS, customer, product and corporate databases, it is difficult to make current, valid data available to analytics platforms and applications. Bad data leads to unreliable results, and service providers cannot afford to be wrong.

As attractive as it would be to replace old systems with new, in reality service providers have existing processes and OSS/BSS solutions that are generally well suited for their intended purpose. Add to that the risk and failure rates of complex big data and analytics projects and service providers are hesitant to make abrupt changes to operational systems and applications.

There are, however, data alignment tools available that search existing sources to

capture current views of customer, revenue and operational data that are then validated and aligned before being presented to the analytics engine. Tools, like the semantic search engine from Ontology, collect structured and unstructured transaction data in its original form while managing data alignment and changes to reduce upfront expense and prevent costly, time-consuming data conversion projects. The resultant data model is flexible and can be readily modified to accommodate customer, network or service changes and new data from additional sources can be added as analytics needs evolve.

There are numerous analytics platforms available in the market. However, where traditionally there would be a need for a data scientist or information engineer to develop models and manipulate data, solutions are coming to market from companies like cVidya, LavaStorm, Comptel, and WeDo Technologies that embed analytics applications in OSS/BSS to target specific operational needs like revenue management, determining churn propensity, and service assurance.

When it comes to analytics applied to the structured data required for revenue management, cost does not appear to be a primary issue. The need is so great and the return is so positive when analytics are applied to fraud management, revenue assurance, and settlements that service providers rarely question the cost of the solution. However, when asked to provide funding for analytics efforts that include unstructured data used for product and offer management, the returns are less visible and the risk to customer experience is greater. As a result, service providers are approaching those efforts more cautiously.

And lest we forget, the application of analytics to data of any kind is not just a technical challenge that can be thrown over the wall to IT. Every service provider has access to the same technology. To deliver value, business and executive leadership must be engaged to ensure that the tools are solving business problems and delivering business results. That requires more than budget; it requires time and resource commitment from subject matter experts, executives and managers across the business. Those individuals must stay involved to ensure that the resultant deployment of analytics and big data meets their needs.