



Wednesday, July 23, 2014

## Looking for the light

### Data mining provides a promising tool in the search for a winning data strategy

BY PHIL CARSON  
SEPTEMBER/OCTOBER 2001

As a scientist at NASA's Jet Propulsion Lab in Pasadena, Calif., Usama Fayyad used data mining techniques a decade ago to discern the difference between distant stars and galaxies and to examine volcanoes on Venus.

Today, as CEO of Kirkland, Wash.-based digiMine, he focuses on matters closer to home. Although a good deal of digiMine's data mining work is applied to consumer behavior on the Web, the firm has launched a Wireless Business Intelligence service for carriers and content providers.

The goal of data mining is to improve upon traditional customer relationship management techniques to predict consumer behavior in voice and wireless data use. The process calls for using powerful, computer-driven algorithms to extract useful information from voluminous records of actual mobile data sessions. Whereas CRM often has served "heat of the moment" demands in customer service call centers, data mining is primarily a strategic tool. The possibilities for the mobile data sector include nothing less than guiding a nascent industry to fulfill its promise.

The confluence of several factors is creating an opportunity for data mining and its wireless patrons. The advent of next-generation networks yields more varied data on how subscribers use the mobile Internet. The sophistication of data mining has grown to take advantage of the myriad variables on usage generated by new networks, devices and services. And in tough economic times exacerbated by the slow uptake of mobile data, carriers more than ever need new tools to shape wireless' next growth opportunity.

### Desperately Seeking... Guidance

Just how important is data mining to the carriers, how well developed is their strategy and how will they use it? Detailed answers are not forthcoming.

Dan Wilinsky, director of media relations at Sprint PCS, only hints at the answers--and the issues--when he responds: "It's too early for Sprint PCS to give its perspective. Our competitors would really like to understand our strategy in this space, so it doesn't benefit us at this time to discuss it."

Alexa Graf, national media relations manager for AT&T Wireless, offers a bit more, though she is careful to stick to generalities. AT&T Wireless performs its own "tracking of customer usage patterns" on both voice and data networks, she says, for completing transactions, billing and measuring network performance. On the carrier's new GPRS network it will track aggregate information on the types of data applications customers are using. "Understanding the market ... is critical to our mobile data strategy," she says. "As customers begin to use wireless data applications, the trends they set will help determine the content we provide. The key is to understand customer behavior to be competitive in the marketplace."

### Developing New Tools

How will the carriers themselves find their way through the currently amorphous and rather flat mobile data service offerings to create a rich and exciting experience? Many believe that data mining will provide one powerful tool in that quest.

Any discussion of data mining ought to begin where the process itself begins: with the framing of the business questions that need answering.

"You can't collect the data unless you know what you want to collect," says Andy Fessel, vice president for wireless data intelligence at Telephia Inc.,

which partners with digiMine on the Wireless Business Intelligence service. "So there is a planning component to what you're going to measure that has to be thought out ahead of time."

He begins with various questions focused on teasing out patterns of actual usage by subscribers. "As a carrier, it is beholden upon you to understand your customers," he says. "'CRM' basically means holding onto who you've already got. Maybe we need to develop a new term, like CDM, Customer Development Management. We need to grow this service. We need to understand how they're using data and how we can build on that."

What are the patterns of usage of wireless data services? What time of day attracts the high-end users? Are consumers making short visits to one or two sites, then they're out because they're busy? Or are they spending a lot of time, looking for information or playing games or communicating? What are the navigational issues?

"Right now carriers are working hard to present what they think is a good experience, but it's essential to track consumers as they go through that experience," Fessel says. "Do they have to click through a lot of content to get where they want to go? Should something be moved up the access list? Clearly there are huge opportunities around consumers' different needs."

Understanding consumers' needs in the mobile data space helps formulate essential business questions that address carriers' needs, Fessel adds.

"How do you segment and understand your consumer groups so you can sell them into your content providers, potential advertisers, as well as understand them in a marketing sense?"

Beyond using data mining to develop and market future mobile data offerings, Fessel argues that the business intelligence being gathered—naturally, he hopes, through digiMine and Telephia's WBI—has additional value to a carrier's alliance partners or clients.

"I'd say the information generated at the carrier level carries with it the necessity of communicating it in some summary fashion/refined form to content partners, advertisers, m-commerce partners, infrastructure suppliers, the investment community as well as comparing and contrasting this information against all sectors of the marketplace, be it domestic or international."

### **Crunching Big Data**

Having defined the essential questions on consumer behavior and the resulting business issues, it's only natural to turn to the nuts and bolts of data mining itself.

Talking to digiMine's Fayyad, you first have to ask how he traveled from discerning answers to astronomical puzzles to teasing out consumer behavior on the Web or wireless networks.

"Data is pretty much data, at the abstract level, whether you are trying to infer the identity of celestial objects or predicting the behavior of wireless customers," he says. "The problems are equally interesting. My theory is: Data is everywhere and our capacity as human beings to deal with it is very limited. So anywhere you have data you have a strong need for data mining."

You also find out that between the Jet Propulsion Lab and digiMine Fayyad spent five years with Microsoft, figuring out how to integrate data mining algorithms with SQL Server to make it more responsive to users.

"It became apparent to me at JPL that there are problems that scientists spend years trying to solve that data mining could solve in a matter of weeks, perhaps days, even hours," Fayyad recalls. "Think what it could do in the corporate world! In my last year at Microsoft I noticed that although corporations know how to store data, they don't know how to store data correctly to manage it for analysis and decision support."

"To provide the solutions businesses want to see, you have to have the infrastructure," Fayyad says. "You have to build the data warehouses, the data mining algorithms working on top of the warehouses, and presenting this work back to humans in a form they can understand. That was the missing piece. That's why we started digiMine. The idea is to say, 'Look. We will capture the data, manage it, warehouse it and deliver the analytics and reporting on top of that.'"

