On the Dark Side of the Coin: Characterizing Bitcoin use for Illicit Activities

Hampus Rosenquist, David Hasselquist, Martin Arlitt, and Niklas Carlsson

Linköping University, Sweden, and University of Calgary, Canada

LINKÖPING UNIVERSITY

Imagine you could exchange any sum of money, worldwide, instantly, and effortlessly...

Imagine you could exchange any sum of money, worldwide, instantly, and effortlessly...

...beyond authorities' control.

Imagine you could exchange any sum of money, worldwide, instantly, and effortlessly...

...beyond authorities’ control.

I can work with that.

What kind of illicit activities is Bitcoin used for?

Scams

Ransomware

Money laundering

Darknet markets

Sextortion

Other...

Bitcoin use for illicit activities is widespread and turns over **large sums of money**.

These activities have increasingly negative societal effects, **affecting large number of victims**, often preying on the weak.

Therefore, we argue that it is important to **shine a light on the Bitcoin patterns** associated with different illicit activities.

Bitcoin use for illicit activities is widespread and turns over **large sums of money**.

These activities have increasingly negative societal effects, **affecting large number of victims**, often preying on the weak.

Therefore, we argue that it is important to **shine a light on the Bitcoin patterns** associated with different illicit activities.

Bitcoin use for illicit activities is widespread and turns over **large sums of money**.

These activities have increasingly negative societal effects, **affecting large number of victims**, often preying on the weak.

Therefore, we argue that it is important to **shine a light on the Bitcoin patterns** associated with different illicit activities.

Contributions

- **High-level characterization** of the transactions received by the Bitcoin addresses reported to the Bitcoin Abuse Database (2017-2022)
    - Aggregate basis
    - Per-category basis
- **Temporal analysis** that captures
    - Long-term trends
    - Weekly patterns (per category)
    - Correlations with the first report date (per category)
- **Analyze the outflow of bitcoins** from reported addresses

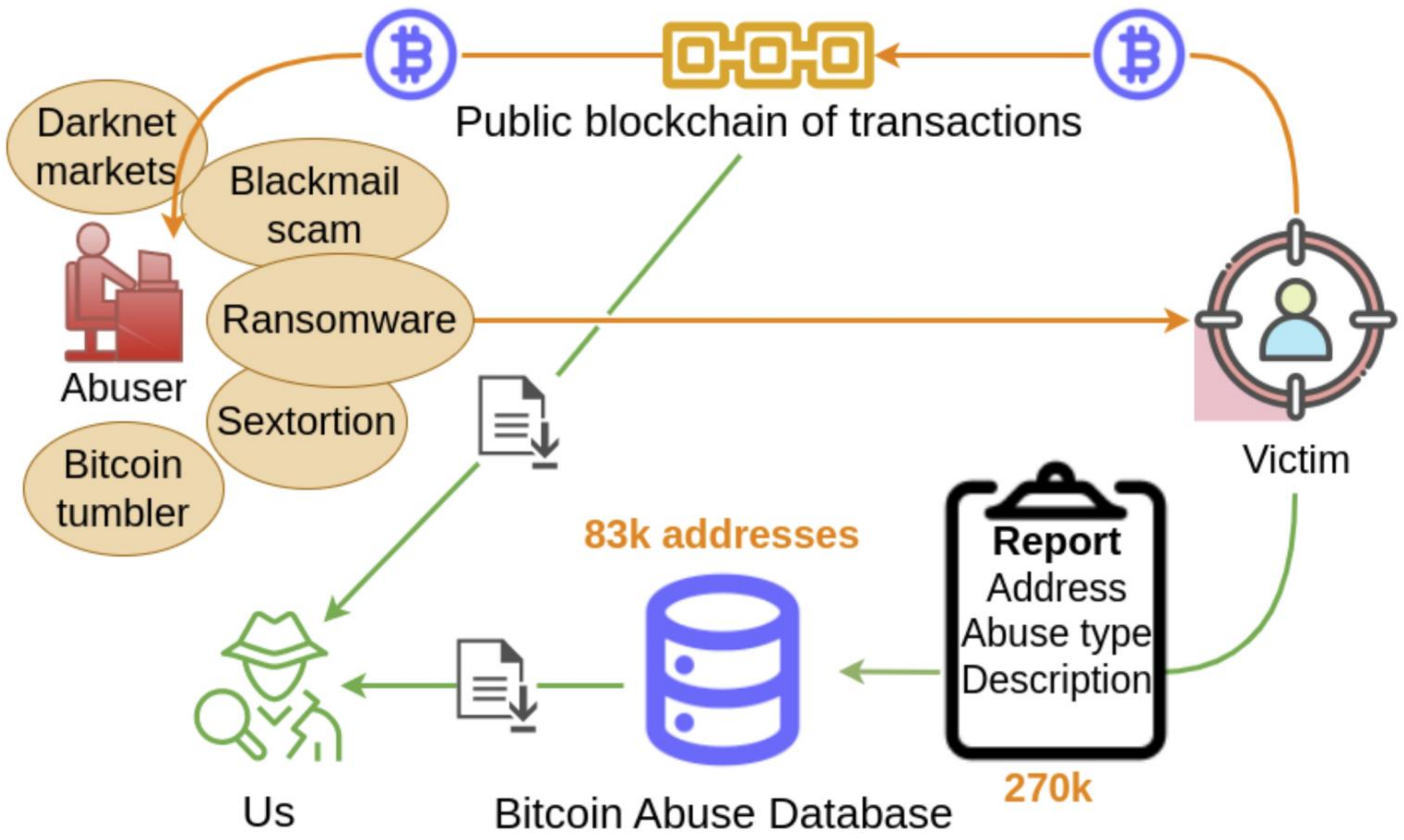
Contributions

- **High-level characterization** of the transactions received by the Bitcoin addresses reported to the Bitcoin Abuse Database (2017-2022)
  - Aggregate basis
  - Per-category basis
- **Temporal analysis** that captures
  - Long-term trends
  - Weekly patterns (per category)
  - Correlations with the first report date (per category)
- **Analyze the outflow of bitcoins** from reported addresses

Contributions

- **High-level characterization** of the transactions received by the Bitcoin addresses reported to the Bitcoin Abuse Database (2017-2022)
  - Aggregate basis
  - Per-category basis
- **Temporal analysis** that captures
  - Long-term trends
  - Weekly patterns (per category)
  - Correlations with the first report date (per category)
- **Analyze the outflow of bitcoins** from reported addresses

Methodology

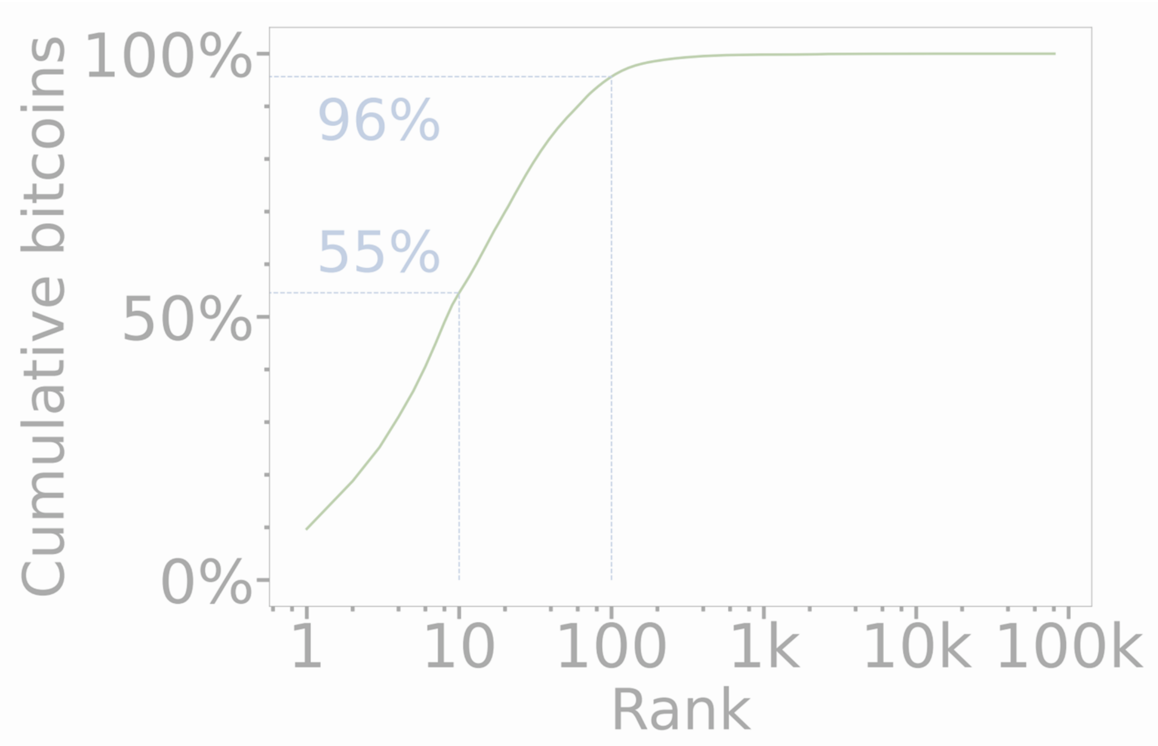


Aggregate High-Level Characterization

- How successful is each address?
- Model of the tail distribution

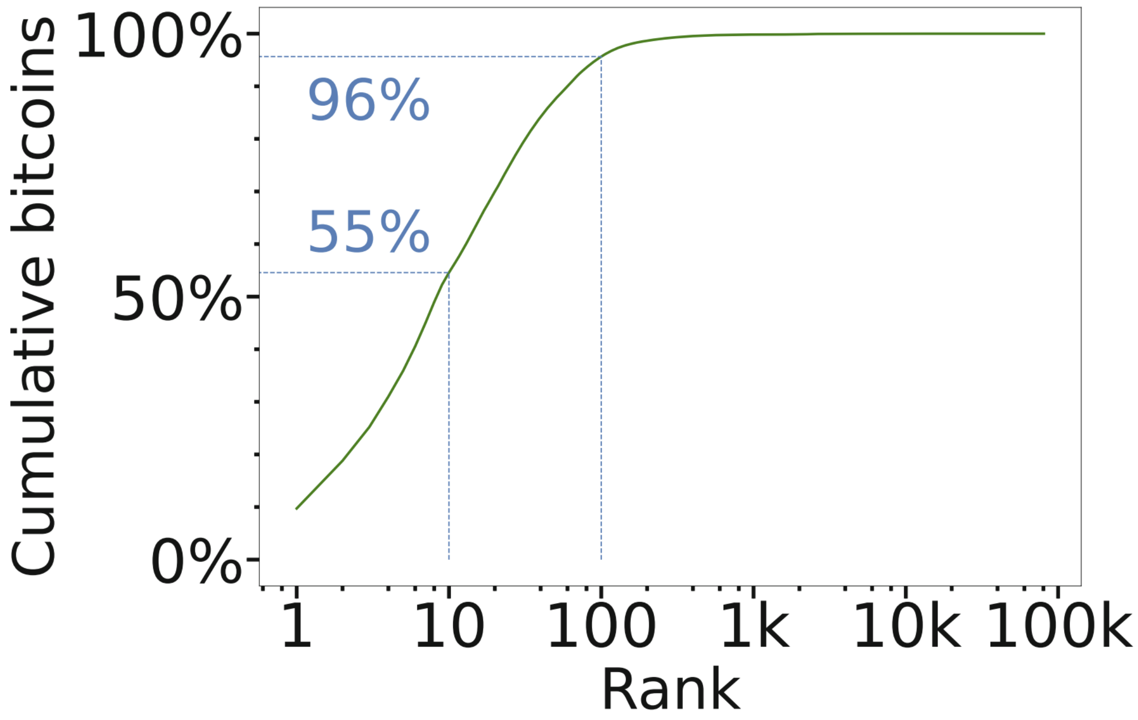
How Successful is Each Address? / High Skew

- All 83k addresses received 31M bitcoins
- Top-10 together received 55%
- Top-100 together received 96%
- Top-1000 together received 99.8%



How Successful is Each Address? / High Skew

- All 83k addresses received 31M bitcoins
- Top-10 together received 55%
- Top-100 together received 96%
- Top-1000 together received 99.8%



How Successful is Each Address? / Big Hitters

Table 4.1: Overview of the top-10 highest receiving reported addresses.

Received [BTC]	Median [BTC]	Category	Description
3,048,040	40.0	Other	Trading investment scam.
2,845,086	18.0	Other	Foreign exchange trading scam, “investment in terror".
2,009,608	25.0	Other	“Investment in terror", begs for treatment money.
1,815,619	800	Other	“Investment in terror".
1,535,341	45.0	Other	“Investment in terror".
1,459,182	160	Ransomware	“Investment in terror".
1,378,975	800	Other	“Investment in terror".
1,259,824	0.50	Other	“Investment in terror".
1,030,376	505	Other	“Inhumane" bank account theft via remote desktop.
724,340	1,150	Other	“Investment in terror", begs for treatment money.

How Successful is Each Address? / Big Hitters

Table 4.1: Overview of the top-10 highest receiving reported addresses.

Received [BTC]	Median [BTC]	Category	Description
3,048,040	40.0	Other	Trading investment scam.
2,845,086	18.0	Other	Foreign exchange trading scam, "investment in terror".
2,009,608	25.0	Other	"Investment in terror", begs for treatment money.
1,815,619	800	Other	"Investment in terror".
1,535,341	45.0	Other	"Investment in terror".
1,459,182	160	Ransomware	"Investment in terror".
1,378,975	800	Other	"Investment in terror".
1,259,824	0.50	Other	"Investment in terror".
1,030,376	505	Other	"Inhumane" bank account theft via remote desktop.
724,340	1,150	Other	"Investment in terror", begs for treatment money.

Top address has received 3M bitcoins ~ $79B

- Comparable to the GDP of Luxemburg
- Trading investment scam

Organized Bitcoin scam group

- Worldwide
- "Investment in terror"

How Successful is Each Address? / Big Hitters

Table 4.1: Overview of the top-10 highest receiving reported addresses.

Received [BTC]	Median [BTC]	Category	Description
3,048,040	40.0	Other	Trading investment scam.
2,845,086	18.0	Other	Foreign exchange trading scam, "investment in terror".
2,009,608	25.0	Other	"Investment in terror", begs for treatment money.
1,815,619	800	Other	"Investment in terror".
1,535,341	45.0	Other	"Investment in terror".
1,459,182	160	Ransomware	"Investment in terror".
1,378,975	800	Other	"Investment in terror".
1,259,824	0.50	Other	"Investment in terror".
1,030,376	505	Other	"Inhumane" bank account theft via remote desktop.
724,340	1,150	Other	"Investment in terror", begs for treatment money.

Top address has received 3M bitcoins ~ $79B

- Comparable to the GDP of Luxemburg
- Trading investment scam called "CapitalBullTrade"

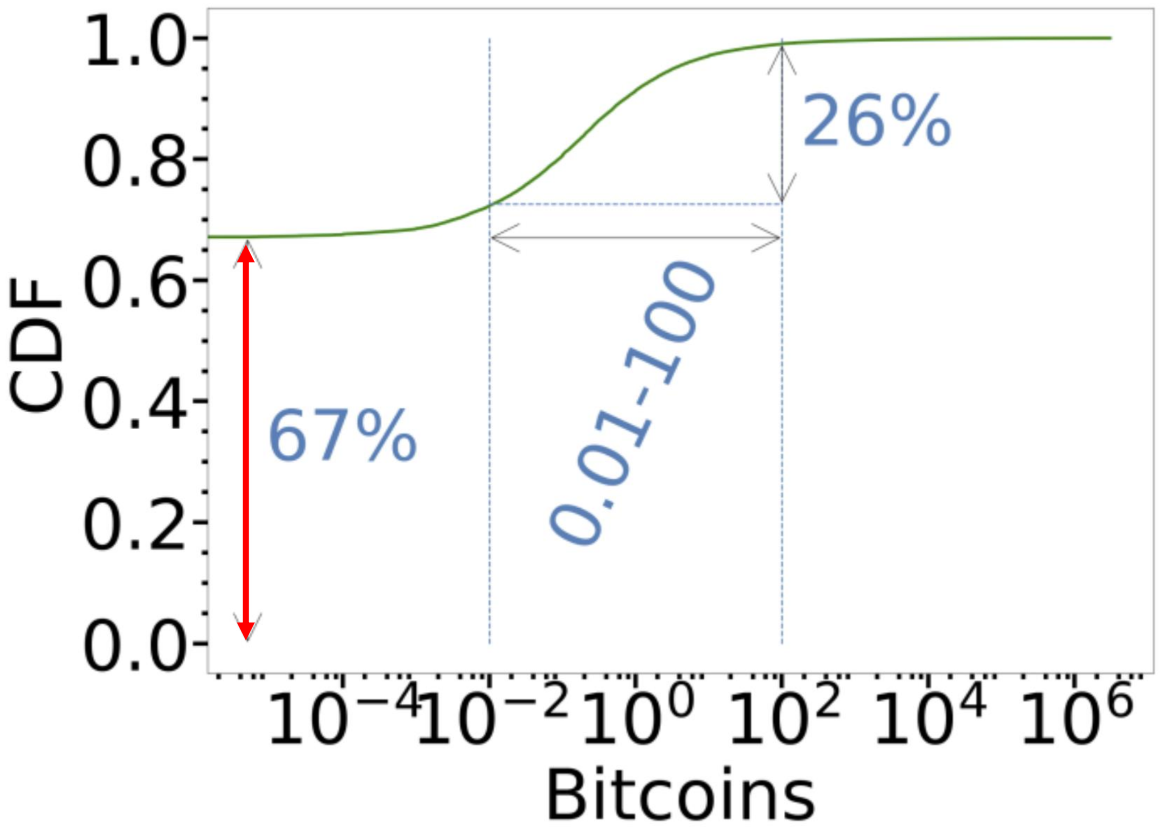
Organized Bitcoin scam group

- Worldwide
- "Investment in terror"

How Successful is Each Address? / The Most Common Cases

67% of the reported addresses received *no bitcoins at all.*

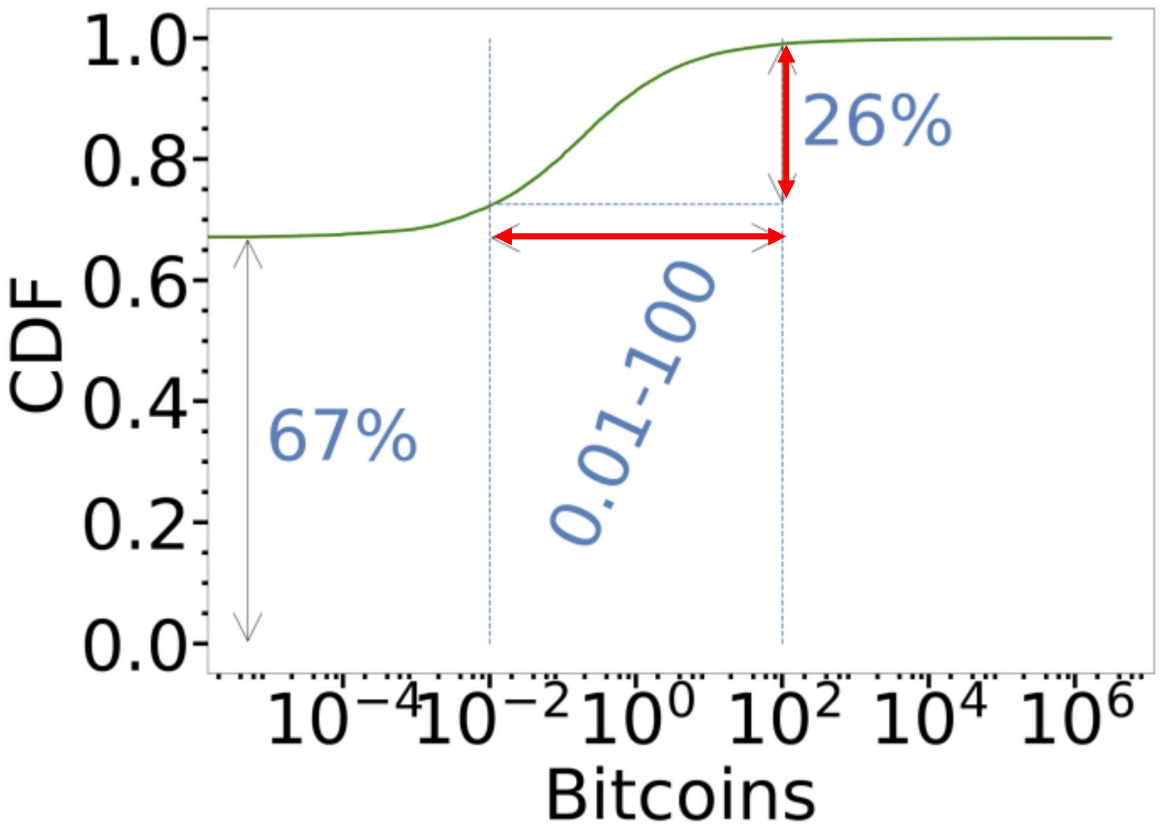
Among the addresses that received some funds, most received 0.01–100 bitcoins (~ $260–$2.6M).



