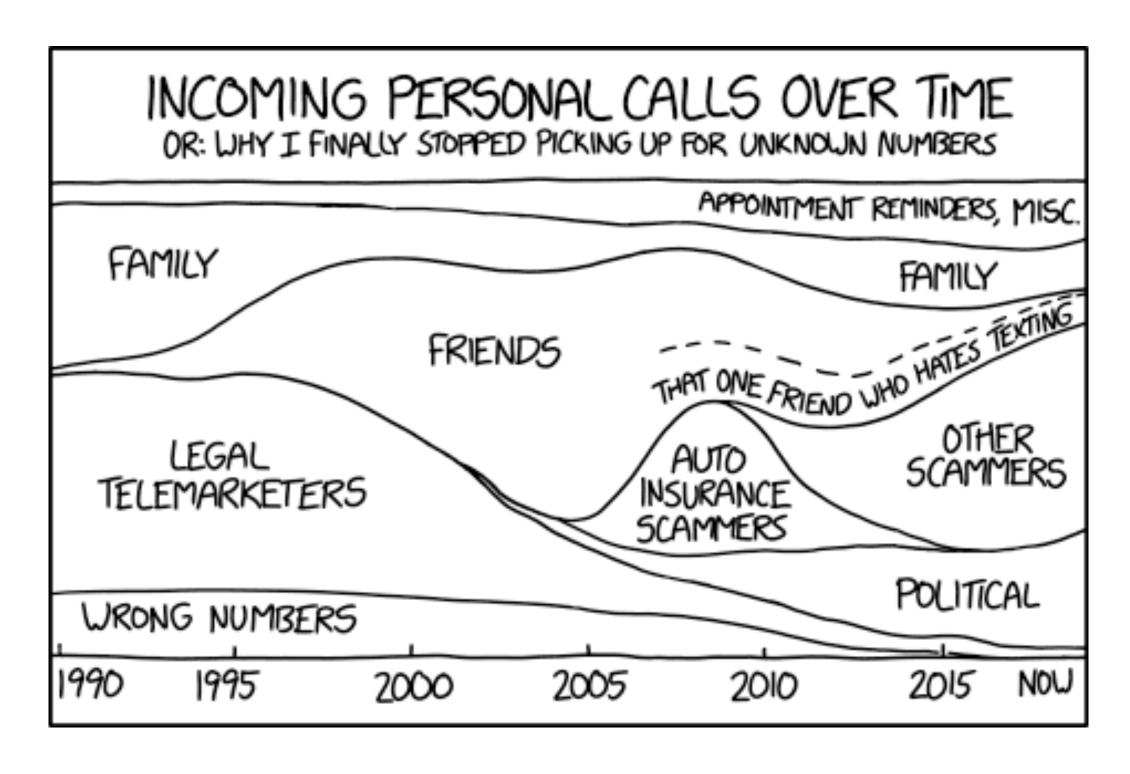
# Who's Calling?

Characterizing Robocalls through Audio and Metadata Analysis

**Sathvik Prasad**, Elijah Bouma-Sims, Athishay Kiran Mylappan, Bradley Reaves North Carolina State University

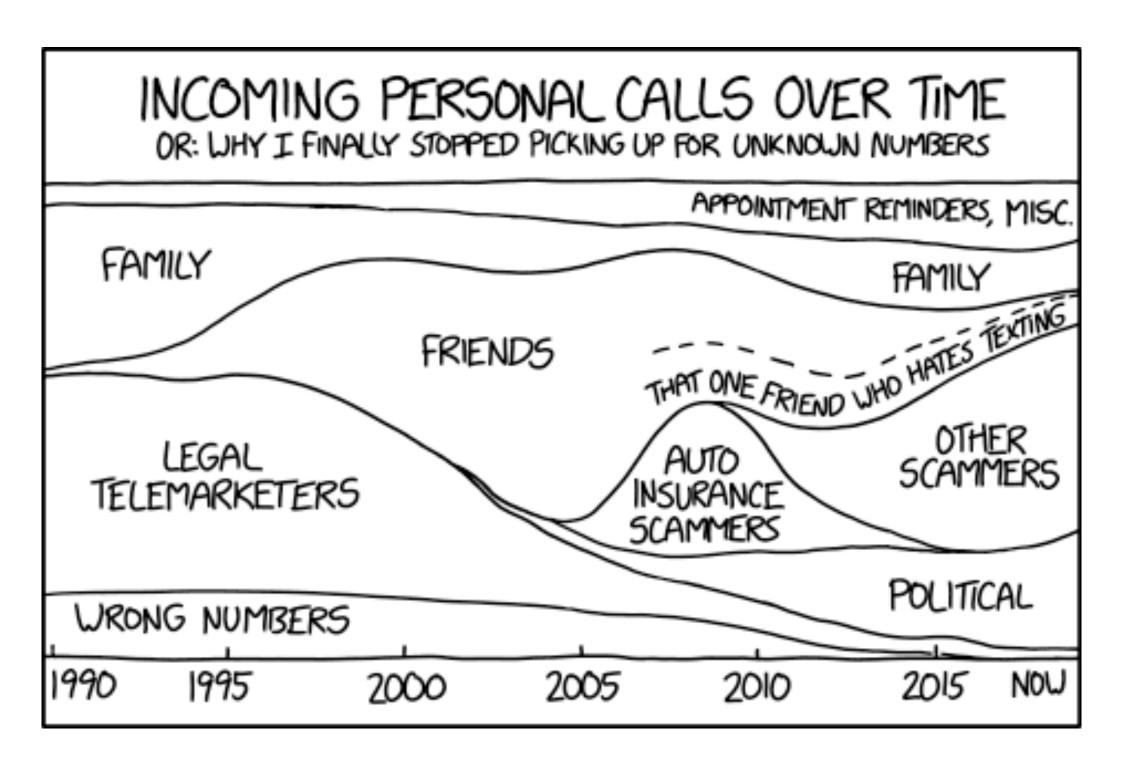


### Some are annoying





### Some are annoying



Some can cause serious harm!

The Washington Post

Democracy Dies in Darkness

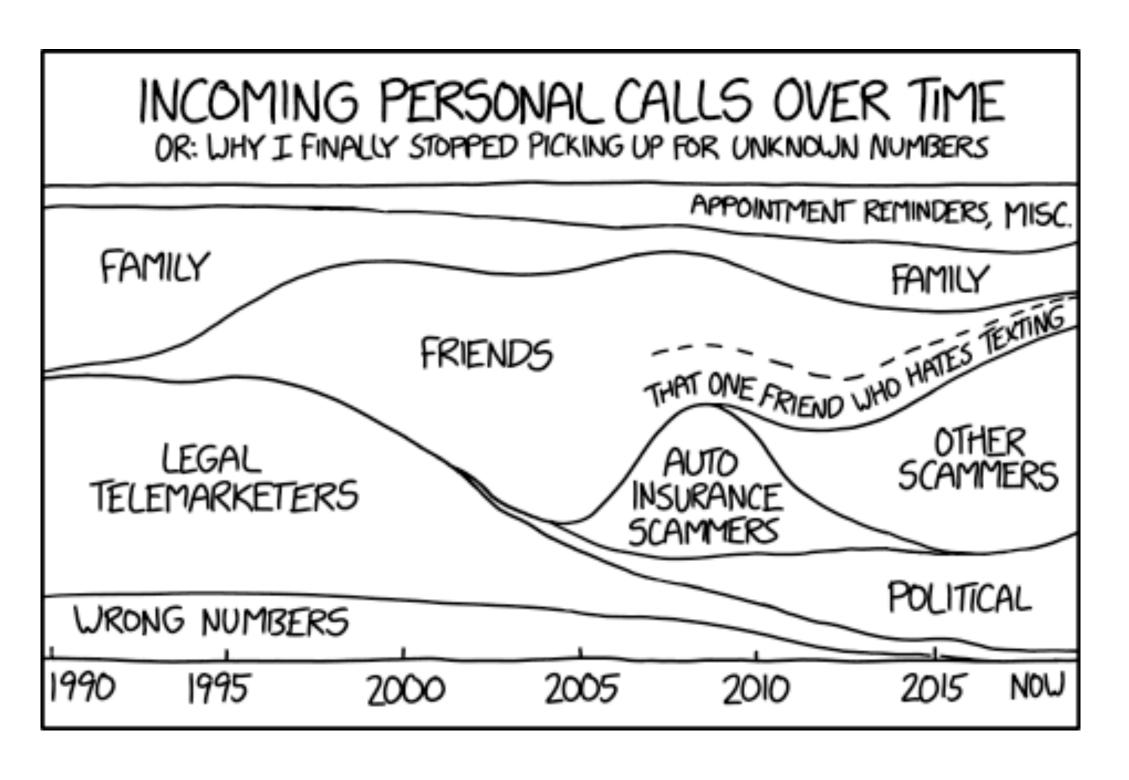
Get 1 year for \$29

That can actually kill somebody': Scam robocalls are pitching fake coronavirus tests to vulnerable Americans





### Some are annoying



Some can cause serious harm!

