Algorithm for data-agnostic event detection from data streams

Technology #cu16056

Event detection from data streams (for example, detection of computer breaches or new distant stars) typically involves a labor-intensive preprocessing step by experts, which is slow, error-prone, and expensive. Additionally, many of these events have low intensity signals that are obscured by random noise. This technology uses “Information Annihilation” (IA), a computational method to detect events by comparing two data streams without prior data interpretation. This technology provides a fast, low-cost method of event detection from any form of data stream, including low intensity data. Such high-sensitivity signal processing has wide applications in many fields, including bioinformatics, astrophysics, communications, and computer security.

Algorithmic comparison of data streams for fast, low cost, sensitive event detection

This technology uses comparison of data streams to accomplish the difficult task of event detection without prior data interpretation. While standard stochastic signal analysis relies on expert pre-processing, this algorithm instead uses a universal metric to assess similarities between data streams. Specifically, every data stream, regardless of source, has an anti-stream of data. The IA step combines the stream and anti-stream to produce randomly generated noise. If the combined data significantly deviate from flat white noise, an event is present. Even events with low signal intensity will deviate from noise after IA, allowing for high sensitivity detection.

A prototype of the technology has been tested and was highly sensitive in a variety of detection tasks, including identification of abnormal heart activity from heart murmur recordings. This technology was also able to detect astronomical events that were missed by traditional stochastic analysis methods.

Lead Inventor:

Hod Lipson, Ph.D.

Applications:

- Signal processing for communications
• Detection of abnormal medical events from monitors (EKG, EEG, heart monitors)
• Identification of astronomical events by processing of electromagnetic radiation
• Signal detection from large-scale bioinformatics data (RNA-seq, gene expression, protein expression)
• Cyber security by detecting malicious events
• Monitoring for system failures
• Detection of environment abnormalities in factories, electricity plants, and other hazardous workspaces

Advantages:

• Applicable to any form of data stream and is useful across many fields
• High sensitivity
• Fast
• Runs in linear time, so applicable to real-time processing
• Does not require expert pre-processing
• Reduces human error in data interpretation
• Identifies low intensity events
• Can be embedded on a chip

Patent Information:

Patent Pending (US 20150242469)

Tech Ventures Reference: IR CU16056

Related Publications:


Inventors

Hod Lipson