How Successful is Each Address? / The Most Common Cases

67% of the reported addresses received *no bitcoins at all.*

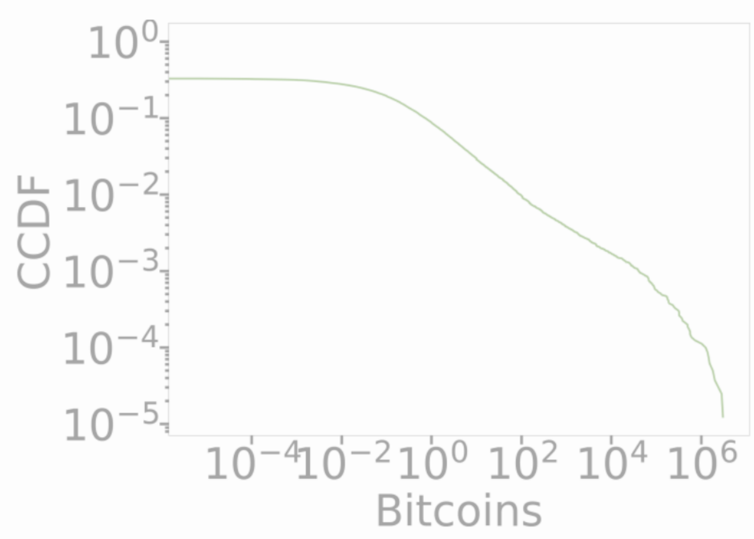
Among the addresses that received some funds, most received 0.01–100 bitcoins (~ $260–$2.6M).



How Successful is Each Address? / Heavy-Tailed Distribution

The high skew previously witnessed suggest a heavy-tailed distribution.

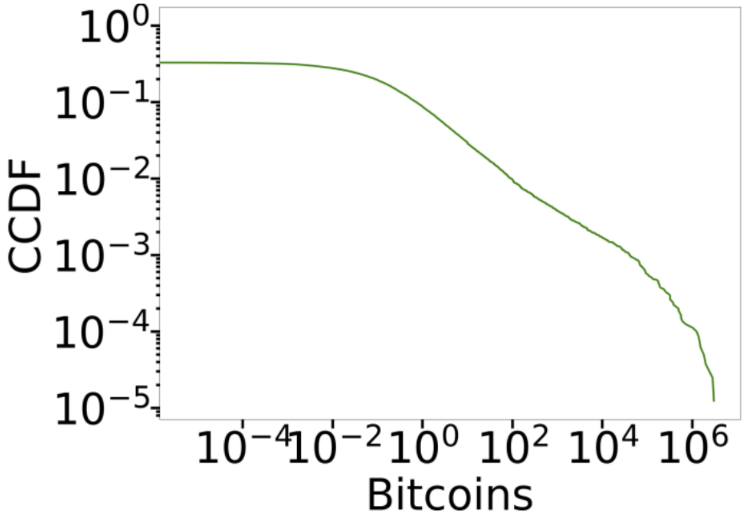
- The CCDF confirms this
- The curvature toward the end suggest that the tail is not a power law



How Successful is Each Address? / Heavy-Tailed Distribution

The high skew previously witnessed suggest a heavy-tailed distribution.

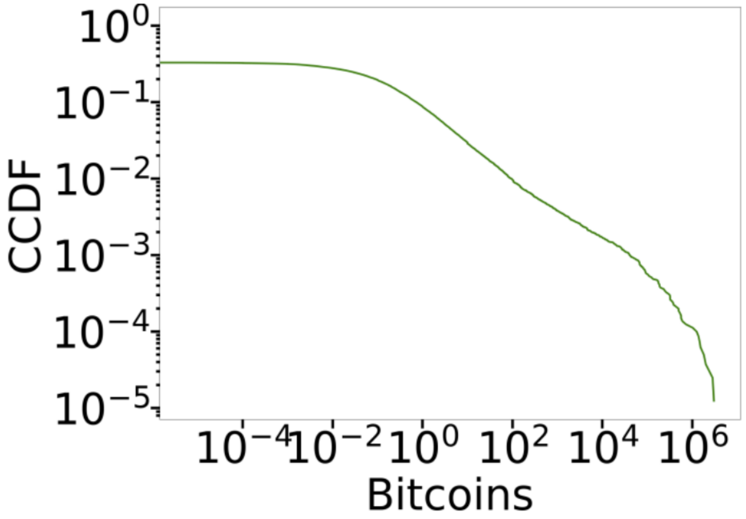
- The CCDF confirms this
- The curvature toward the end suggest that the tail is not a power law



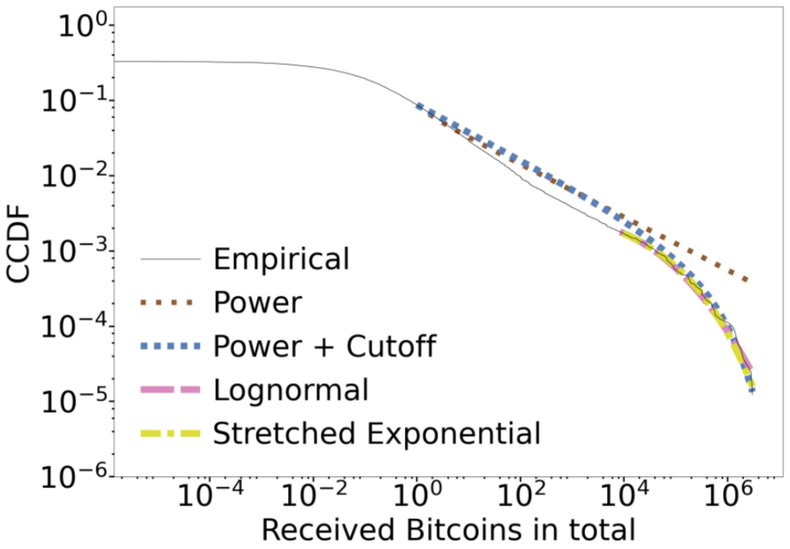
How Successful is Each Address? / Heavy-Tailed Distribution

The high skew previously witnessed suggest a heavy-tailed distribution.

- The CCDF confirms this
- The curvature toward the end suggest that the tail is not a power law

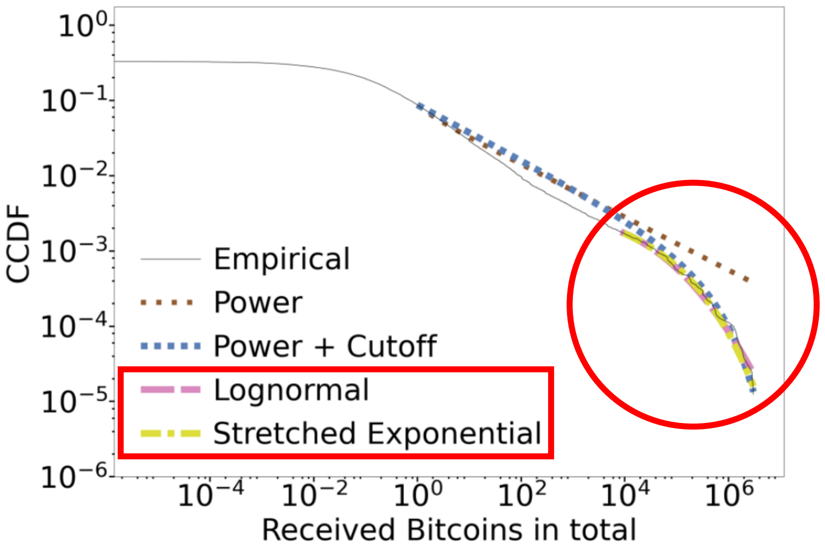


Model of the Tail Distribution



Distribution	$f(x)$	$x_{\min}$	Shape parameter(s)	KS
Power law	$f(x) = Cx^{-\alpha}$	1	$\alpha = 1.35$	0.057
Power + Cutoff	$f(x) = Cx^{-\alpha}e^{-x/\beta}$	1	$\alpha = 1.36, \beta = 4.71 \cdot 10^{-7}$	0.099
Lognormal	$\frac{1}{x} \cdot e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}$	8,372	$\mu = 9.72, \sigma = 2.17$	0.035
Stretched Exponential	$\lambda\beta x^{\beta-1}e^{-\lambda x^{\beta}}$	9,444	$\beta = 0.30$	0.036

Model of the Tail Distribution

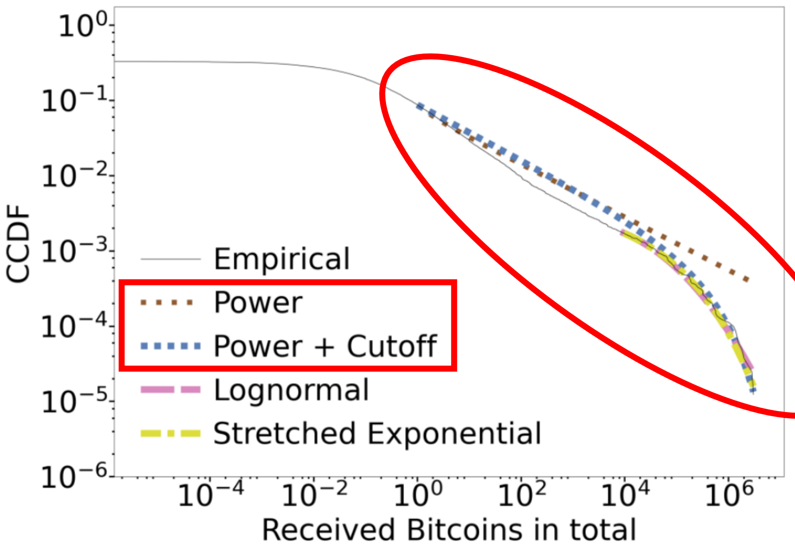


Lognormal and stretched exponential gave best fit for their range, but only cover the very end of the tail.

The power-law with exponential cutoff fits best visually, but the pure power-law has a lower Kolmogorov-Smirnow distance.

Distribution	$f(x)$	$x_{\min}$	Shape parameter(s)	KS
Power law	$f(x) = Cx^{-\alpha}$	1	$\alpha = 1.35$	0.057
Power + Cutoff	$f(x) = Cx^{-\alpha}e^{-x/\beta}$	1	$\alpha = 1.36, \beta = 4.71 \cdot 10^{-7}$	0.099
Lognormal	$\frac{1}{x} \cdot e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}$	8,372	$\mu = 9.72, \sigma = 2.17$	0.035
Stretched Exponential	$\lambda\beta x^{\beta-1}e^{-\lambda x^{\beta}}$	9,444	$\beta = 0.30$	0.036

Model of the Tail Distribution



Lognormal and stretched exponential gave best fit for their range, but only cover the very end of the tail.

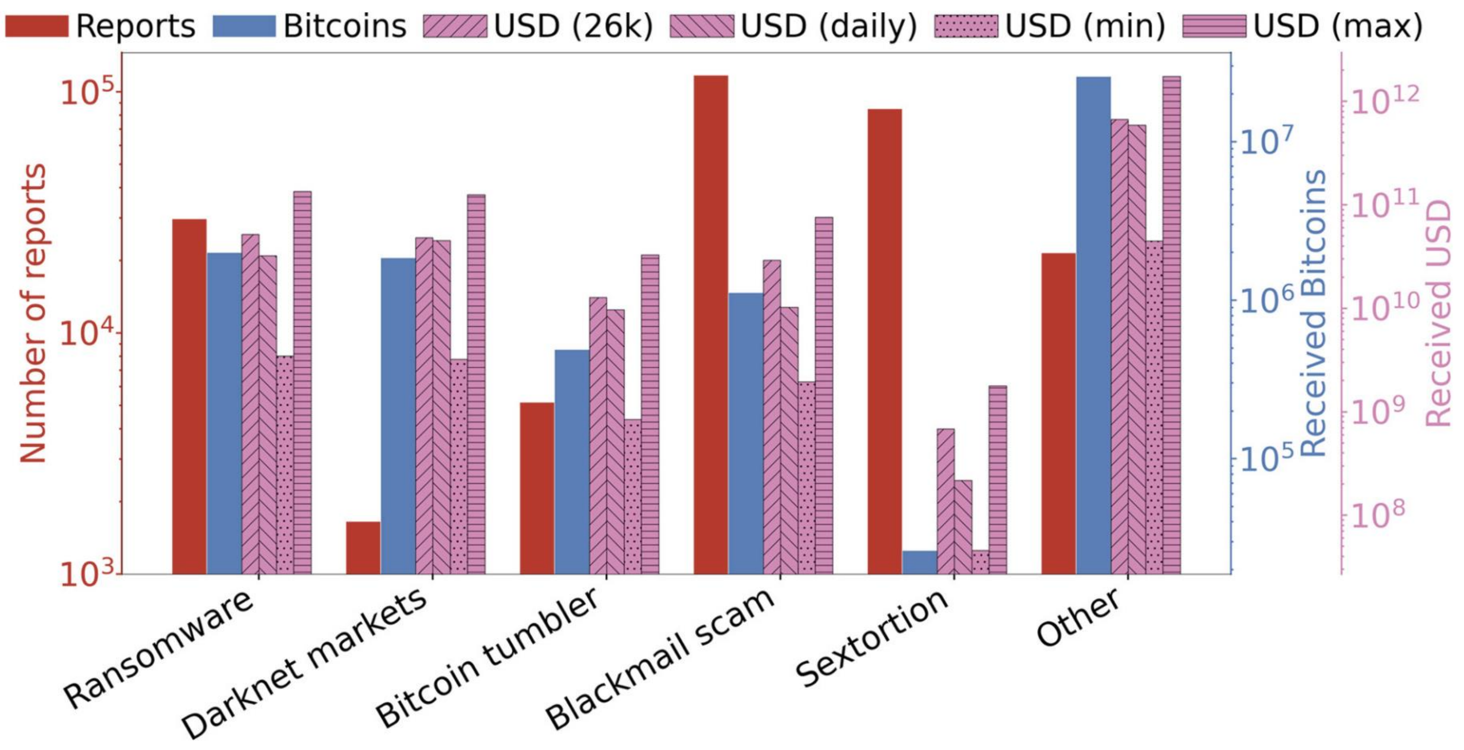
The power-law with exponential cutoff fits best visually, but the pure power-law has a lower Kolmogorov-Smirnow distance.

Distribution	$f(x)$	$x_{\min}$	Shape parameter(s)	KS
Power law	$f(x) = Cx^{-\alpha}$	1	$\alpha = 1.35$	0.057
Power + Cutoff	$f(x) = Cx^{-\alpha}e^{-x/\beta}$	1	$\alpha = 1.36, \beta = 4.71 \cdot 10^{-7}$	0.099
Lognormal	$\frac{1}{x} \cdot e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}$	8,372	$\mu = 9.72, \sigma = 2.17$	0.035
Stretched Exponential	$\lambda\beta x^{\beta-1}e^{-\lambda x^{\beta}}$	9,444	$\beta = 0.30$	0.036

Category-Based High-Level Characterization

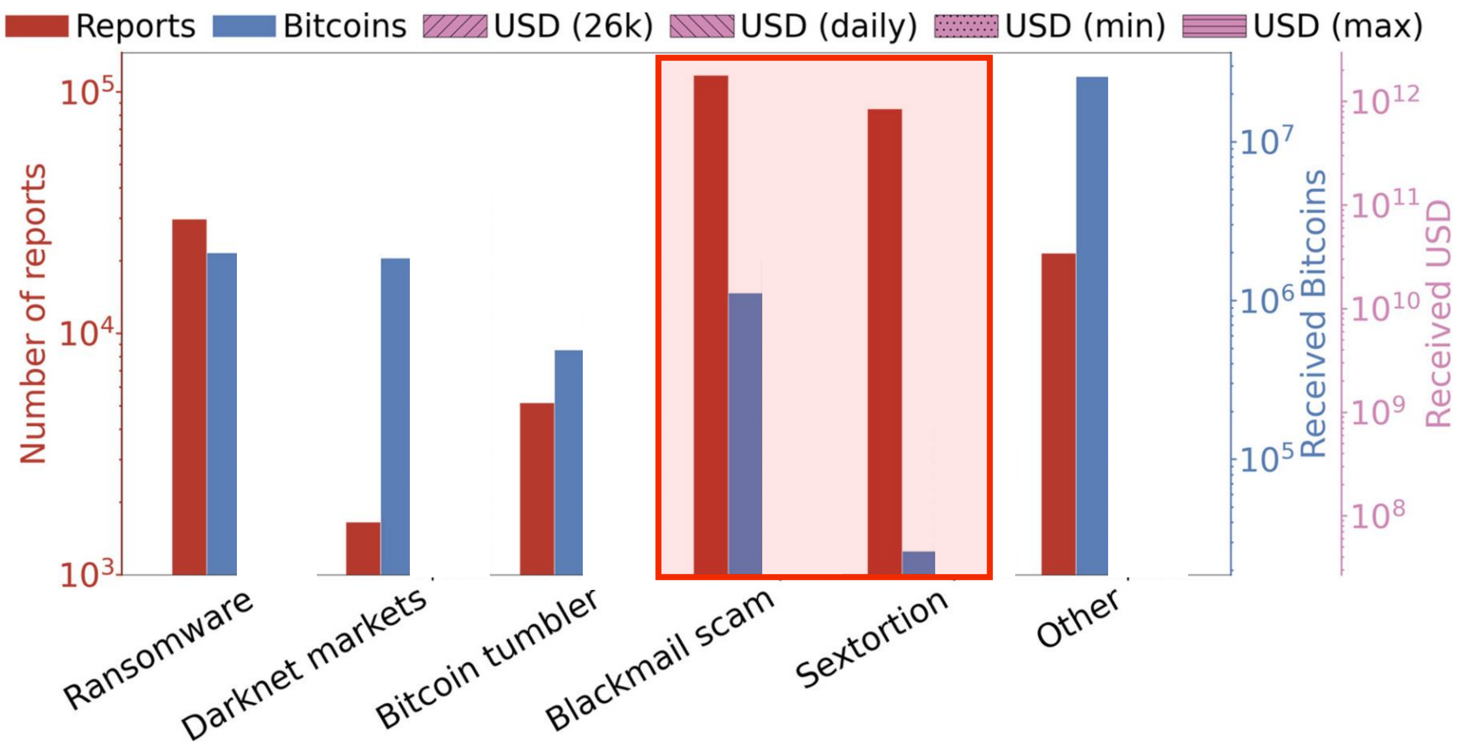
- High-level comparison
- Transactions-based analysis
- Report frequencies

