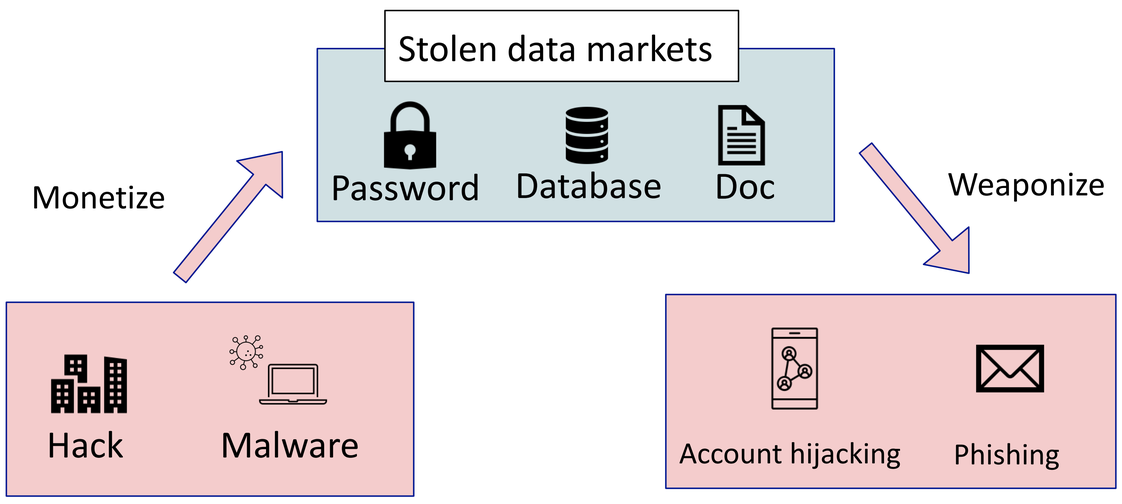
Stayin' Alive: How Global Stolen Data Markets Thrive on Telegram

Tina Marjanov, Taro Tsuchiya, Konstantinos Ioannidis, Jack Hughes, Nicolas Christin, Alice Hutchings

Background

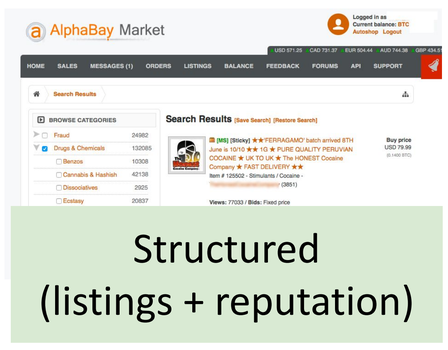
Why do we study stolen data markets?



Stolen data markets monetize breached data and facilitate downstream cybercrimes.

Where do cybercriminals buy them?

Darknet Marketplaces



Structured
(listings + reputation)

Telegram Messenger

Unstructured
(group + private chats)

The communities are moving to Telegram. How can we moderate chat-based platforms?

Data

How can we keep track of the channels?

Experts' Keywords (e.g.,"combo")

Darkweb forum links

+ Snowball

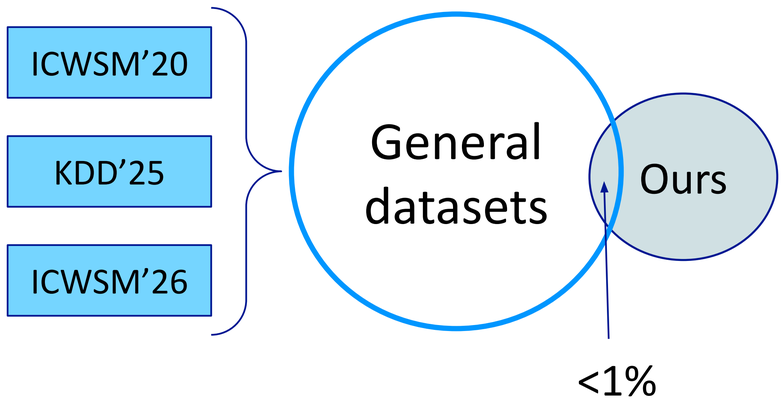
Manual curation

1,500+ groups

14+ million messages

3.6+ million files
(e.g., text, image, csv)

208,000+ groups



The communities are disjoint from the rest of the Telegram ecosystem.