That's where digiMine's partnership with Telephia comes in. Telephia provides a suite of business intelligence services for wireless carriers, knows the industry's language and concerns and can provide the industry-specific analytics that cap off digiMine's data mining work.

At the heart of that work lies algorithms. Fayyad, who has a Ph.D. in

computer engineering, warms to his own definition, which is deceptively simple. "An algorithm is a sequence of instructions, applied on a large scale, with criteria in the algorithm that says, 'Aha! These correlations might be an interesting possibility, so let's explore from there.' So it's a simple mechanical procedure with some criteria to tell it where to look next. Then you have to simplify the model to generate explanations we can understand.

"The sheer amount of data overwhelms, and smaller samples don't generate useful conclusions," he says. "So the question is: Can we train a tool to look for what we want to know?"

Algorithms run through a carrier's usage data on potentially millions of phone calls or data sessions, encompassing huge numbers of variables, looking for correlations, patterns, regions that appear similar to other regions, hidden relationships, predictive models. Common business questions to which data mining can provide answers include, at one end: Are there usage patterns reflective of potential churners? What are the most-used services? At the higher end: Are there common subscriber clusters to suggest market segmentation? How do subscribers use a service and, based on their use of current options, what others might appeal to them?

In exploring patterns of human behavior, Fayyad soon discovered that past behavior is the best predictor of future behavior and more accurate than demographic or psychographic information. Thus the advantage to data mining actual usage records. And as for voice vs. data, the latter clearly captures his imagination.

"Voice use is pretty boring because all you know is when the call happened, from this point to this point, and its duration," he says. "With data use, you find richer things, like is this customer trying to buy tickets, find sports scores, checking news, stock quotes? What do they find useful and compelling? The carriers offer all these services but they have no idea where users are spending their time."

### **Mining's Business Context**

Data mining has been a fundamental business tool for nearly a decade, though its use is not yet pervasive, according to "Turning Data Into Dollars," a new report from Forrester Research. About 23 percent of telecommunications firms overall have completed implementing some form of it—matching the average among the "Global 3500," the world's largest corporations with more than \$1 billion in annual revenue.

Joe Butt, senior analyst at Forrester and author of "Turning Data Into Dollars," found, however, that data mining efforts often fall short due to four "C"s: complexity, complacency, coordination and cost. In a nutshell, the mining of "big data"—such as that generated by major wireless carriers tracking consumer usage patterns—requires costly, complex infrastructure run by highly qualified personnel; it has not attracted the attention of business executives who must formulate the questions data can answer, and executives and IT managers have not coordinated their efforts.

"Business people have to lead the discussion, even though there's obviously a heavy IT component," Butt says.

Most data mining is directed at forecasting behavior in a highly segmented way, he continues. "The segmentation of product/service offerings is a crucial business advantage." That dovetails neatly with carriers' and content providers' recognition that the mobile data experience will have to cater to a vast diversity of subscribers and their disparate needs.

Butt rattles off the names of many companies that provide outsourced data mining services, with a number of them—MicroStrategies, Lightbridge, Experian and digiMine—offering experience and expertise in the wireless field. The size of this industry segment is difficult to measure, however, because most of these firms work in both the wired and wireless domains.

"DigiMine is the only firm we know of that has said, 'Hey, we're going to make a business out of this,'" Butt says. "Even though they're doing mostly Web-generated data, they're saying they can crunch any data you want."

WBI launched in June, yet digiMine is not yet crowing about the response. "We have signed a few customers, including a major wireless carrier," says Joel Sider, spokesman for digiMine. "That's pretty much all I can tell you."

### **Privacy Hinges on Identity**

As for the idea that firms are out there, offering to "crunch any data you want," makes some uneasy. For David Sobel, general counsel at the Electronic Privacy Information Center, the issues surrounding data mining in the mobile data space aren't much different from other technologies, wireless or otherwise. He is adamant that consumers have a clearly

presented opportunity to opt-in to being located or tracked in other ways; that carriers avoid the linkage of behavioral data with an identifiable individual and that, regardless of safeguards, the mere existence of data on an individual's behavior is a violation of privacy.

"Once it's collected, regardless of whatever assurances are provided, it's sitting there waiting for law enforcement to retrieve it with a subpoena, for attorneys in civil litigation who want to know where you've been on a particular day. So once it exists, from our perspective, you've got a privacy problem. If information is anonymous, collected in the aggregate, or applied in the aggregate, you've eliminated the problem."

AT&T Wireless' Graf emphasizes her carrier's respect for privacy. "There has been a strongly negative reaction from the public as to what kind of information and what level of detail about behavior patterns is being tracked," she says. "So following customer usage patterns on a more general and aggregate level allows us to avoid the pitfalls of upsetting our customers with invasive marketing."

"The basic issue is giving the individual control over information about themselves," Sobel concludes. "This organization tends to think of privacy as a basic human right."

Despite digiMine, AT&T Wireless and Telephia's assurances that they intend to use data mining only in aggregate, Fayyad acknowledges that the data on individuals is there, ultimately allowing market segmentation to extend down to the individual level. "That's possible now," he says, brushing off privacy concerns. "How it's used is always going to be a business decision. It's not a technology issue. The question becomes, 'Can we get to a level where that targeting is so accurate, you'll pay attention?' Then it has real value."

At least that is Fayyad's vision of data mining's ultimate value to subscribers. As for its value to carriers, that depends on how they formulate their business questions and how they use resulting analytics to shape the mobile data industry. For now, that's a competitive secret. At this juncture, you may as well try to decipher the stars.