Terror Web   08.03.06

The Internet's popularity for connecting people has a dark side. It's also become a major planning tool for terrorists. As this ScienCentral News video explains, one computer scientist is using artificial intelligence to "connect the dots" between terrorist groups.

A Portal to the Underworld

Stepped up security measures at our nation's border crossings and transportation hubs may be giving terrorists fewer places to hide. But computer scientist Hsinchun Chen says they're chatting away in cyberspace.

Video

"There is the dark side of the Web," says Chen, director of the Artificial Intelligence Lab at The University of Arizona's Eller College of Management. "The Web has been used by a lot of people hiding in the dark."

He says that similar to colleges and universities offering courses online, terrorist organizations have found a whole new channel for their ideas. "Those terrorist movements have moved to the Internet, using the Internet to recruit new members, to train them, to educate them," Chen says.

Following the September 11th attacks, Chen created the "Dark Web Portal," the largest collection of online terrorist information in the world.

It uses various artificial intelligence techniques to scan the Web for possible terrorist content, and then analyze and report on that activity. The program automatically sifts through terrorist and extremist Web pages, forum pages, and attachments.

Currently, Chen says, they have almost 2 terabytes of data. Bytes are measures of computer data, and a terabyte is about a trillion bytes, or 1024 gigabytes. Many household computers today have a capacity of a few dozen gigabytes.
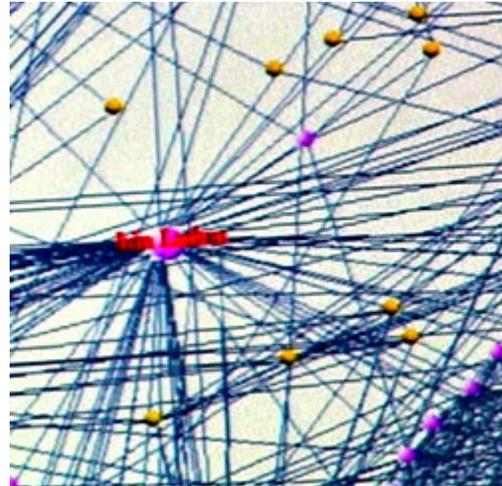
The Dark Web Portal does a content analysis by reading Web sites and online forums and tracking what ideas are taking off. "This is really like disease surveillance," Chen says. "Infectious disease that spreads, you have to catch it when it happens."

The system also does a link analysis by figuring out which Web sites connect to each other. While research has already been done in this area, much of it has been

piecemeal, Chen says, and this is the first project to collect such a large amount of data with such fast processing speeds. "Instead of spending months studying a collection of thousands of documents within a matter of minutes we can analyze hundreds of thousands, even millions of documents," he says.

Connections to bin Laden

Chen can use the system to illustrate clusters of terrorist groups. The team was able to visually map the network of September 11th suspects, demonstrating how they can pinpoint the ringleaders in terrorism. The graphic looks like a complicated Web of connections consisting of hubs and rings with lines radiating out of the each hub. "Using the mapping technique we were able to show that bin Laden is in the center and there are four major groups of people affiliated with him," he says.

This kind of information could be useful for intelligence analysts and people who study terrorism. "The power in collecting and visualizing [this information]," Chen says, "is you can shorten the monitoring time for analysis."

Chen says that his program scans sites that are completely open to the public, and he does not have a security clearance for top secret information. Some terrorist and extremist information, he says, can simply be found within discussion boards on popular sites.

However, some computer privacy experts are concerned by the system's broad reach and say that since this technology is used by the government, checks should be in place.

Lance Hoffman is a professor at the George Washington University, and has worked in computer security and privacy for the last 30 years. "The system produces potential representations of networks," he says. "These may or may not be accurate ... This is new groundbreaking research. You're not going to get things right all the time."

Hoffman says that this program should be just one tool in the fight against terrorism, and that the government should be aware of concerns for civil liberties. "We can't do the research first and ask questions later. I would be much more comfortable if the government, without giving classified details, could show how this would all fit together," he says.

Chen's code is open-sourced, meaning that his research is publicly available to computer programmers. While Chen stresses that his main area of research is developing better

methods of data collection, he says that the government uses parts of his technology for online surveillance. He agrees that checks should be in place to protect civil liberties and says his ongoing research is improving the accuracy of the system.

Chen's work was featured in the July 2006 edition of *Discover* magazine and published in the IEEE International Conference on Intelligence and Security Thematics from 2003 to 2006. His research is funded by the National Science Foundation, the Library of Congress, and several intelligence agencies of the U.S. federal government.