

# Finding patterns of importance in a deluge of data

9 January 2007

---

Dartmouth engineers George Cybenko and Vincent Berk think that PQS, or process query systems, are the way to go to make sense of the huge volume of data we collect each day from computer network monitors, video surveillance cameras, financial transaction records, databases of email exchanges, etc. The duo present their case in a paper published in this month's *IEEE Computer*, the flagship magazine of the Institute of Electrical and Electronics Engineers' Computer Society.

"PQS closes the gap between gathering a tremendous amount of valuable data and figuring out what the data means," says Cybenko, the Dorothy and Walter Gramm Professor of Engineering at Dartmouth's Thayer School of Engineering.

PQS has been an evolving algorithmic and software framework for the past few years. Cybenko and Berk say that PQS is a useful and incredibly powerful tool to quickly analyze credit reports for ID theft, discover attacks on computer networks, and measure activity at, say, national borders, mall parking lots, or wildlife refuge areas. According to Cybenko, "PQS can do for discrete, categorical data analysis problems what classical times series analysis did for finance and control systems where the data are numerical."

It is based on the premise that sensed environments, be they computer networks, email traffic, or high-security buildings, all consist of processes with distinct states, dynamics, and observables. PQS works to detect and understand the changes or irregularities in these processes. The PQS software is easily installed with the sensor equipment to collect, monitor, and sort out a great deal of data.

"I think the most interesting application of PQS to date is in network security monitoring," says Vincent Berk, research associate and lecturer at

Dartmouth's Thayer School. "Network administrators have many options when it comes to monitoring tools, however none of them are integrated; and, while all of them produce useful information, it's often in hugely impractical quantities. PQS brings together the information, and effectively focuses on the most important issues first. To my knowledge there has not been a new software technology that is this versatile since the introduction of relational databases."

Source: Dartmouth College

APA citation: Finding patterns of importance in a deluge of data (2007, January 9) retrieved 20 September 2019 from <https://phys.org/news/2007-01-patterns-importance-deluge.html>

*This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.*