Search engine mashup

July 6 2007

A mashup of two different types of web search tools could make find the useful nuggets of information among all the grit on the Internet much easier.

We have all experienced the inconvenience of trying to search for something on the World Wide Web and the search engine spewing back thousands of results, none of which look very relevant. The addition of new addresses, such as Europe's ".EU" domain mean that the number of web pages available will continue to grow and the problem of taming all that information will only get worse.

To help solve the problem, information scientists Liu Wei and Chen Junjie of the Taiyuan University of Technology, in Shanxi, China, have brought together to distinct types of computer software to help them build a new search engine that can intelligently crawl other search engines. They describe their new search robot in Inderscience's International Journal of Agent-Oriented Software Engineering.

"Traditional search engines cannot cope easily with this rapid expansion of information resources, explains Junjie.

He and his colleagues turned to the concept of search agent. Search agents, are intelligent virtual robots that can scan data very quickly looking for keywords and assessing the context of their findings. For instance, returning results related only to apples rather than Apple computers, when a fruit farmer searches for information on tasty new strains. There are several tools available to computer programmers for

creating such agents. The researchers have then combined the search agent idea with another technology, the so-called meta search engine.

Meta searches involve scanning information, not from a single source, the Google or Yahoo index say, but from all available sources. Sites such as Bytesearch, MetaCrawler and Ixquick are well-known meta search tools.

The team has developed a new intelligent search agent and combined it with a meta search tool. The intelligent agent can determine the context of the user's search terms and choose appropriate search engines to scan. It then retrieves the most relevant results. Junjie explains that this approach, boosts the precision rate and the recall rate of traditional search engines, and fulfil users' query requests well.

Team member Liu Wei provides an example of how the system might work in practice. Suppose you are interested in a specific football team," he suggests, "our meta-search engine would find your user profile containing your interest in football, when you enter a keyword such as "football", the information on your favourite team from the various search engines will be retrieved preferentially." This not only boosts the precision rate and the recall rate of the meta-search engine but also assists with finding other more obscure information based on context.

Source: Inderscience Publishers