

## Testing Intrusion Detection Systems: Issues Without Answers

John McHugh (CERT/CC)  
jmchugh@cert.org

### Abstract

In 1998 (and again in 1999), the Lincoln Laboratory of MIT conducted a comparative evaluation of Intrusion Detection Systems (IDSs) developed under DARPA funding. While this evaluation represents a significant and monumental undertaking, there are a number of issues associated with its design and execution that remain questionable. The difficulties associated with this evaluation have been the subject of several papers and a number of presentations. As a result of our investigations of Lincoln's efforts, we have been attempting to develop an appropriate framework in which similar, but meaningful and useful, evaluations can be performed. This talk will contrast our proposed approach with the work that Lincoln performed (and is continuing to perform). Our primary conclusion for signature based systems are that we simply do not know enough to generate appropriate artificial background data for false alarm evaluation, but that there are systematic approaches to measuring true positive and negative performance, under both ideal and appropriate environmental stress conditions. The situation is much less clear for with respect to anomaly based systems since the relationships between anomalous and intrusive behavior are poorly understood. In both areas, there is a paucity of theory that can be applied to the problem and we feel that the ad hoc and intuitive approaches that characterize today's efforts may be nearing their limits.