

## Tor ネットワークのクローラシステムの開発と違法商品販売サイトの調査

鳥居 洸希 †

菊池 浩明 †

明治大学総合数理学部 †

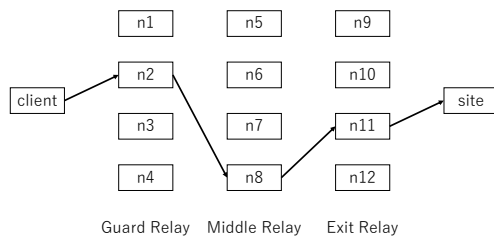


図1 Tor 原理図

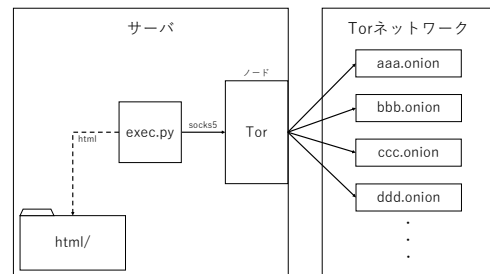


図2 システム構成図

## 1 はじめに

近年のプライバシーに対する意識の高まりから、匿名通信の重要性はますます高まっている。中でも、匿名通信システム Tor(The Onion Router) は、発信元の情報を隠したままのサーバ運営が可能であり、広く用いられている。本来、匿名通信はプライバシー保護の目的で利用されるものであるが、不正な目的で利用されるケースも多い [1]。

そこで本研究では、ダークウェブ上の違法物品販売サイト上で売られている物品の種類と存在期間に関する調査を行う。網羅的な調査を効率的に行うため、システムの自動化を試みる。

## 2 Tor の匿名性

Tor ネットワーク内のウェブサイトは、特有の.onion という TLD(Top Level Domain) を持つ。この TLD のサイトには、通常のブラウザではアクセスすることが出来ず、Tor ブラウザが必要になる。

図1に Tor の原理図を示す。匿名性を確保するために、いくつかの中継サーバを経由し、目的のウェブサイトアクセスする。これらの中継サーバ(リレー)は、クライアントに近い方から Guard, Middle, Exit と呼ばれている。図1の例では、 $n_{11}$ ,  $n_8$ ,  $n_2$  の順に多重に暗号化がなされている。これらのリレーは一定時間ごとに接続先が変わり、これによって匿名性を高めている。

## 3 クローラシステムの開発

本研究では、Tor ネットワークにアクセスするために、Tor をインストールした Linux サーバと Python を用いた。図2に本研究で開発したクローラシステムの構成図

を示す。クローラは、`crontab` で定期的に自動実行される。第一引数に指定された URL を目的のサイトとして、`socks5` を用いて Tor ネットワークに接続する。`html` をサーバに保存し、必要なデータの抽出を行う。

## 4 違法商品販売サイトの調査

## 4.1 概要

今日の Tor サイトのほとんどは、CAPTCHA を用いてクローラを防止している。そこで、Tor ブラウザを使用し手動でデータ取得をしたウェブサイトも含めて調査を行なった。

表1は本研究で調査したサイトの一覧とデータである。本システムを用いた自動取得は2018年12月上旬、手動での取得は2018年9月上旬である。表2にカテゴリごとの情報を示す。ここで、Drugs の平均価格については、販売されている量が異なり集計が困難なため割愛している。また、表1と表2の平均価格は1 EUR を1.14 USD, 1 BTC を3756.93 USD として USD に換算した値である。

## 4.2 違法商品検索システム

6つのサイトにて扱われている2094件の違法商品の種類や傾向を明らかにするため、商品の検索システムを開発した。本システムの SQL テーブルスキーマを表3に示す。PHP を用いてデータの可視化を提供している。本サイト\*ではデータの検索が可能である。検索結果の例を図3に示す。検索結果が所属するサイトとカテゴリの割合も表示する。

## 4.3 調査結果

表1に調査結果を示す。(ただし、Tochka Free Market は URL 構成が変わっていたため調査の対象から外した。)Apollon Market を除く全てのサイトは Drugs に属

Development of Tor network crawler system and survey of illegal product sales site

†Hiroki Torii, Hiroaki Kikuchi, School of Interdisciplinary Mathematical Science, Meiji University.

