Binjuice: Fast Location of Similar Code Fragments Using Semantic Juice

There is a growing need for the comparison of binary executables in applications such as threat-detection via malware analysis and copyright infringement. For malware analysis in particular, finding matching binaries involves searching extremely large collections (e.g., millions) of malware. Additionally, it is important that the comparison algorithms account for changes due to code evolution, changes in compiler optimizations, and post-compile obfuscations. Various technologies have been developed for these purposes; however most suffer from low sensitivity, scalability and robustness. In the current invention, researchers at the University of Louisiana at Lafayette introduce the concept of ‘juice’ — an algebraic generalization of the denotational semantics of a program. The juice captures the essential relations established by a piece of code, independent of choices of registers and literal constants. The juice then serves as a template of the code that is invariant against choices made by compilers or by code obfuscation tools and permits fast-matching of related code.

**KEY ASPECTS OF THE TECHNOLOGY:**

- Very fast matching and scalable to large collections; permits near real-time threat detection;
- Can serve as an analytical node within the intelligence community;
- Development funded by DARPA and AFOSR;
- Based on patent-pending technologies.

UL Lafayette specializes in Research for a Reason. We recognize that the current technology may by brought into practical use for public benefit and yield economic value. Accordingly, we currently seek a commercial partner interested in commercial development of this technology via licensing and/or collaborative research partnerships. To learn more about this research and/or partnership opportunities please contact Seth Boudreaux, Associate Director of Innovation Management, via the info provided below.