The dawn of web 3.0: website mapping and vulnerability scanning in 3D, just like you saw in the movies

Teal Rogers and Alejandro Caceres

Intro

- Teal has experience with 3D visualizations and organizing massive amounts of data.
- Alex has a background in distributed computing, network reconnaissance, and vulnerability detection.

The Problem

- The internet contains a massive amount of data that is extremely interconnected, however we lack any good solutions for visualizing these connections. Classical approaches to showing that much data just don't work. You either have to eliminate too much data or else allow everything to become too confusing.
- This is where 3D comes in. By organizing our data in 3D we finally have a platform for displaying the invisible structure of the web. Now that we finally have a way to structure the web visually and intuitively, we can start adding data to that picture. For instance by allowing everyone (everyone who uses our software at least) to see just how many sites are riddled with security vulnerabilities.

The Problem

- I have been analyzing the security of sites on the internet for some time now. Most websites on the Internet are a complete mess. My PunkSPIDER project discovered this when I unleashed my distributed fuzzer on the entire Internet and started cataloging the results.

How we built it

- The first thing we needed to do was what Google already did so well -- collect links on the Internet and keep our index up to date. One of our requirements was that this metadata include extensive information on the vulnerabilities of a website. In order to find this, we are performing thorough, but minimally invasive, application-level vulnerability scanning against every site we crawl.
- We are leveraging the open-source and free Apache Nutch project along with some custom built Nutch plugins to help us out with this. Nutch is an extremely powerful Hadoop cluster-based distributed web spider.

How we built it

- We built a custom, distributed web application fuzzer to find vulnerabilities as fast as we can spider. By using Alex's experience in building high-speed, distributed web app fuzzers (see PunkSPIDER) we were able to build a custom one for this project relatively quickly. Application vulnerability detection is an integral part of the back-end workflow of this project and is in fact, built directly into our web spidering efforts.
- The back-end's goal is to make website security data an integral part of high speed crawling, therefore allowing us to make this an integral part of the visual metadata in the 3D engine.

How we built it

- All of the structures you see here are organic. Pages repel each other, links between pages pull them closer together, and every page floats to its own level based on how many hops it is from the home page. Using these basic physical principals each site creates its own unique structure based on how its links are structured.

Where to go

- This is just the beginning, we have a lot more to add to our view of web 3.0, and we want your help. If you're interested, come to trinarysoftware.com or hyperiongray.com, try the software for yourself, and add yourself to the mailing list. We will be giving away free beta access to everyone on the mailing list in a few weeks, and we want your input on where you would like to see web 3.0 go from here.
- If you want to hear more about Alex's distributed network reconnaissance and attack tools, he is giving a speech about them in track 1 at 3:00 today.