High-Level Comparison



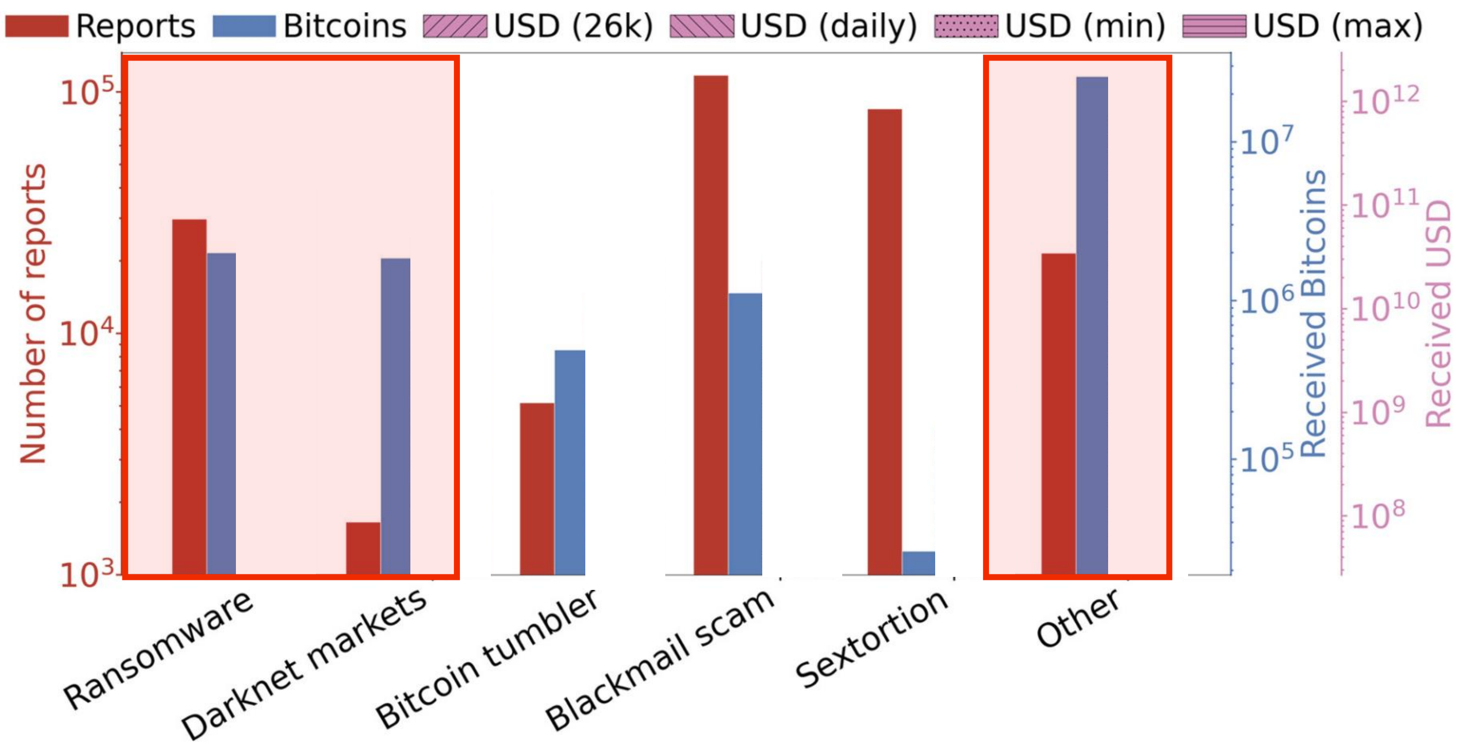
Note: log-scale

High-Level Comparison



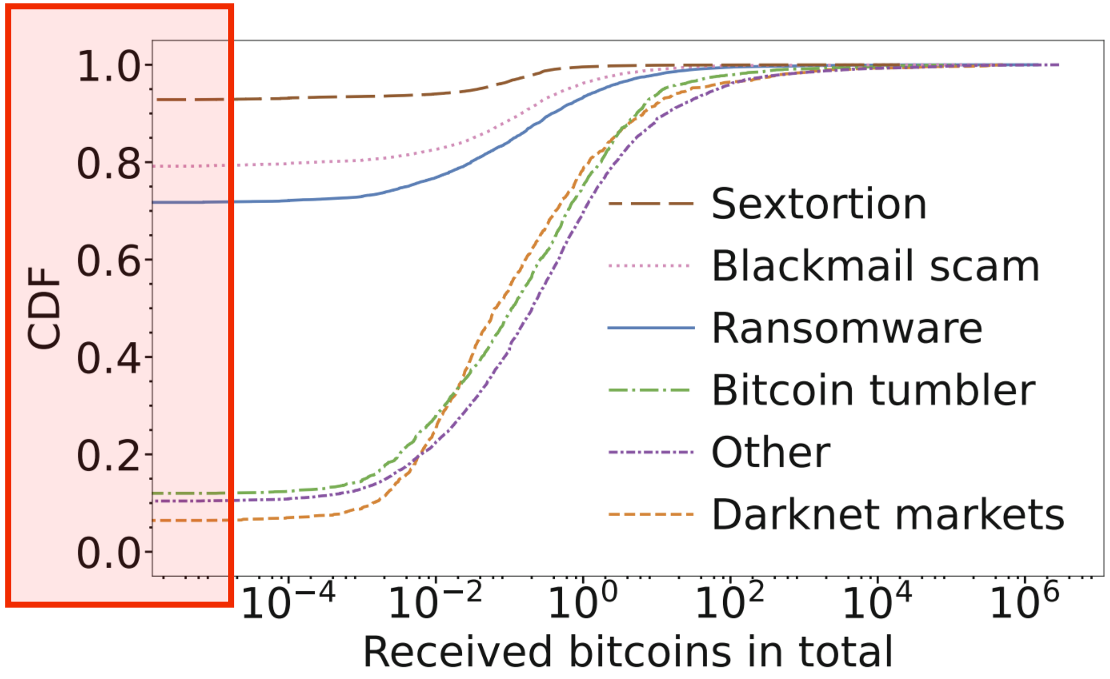
- *Blackmail scam* and *Sextortion* are the most highly reported (red) but also among the three lowest receiving categories (blue).
- *Ransomware*, *Darknet markets*, and particularly *Other* receive much more (blue), but are reported less (red).

High-Level Comparison



- *Blackmail scam* and *Sextortion* are the most highly reported (red) but also among the three lowest receiving categories (blue).
- *Ransomware*, *Darknet markets*, and particularly *Other* receive much more (blue), but are reported less (red).

Fraction of Addresses Not Attracting Any Funds



Fraction receiving no bitcoins:

• Sextortion	93%
• Blackmail scam	79%
• Ransomware	72%
• Bitcoin tumbler	12%
• Other	10%
• Darknet markets	6%

Distribution comparisons

- Per-category basis, the CCDFs become significantly more power-law-like
- Power-law fitting confirms this
- The slopes are similar, instead the difference lies in the relative shift to each other

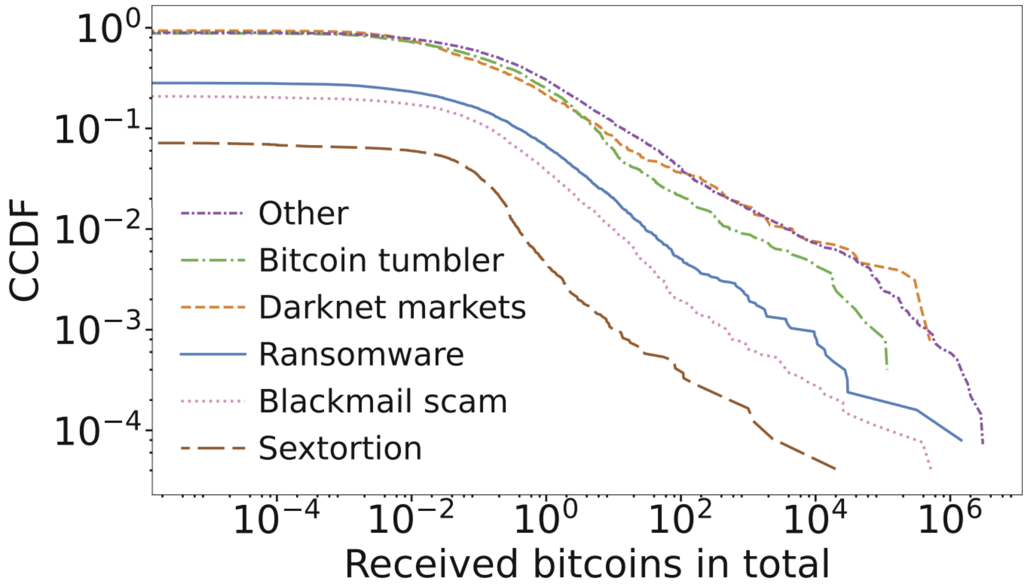


Table 4. Power-law fitting of per-category CCDFs.

Category	Slope estimate		Confidence interval
	$x_{\min}$	α (σ)	95%
Sextortion	1	1.423 (0.041)	$\alpha \pm 0.000518$
Blackmail	1	1.419 (0.013)	$\alpha \pm 0.000161$
Ransomw.	1	1.388 (0.013)	$\alpha \pm 0.000234$
Darknet	1	1.309 (0.019)	$\alpha \pm 0.00101$
Tumbler	1	1.391 (0.016)	$\alpha \pm 0.000612$
Other	1	1.329 (0.005)	$\alpha \pm 0.0000851$

Distribution comparisons

- Per-category basis, the CCDFs become significantly more power-law-like
- Power-law fitting confirms this
- The slopes are similar, instead the difference lies in the relative shift to each other

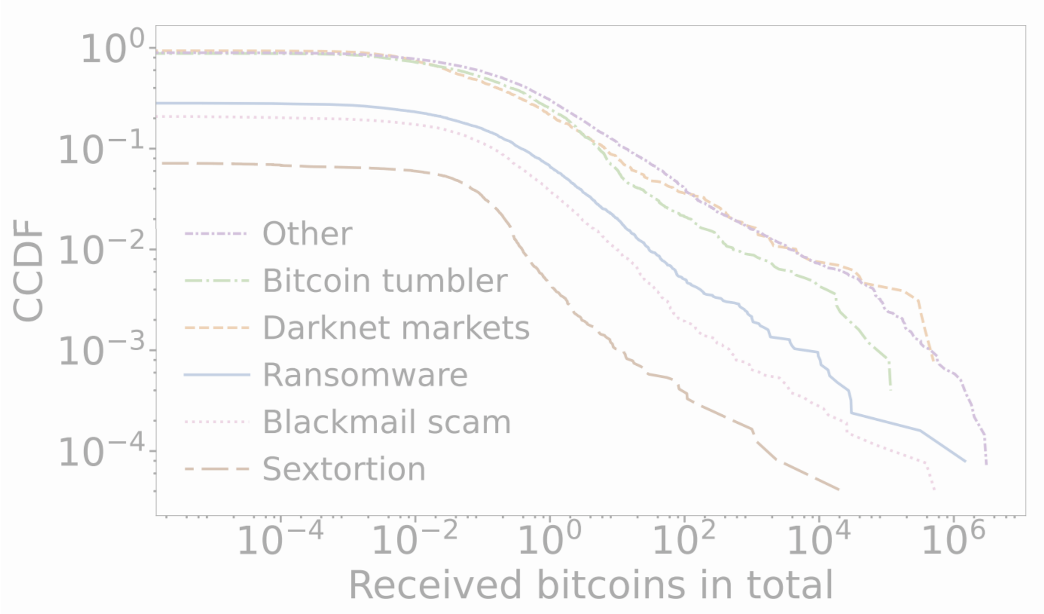


Table 4. Power-law fitting of per-category CCDFs.

Category	Slope estimate		Confidence interval
	$x_{\min}$	α (σ)	95%
Sextortion	1	1.423 (0.041)	$\alpha \pm 0.000518$
Blackmail	1	1.419 (0.013)	$\alpha \pm 0.000161$
Ransomw.	1	1.388 (0.013)	$\alpha \pm 0.000234$
Darknet	1	1.309 (0.019)	$\alpha \pm 0.00101$
Tumbler	1	1.391 (0.016)	$\alpha \pm 0.000612$
Other	1	1.329 (0.005)	$\alpha \pm 0.0000851$

Distribution comparisons

- Per-category basis, the CCDFs become significantly more power-law-like
- Power-law fitting confirms this
- The slopes are similar, instead the difference lies in the relative shift to each other

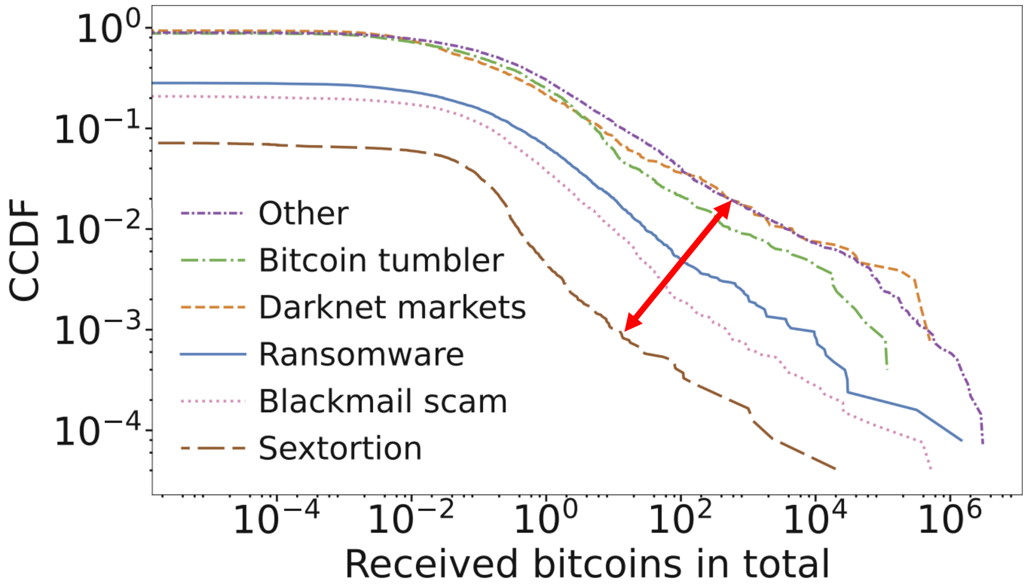
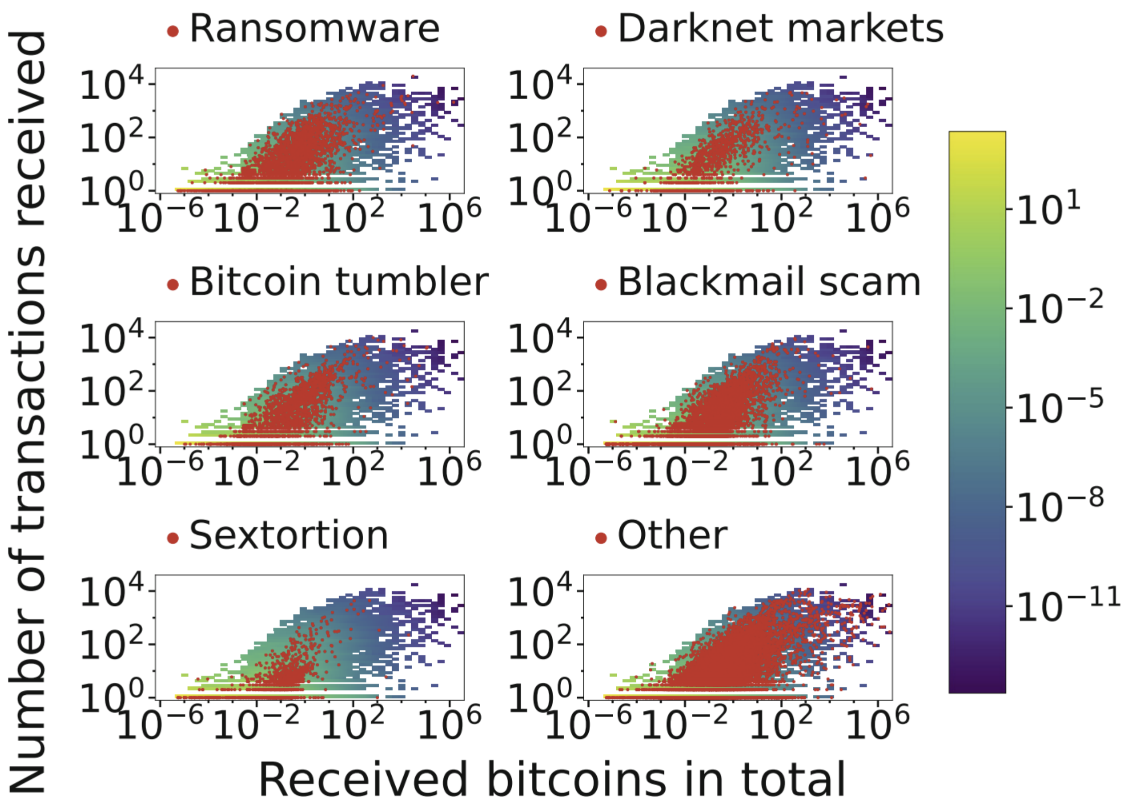


Table 4. Power-law fitting of per-category CCDFs.

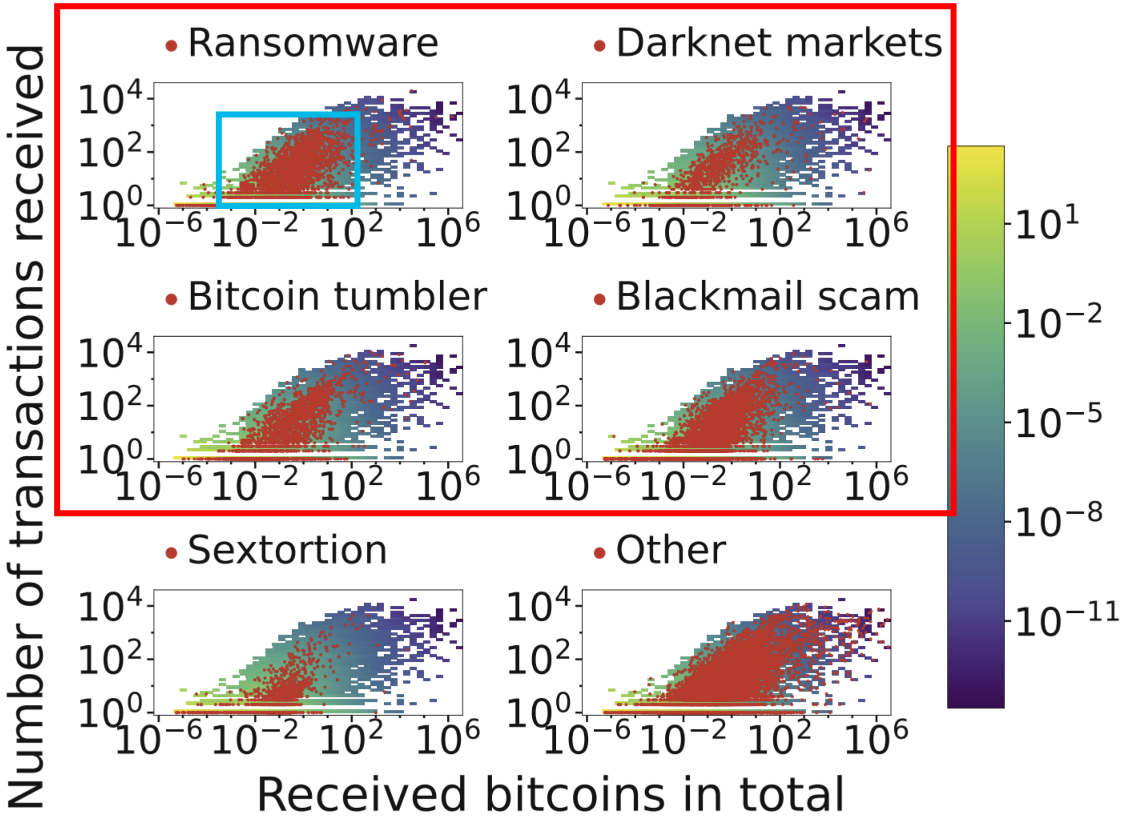
Category	Slope estimate		Confidence interval
	$x_{\min}$	α (σ)	95%
Sextortion	1	1.423 (0.041)	$\alpha \pm 0.000518$
Blackmail	1	1.419 (0.013)	$\alpha \pm 0.000161$
Ransomw.	1	1.388 (0.013)	$\alpha \pm 0.000234$
Darknet	1	1.309 (0.019)	$\alpha \pm 0.00101$
Tumbler	1	1.391 (0.016)	$\alpha \pm 0.000612$
Other	1	1.329 (0.005)	$\alpha \pm 0.0000851$

Transactions-Based Analysis



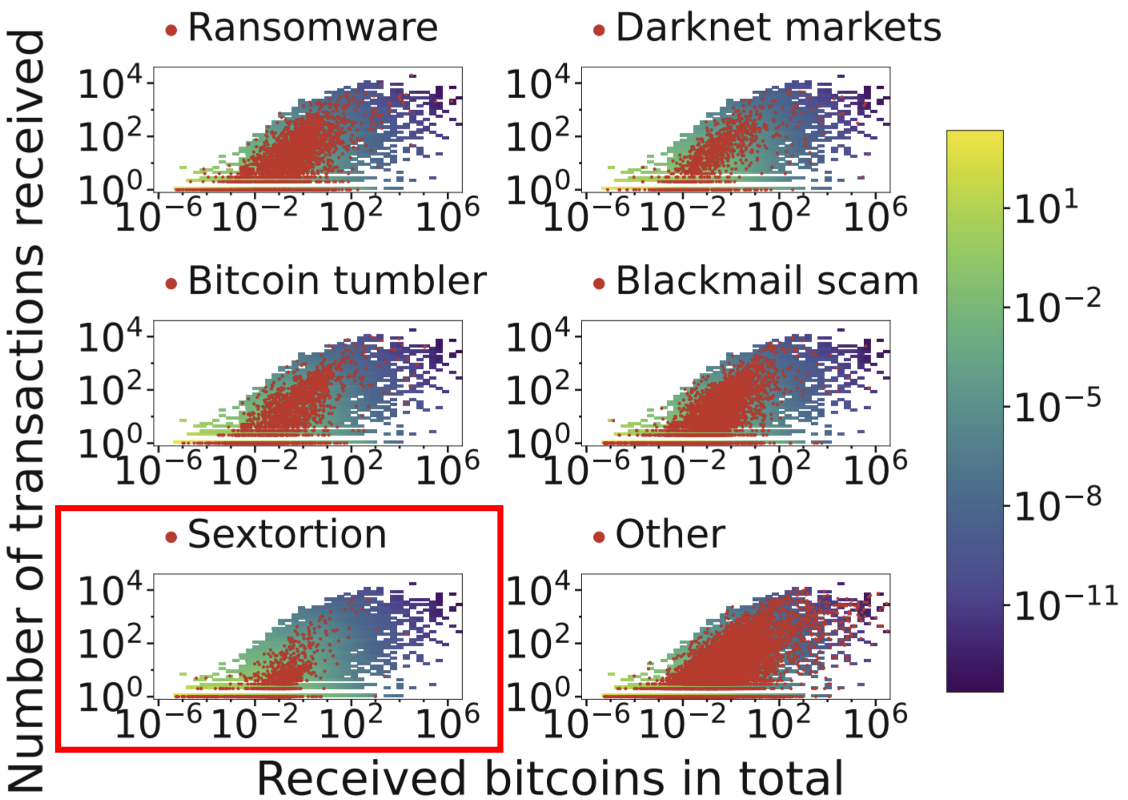
Transactions-Based Analysis

- The first four are similar, with a clear cluster receiving up-to 100 bitcoins spread over up-to 1K incoming transactions
- Sextortion addresses receive fewer bitcoins and fewer transactions
- "Other" includes many of the highest receiving addresses