The Washington Post

Democracy Dies in Darkness

Get 1 year for \$29

That can actually kill somebody': Scam robocalls are pitching fake coronavirus tests to vulnerable Americans

The Washington Post

Democracy Dies in Darkness

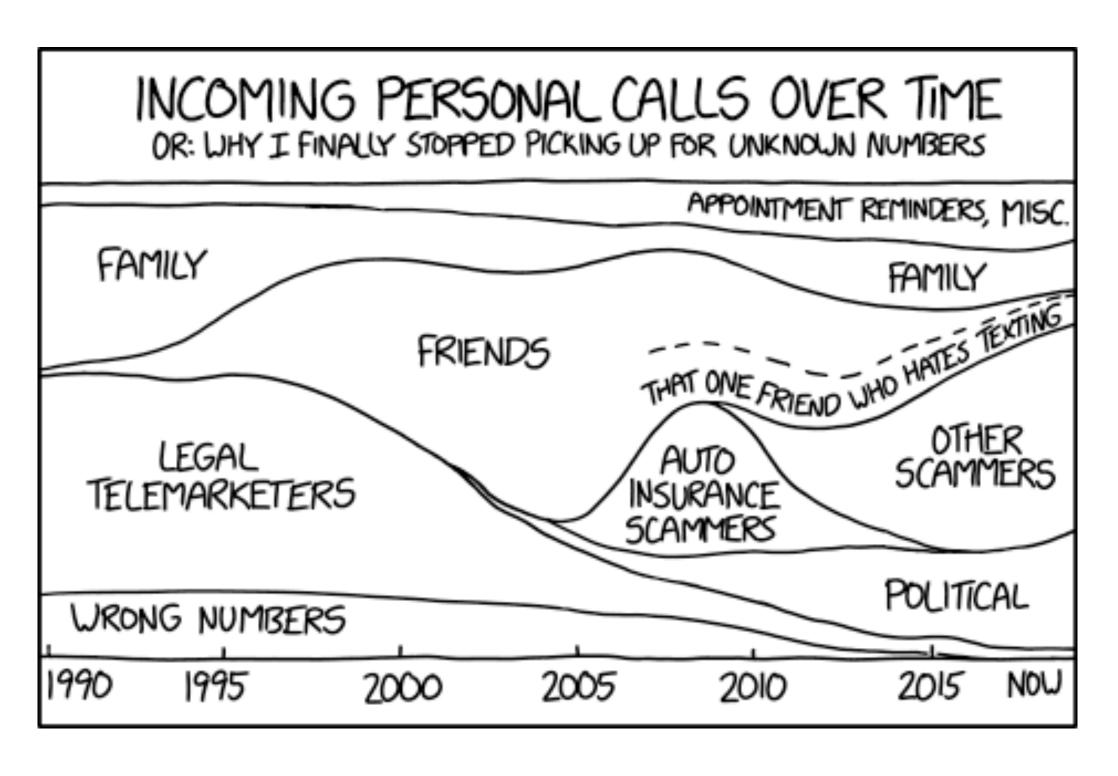
Get 1 year for \$29

Suspicious robocalls pitching health insurance surged in October. Now, the feds are scrambling to act.

https://xkcd.com/2053/



### Some are annoying



Some can cause serious harm!

The Washington Post

Democracy Dies in Darkness

Get 1 year for \$29

That can actually kill somebody': Scam robocalls are pitching fake coronavirus tests to vulnerable Americans

The Washington Post

Democracy Dies in Darkness

Get 1 year for \$29

Suspicious robocalls pitching health insurance surged in October. Now, the feds are scrambling to act.

THE WALL STREET JOURNAL.

Subscribe

Home World U.S. Politics Economy Business Tech Markets Opinion Life & Arts Real Estate WSJ. Magazine

FCC Seeks \$225 Million Fine From Telemarketer Accused of Making 1 Billion Robocalls

https://www.wsj.com/articles/fcc-seeks-225-million-fine-from-telemarketer-accused-of-making-1-billion-robocalls-11591730860 https://www.washingtonpost.com/technology/2019/11/19/suspicious-robocalls-pitching-health-insurance-surged-october-now-feds-are-scrambling-act/https://www.washingtonpost.com/technology/2020/03/19/robocalls-coronavirus-test/



#### **NC STATE** UNIVERSITY

## Primary Questions

. Is the robocalling problem getting worse?



- I. Is the robocalling problem getting worse?
- 2. By answering robocalls, will you receive more robocalls?

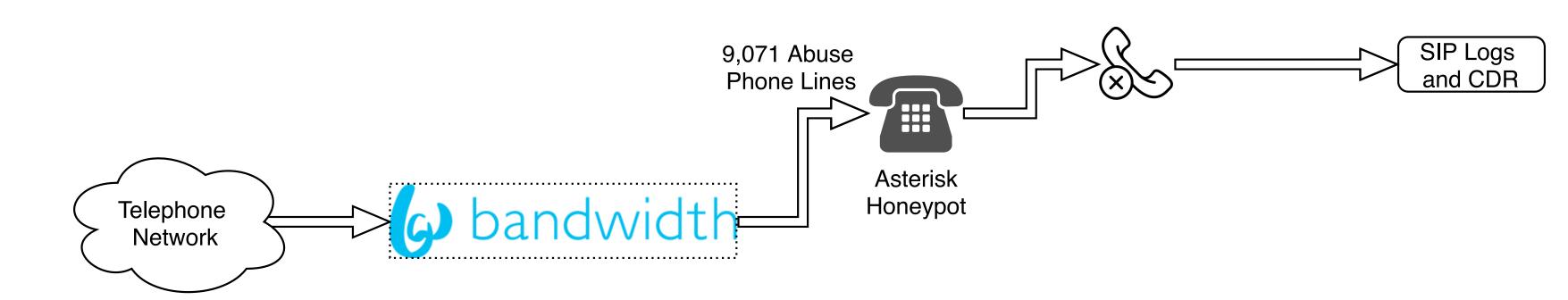


- I. Is the robocalling problem getting worse?
- 2. By answering robocalls, will you receive more robocalls?
- 3. Who is calling you?