\*<https://windy.mind.meiji.ac.jp/%7Etori/catalog/>

表1 調査したサイトの情報

サイト名	主要カテゴリ	商品数	平均価格	取消し数	登録	CAPTCHA	通貨
Apollon Market	Digital Goods	448	22.21	293	有	有	BTC
BERLUSCONI MARKET	Drugs	535	500.86	10	有	有	EUR
Dream Market	Drugs	800	280.3	134	有	有	BTC
ElHerbolario's Shop	Drugs	13	-	0	無	無	EUR
Green Road	Drugs	150	42.54	0	有	有	BTC
Tochka Free Market	Drugs	83	375332.41	-	有	有	USD
Wallstreet Market	Drugs	78	76.47	3	有	有	USD, EUR

表2 カテゴリ

カテゴリ	商品数	サイト数	平均価格
Drugs	846	7	-
Digital Goods	356	5	379.51
Services	199	5	38.87
Fraud	117	3	58.16
Guides & Tutorials	80	4	38.92
Weapons	54	1	2165.13
Hosting & Security	52	3	43.33
Counterfeit Items	51	2	645.58
Software & Malware	50	4	66.88
Carded Items	33	2	166.22
Other	243	5	-

表3 テーブルスキーマ

カラム名	説明	型
id	特有の ID	INTEGER
url	ドメイン以下の URL	TEXT
name	商品名	TEXT
price	値段	TEXT
category	カテゴリ	TEXT
subcategory	サブカテゴリ	TEXT
site_id	サイト ID	INTEGER

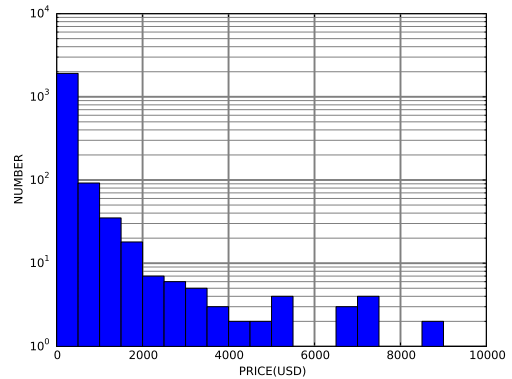


図4 価格分布

4.4 考察

Apollon Market の商品が群を抜いてなくなっていることについて、これは Digital Goods に属している商品が多いことに関係があると考えられる。ハードウェアやソフトウェアの商品もいくらかあったが、このカテゴリに属する商品の多くは、詐欺やクラッキングに関するノウハウが多かった。これらのノウハウ自体が詐欺商品として売られており、入れ替わりが早いのではないかと考える。

5 おわりに

本研究では、2107 件の商品の情報を取得し、違法商品販売サイトで取引される商品の種別や値段を調査した。調査の結果、主要カテゴリは Drugs であることが明らかとなった。また、商品の約半数は 500 USD 以下で取引されることが明らかとなった。

今後は、アクセスの方法を変えるなどして CAPTCHA を突破することを課題とする。

参考文献

- [1] トレンドマイクロ セキュリティブログ, (https://blog.trendmicro.co.jp/archives/8492, 2018 年 12 月 18 日参照.)
- [2] 大中彩香, “Tor ネットワークのない違法商品販売サイトの調査”, 明治大学 2017 年度卒業論文, 2017.
- [3] M. Leech, et al., “SOCKS Protocol Version 5”, Internet RFC 1928, 1996.

する商品が 1 番多かった。また、多くのサイトは 2 ヶ月間で数%~十数%の商品のみなくなっている。しかし Apollon Market だけは 448 件中 293 件と、約 65%の商品がなくなっている。

図4に商品の価格帯を示す。500 USD 以下で取引される商品が約半分を占め、2000 USD 以上するものは 50 件にも満たない。1000 USD を超える商品には、武器が多かった。

図3 サイト表示例(一部)