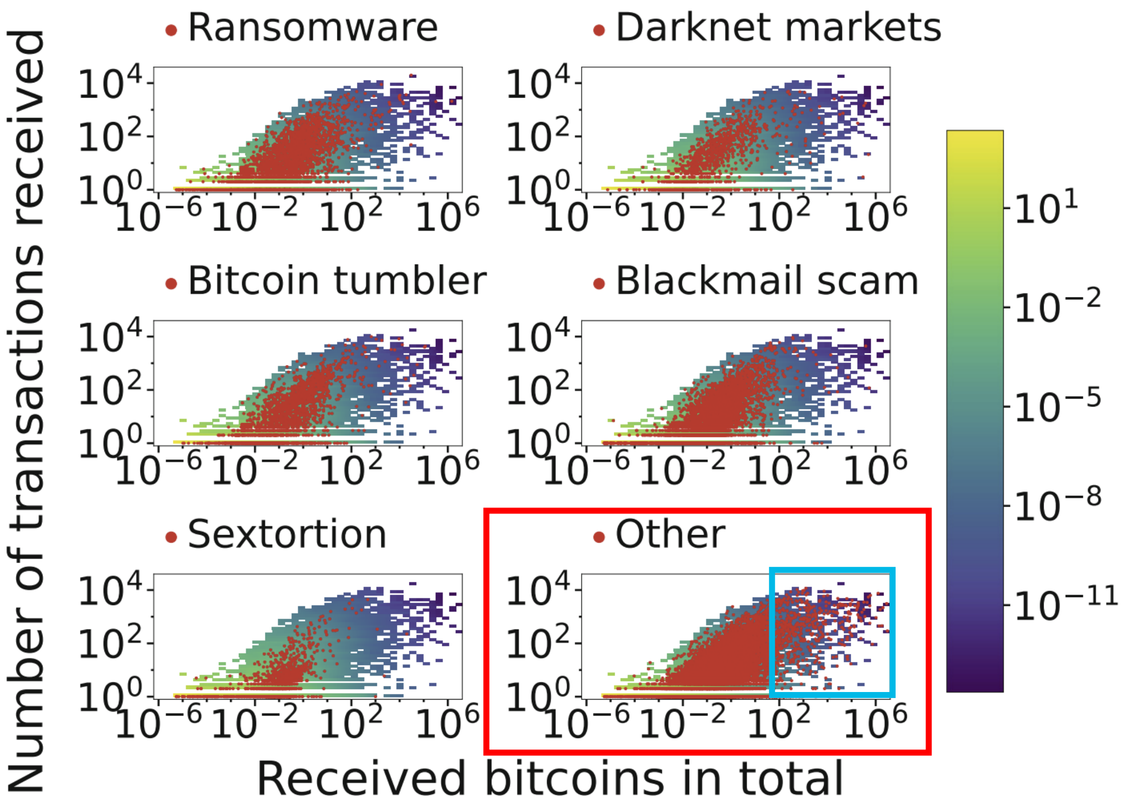
Transactions-Based Analysis

- The first four are similar, with a clear cluster receiving up-to 100 bitcoins spread over up-to 1K incoming transactions
- Sextortion addresses receive fewer bitcoins and fewer transactions
- “Other” includes many of the highest receiving addresses

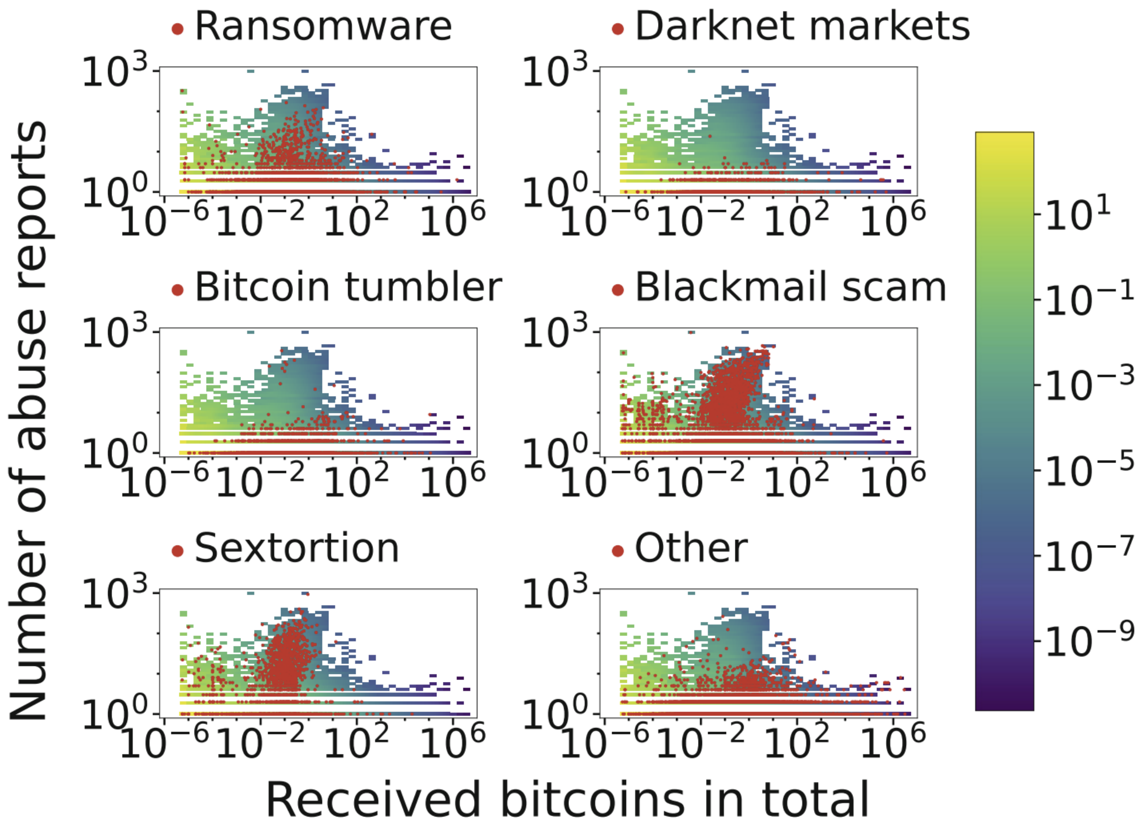


Transactions-Based Analysis

- The first four are similar, with a clear cluster receiving up-to 100 bitcoins spread over up-to 1K incoming transactions
- Sextortion addresses receive fewer bitcoins and fewer transactions
- “Other” includes many of the highest receiving addresses

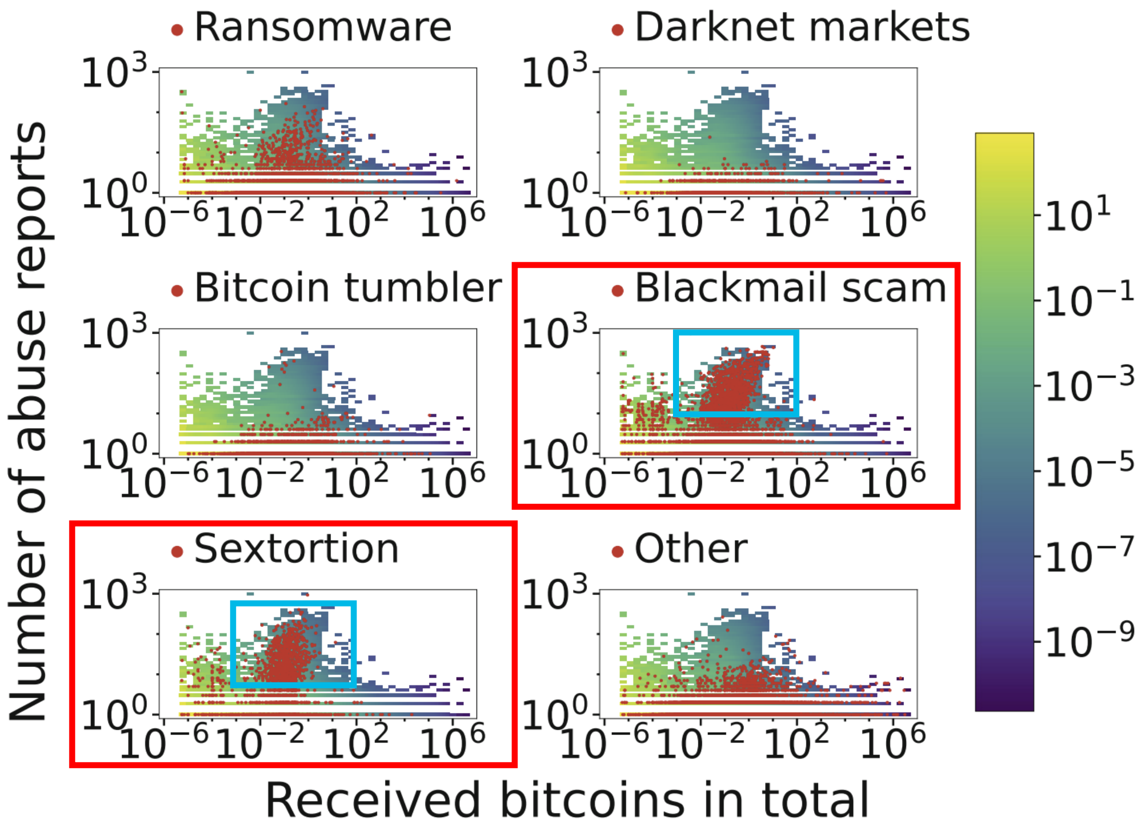


Report Frequencies



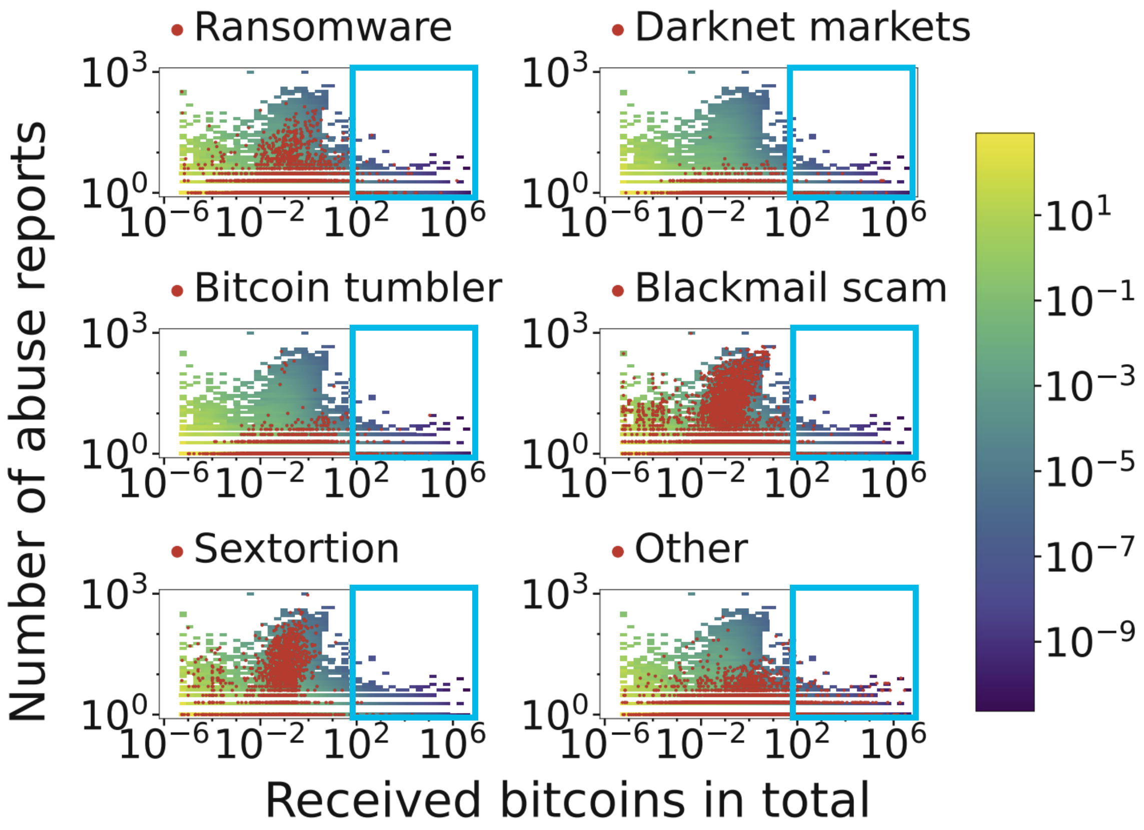
Report Frequencies

- Low-effort attacks targeting many users (Blackmail scams and Sextortions) have much higher report frequency
- Addresses best at attracting funds are not highly reported



Report Frequencies

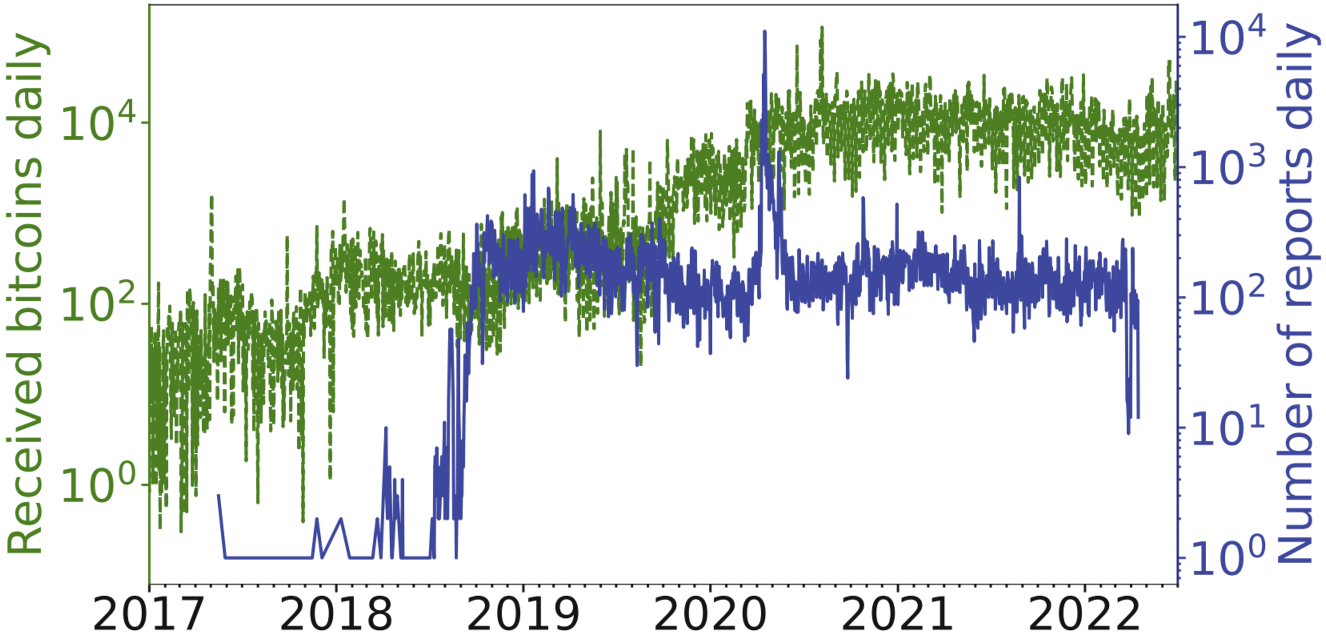
- Low-effort attacks targeting many users (Blackmail scams and Sextortions) have much higher report frequency
- Addresses best at attracting funds are not highly reported



Temporal Analysis

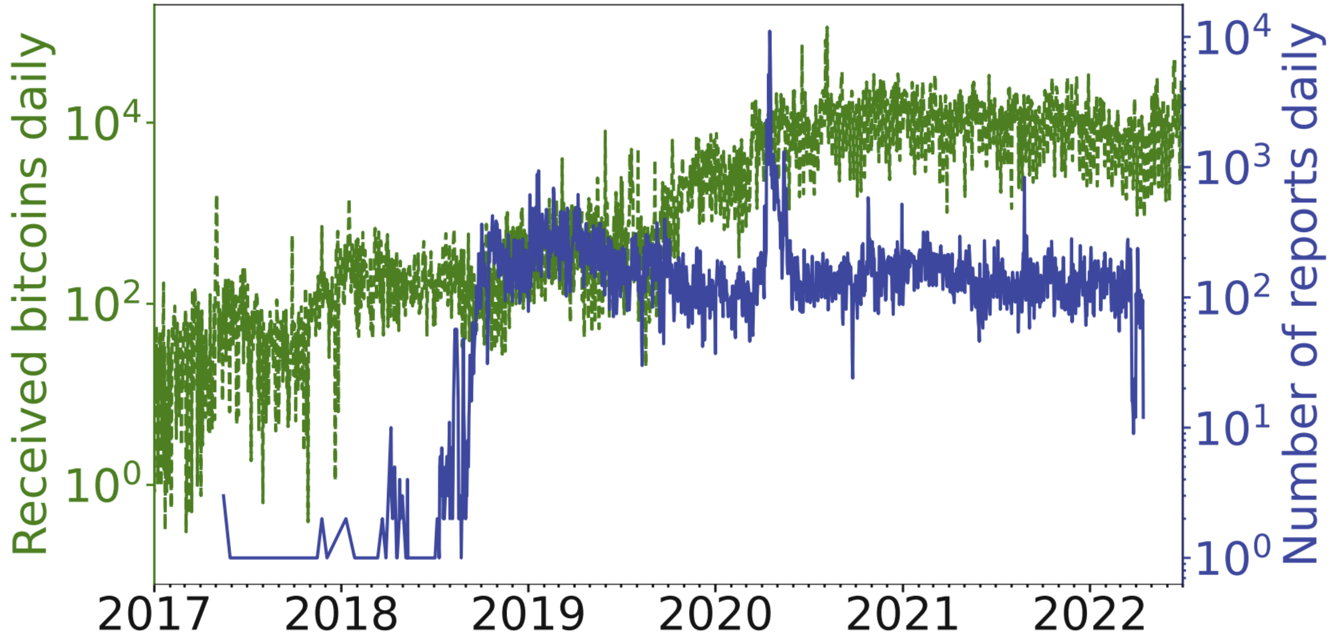
- Longitudinal timeline
- Time of the week
- Initial report date analysis

Longitudinal Timeline / High-Level



- Bitcoin Abuse Database was created in 2017 and gained popularity in late 2018. Remain relatively steady at an order of 100's per day.
- A substantial (roughly 100x) increase to 10,000 bitcoins received per day, over a three-year period (2019–2022).

Longitudinal Timeline / High-Level

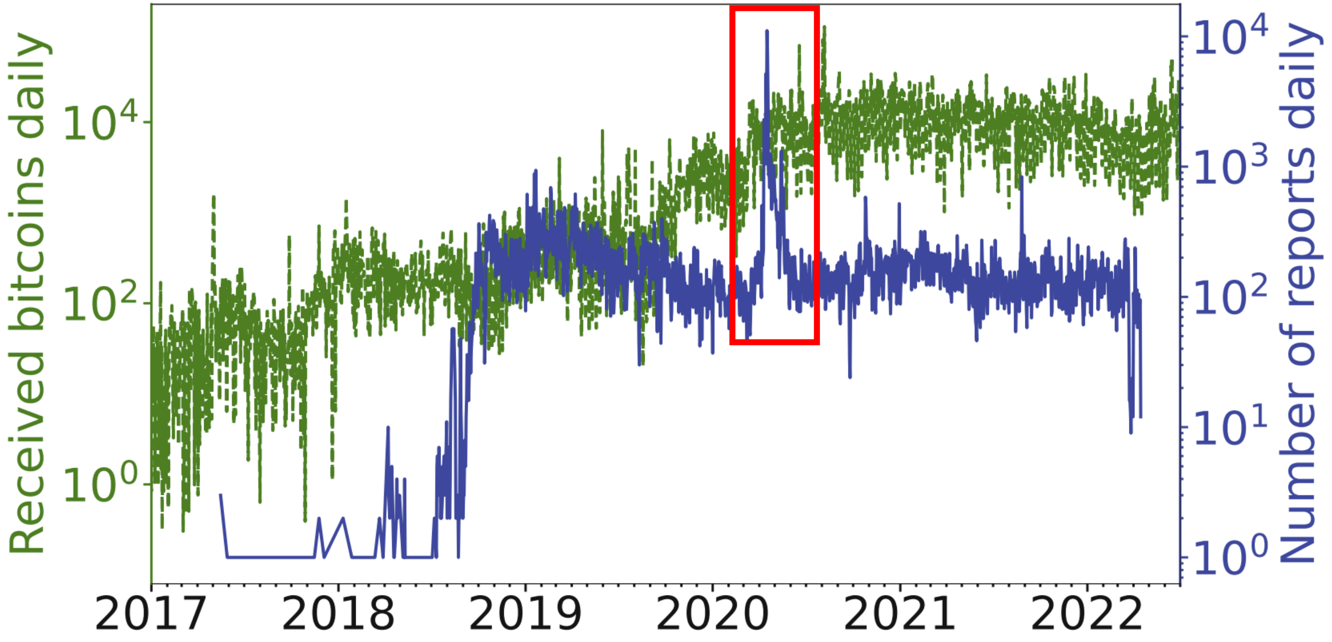


- Bitcoin Abuse Database was created in 2017 and gained popularity in late 2018. Remain relatively steady at an order of 100's per day.
- A substantial (roughly 100x) increase to 10,000 bitcoins received per day, over a three-year period (2019–2022).