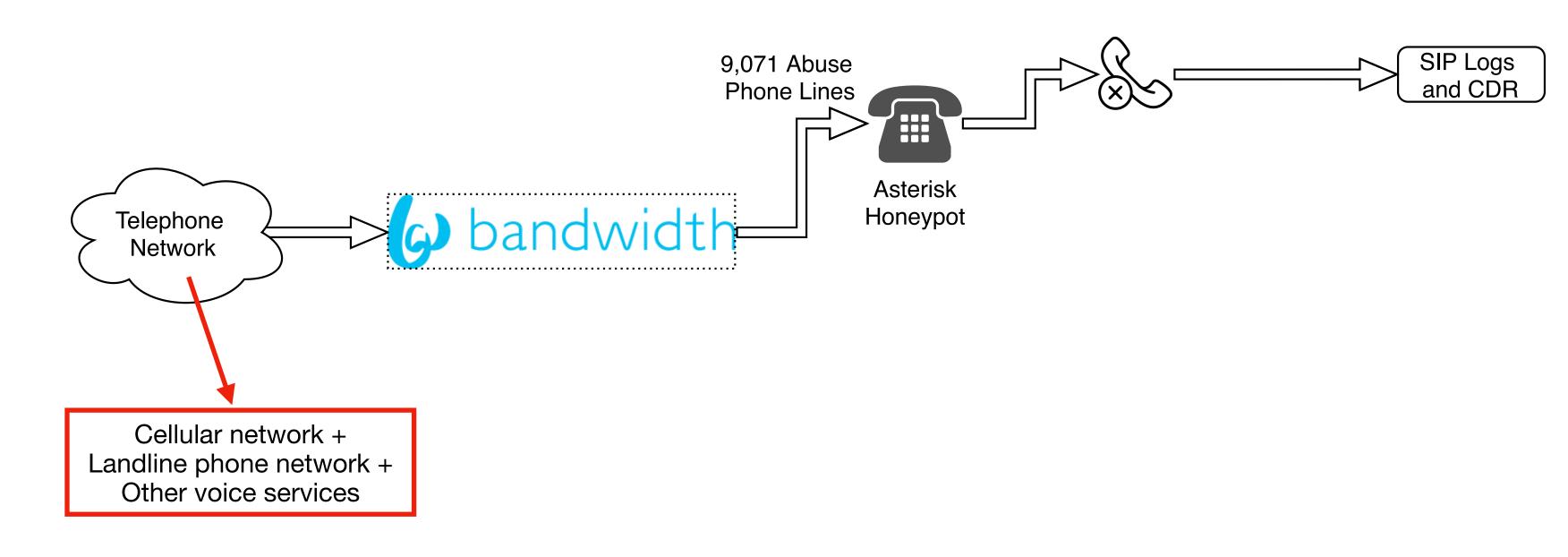
- I. Is the robocalling problem getting worse?
- 2. By answering robocalls, will you receive more robocalls?
- 3. Who is calling you?
- 4. What strategies do robocalls use to entice their victims?





