Decision Tree Rule Induction for Detecting Covert Timing Channels in TCP/IP Traffic Félix Iglesias, Valentin Bernhardt, Robert Annessi, Tanja Zseby

1. A set of statistical methods called DAT (Descriptive Analytics of Traffic) has been previously proposed as a general approach for detecting covert channels. In this paper, we implement and evaluate DAT detectors for the specific case of covert timing channels. Additionally, the paper proposes machine learning models to induce classification rules and enable the fine parameterization of the DAT detector.
2. In this work, the focus is on timing channels.
3. Main goals
   - To deeply evaluate DAT detectors for the specific case of covert timing channels. To this end, a testbed for generating and testing covert channels has been created.
   - To establish a methodology for tuning DAT detector parameters and rules based on Decision Trees learners.
   - To generalize a set of constituent rules for DAT detectors in the scope of covert timing channels, thus enabling the incorporation of DAT detectors in future IDS.
4. A DAT detector consists of three well-differentiated phases pre-processing (P), feature extraction and transformation (T), and flow labeling and covert channel detection (D).
5. The first phase takes raw traffic captures as inputs (e.g., PCAP files). Traffic captures are parsed, and packet vectors are formed with meaningful, homogeneous information for the subsequent analysis.
6. In the second step, packet vectors are transformed into flows. At the same time, pre-defined calculations and estimations are conducted on the flows. Finally, new two-level structured OD-flow vectors (origin-destination flow vectors) are created.
7. A complete implementation of DAT detectors is expected to analyze traffic according to four different blocks or steps:
   - A Packet Compliance Checker, which, according to a set of fixed policies, detects traffic that is corrupted or does not comply with standard practices.
   - The Intra-field Analysis block, which checks TCP/IP header fields separately by examining the corresponding OD-flow vector values.
   - The Inter-field Analysis block, which considers combinations of OD-flow vector values in different TCP/IP header fields.
   - A ML-based Detector (machine-learning-based), which compares ODflows by using a library that contains representative patterns (footprints) of OD flows with covert channels. Such patterns are linked to known, published techniques.
8. Eight covert timing channels have been implemented for the conducted experiments based on packet inter-arrival times (iats).
9. Experiments were conducted, the objective of which was to create and refine DAT rules for the detection of covert timing channels. For this purpose, training and testing processes were performed with known—i.e., labeled—datasets. The training process consisted of three main phases: (1) dataset generation, (2) model obtaining, and (3) rule generalization.
10. The generalized model gave the following performance metrics

-Accuracy 99.18%
- Precision 90.95%
- Recall 95.95%