Longitudinal Timeline / Noteworthy Spikes

The biggest spike was on April 16th, 2020.

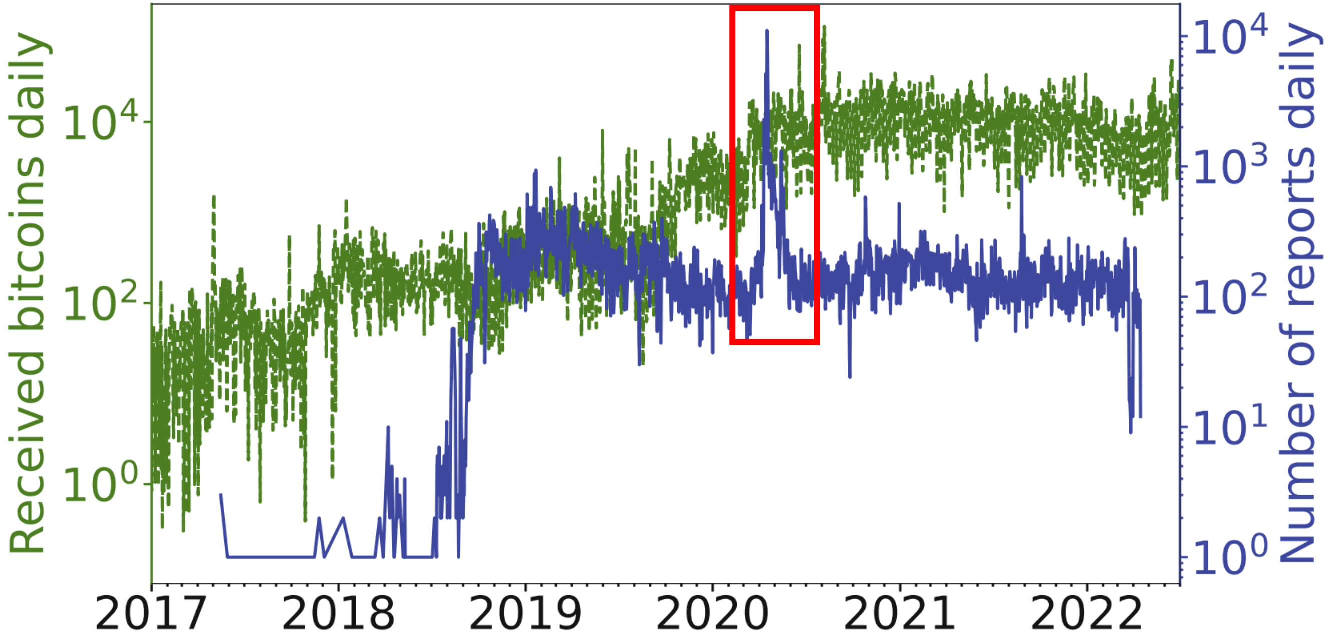
- 11K reports (roughly 100x daily average)
- Both US and Australian governments warn about a particular style of scam email around the same time
- A lot of the reports are clearly talking about the same type of attack



Longitudinal Timeline / Noteworthy Spikes

The biggest spike was on April 16th, 2020.

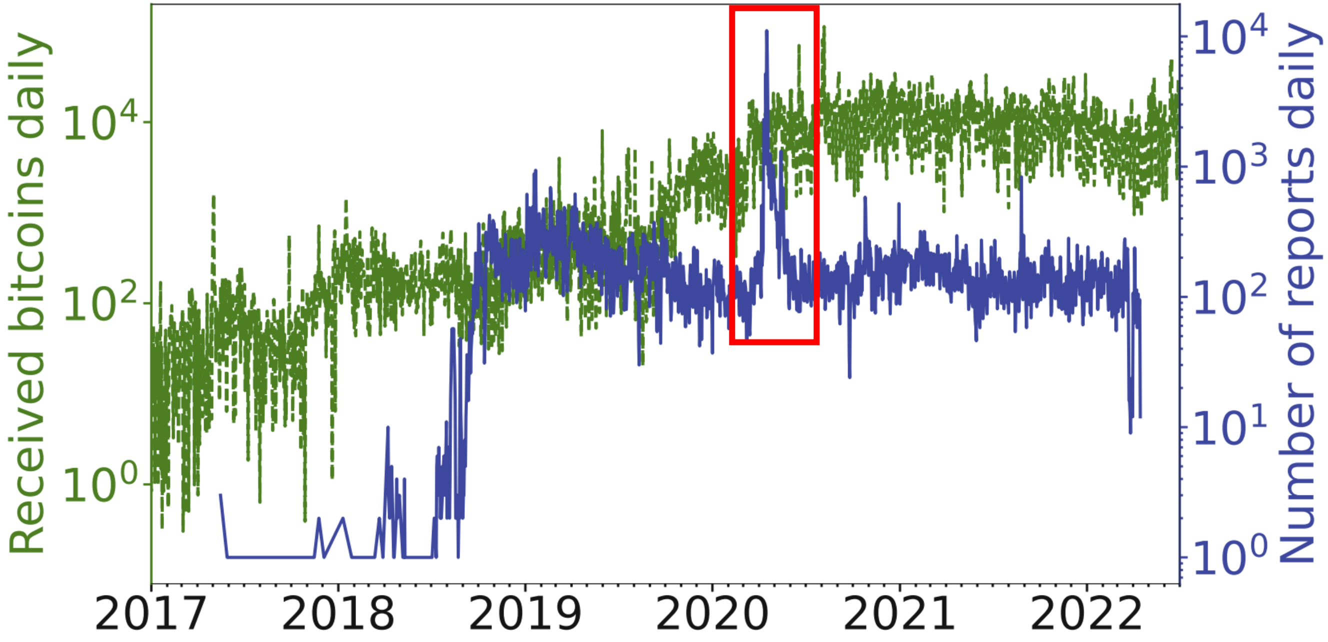
- 11K reports (roughly 100x daily average)
- Both US and Australian governments warn about a particular style of scam email around the same time
- A lot of the reports are clearly talking about the same type of attack



Longitudinal Timeline / Noteworthy Spikes

The biggest spike was on April 16th, 2020.

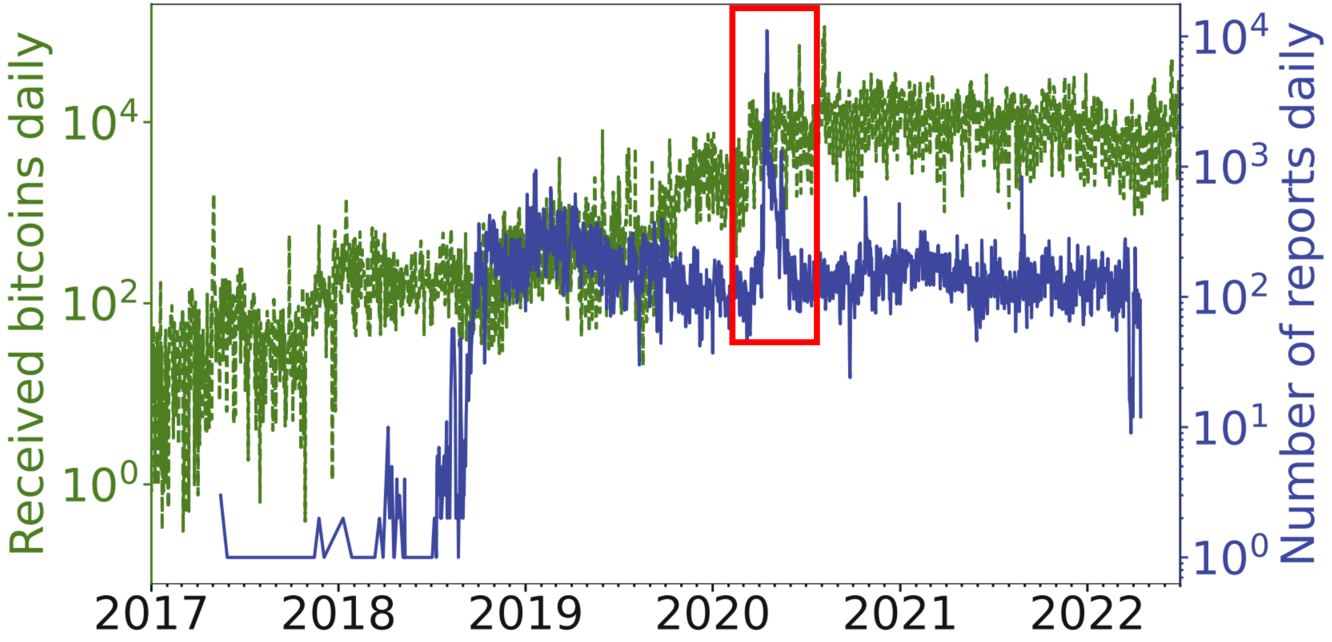
- 11K reports (roughly 100x daily average)
- Both US and Australian governments warn about a particular style of scam email around the same time
- A lot of the reports are clearly talking about the same type of attack



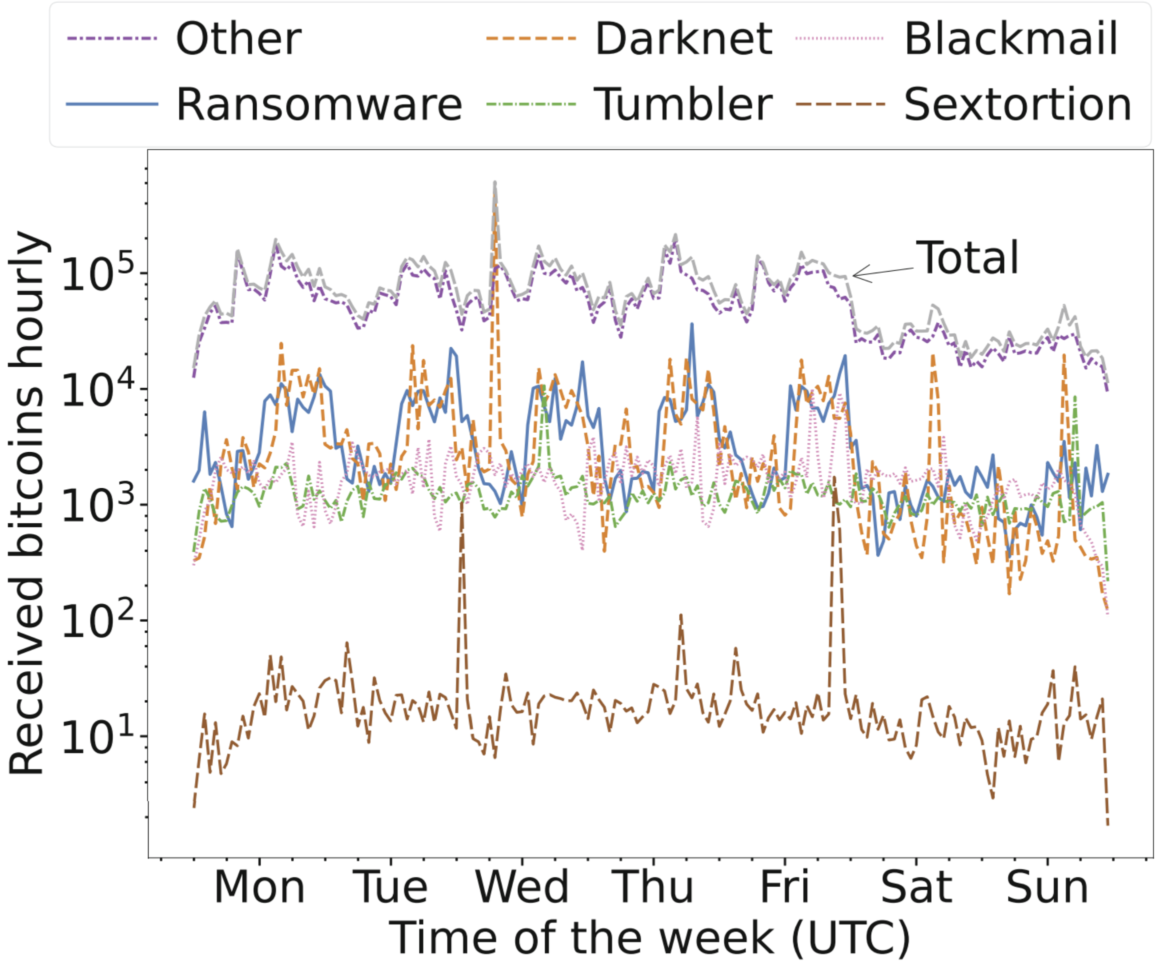
Longitudinal Timeline / Noteworthy Spikes

The biggest spike was on April 16th, 2020.

- 11K reports (roughly 100x daily average)
- Both US and Australian governments warn about a particular style of scam email around the same time
- A lot of the reports are clearly talking about the same type of attack

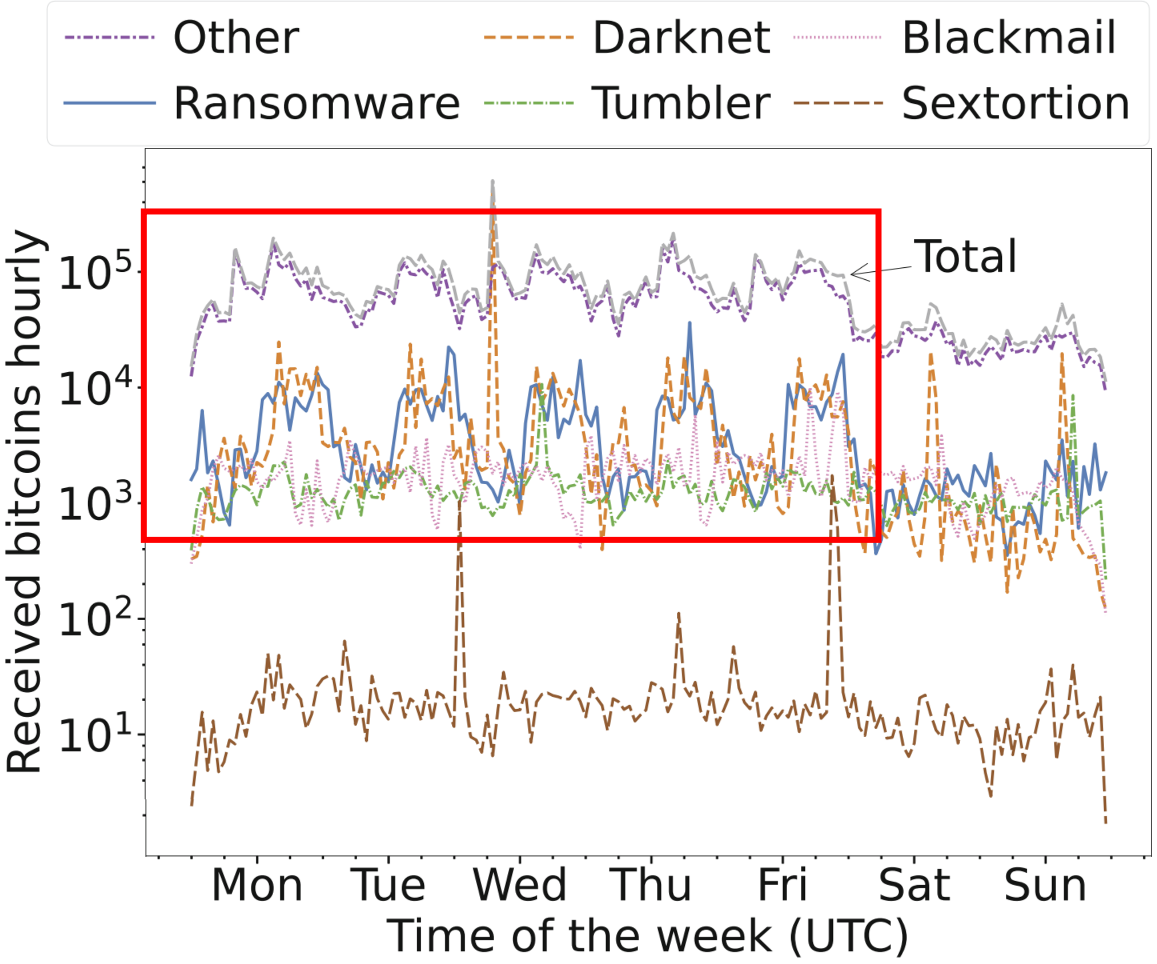


Time of the Week



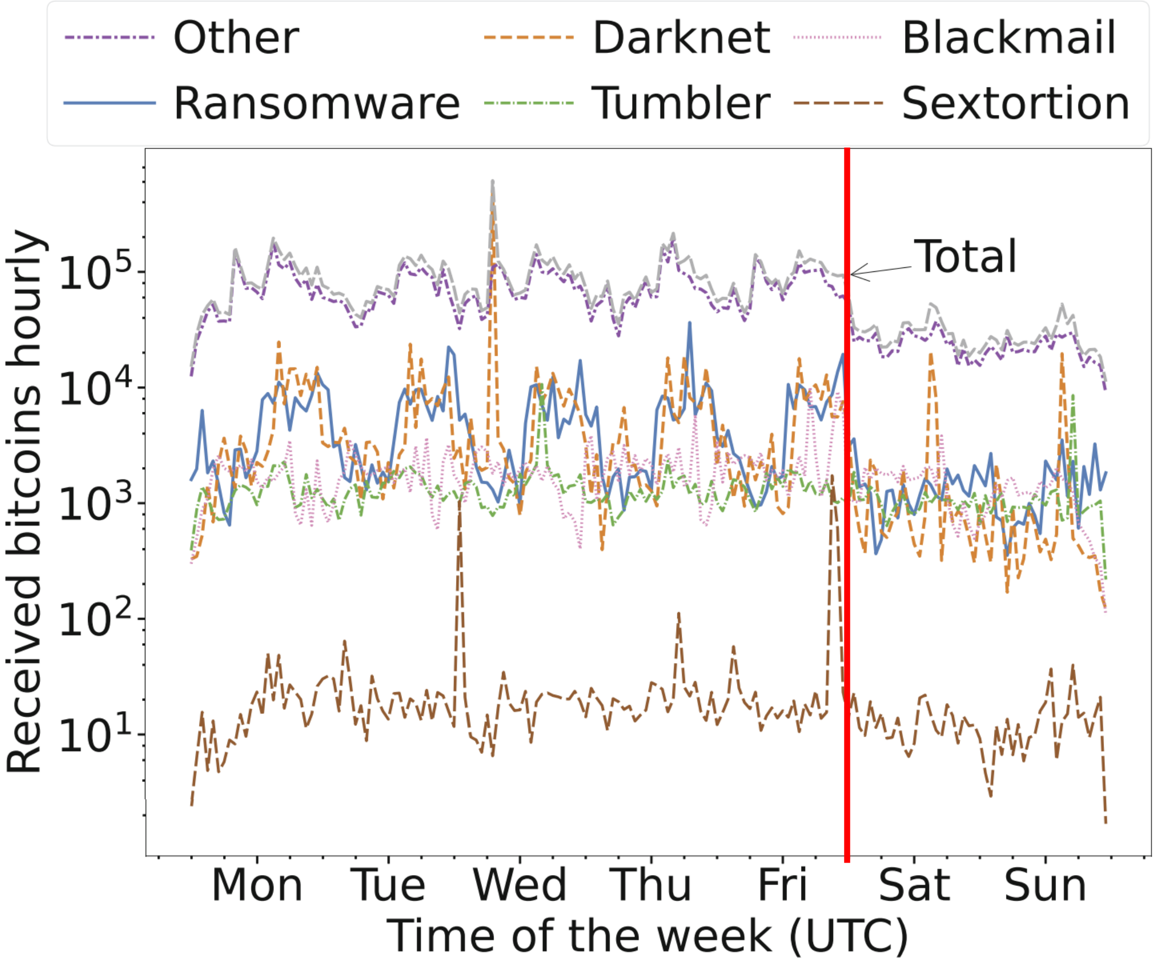
Time of the Week

- Much bigger volumes daytime/evenings (UTC)
- More funds being transferred during weekdays than weekends
- Bitcoin tumbler has the least pronounced pattern, possibly suggesting some level of automation
- The spikes are caused by large individual transactions



