KOJAK Link Discovery Tools

Anomaly Detection in Large Semantic Graphs

Benefit: *Completely automatic detection of anomalous threat activity in large and complex datasets. UNICORN can "find things without knowing what one is looking for." It is domain-independent, unsupervised and data-driven which eliminates bias and the need for training examples. This enables rapid application to new datasets and addresses the evolutionary aspect of threat behavior.*

Mission

Finding evidence of threat individuals and threat activity in large databases is an important but difficult problem for intelligence analysts. Data is often complex, spread over many databases and contains mostly benign activity that makes it difficult to find, focus on and visualize the important parts of the data. Moreover, threat behavior is rare and threat actors constantly evolve and adapt, therefore, looking for patterns of past behavior might miss important clues.

Solution

KOJAK UNICORN uses anomaly detection to find anomalous entities in large and complex datasets that are best represented by *semantic graphs* (a representation scheme commonly used by intelligence analysts). Anomalies can indicate threat behavior even if no particular threat pattern is known or applicable, since threat behavior is unusual and covert and often carried out by actors who try to disguise themselves but might get things wrong.

UNICORN summarizes the graph neighborhood of a node to a certain depth into a *semantic profile* and then looks for nodes with the most abnormal profiles. This method can successfully detect threat actors hidden in large datasets (see Figure 1). KOJAK Simplifier uses the same technology plus normalization and abstraction operators to simplify large semantic graphs onto their essence, allowing analysts to focus on and visualize only the important parts of the data.

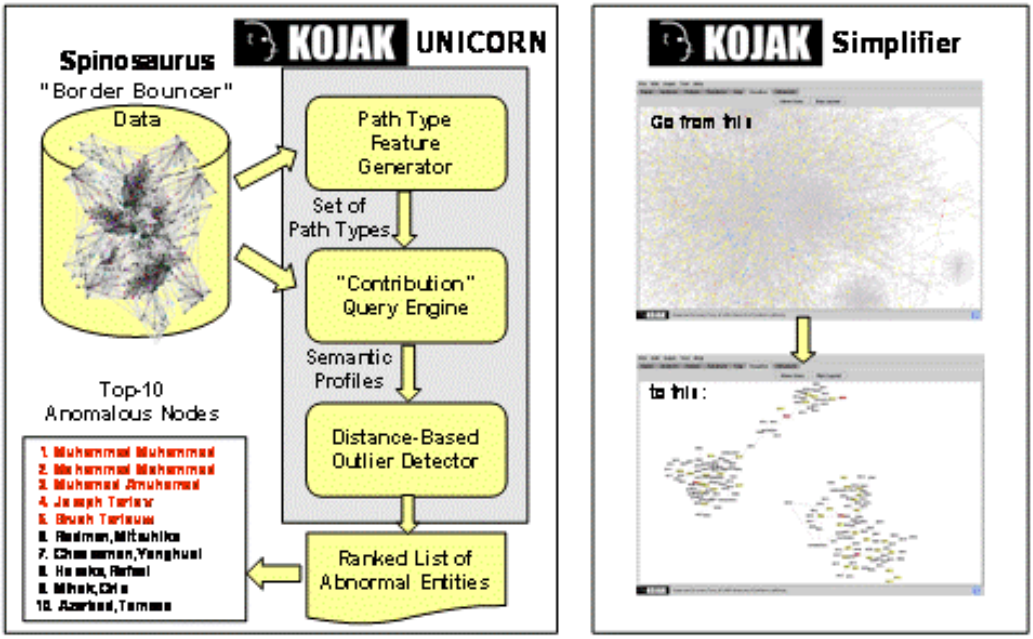

Figure 1: *Left: KOJAK UNICORN automatically discovers threat individuals hidden in the large synthetic Spinosaurus dataset develped by PNNL. Right: KOJAK Simplifier extracts the essence of a complex simulated dataset in the domain of Russian organized crime.*

Funded by: *DHS; prior funding for the development of KOJAK was provided by DARPA, AFRL, DHS and ITIC.*

Early Development — **Lab Prototype** — **Commercial Product**

For more information, contact:
Hans Chalupsky, (310) 448-8745, hans@isi.edu
(www.isi.edu/isd/LOOM/kojak)