Fighting against paedophile activities in the KAD P2P network

Thibault Cholez, Isabelle Chrisment, Olivier Festor

Fighting against paedophile activities in the KAD P2P network

Thibault Cholez, Isabelle Chrisment and Olivier Festor

MADYNES - INRIA Nancy Grand Est, France
{thibault.cholez,isabelle.chrisment,olivier.festor}@loria.fr

Abstract. In this poster, we present a solution to fight against paedophile activities in KAD. Our distributed architecture can monitor and act on paedophile contents in a very efficient way by controlling keywords and files. Early results on the real network demonstrate the applicability of our approach.

Keywords: Honeypot, KAD, DHT

1 Motivation

KAD is a part of the **eMule** software and one of the major file sharing P2P networks ($\sim$ 3 millions of simultaneous users). KAD uses a structured architecture called Distributed Hash Table (DHT) to allow users to retrieve a specific file from keywords and the possible sources for a file. Observing users and controlling contents in KAD are real technical challenges:

- Each file and keyword is published on dozens of peers on the DHT, in order to keep the information available.
- As paedophile contents can be referenced through normal keywords, monitoring only files can lead to false positive (normal users considered as paedophiles).
- Attracting users with Honeypots (fake files) is resource consuming because popular files need to show a high number of sources.
- Recent protection mechanisms inserted in KAD mitigate the Sybil attack (insertion of many fake peers from a single computer to disturb the network).

2 Our features to fight against paedophile activities

Thanks to our specific distributed architecture exploiting some KAD weaknesses, we can provide several features helping to study and fight against paedophile activities on that network. Being given the hash of specific contents, we can do:

- **Passive monitoring:** we transparently monitor all the requests sent to the targeted contents in the network. We can discover all the new published files for a given keyword and all the peers sharing a file.

– **Eclipsing content:** we eclipse entries of the DHT to remove the targeted contents from the network and prevent users from accessing it.
– **Index poisoning:** we poison the DHT references with very attractive fake files showing a high number of sources.
– **Promoting Honeypots:** we attract the final download requests for the controlled files towards our Honeypots.

By attracting all the publications and searchs of paedophile contents (keywords and files), our architecture can assess and control the paedophile behavior from the initial search of keyword to the final download.

3 Experiments

To test our solution, we eclipsed the good references for the keyword "spiderman" and poisoned them with 4 fake files for one day. The results in figure 1 and 2 show that our architecture is effective and the importance to control the number of sources to build an efficient Honeypot. Our upcoming work consists to deploy our architecture to specifically study and fight against paedophile activities.

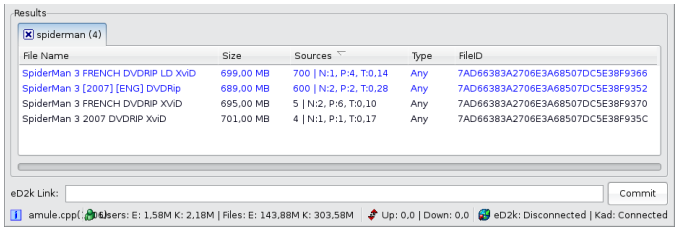

Fig. 1. Results of a search for "spiderman" under eclipse and poison (4 fake files)

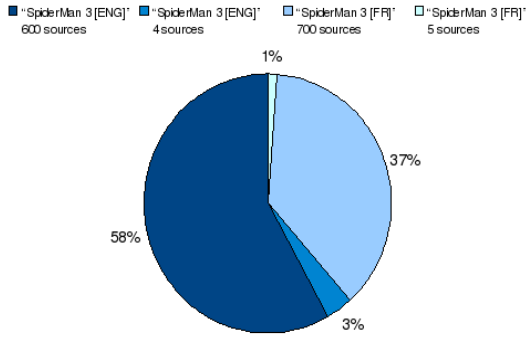

Fig. 2. Proportion of download requests received for each fake file

Acknowledgment: This work is funded by the French ANR Research Project MAPE(Measurement and Analysis of Peer-to-peer Exchanges for pedocriminality fighting and traffic profiling), under contract ANR-07-TLCOM-24.