



Contribution ID: 543

Type: **Talk**

CerebRAN: A Dynamic Behavioural Dataset for Ransomware Detection

Monday, 1 December 2025 11:00 (30 minutes)

Abstract: Ransomware remains a significant cyber threat, yet research is often hampered by a lack of modern, balanced datasets. This study proposed CerebRAN, a new dataset made from dynamic analysis of ransomware (400 samples) and goodware (399 samples). We provide a detailed methodology from the sample collection to the extraction of features using Cuckoo Sandbox on a Windows 7 operating system. To validate the usability of CerebRAN, we performed machine learning experiments using Random Forests and Logistic Regression using the Recursive Feature Elimination with Cross-Validation (RFECV) technique. The results we obtained from the experiments show that the Random forests were the superior classifier on CerebRAN scoring accuracy of 0.9625, precision of 0.9628, recall of 0.9625 and f1-score of 0.9625. Logistic regression scored an accuracy of 0.9562, precision of 0.9563, recall of 0.9563 and an F1-score of 0.9562. Random forests outperformed Logistic regression using an optimum 48 features while logistic regression used 174 features. This experiment highlighted how effective and valuable CerebRAN is for the development of robust detection tools. The dataset and sample metadata are publicly available on GitHub.

Presenting Author

Email

Student or Postdoc?

Institute

Registered for the conference?

CHPC User

CHPC Research Programme

Presenters: MUSIIWA, Emmanuel (Rhodes University); MOYO, Moses; CHINDIPHA, Stones (Rhodes University)

Session Classification: ISSA