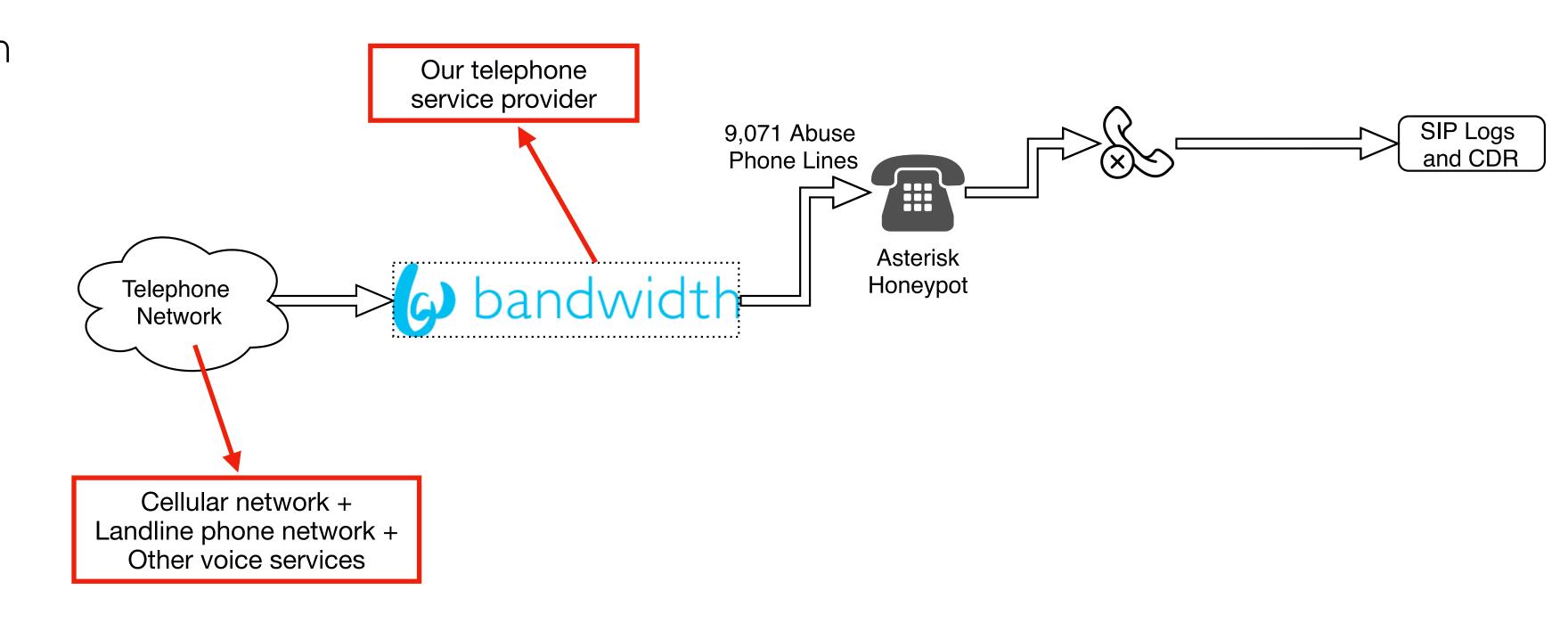






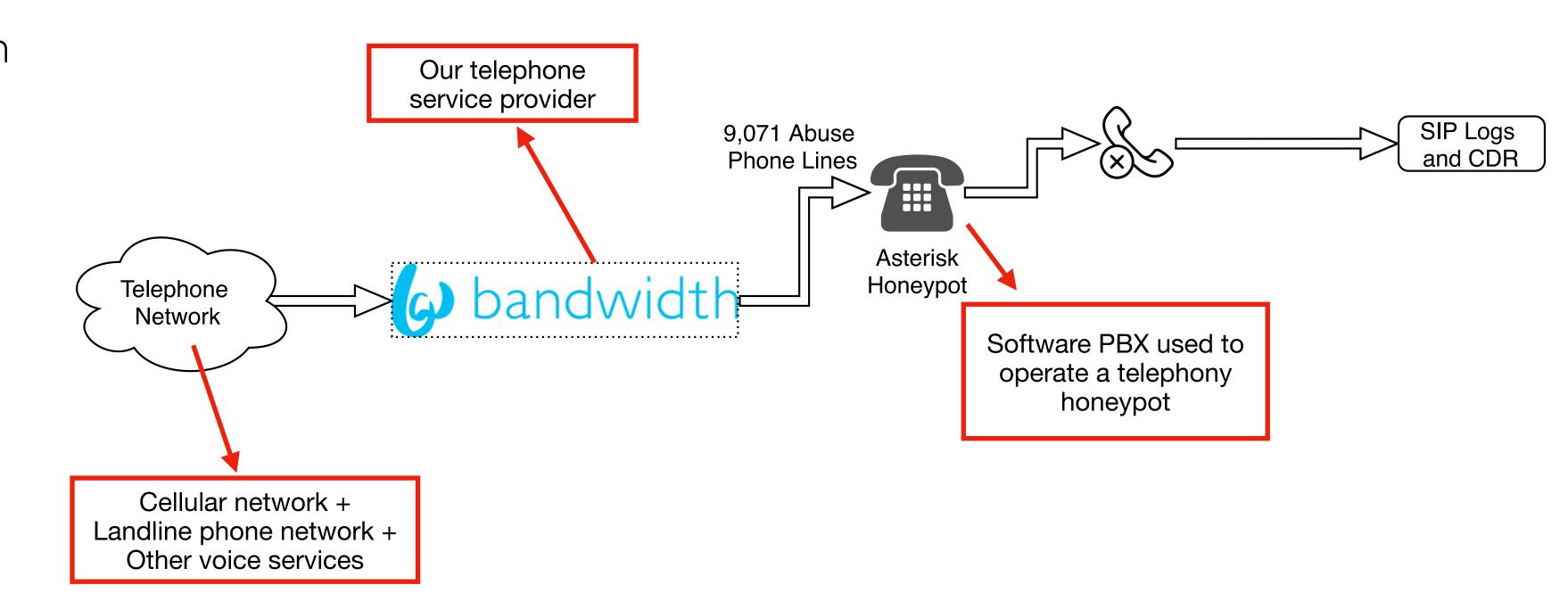






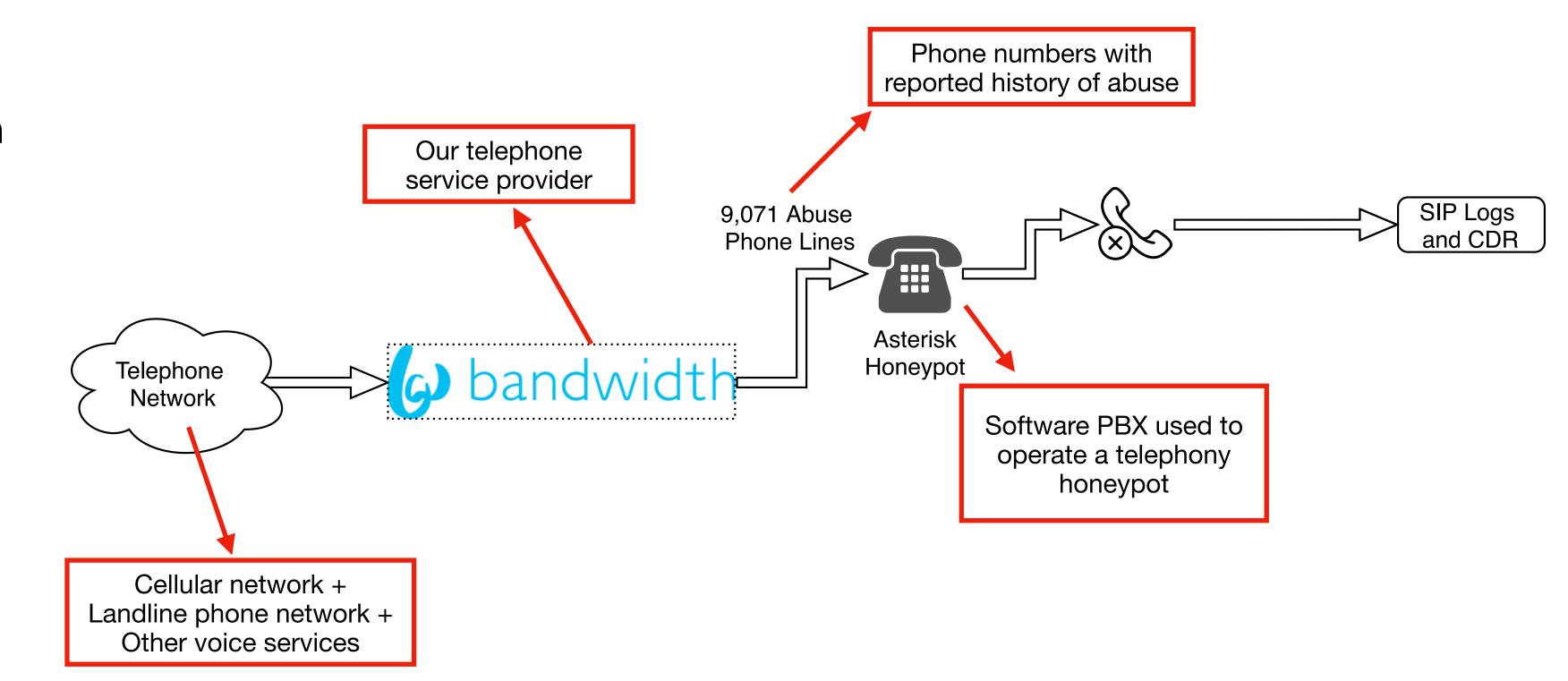






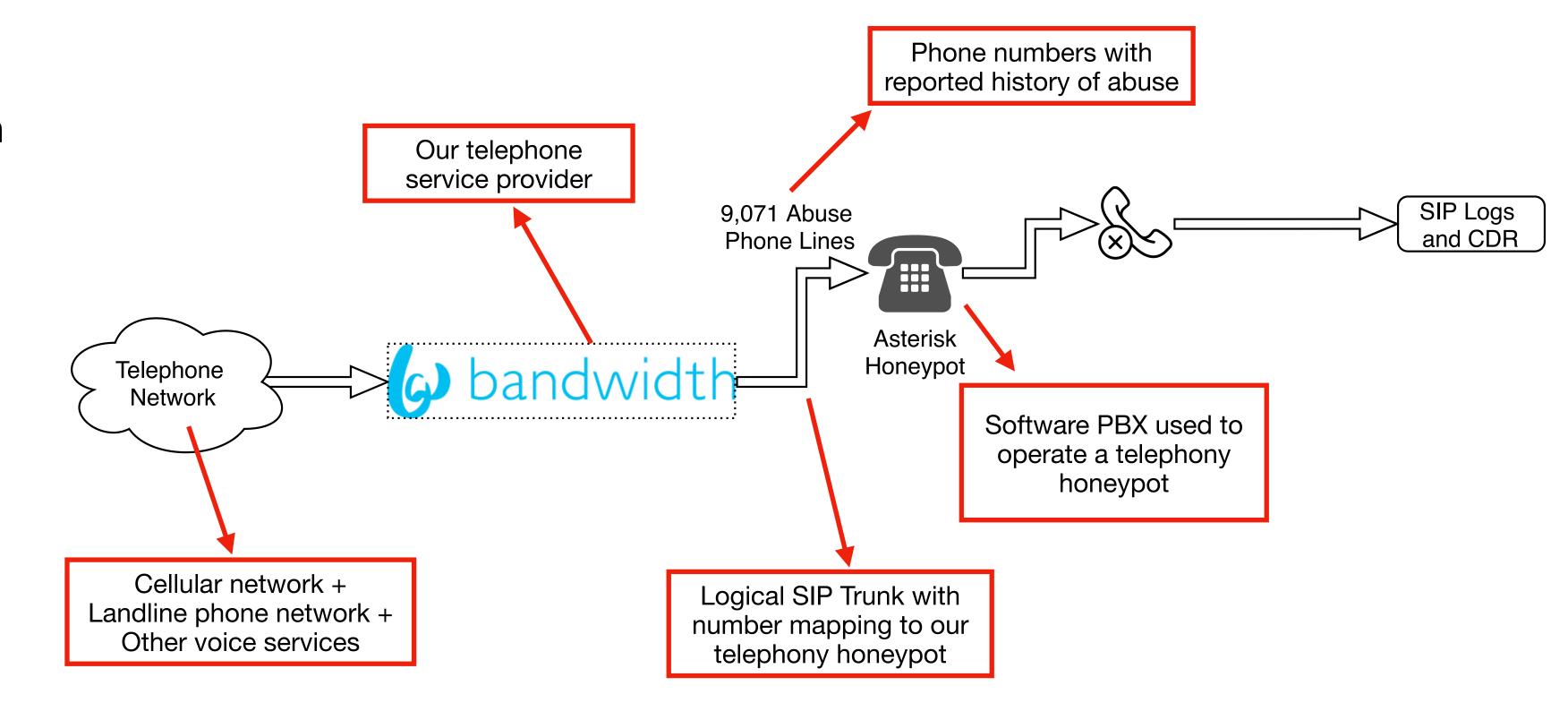






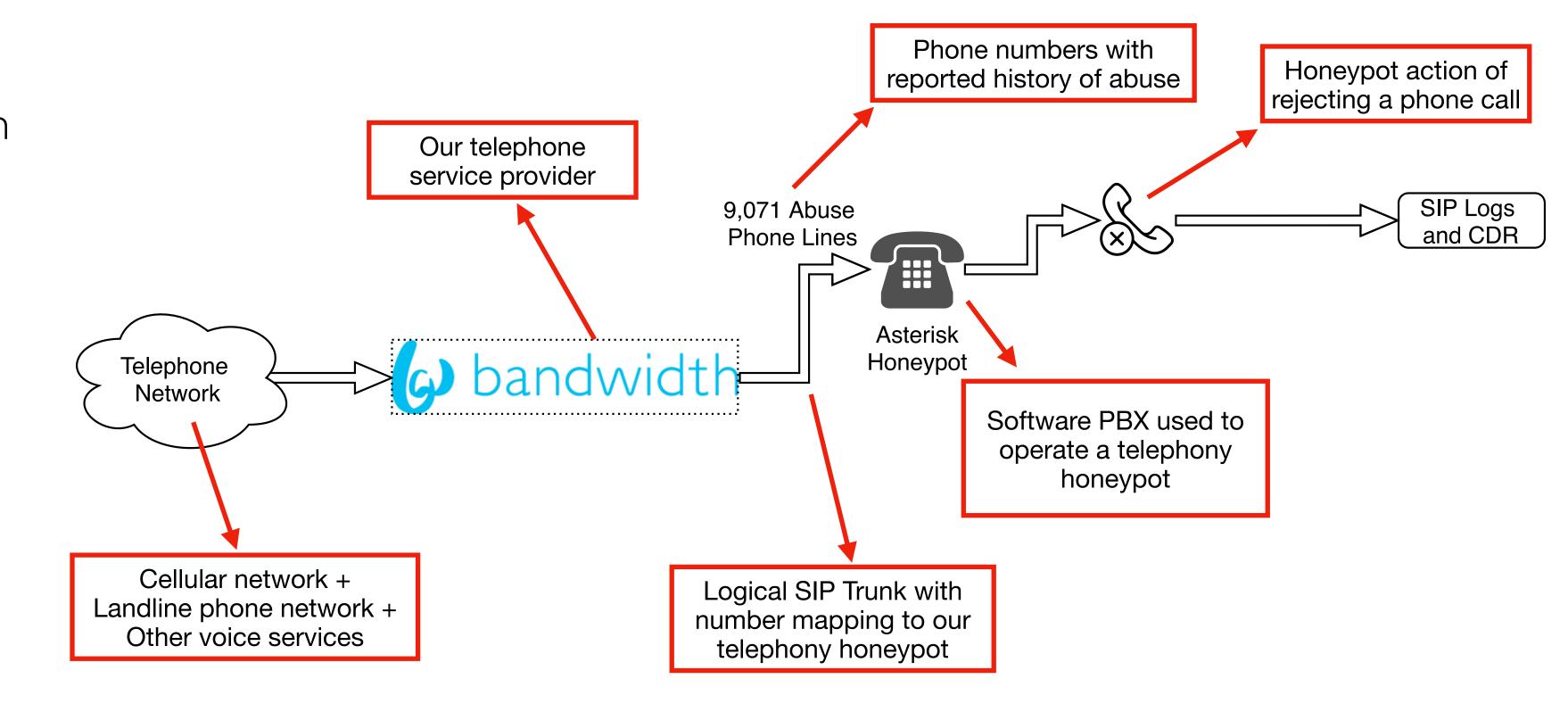






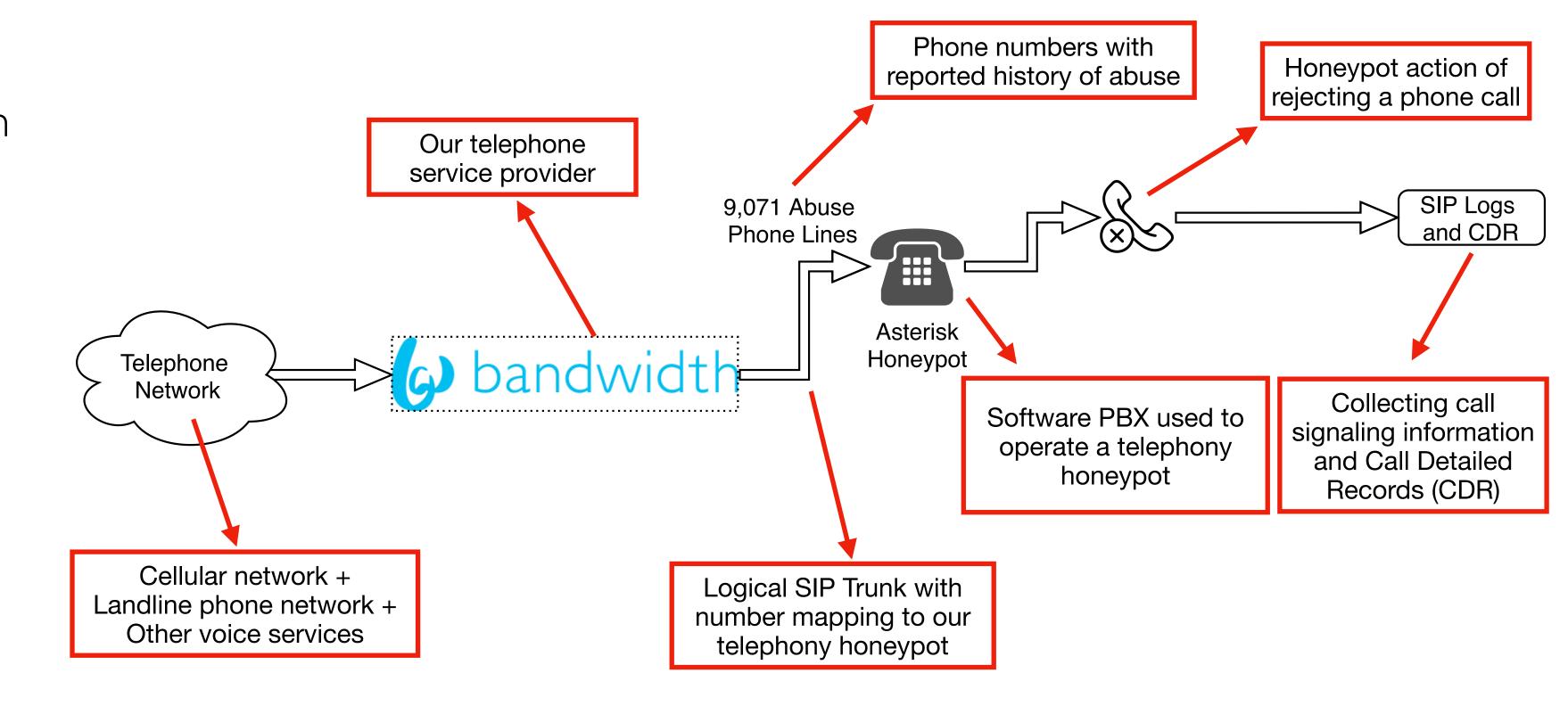








