

Ideas & Trends; Enron Offers An Unlikely Boost To E-Mail Surveillance

By GINA KOLATA (NYT) 1328 words

Published: May 22, 2005

AS an object of modern surveillance, e-mail is both reassuring and troubling. It is a potential treasure trove for investigators monitoring suspected terrorists and other criminals, but it also creates the potential for abuse, by giving businesses and government agencies an efficient means of monitoring the attitudes and activities of employees and citizens.

Now the science of e-mail tracking and analysis has been given a unlikely boost by a bitter chapter in the history of corporate malfeasance -- the Enron scandal.

In 2003, the Federal Energy Regulatory Commission posted the company's e-mail on its Web site, about 1.5 million messages. After duplicates were weeded out, a half-million e-mails were left from about 150 accounts, including those of the company's top executives. Most were sent from 1999 to 2001, a period when Enron executives were manipulating financial data, making false public statements, engaging in insider trading, and the company was coming under scrutiny by regulators.

Because of privacy concerns, large e-mail collections had not previously been made publicly available, so this marked the first time scientists had a sizable e-mail network to experiment with.

"While it's sad for the people at Enron that this happened, it's a gold mine for researchers," said Dr. David Skillicorn, a computer scientist at Queen's University in Canada.

Scientists had long theorized that tracking the e-mailing and word usage patterns within a group over time -- without ever actually reading a single e-mail -- could reveal a lot about what that group was up to. The Enron material gave Mr. Skillicorn's group and a handful of others a chance to test that theory, by seeing, first of all, if they could spot sudden changes.

For example, would they be able to find the moment when someone's memos, which were routinely read by a long list of people who never responded, suddenly began generating private responses from some recipients? Could they spot when a new person entered a communications chain, or if old ones were suddenly shut out, and correlate it with something significant?

There may be commercial uses for the same techniques. For example, they may enable advertisers to do word searches on individual e-mail accounts and direct pitches based on word frequency.

"Will you let your e-mail be mined so some car dealer can send information to you on car deals because you are talking to your friends about cars?" asks Dr. Michael Berry, a computer scientist at the University of Tennessee who has been analyzing the data.

Working with the Enron e-mail messages, about a half-dozen research groups can report that after just a few months of study they have already learned that they can glean telling information and are refining their ability to sort and analyze it.

Dr. Kathleen Carley, a professor of computer science at Carnegie Mellon University, has been trying to figure out who were the important people at Enron by the patterns of who e-mailed whom, and when and whether these people began changing their e-mail communications when the company was being investigated.

Companies have organizational charts, but they reveal little about how things really work, Dr. Carley said. Companies actually operate through informal networks, which can be revealed by analyzing "who spends time talking to whom, who are the power brokers, who are the hidden individuals who have to know what's going on," she said.

With the Enron data, Dr. Carley continued, "what you see is that prior to the investigation there is this surge in activity among the people at the top of the corporate ladder." But she adds, "as soon as the investigation starts, they stop communicating with each other and start communicating with lawyers." It showed, she says, "that they were becoming very nervous."

The analyses also found someone so junior she did not show up on organization charts but who, whichever way the e-mail data was mined, "shows up as a person of interest," Dr. Skillicorn said, in the language of intelligence analysts. In the investigation of a terror network, pinpointing such a person could be of enormous significance.

Dr. Berry said the e-mail traffic patterns tracked major events, like the manipulation of California energy prices. "We could see how things built up right before the bankruptcy," he said.

There were e-mail surges with each crisis, pointing to a problem that was consuming Enron employees. And in each crisis, there were features of certain e-mail messages -- word choices, routing patterns -- that allowed the computer scientists to isolate them from the morass of irrelevant personal or business messages.

One thing that didn't show up when the researchers screened for changes in word use was guardedness, said Dr. Skillicorn, a failure that was revealing in itself. Ordinarily, he said, when people are being deceptive they are more self-conscious, and their word use becomes simpler, as though they are trying too hard to sound natural.

But that apparently never occurred at Enron because its employees remained unconcerned while they engaged in illegal activity. "It wasn't a case of keeping a low profile," Dr. Skillicorn said. "They didn't worry about the story they were telling."

The scientists who are studying the Enron data said they assumed intelligence agencies are doing similar classified analyses on international e-mail traffic. Since World War II, a five-nation consortium of the United States, Canada, Britain, Australia and New Zealand have cooperated in a vast communications collection and analysis program called Echelon, for example, one that has assumed increasing importance since the terror attacks of Sept. 11, 2001.

No one in the unclassified world knows precisely what is being done with the Echelon data. But, Dr. Berry said, surveillance in the civilian world could one day have troubling consequences. It could allow companies, without ever actually infringing on e-mail conversations, to track employee attitudes and activities closely and easily.

"They can monitor discussions without actually isolating individuals," Dr. Berry said. "They can assess morale. If they make a cut in salaries, how long does the unhappiness go on? You could track topics and get a sense of how people are responding to policies and flag potential hot spots." Or, he said, managers might be able to learn which people have too much time on their hands.

And, as Dr. Skillicorn notes, if you try to write bland e-mail messages with hidden communications, chances are the programs will pick those out, too.

"It's clearly Orwellian," Dr. Berry said. "And I know that freaks people out."

Chart: "On Their Minds in 2001"

An analysis of word usage in 37,557 employee e-mails revealed patterns that mirrored events at Enron. The company collapsed in December of 2001.

Each dot represents one large-scale discussion of these topics among Enron e-mailers.

BUSINESS TALK -- The company's manipulation of the California energy market remained a hot topic throughout the year. The failed merger with Dynegy and ultimate bankruptcy dominated later.

SEASONAL CHITCHAT -- Talk of college football gave way to fantasy football discussions as the firm neared its end. Common terms included "longhorn," "recruiting," "oklahoma," "injury" and "qb."

Chart shows number of large-scale discussions on following topics among Enron e-mailers:

California energy

Dynegy merger/bankruptcy

Texas football

Fantasy football

Finding Patterns In Corporate Chatter

Computer scientists are analyzing about a half million Enron e-mails. Here is a map of a week's e-mail patterns in May 2001, when a new name suddenly appeared. Scientists found that this week's pattern differed greatly from others, suggesting different conversations were taking place

that might interest investigators. Next step: word analysis of these messages.

(Sources by Dr. Carey E. Priebe and Youngser Park, Johns Hopkins University)