Locality Sensitive Hashing for Security Applications

Jon Oliver[1]
[1]*Trend Micro*

Security solutions based on cryptographic hashes is no longer effective. There is a remarkable class of hashing functions which the security industry and researchers need to consider: Locality Sensitive Hashing.

Locality Sensitive Hashes are a scalable fuzzy hash that have been used by the search engines and internet giants for some years. We have been actively researching and using these hashes. Jon has invented and provided an open sourced model for a Locality Sensitive Hash, TLSH, that is well suited for modern security applications.

In this session, we will give an overview of these hashes, and describe how they can be used to quickly search very large datasets for similar files. In addition, we look at the ways that such solutions can be attacked by bad actors.

We give two cases of their use:
(i) How they can be used for pre-processing data sets for machine learning purposes;
(ii) How they can be used to enhance whitelists so that they are dynamic, and able to keep up with rapidly changing legitimate files in a secure way.

From this session, you will gain a deeper knowledge of advanced security techniques that organisations and researchers need to keep abreast of.