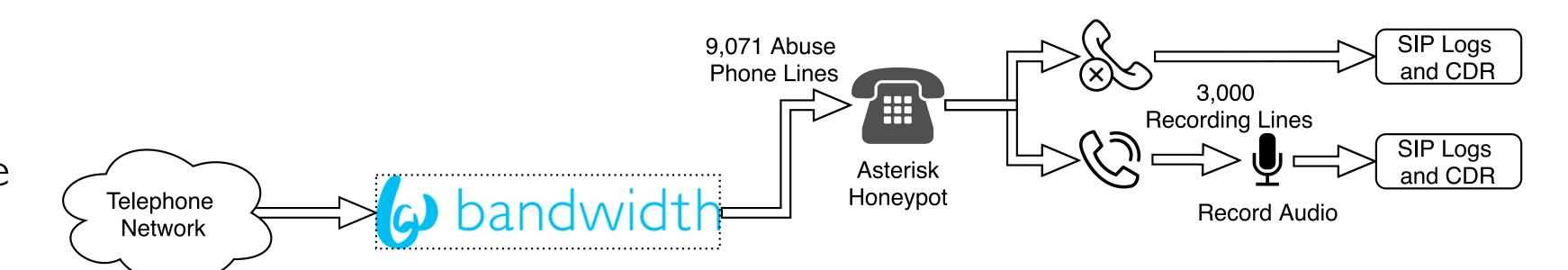


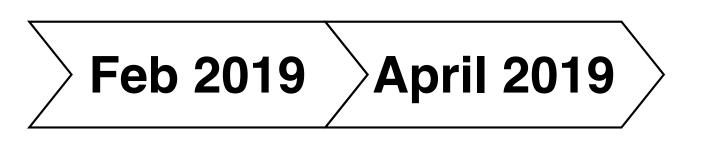






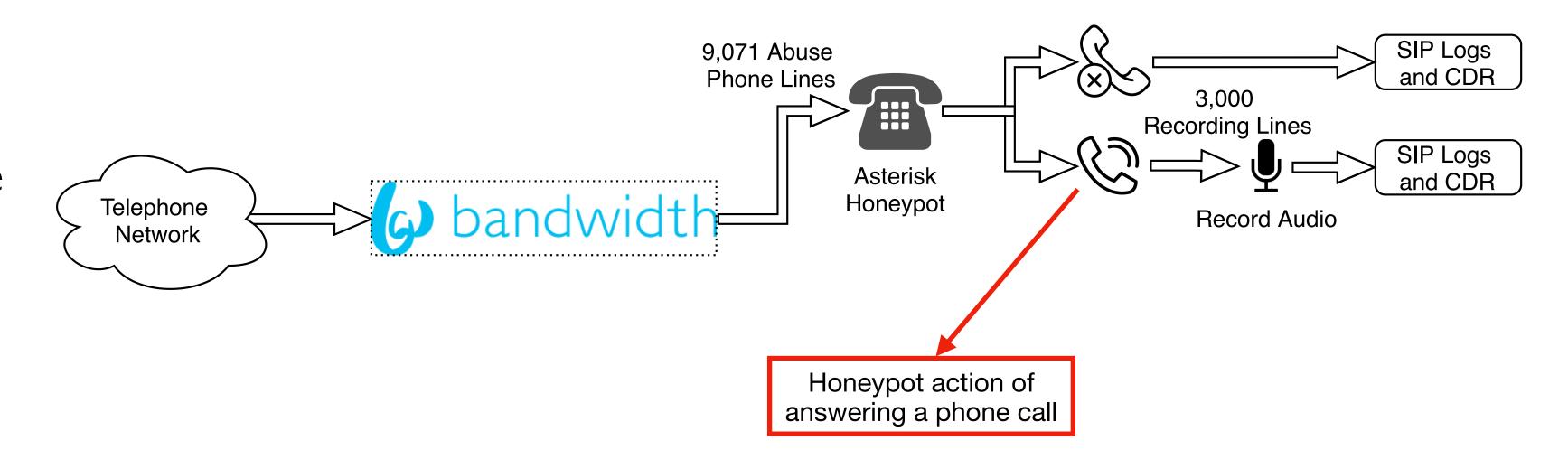
- Telephony honeypot setup to monitor unsolicited phone calls in the United States
- 3,000 randomly selected phone numbers answer phone calls