

Details of Project:

Title:	A Stateful Page-Traversing HTML Parser
Objective:	To design and develop modules with state-keeping and associative capabilities for parsing HTML Webpages.
Brief Description of Project:	<p>This project is part of a greater effort to profile and quantify the effectiveness of key words that are used as search terms in modern Web search engines. The intern will be expected to design and develop the following modules:</p> <ol style="list-style-type: none">(1) A HTML parsing engine with capabilities to identify, extract and analyze embedded hyperlink information(2) A HTML chaining engine capable of associating multiple Webpages together based on specific user-defined criteria(3) A centralized state engine with state-keeping capabilities to track user activity over multiple HTTP sessions <p>At the end of this project, the intern can expect to gain knowledge in a wide range of Web application development frameworks and tools. He/she can also expect to gain rich experience in Linux application development, HTML programming, Perl scripting, BRO script development and analysis, and in using various network traffic analysis tools.</p>
Skills Required:	<ol style="list-style-type: none">(1) Keen interest in Web application development(2) C/C++ programming experience preferred(3) Self-motivated and willing to work independently
Faculty (major):	Computer Engineering, Electrical Engineering, Electronics Engineering, Communications Engineering, Infocommunications Engineering
Duration:	6 months
Individual/Team Project:	Individual Project (one student)