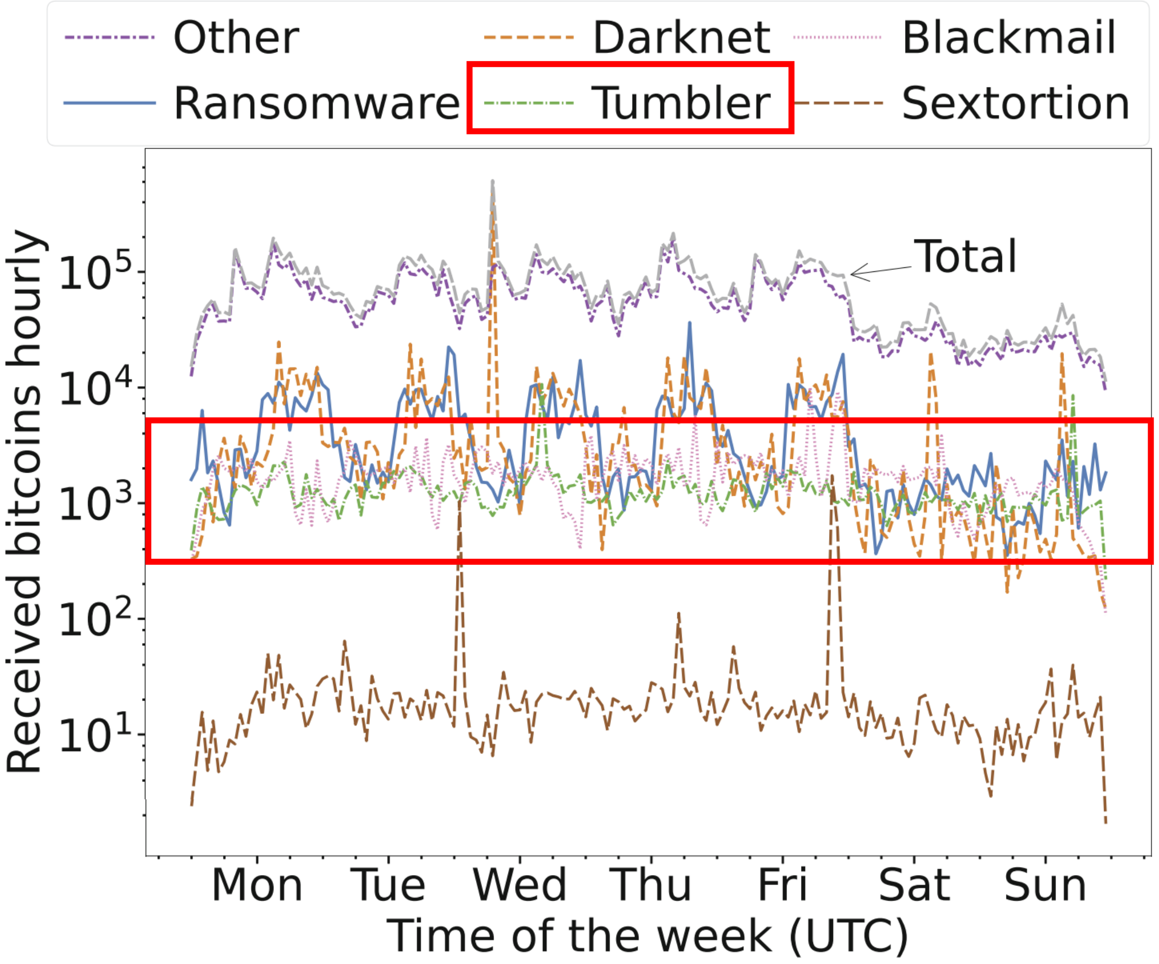
Time of the Week

- Much bigger volumes daytime/evenings (UTC)
- More funds being transferred during weekdays than weekends
- Bitcoin tumbler has the least pronounced pattern, possibly suggesting some level of automation
- The spikes are caused by large individual transactions



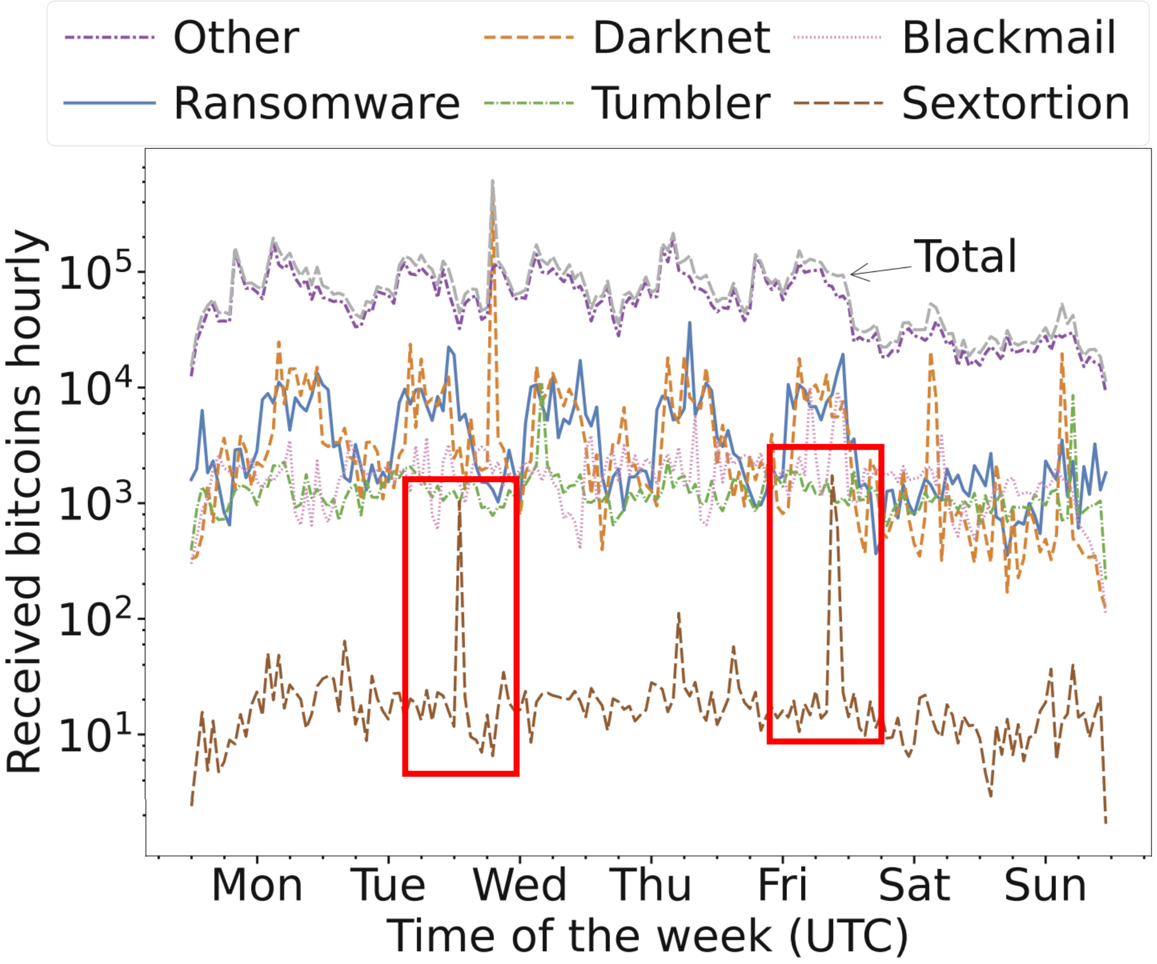
Time of the Week

- Much bigger volumes daytime/evenings (UTC)
- More funds being transferred during weekdays than weekends
- Bitcoin tumbler has the least pronounced pattern, possibly suggesting some level of automation
- The spikes are caused by large individual transactions

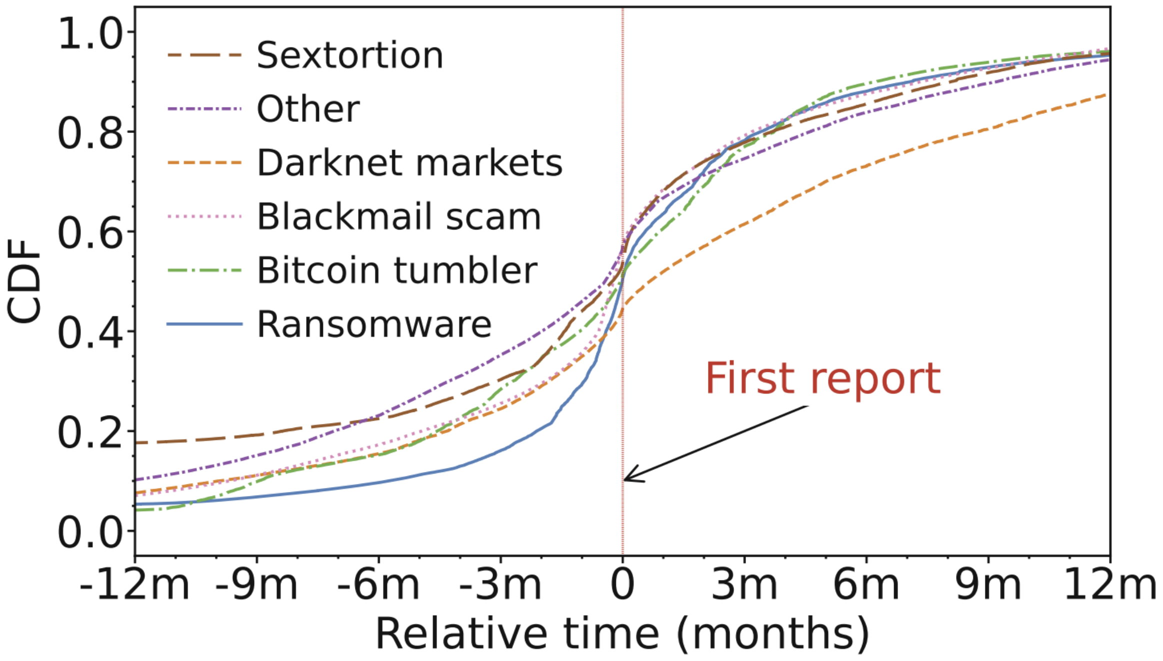


Time of the Week

- Much bigger volumes daytime/evenings (UTC)
- More funds being transferred during weekdays than weekends
- Bitcoin tumbler has the least pronounced pattern, possibly suggesting some level of automation
- The spikes are caused by large individual transactions

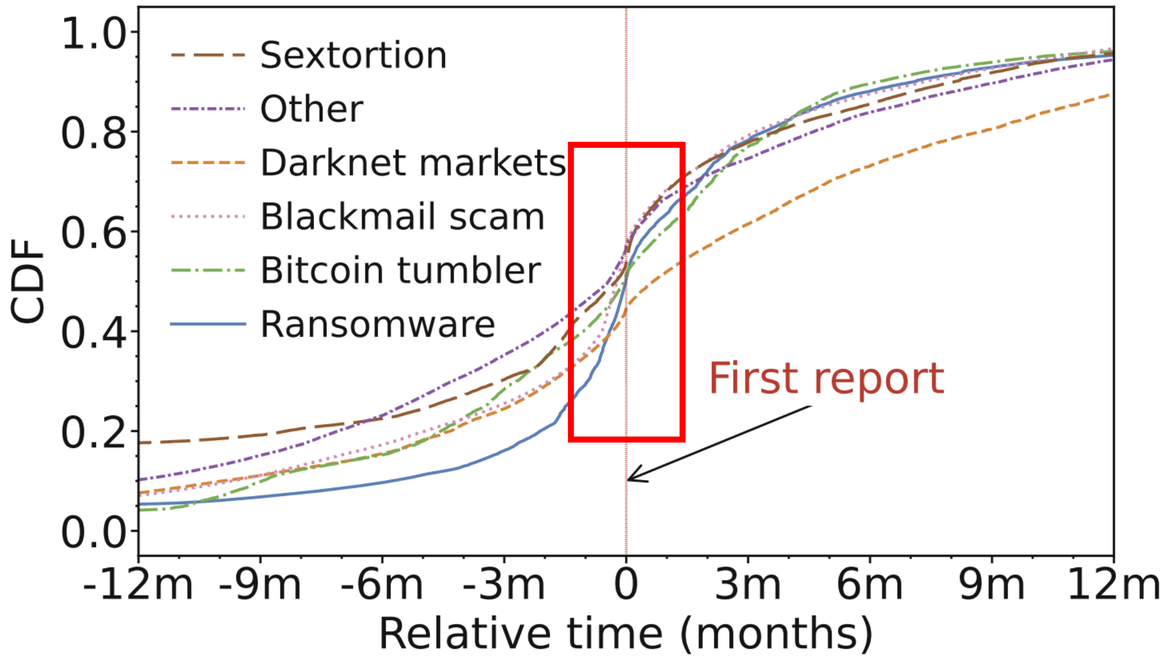


Initial Report Date Analysis



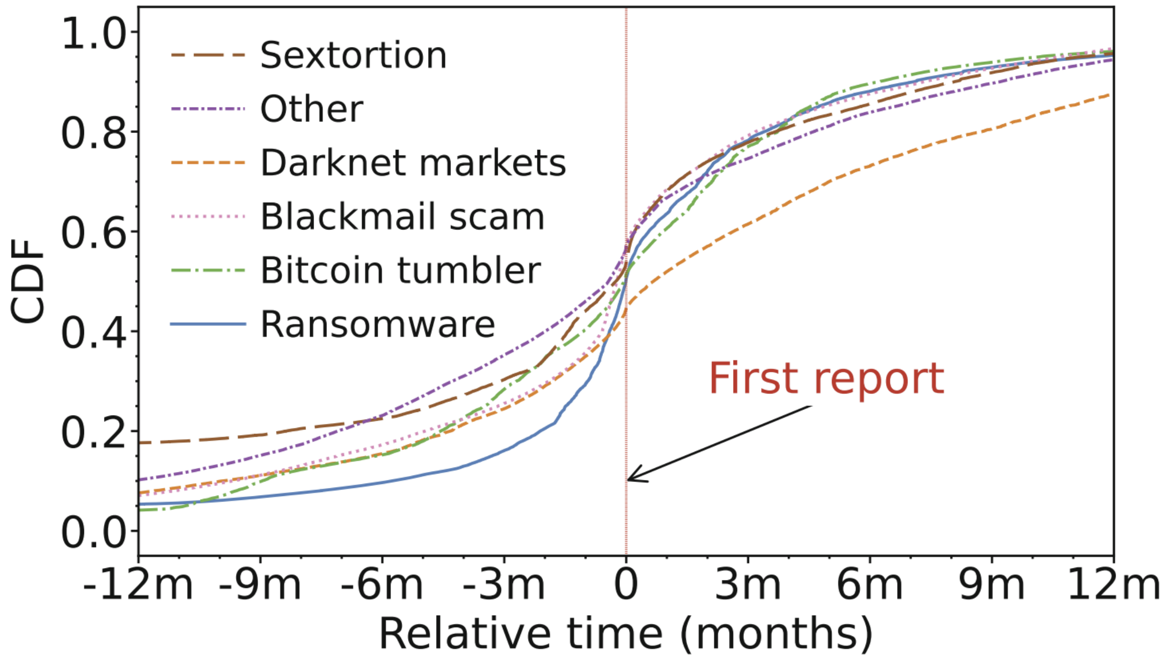
Initial Report Date Analysis

- Addresses are reported around the time that their incoming transaction count is high
- Indicating, the first report often is made around the time of the abuse's highest activity
- This may be a reflection of a significant portion of the addresses only being used for specific attacks
- Darknet markets follow this pattern the least, Ransomware the most



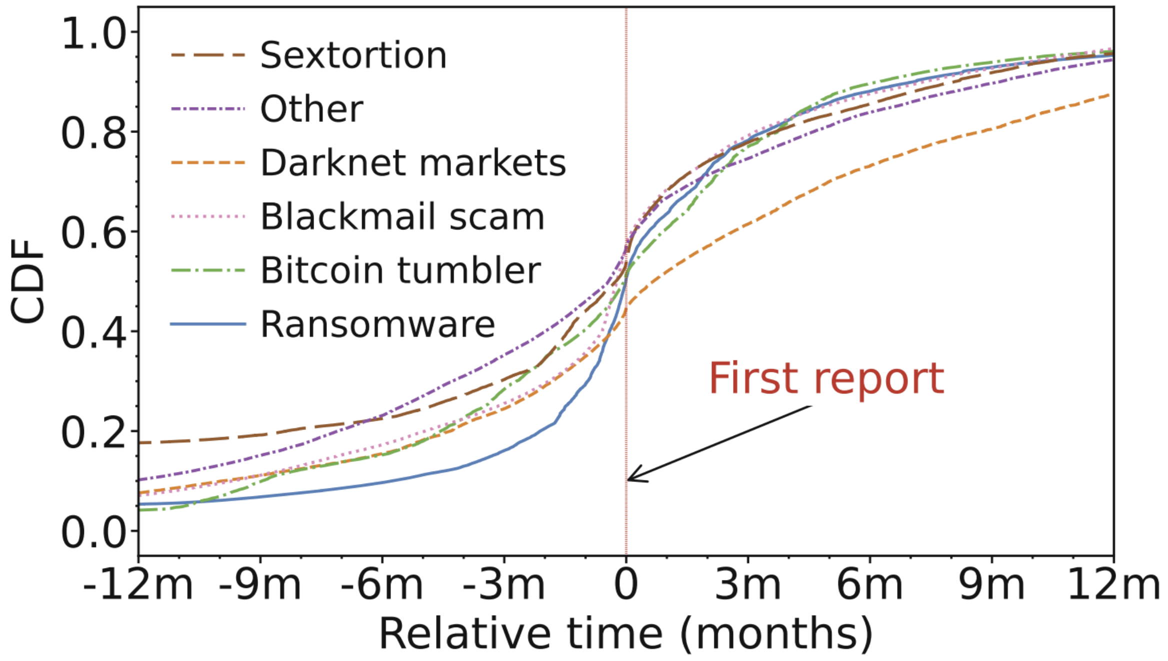
Initial Report Date Analysis

- Addresses are reported around the time that their incoming transaction count is high
- Indicating, the first report often is made around the time of the abuse's highest activity
- This may be a reflection of a significant portion of the addresses only being used for specific attacks
- Darknet markets follow this pattern the least, Ransomware the most



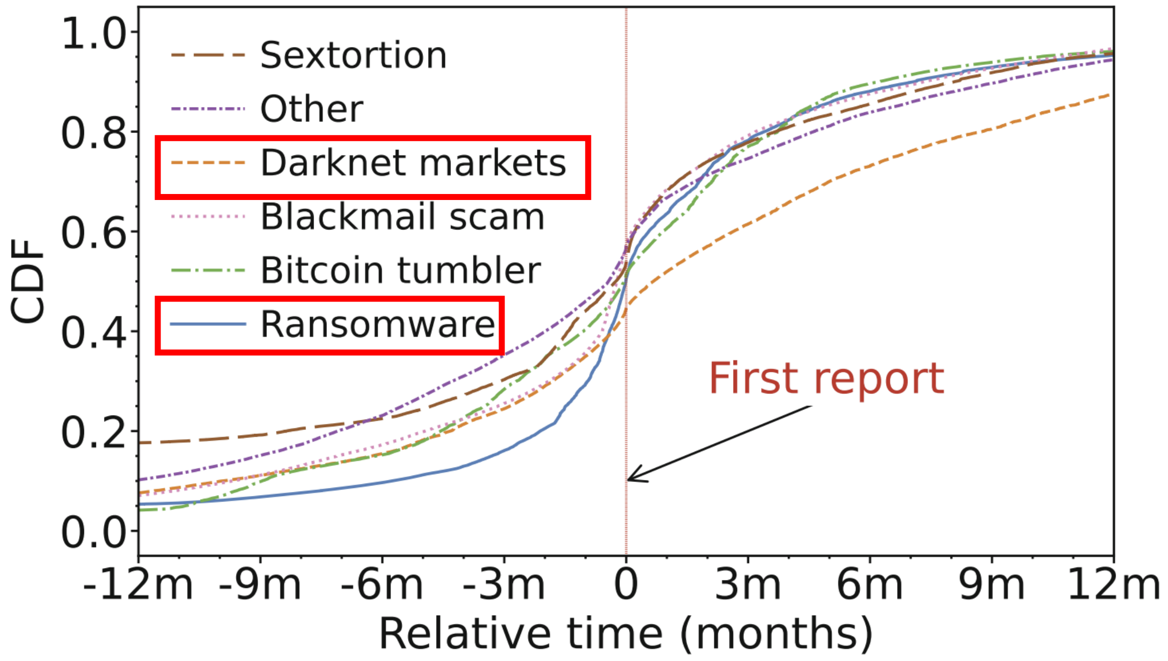
Initial Report Date Analysis

- Addresses are reported around the time that their incoming transaction count is high
- Indicating, the first report often is made around the time of the abuse's highest activity
- This may be a reflection of a significant portion of the addresses only being used for specific attacks
- Darknet markets follow this pattern the least, Ransomware the most



