Verificação de software usando aprendizado de máquina

Artigo: Automated Analysis of Source Code Patches using Machine Learning Algorithms

Autores: Antonio Castro Lechtaler, Julio César Liporace, Marcelo Cipriano, Edith García, Ariel Maiorano, Eduardo Malvacio, Néstor Tapia.

O código desse programa é baseado nos critérios de ferramentas autônomas de análise de *malware*, como Cuckoo Sandbox por exemplo. Também se utiliza de técnicas de inteligência artificial, especialmente *text mining* e aprendizado de máquina.

As técnicas de aprendizado de máquina no projeto são:

Processamento de linguagem natural (NLP); Classificação supervisada, que consiste no desígnio de *tags* para categorizar entrada ou inputs. Análise automática de Patches.

Para fazer essa categorização foi baixado o banco de dados público de bugs da Red Hat e OpenBSD. A partir destes, os classificadores foram “treinados” e podem ser usados de um plugin para estimar automaticamente a *tag* mais apropriada para cada patch analisado.

O software livre para download é uma aplicação web em Python que usa o Web Django *framework*.

Artigo: Static Code Features for a Machine Learning based Inspection - An approach for C

Autor: Hannes Tribus.

O trabalho foca em criar mecanismos para facilitar a aplicação de aprendizado de máquina na verificação de código. O pesquisador se propõe a resolver certos problemas nesse ambiente e desenvolve dois códigos para trabalhar em conjunto, um para decompor o código e o outro para simplificar a detecção de instâncias relevantes para o treinamento dos verificadores.