Findings

What are the main stolen data types?

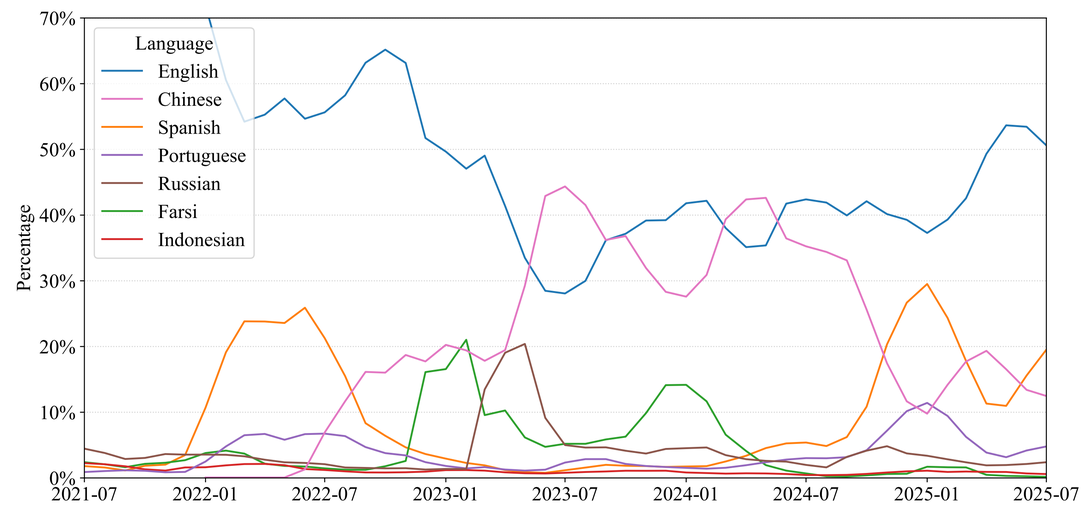
- Use human annotation + LLMs to classify channel data types
- Most channels specialize 1 or 2 data types

Credentials (22%), Infostealer Logs (21%)
Databases (18%), Personal Info (14%)
Carding (19%), Personal Doc (4%)

The markets are segmented to support different cybercrimes.

What languages do they use?

Half of the channels are not in English



High concentration of credit card data by country

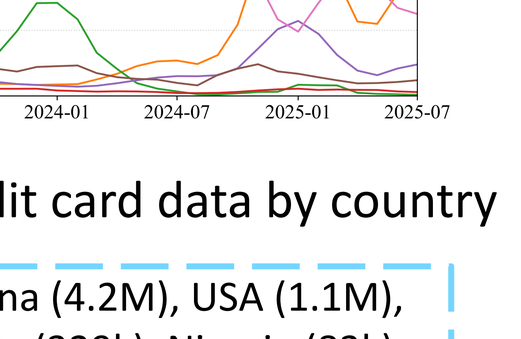
BINs

Argentina (4.2M), USA (1.1M),
Australia (329k), Nigeria (83k) …

The communities increasingly target non-English speaking population and sourcing products globally.

Do they employ any evasion techniques?

Gateway channels
- Don't sell stolen data
- Link other channels



Linked channels
- 91% private
- Survive longer and grow faster (based on our statistical models)

Moderating gateway channels may be effective.

UNIVERSITY OF CAMBRIDGE
CyLab Carnegie Mellon University Security and Privacy Institute