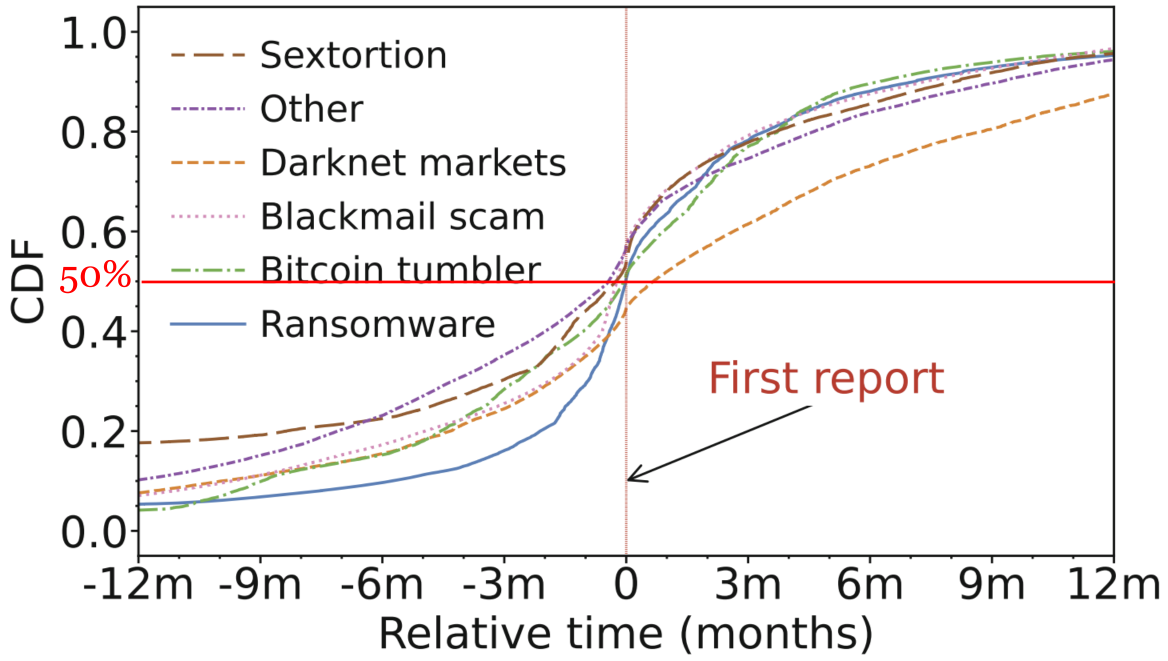
Initial Report Date Analysis

- Addresses are reported around the time that their incoming transaction count is high
- Indicating, the first report often is made around the time of the abuse's highest activity
- This may be a reflection of a significant portion of the addresses only being used for specific attacks
- Darknet markets follow this pattern the least, Ransomware the most



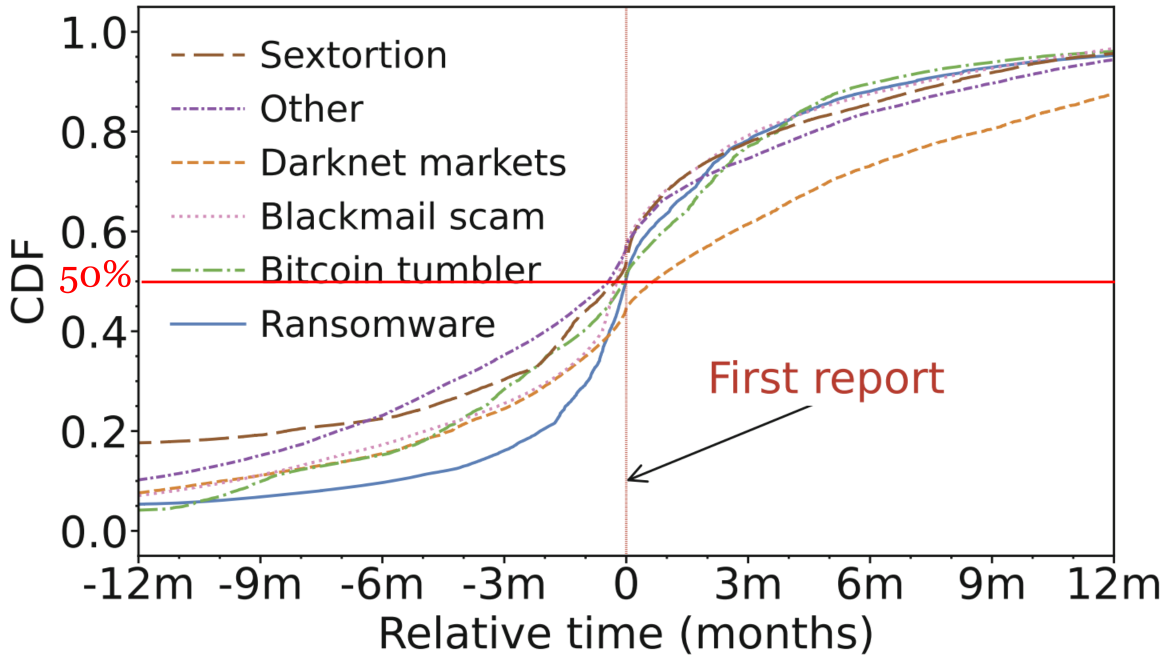
Initial Report Date Analysis

- Most transactions take place before the first report (except Darknet markets)
- Suggesting limited effectiveness to using such reports to "ban" addresses



Initial Report Date Analysis

- Most transactions take place before the first report (except Darknet markets)
- Suggesting limited effectiveness to using such reports to "ban" addresses

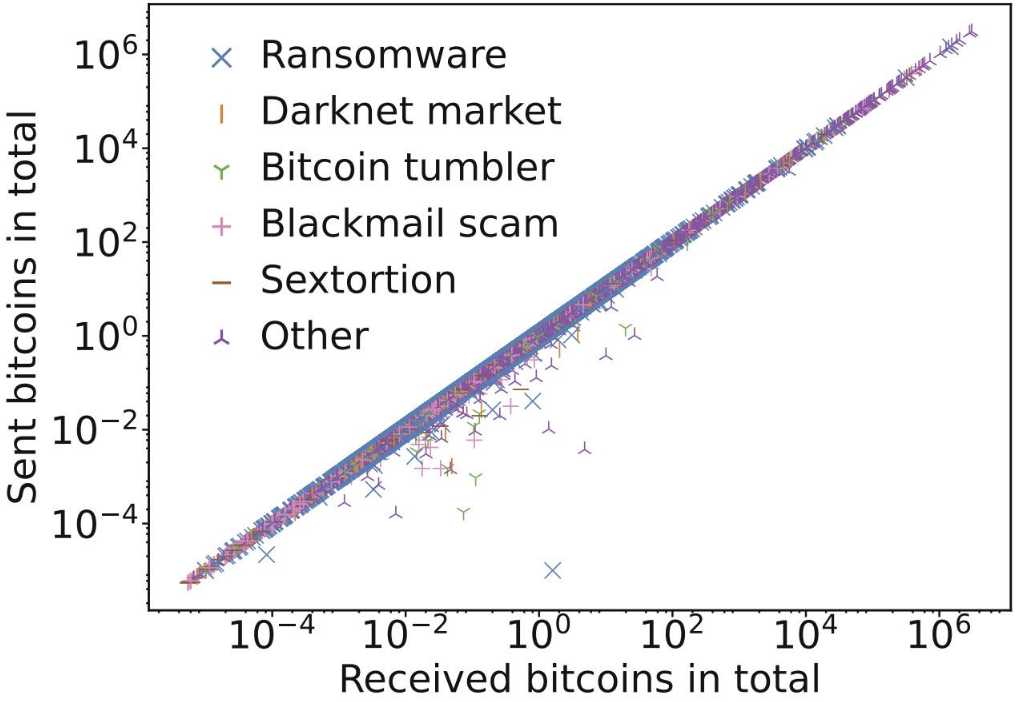


Following the Money

- Why
- Methodology
- One-step concentration or dispersion
- Multi-step analysis

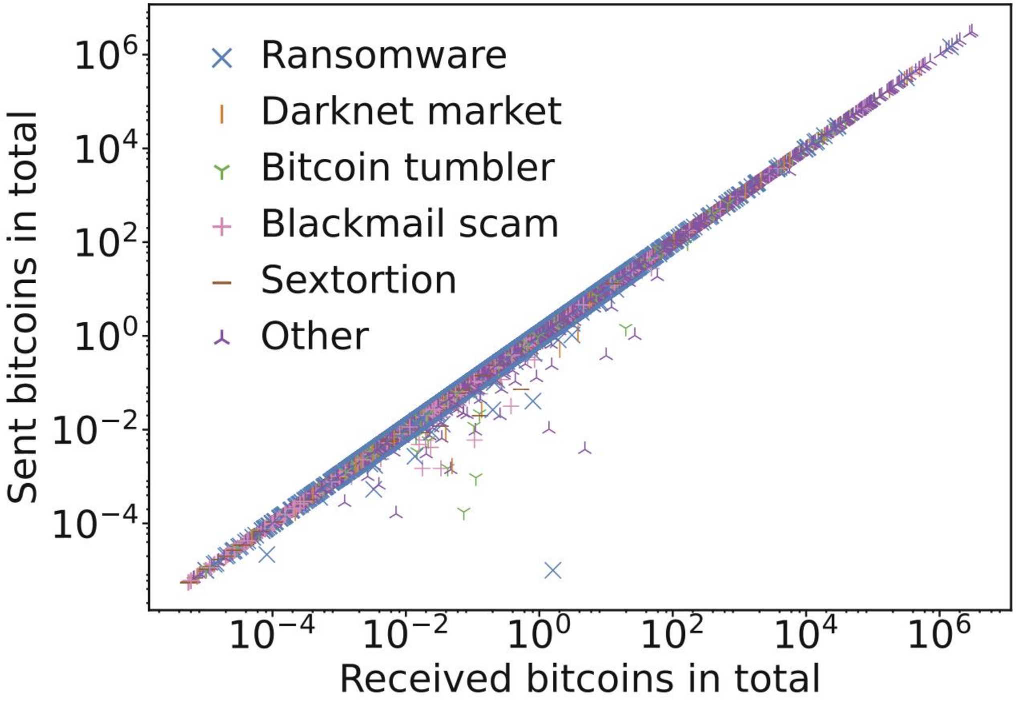
Why?

- Nearly all reported addresses have sent as many bitcoins as they received, **leaving a balance of zero**
- Suggesting these addresses are typically not used to store their gains



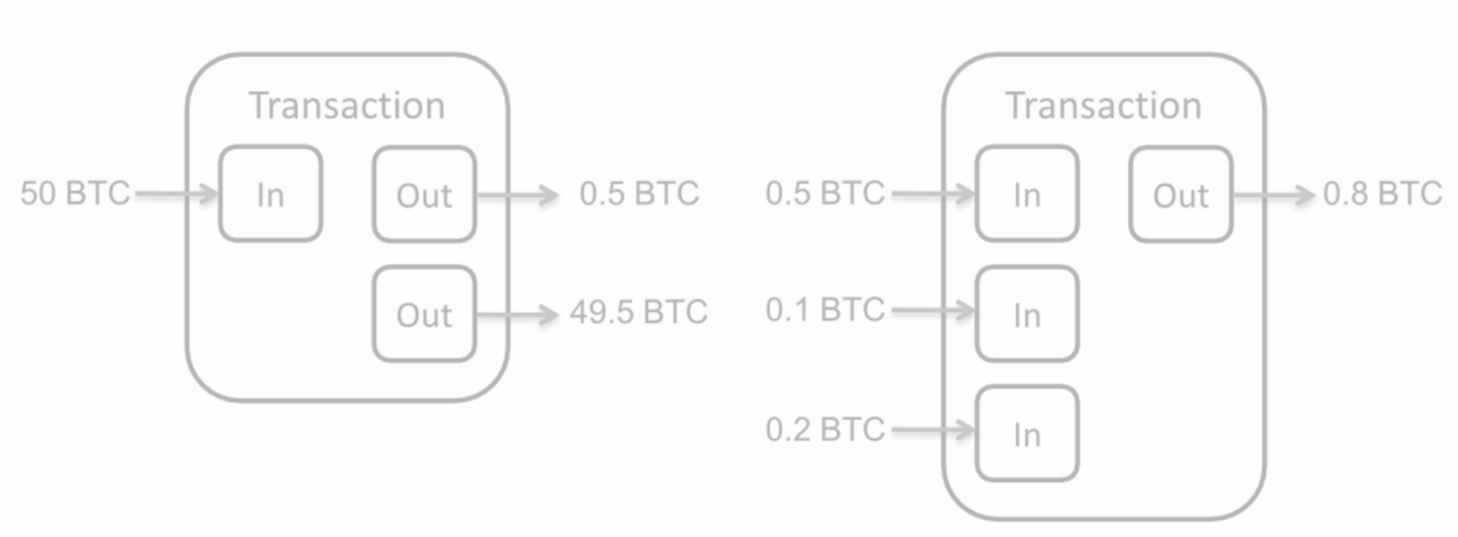
Why?

- Nearly all reported addresses have sent as many bitcoins as they received, **leaving a balance of zero**
- Suggesting these addresses are typically not used to store their gains



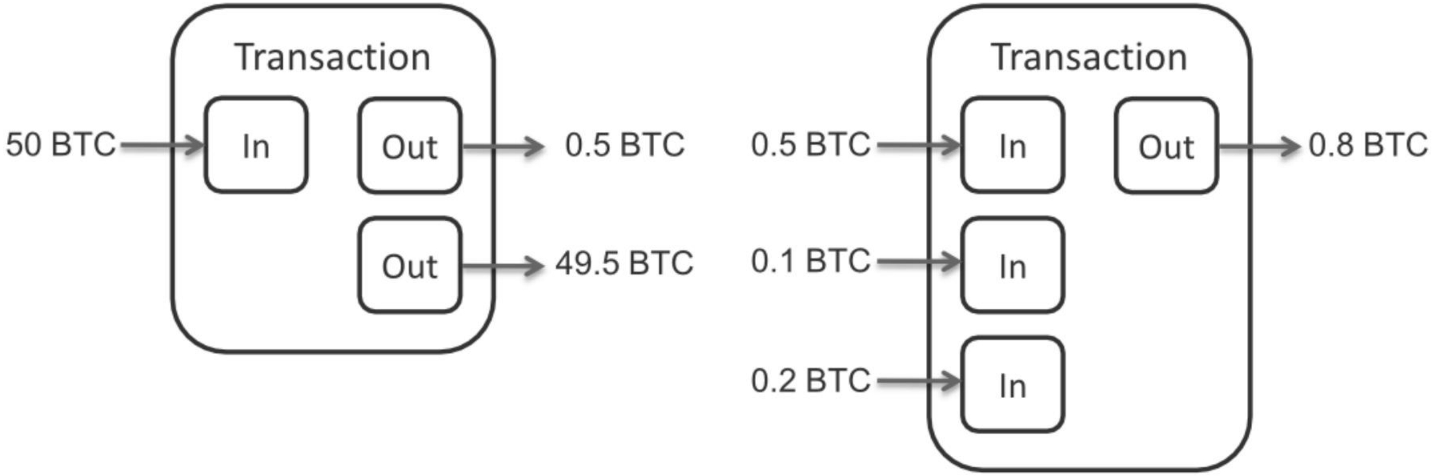
Methodology

- **Scope of analysis:** Studying and comparing potential concentration or dispersion for different abuse categories
- **Challenging case and our solution:** Transactions may have multiple inputs *and* outputs – a melting pot. For this part of the analysis, we only use transactions where the sender and receiver is known.



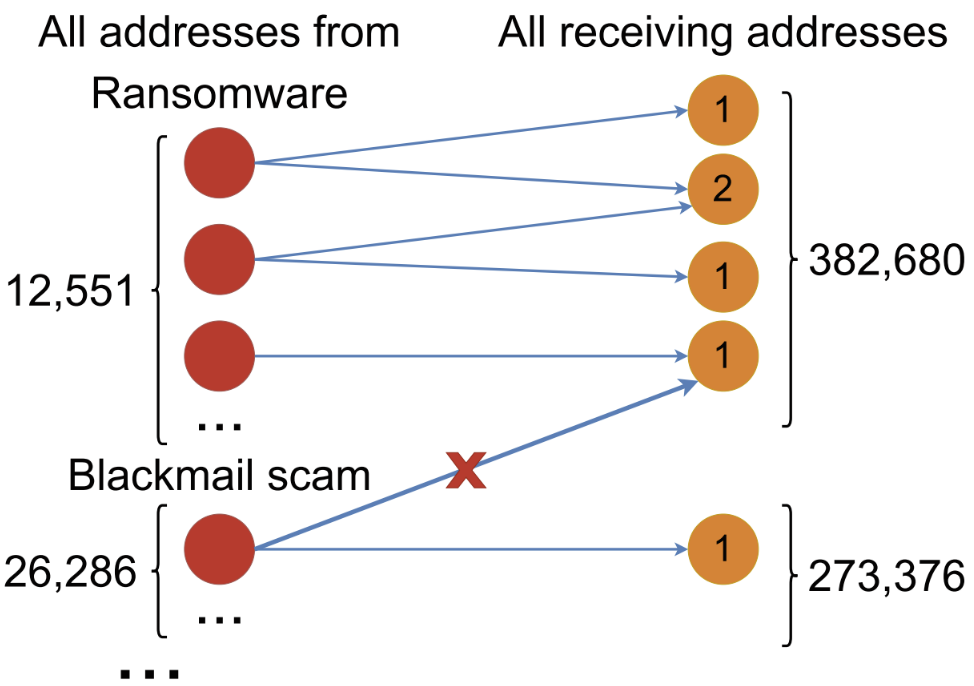
Methodology

- **Scope of analysis:** Studying and comparing potential concentration or dispersion for different abuse categories
- **Challenging case and our solution:** Transactions may have multiple inputs *and* outputs – a melting pot. For this part of the analysis, we only use transactions where the sender and receiver is known.

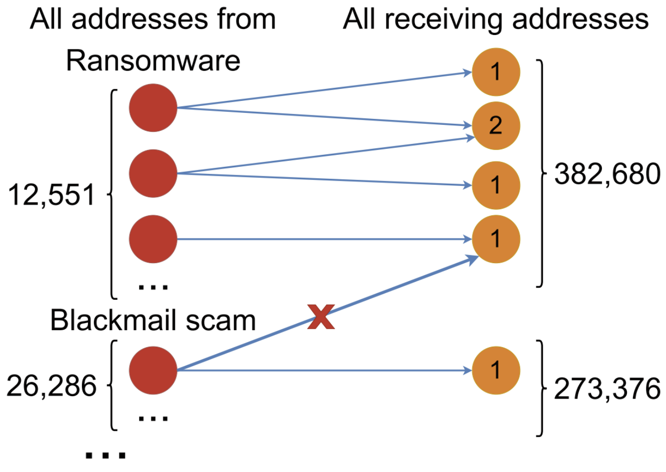


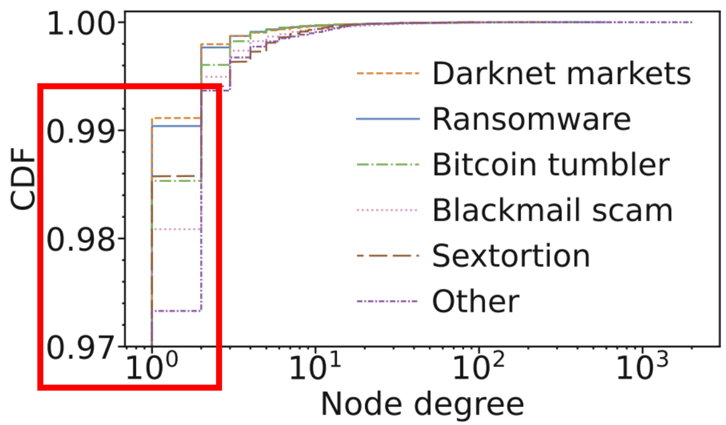
LINKÖPING UNIVERSITY

One-Step Concentration or Dispersion

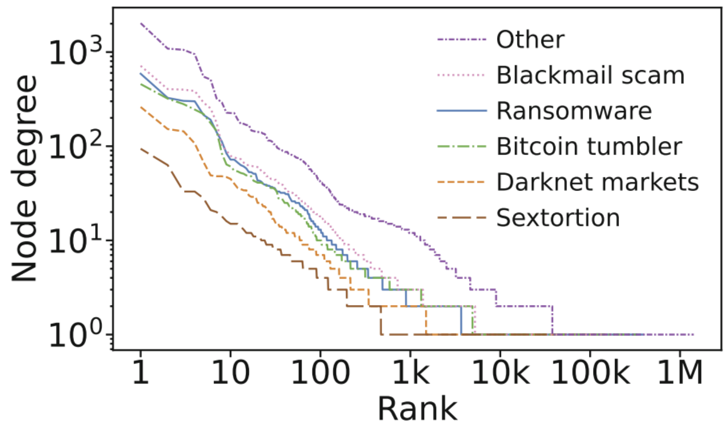


One-Step Concentration or Dispersion





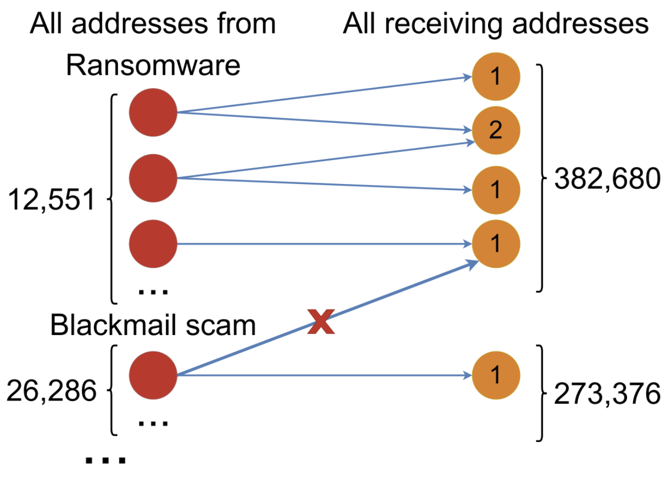
(a) CDF

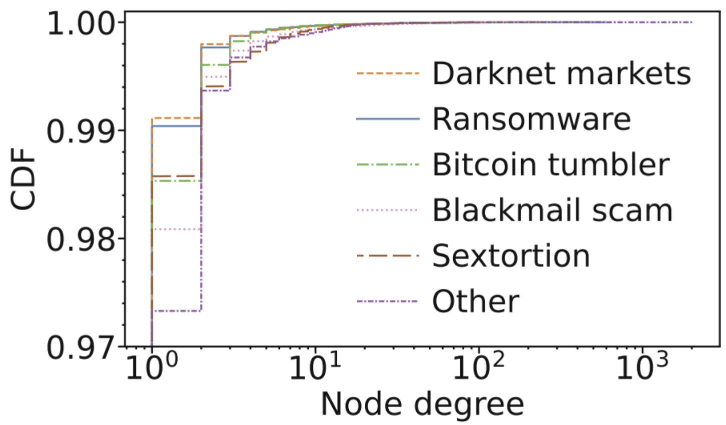


(b) Rank plot

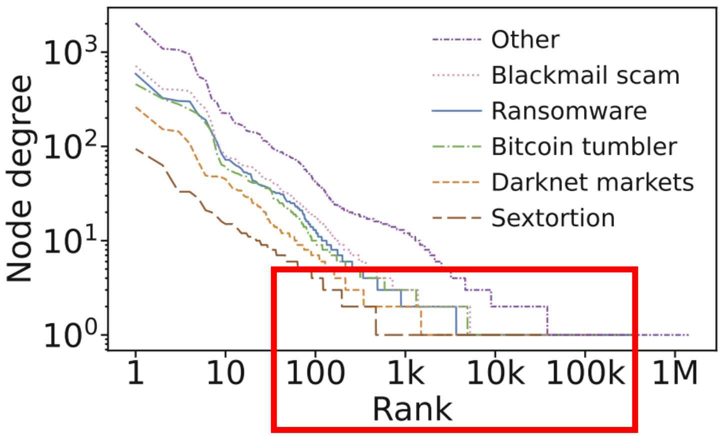
- The overwhelming majority are not visibly related when tracing the money one-step (97-99%)
  - Suggesting high dispersion
- However, all categories has at least one address with a node in-degree over 100
- Significant difference between categories