while other numbers reject phone calls







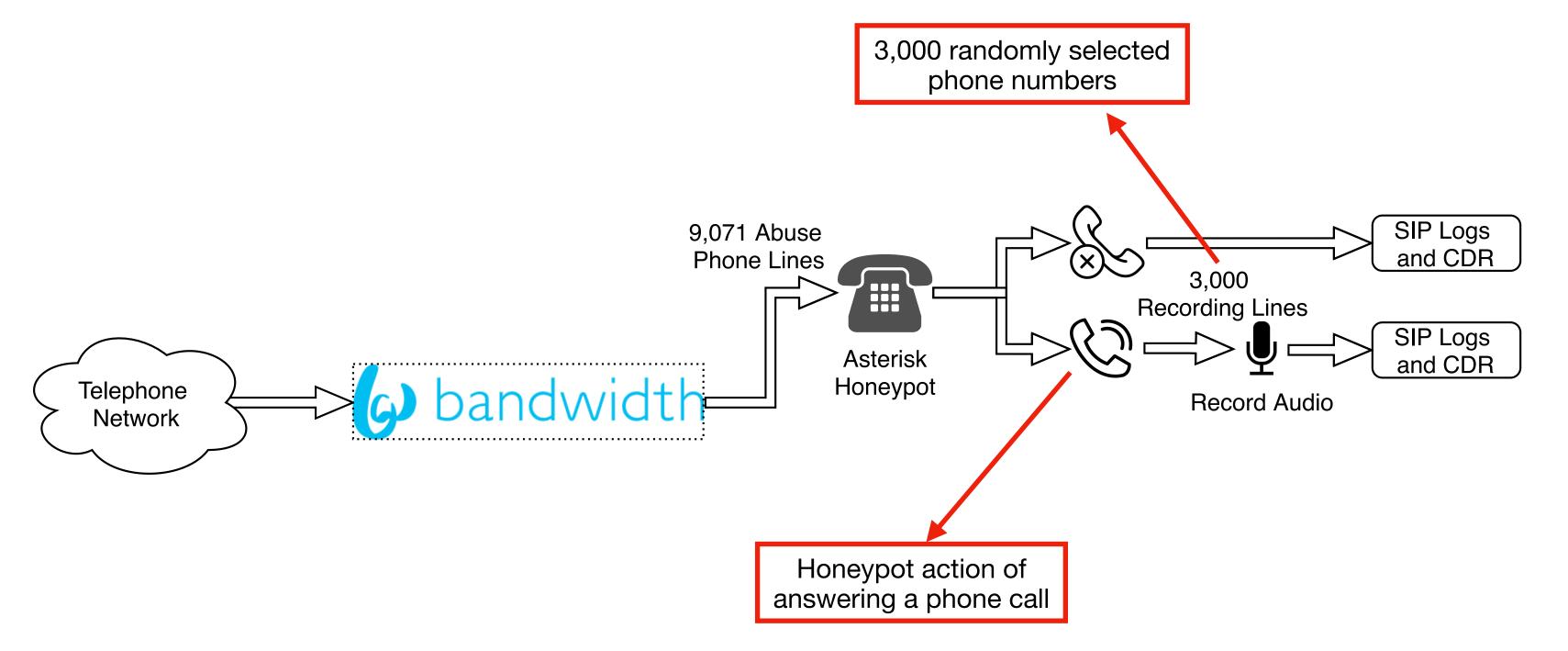
- Telephony honeypot setup to monitor unsolicited phone calls in the United States
- 3,000 randomly selected phone numbers answer phone calls while other numbers reject phone calls







