Further Analysis of an Off-Line Intrusion Detection System: An Expanded Case Study in Multi-Objective Genetic Algorithms

Pedro A. Diaz-Gomez and Dean F. Hougen*
Robotics, Evolution, Adaptation, and Learning Laboratory—REAL Lab
School of Computer Science, University of Oklahoma, OK, USA

One computer security mechanism used currently is the *Intrusion Detection System* (IDS). *Off-line* IDSs search audit trail registries looking for user activities that match patterns of events known as attacks. Because such search is NP-complete, heuristic methods will need to be employed as databases of events and attacks grow. Genetic Algorithms (GAs) have been widely used as heuristic search methods. However, balancing the need to detect all possible attacks in an audit trail with the need to avoid warnings of attacks that do not exist is a challenge, given the scalar fitness values required by GAs. We present an expanded case study of *GASSATA*—a Genetic Algorithm as an Alternative Tool for Security Audit Trail Analysis (Mé, 1998), a previously proposed GA-based IDS that shows this difficulty with respect to its fitness function. Previously we proposed a new method to overcome it. Here we compare our method to an alternative of using a different selection mechanism.

GASSATA is an off-line IDS (Mé, 1998) with fitness function $F(I) = \alpha + \sum_{i=1}^{N_a} W_i \cdot I_i - \beta * T^2$, where I is the hypothesis vector, α maintains $F(I) > 0$ in order to retain diversity in the population (using proportional probability selection), N_a is the number of known attacks, W is the weighted vector that reflects the risk of each attack, β provides a slope for the penalty function, and *T* is the number of times for which $(AE \cdot I)_i > OV_i$, where *AE* is the attack-event matrix that shows which events are required for each attack, and *OV* is the observed vector of events.

Mé (1998) reports good results with *GASSATA* but our experience has been that the system often generates false positives and negatives (Diaz-Gomez & Hougen, 2005). For this reason, based on experimental results we proposed a new fitness function: $F(I) = N_e - T'$, where N_e corresponds to the number of events, and T' corresponds to the number of times that $(AE \cdot I)_i > OV_i$, *for each attack* I_i. That is, if a hypothesized attack I_i *considered alone*, would cause $(AE \cdot I)_i > OV_i$ for some i, and another hypothesized attack I_j *considered alone*, would also cause $(AE \cdot I)_i > OV_i$, then T' would have a value of 2 (whereas T would have a value of 1). Now, the better the hypothesized vector I, the smaller T' is, and of course, $F(I) \rightarrow N_e$, the maximum. To avoid false negatives, we added a mechanism that takes the union of all newly hypothesized attacks that are consistent with the existing aggregate solution set.

The results found with the new fitness function and mechanism are that *there are no false positives* and that the number of *false negatives decreases dramatically* compared to the results we saw previously (Diaz-Gomez & Hougen, 2005). This time 70 runs were performed—10 repetitions each for 7 different cases—and only one time a false negative was present.

Here we repeat this experiment in order to compare proportional probability with tournament selection, and we obtain similar results: false positives and false negatives with *GASSATA's* fitness function; no false negatives or false positives with the fitness function and mechanism we suggested in order to solve this problem.

This research shows some difficulties in providing accurate values to parameters in the fitness function suggested in *GASSATA* (Mé, 1998) and proposes a solution independent of variable parameters making the fitness function to solve this particular problem quite general and independent of the audit trail data. Our solution has proved more effective than both the original and a variant of the original using tournament selection.

References

Diaz-Gomez, P., and Hougen, D. 2005. Improved off-line intrusion detection using a genetic algorithm. To appear in *Proceedings of the Seventh International Conference on Enterprise Information Systems*.

Mé, L. 1998. GASSATA, a genetic algorithm as an alternative tool for security audit trail analysis. In *First International Workshop on the Recent Advances in Intrusion Detection*.

---

*E-mail: `pdiazg@ou.edu` and `hougen@ou.edu`