

Generic Dark-Market-Crawler

Degree programme : BSc in Computer Science | Specialisation : IT Security
Thesis advisor : Prof. Dr. Emmanuel Benoist
Expert : Jean-Marie Leclerc

The Generic Dark-Market-Crawler (GEN-DMC) is a Software with the purpose to collect and analyze data from Dark Markets in a generic way. Dark Markets are illegal marketplaces found in the Dark Net. They sell all sorts of illegal goods such as drugs, stolen credit card details, official documents, pornography, and sometimes even weapons. The GEN-DMC downloads raw HTML-pages and extracts the important information to mirror the database behind a Dark Market.

Dark Markets

Dark Markets are digital marketplaces found in the Dark Net. They are in a lot of ways similar to other online platforms that are built for trading such as eBay or Amazon. The key difference is the anonymity that is required for a Dark Market to exist as well as the goods that are sold. Most of the goods that are sold on Dark Markets can not be acquired in a legal way.

Programmatic Architecture

Each of these Dark Markets are hosted by different individuals, which all build their websites in a distinct manner. Therefore each of those marketplaces has a unique architecture. The differences lie in the routes (links), the levels of categories, and the product and vendor detail pages.

Stability

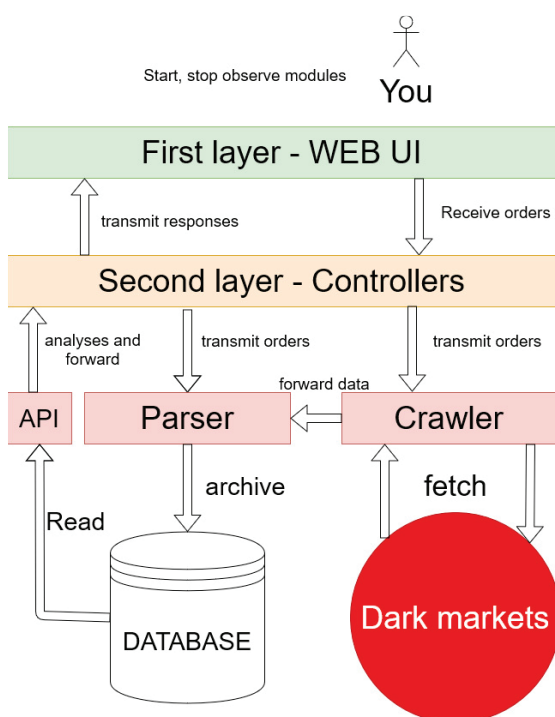
On average, a Dark Market is active for around 9 months. The markets close for a number of reasons. There is the so-called «exit-scam» in which the operators close the market down abruptly taking all the money that is currently held on the site. Moreover, there is closure on mutual agreement and closure due to law enforcement targeting a site. There are also cyber-wars going on between the sites which end in de-anonymization or denial of service and therefore closure.

Motivation

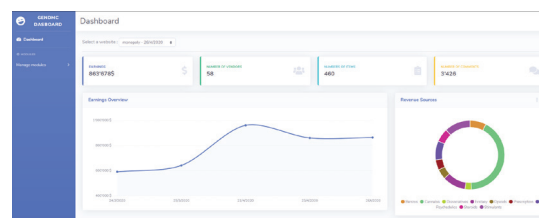
The motivation behind the GEN-DMC is to create a software to analyse Dark Markets without having to depend on the marketplaces. By modularization, the software can easily be configured to analyse given Dark Markets.

System Architecture

The application consists of four parts. First, there is the Cockpit that interacts with the user and commands the application. Secondly, there is the Crawler, which is responsible to download raw HTML-files from the markets. The Crawler is configured through modules. The modules specify which pages to download and how to react in case of errors. The third part is the Analyser, which is responsible to extract the important data, unify and save it. The Analyser is also configured through modules, the modules specify which information and how it has to be extracted. The last part is an API, that is used to access the data that has been acquired.



System architecture



GEN-DMC Dashboard



Nicolas Pantazis
pantazis.nicolas@gmail.com



Jordan Bonnie Wilkinson
079 513 45 96
jordan.wilkinson@gmx.ch