Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by **Battelle** *Since 1965*

Green Hornet
Visualization for Large Scale Graphs

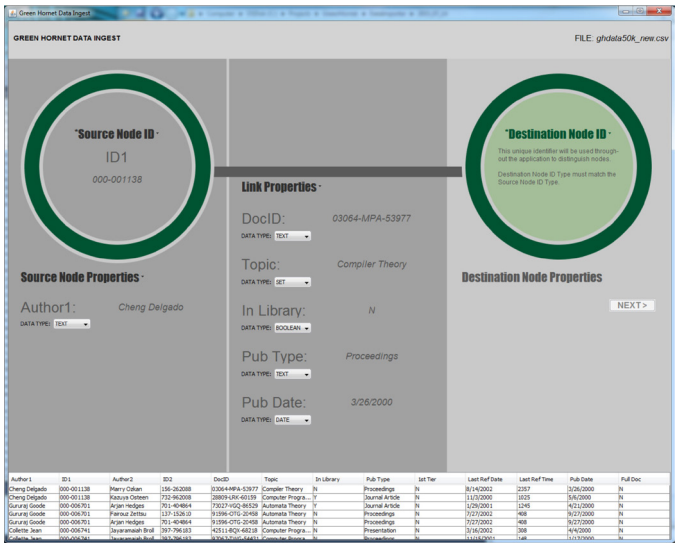
Green Hornet is a visual analytics tool for large graph exploration. Developed by Pacific Northwest National Laboratory (PNNL), Green Hornet is designed to allow users to explore graphs at a scale larger than any previously available tool, while also being easy enough for a novice to use. Green Hornet enables users to visually explore any kind of data that can be represented as a series of nodes and links, along with its associated metadata.

CHALLENGE

Analysts often find themselves working with data in the shape of a graph or a network. This includes such diverse domains as cyber systems, social networks, power grids and biological networks. Often, graphs can grow to a tremendous size, far too large for many of the most popular graph visualization tools to handle. Even when the tools are capable of displaying such large graphs, the resulting visualization is often too large and complex for the analyst to interact with and make sense of. Also, for those unfamiliar with graph theoretic approaches, these tools are often difficult to use.

SOLUTION

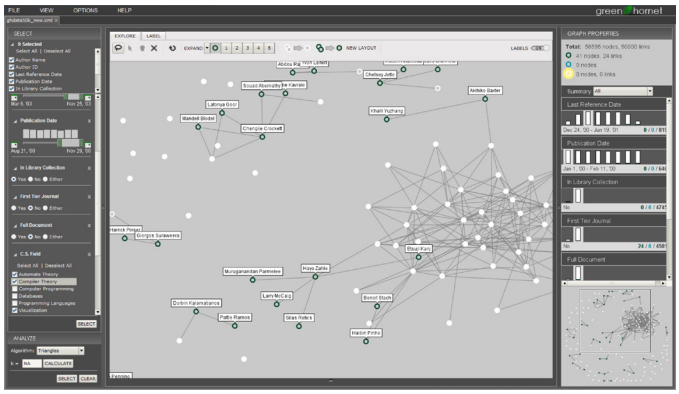
Green Hornet addresses this challenge through a unique multi-scale approach. Instead of trying to represent every individual node in the graph, it groups nodes together into larger "supernodes," which are constructed by grouping closely interconnected nodes. It shows the user a much smaller and more manageable number of supernodes representing the entire graph, instead of displaying millions of nodes on the screen at once. This gives the user an initial global view of the graph that is not visually overwhelming, but still captures important characteristics of its overall structure.



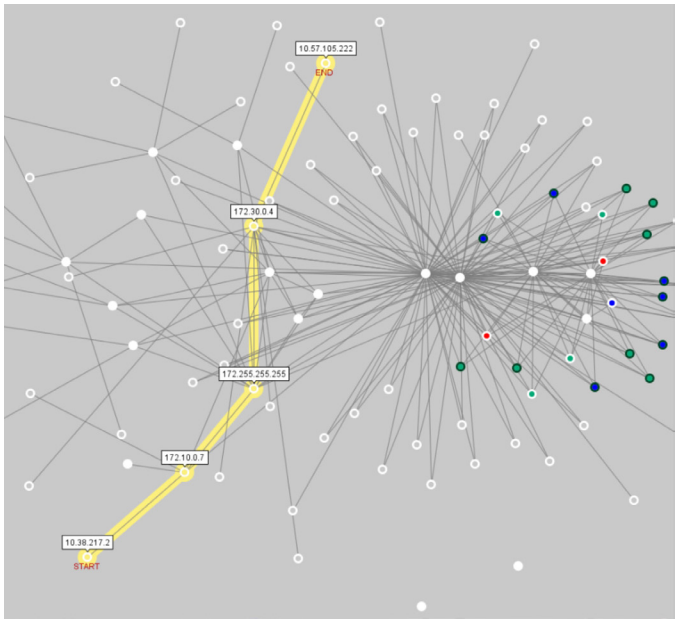
Custom data loader allows for easy ingest of any CSV data into a graph.

Individual nodes and links of interest can be investigated by a process of selection and expansion. Nodes and links can be selected by searching and filtering on their associated attributes and metadata. These individual nodes and edges can then be extracted from their supernodes to show their connections. Likewise, their nearest neighbors can also be extracted, providing a view of their local neighborhood within the context of the large-scale global view of the graph. This unique ability allows users the option to look at both the small-scale and large-scale views of the graph simultaneously.

In addition to being able to search for nodes based on attributes and metadata, nodes and links can also be searched based on their topological features. A number of graph analytic capabilities are available, including shortest path, vertex similarity and community detection. Using these techniques additional nodes of interest can be located and extracted from their associated



Nodes of interest are selected by filtering on data attributes. Selected nodes can be easily extracted from their supernodes.



Green Hornet visualizes the shortest path between two nodes (highlighted in yellow) within a fictional data set.

supernodes. This way, additional individual nodes of interest can be explored.

Green Hornet is designed to be user-friendly, even those with limited graph analytic experience can quickly begin to use it. A data importer allows users to quickly and easily import any data in a CSV format and convert it to a graph. All interactions with Green Hornet are designed to be easy and intuitive, so that even casual users can use the software with ease. A web-based version is currently in development for even easier deployment.

IMPACT

Green Hornet empowers users to visualize and navigate data in ways that were previously infeasible. By gaining the ability to drill down and explore items of interest within the big picture of data, users are more able to interact and analyze their data completely. Gaining insights from large graphs no longer needs to be an extreme challenge, but can now be done easily and intuitively.

ABOUT PNNL

Interdisciplinary teams at PNNL address many of America's most pressing issues in energy, the environment and national security through advances in basic and applied science. Founded in 1965, PNNL employs 4,400 staff and has an annual budget of nearly $1 billion. It is managed by Battelle for the U.S. Department of Energy's Office of Science.

For more information contact:

Kris Cook
Project Manager
kris.cook@pnnl.gov
(509) 375-6347

Patrick Mackey
Technical Lead
patick.mackey@pnnl.gov
(509) 372-6778

Pacific Northwest NATIONAL LABORATORY

U.S. DEPARTMENT OF ENERGY

Proudly Operated by Battelle *Since 1965*