One-Step Concentration or Dispersion





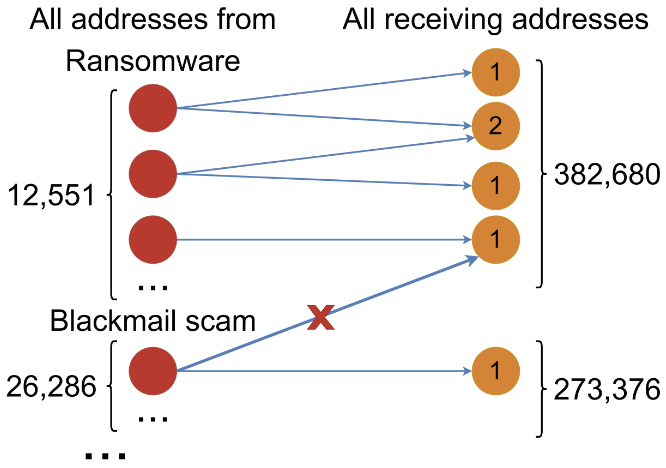
(a) CDF

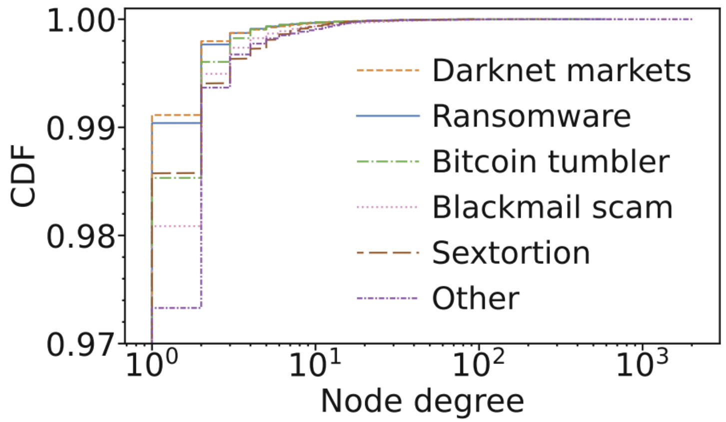


(b) Rank plot

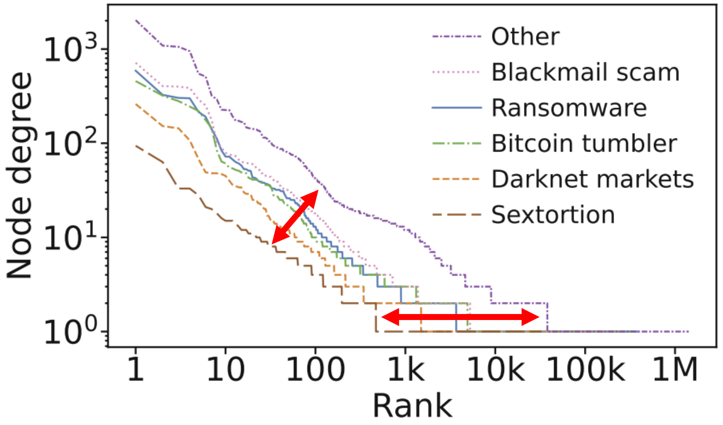
- The overwhelming majority are not visibly related when tracing the money one-step (97-99%)
  - Suggesting high dispersion
- However, all categories has at least one address with a node in-degree over 100
- Significant difference between categories

One-Step Concentration or Dispersion





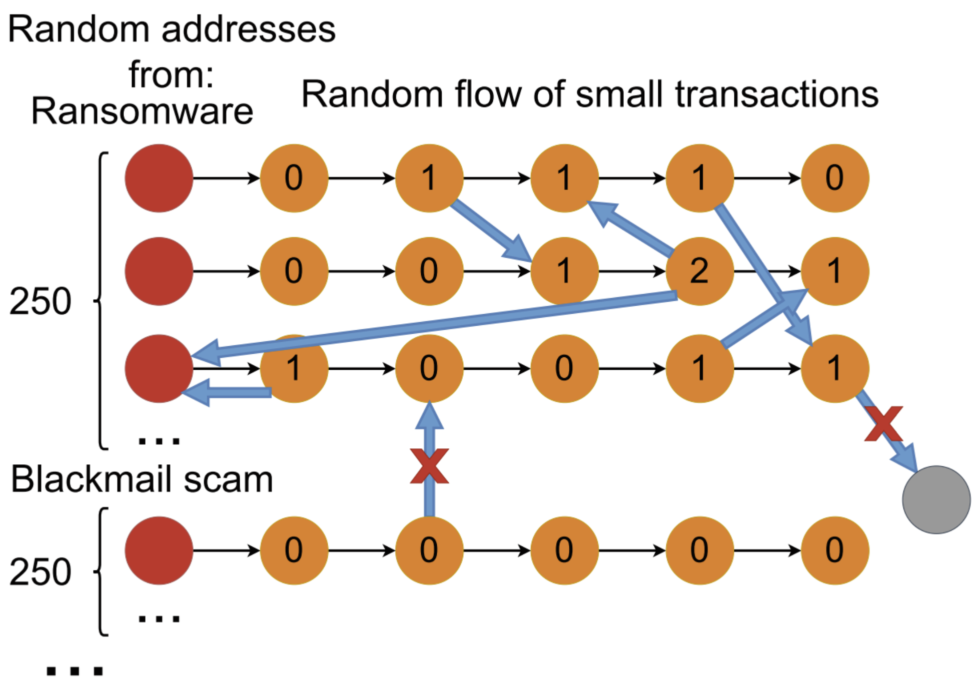
(a) CDF



(b) Rank plot

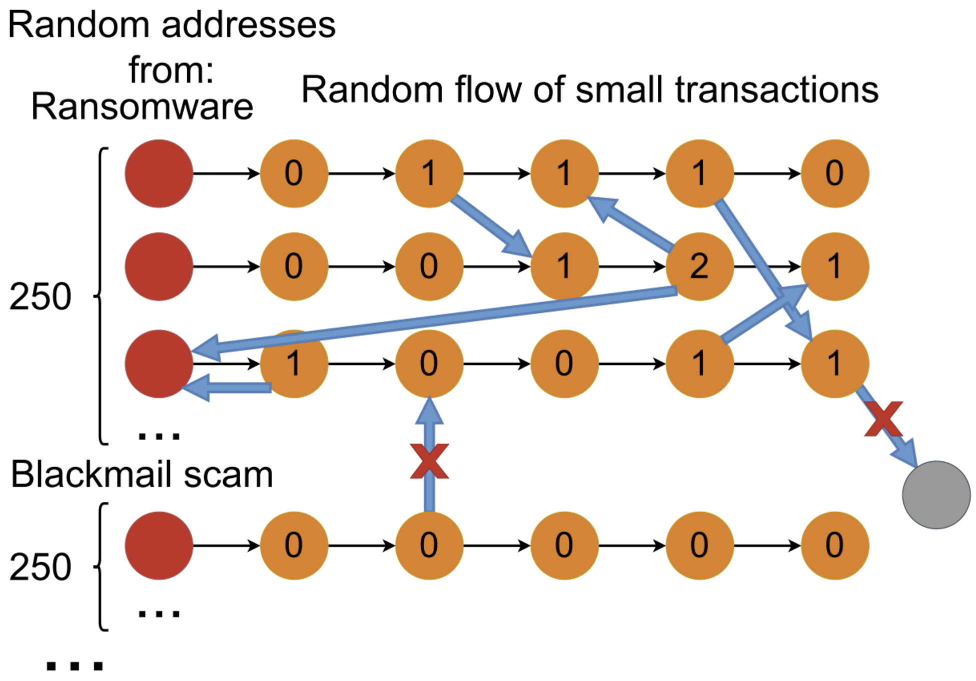
- The overwhelming majority are not visibly related when tracing the money one-step (97-99%)
  - Suggesting high dispersion
- However, all categories has at least one address with a node in-degree over 100
- Significant difference between categories

Multi-Step Analysis



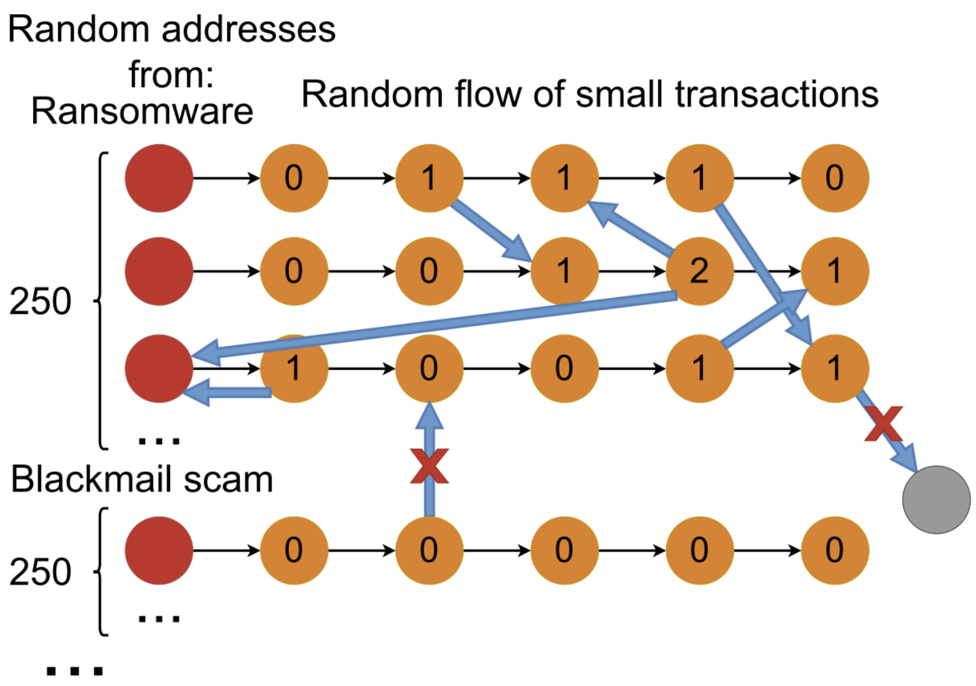
- Is money shuffled around a couple of steps only to collect a few steps later?
- We tracked an equal amount of "penny flows" as bitcoins were moved five steps deep
- 250 random reported address from each category
- Random "penny flow" from each

Multi-Step Analysis



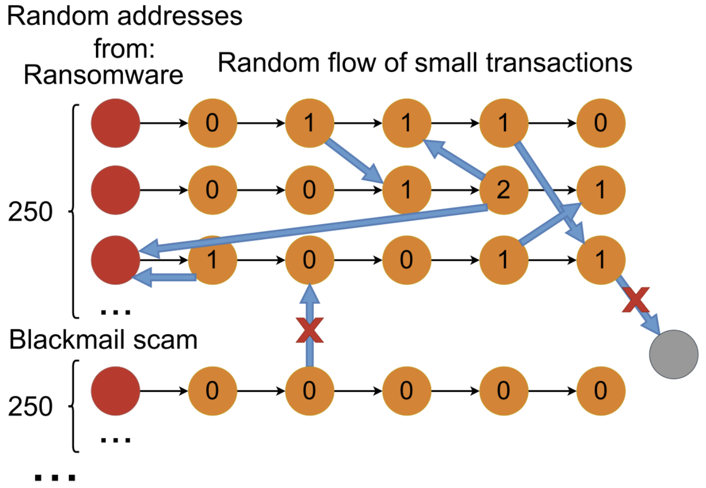
- Is money shuffled around a couple of steps only to collect a few steps later?
- We tracked an equal amount of "penny flows" as bitcoins were moved five steps deep
- 250 random reported address from each category
- Random "penny flow" from each

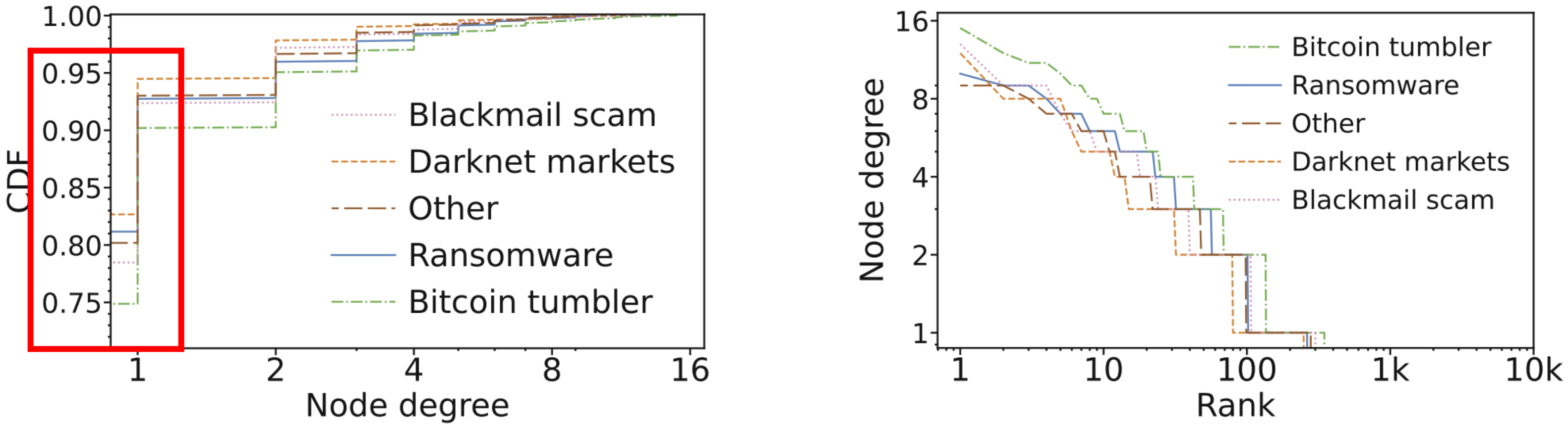
Multi-Step Analysis



- Is money shuffled around a couple of steps only to collect a few steps later?
- We tracked an equal amount of “penny flows” as bitcoins were moved five steps deep
- 250 random reported address from each category
- Random “penny flow” from each

Multi-Step Analysis

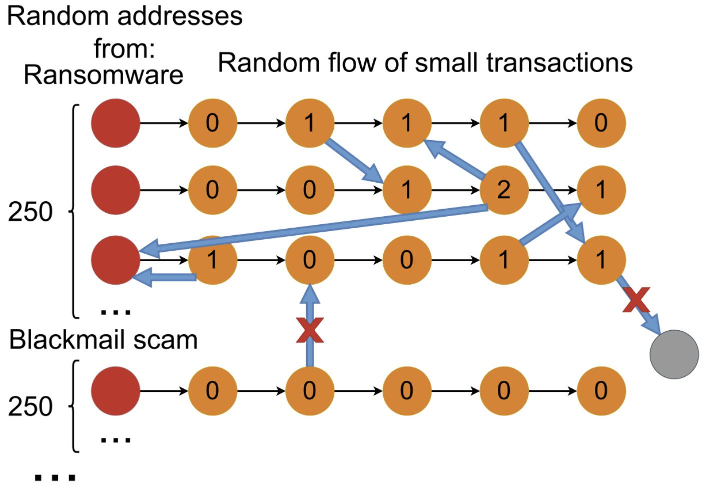


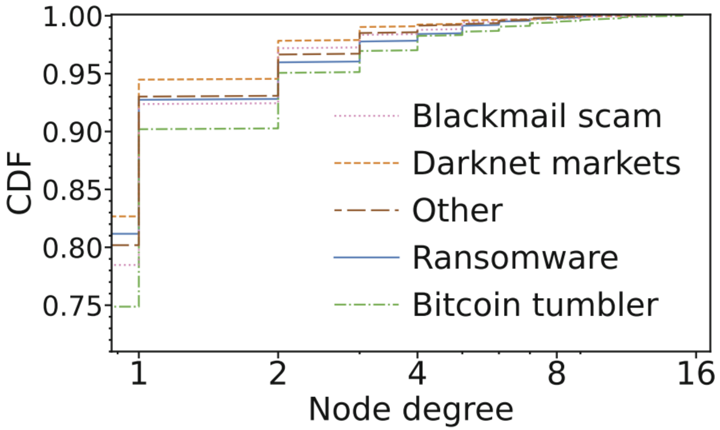


(a) CDF (b) Rank plot

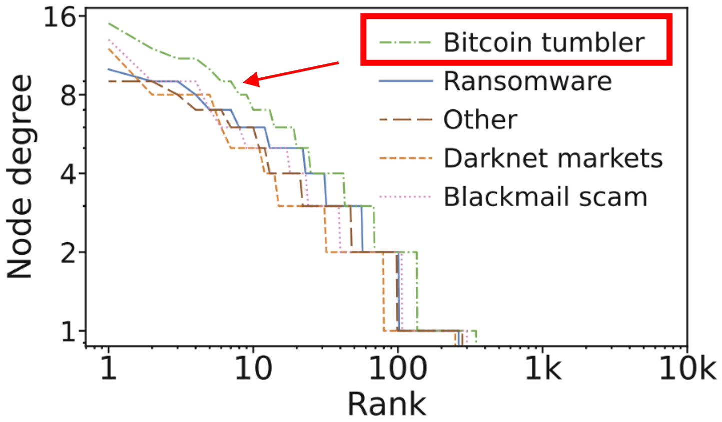
- Fewer addresses with the minimum in-degree, compared to the one-step analysis (from 97-99% to 83-89%)
- Bitcoin tumbler stands out with higher node degrees. Suggesting fewer actors involved in tumbling.
  - Perhaps because of the higher effort required, compared to Blackmails scams (which has the lowest node degree)

Multi-Step Analysis





(a) CDF



(b) Rank plot

- Fewer addresses with the minimum in-degree, compared to the one-step analysis (from 97-99% to 83-89%)
- Bitcoin tumbler stands out with higher node degrees. Suggesting fewer actors involved in tumbling.
  - Perhaps because of the higher effort required, compared to Blackmails scams (which has the lowest node degree)

Contributions

- **High-level characterization** of the transactions received by the Bitcoin addresses reported to the Bitcoin Abuse Database (2017-2022)
  - Aggregate basis
  - Per-category basis
- **Temporal analysis** that captures
  - Long-term trends
  - Weekly patterns (per category)
  - Correlations with the first report date (per category)
- **Analyze the outflow of bitcoins** from reported addresses

Contributions

- **High-level characterization** of the transactions received by the Bitcoin addresses reported to the Bitcoin Abuse Database (2017-2022)
  - Aggregate basis
  - Per-category basis
- **Temporal analysis** that captures
  - Long-term trends
  - Weekly patterns (per category)
  - Correlations with the first report date (per category)
- **Analyze the outflow of bitcoins** from reported addresses

Contributions

- **High-level characterization** of the transactions received by the Bitcoin addresses reported to the Bitcoin Abuse Database (2017-2022)
  - Aggregate basis
  - Per-category basis
- **Temporal analysis** that captures
  - Long-term trends
  - Weekly patterns (per category)
  - Correlations with the first report date (per category)
- **Analyze the outflow of bitcoins** from reported addresses

On the Dark Side of the Coin: Characterizing Bitcoin use for Illicit Activities

LINKÖPING UNIVERSITY

Hampus Rosenquist, David Hasselquist, Martin Arlitt, and Niklas Carlsson

Linköping University, Sweden, and University of Calgary, Canada