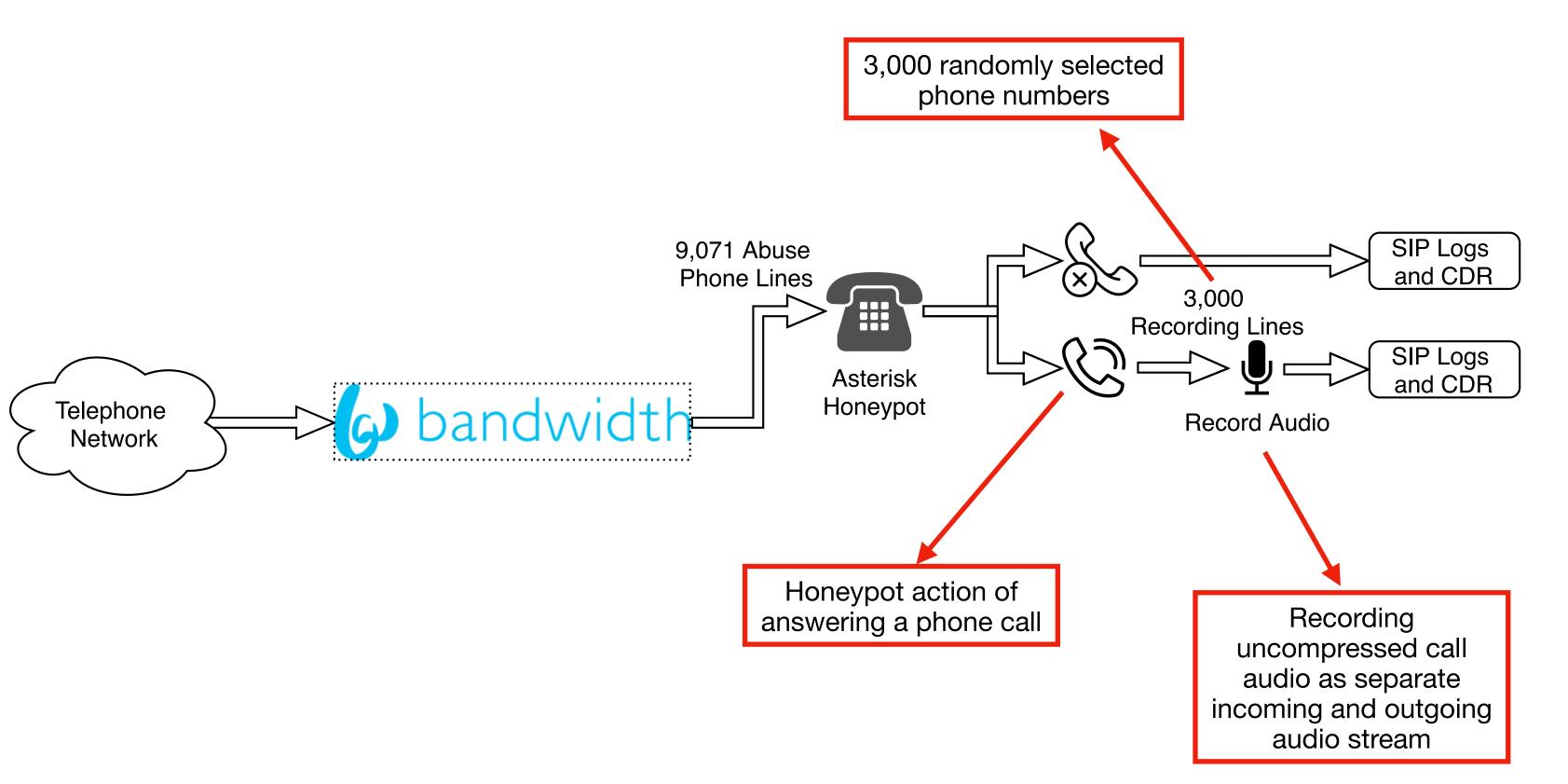
- Telephony honeypot setup to monitor unsolicited phone calls in the United States
- 3,000 randomly selected phone numbers answer phone calls while other numbers reject phone calls







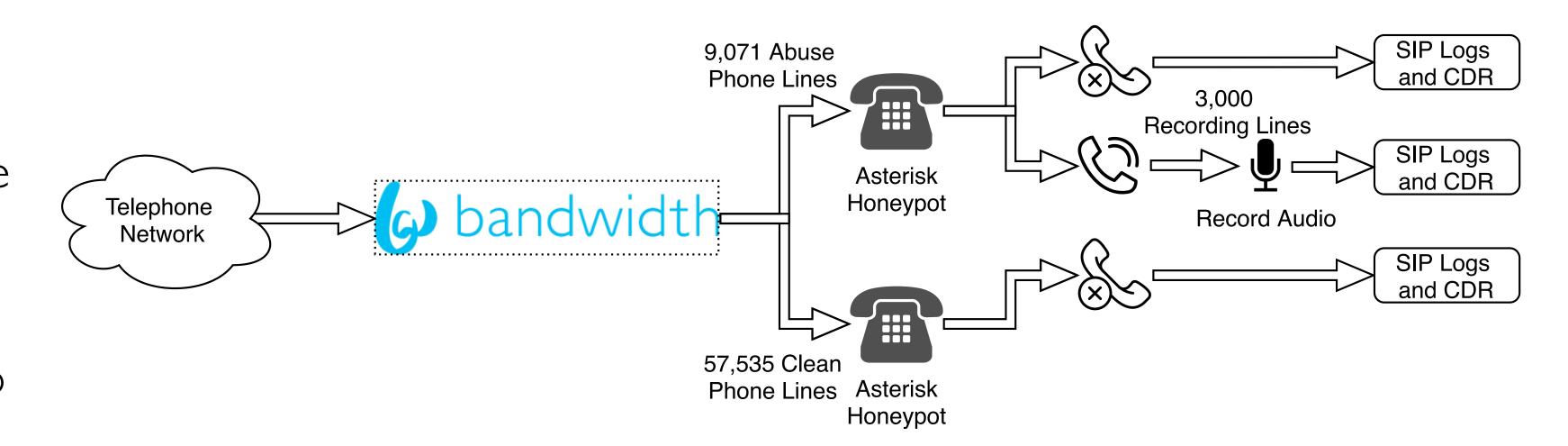
- Telephony honeypot setup to monitor unsolicited phone calls in the United States
- 3,000 randomly selected phone numbers answer phone calls while other numbers reject phone calls







- Telephony honeypot setup to monitor unsolicited phone calls in the United States
- 3,000 randomly selected phone numbers answer phone calls while other numbers reject phone calls
- Over 57,000 clean phone lines with no history of abuse added to our honeypot





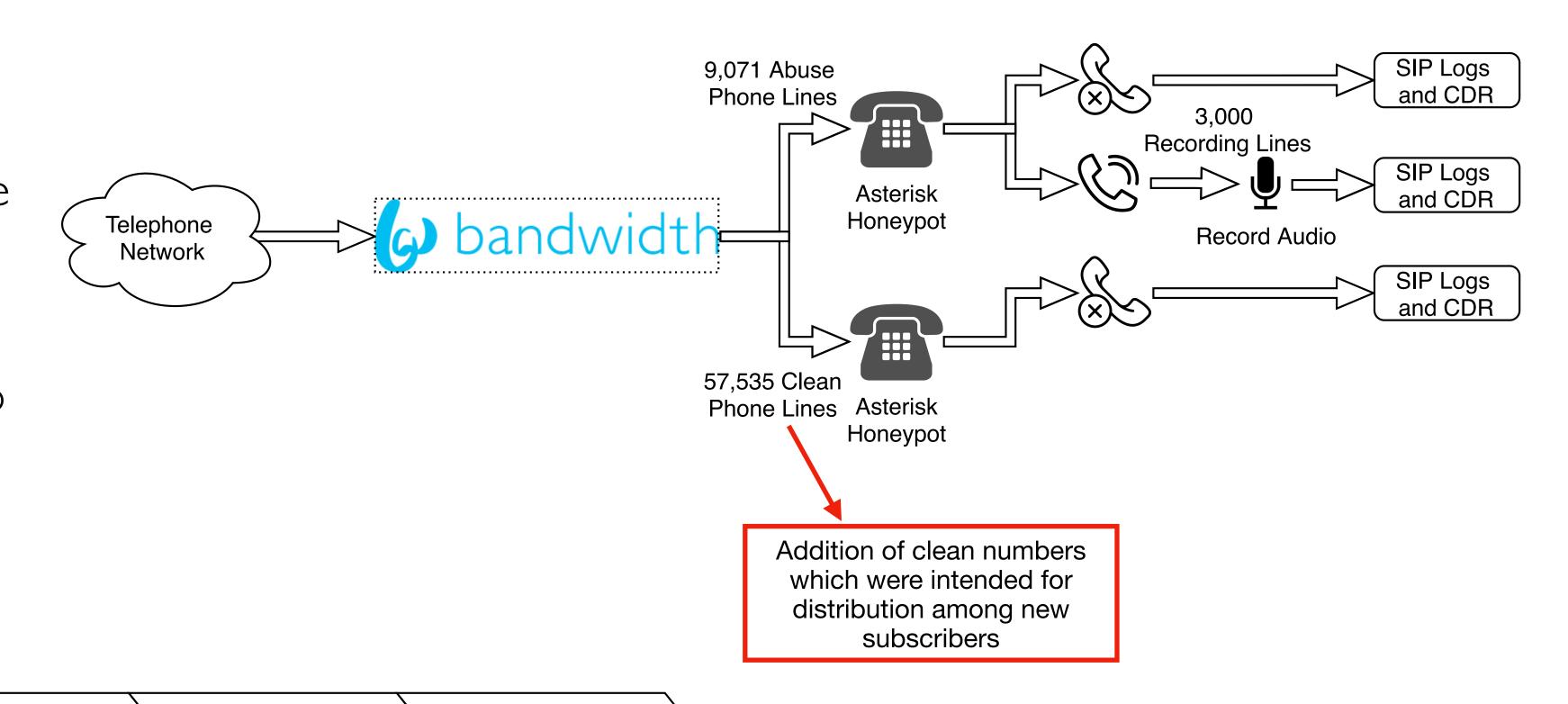


**April 2019** 

**Aug 2019** 

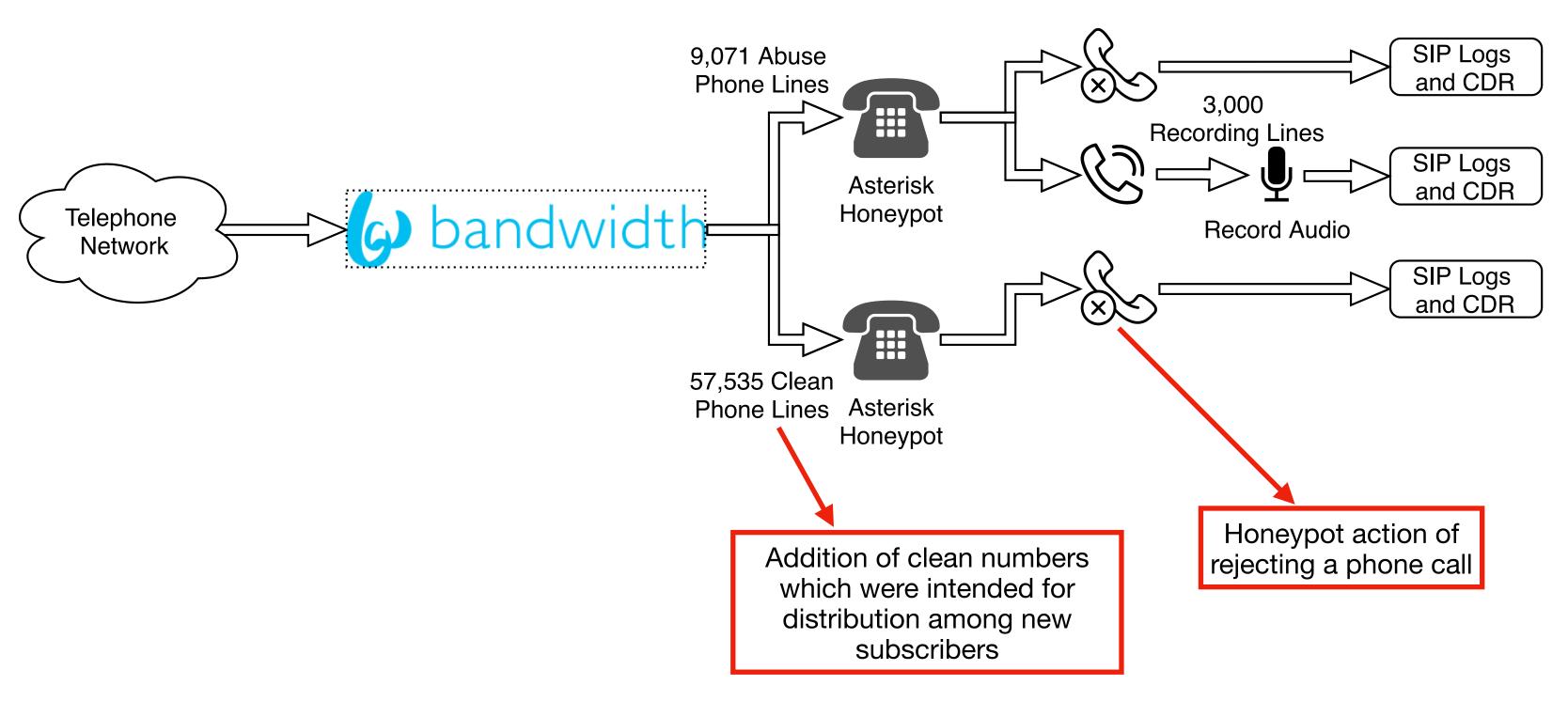
Feb 2019

- Telephony honeypot setup to monitor unsolicited phone calls in the United States
- 3,000 randomly selected phone numbers answer phone calls while other numbers reject phone calls
- Over 57,000 clean phone lines with no history of abuse added to our honeypot





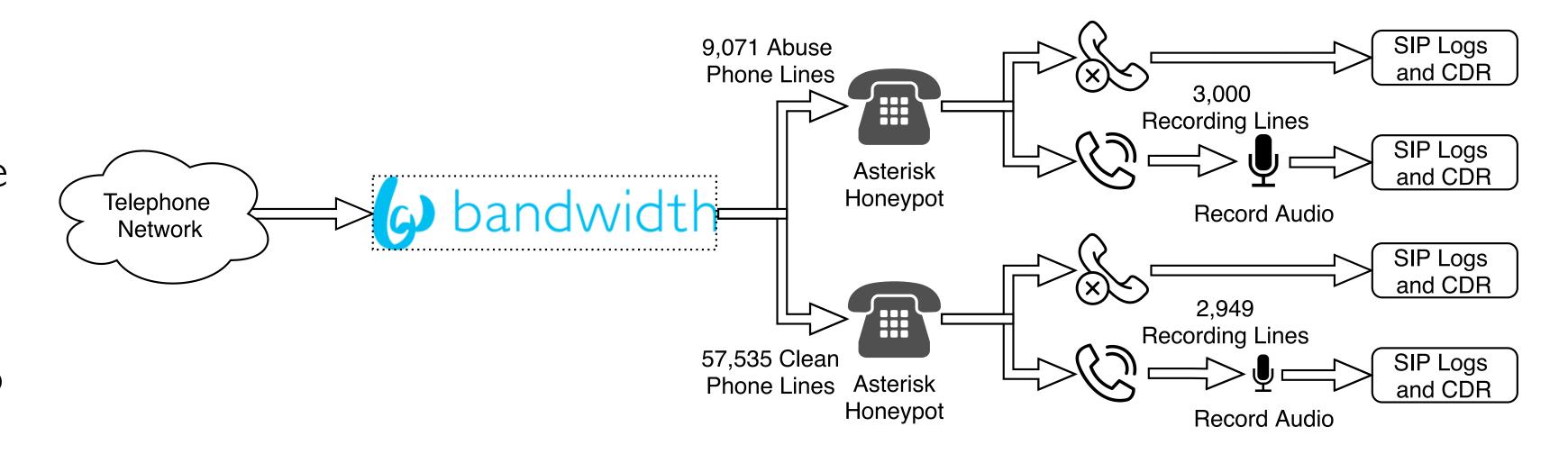
- Telephony honeypot setup to monitor unsolicited phone calls in the United States
- 3,000 randomly selected phone numbers answer phone calls while other numbers reject phone calls
- Over 57,000 clean phone lines with no history of abuse added to our honeypot



Feb 2019 April 2019 Aug 2019



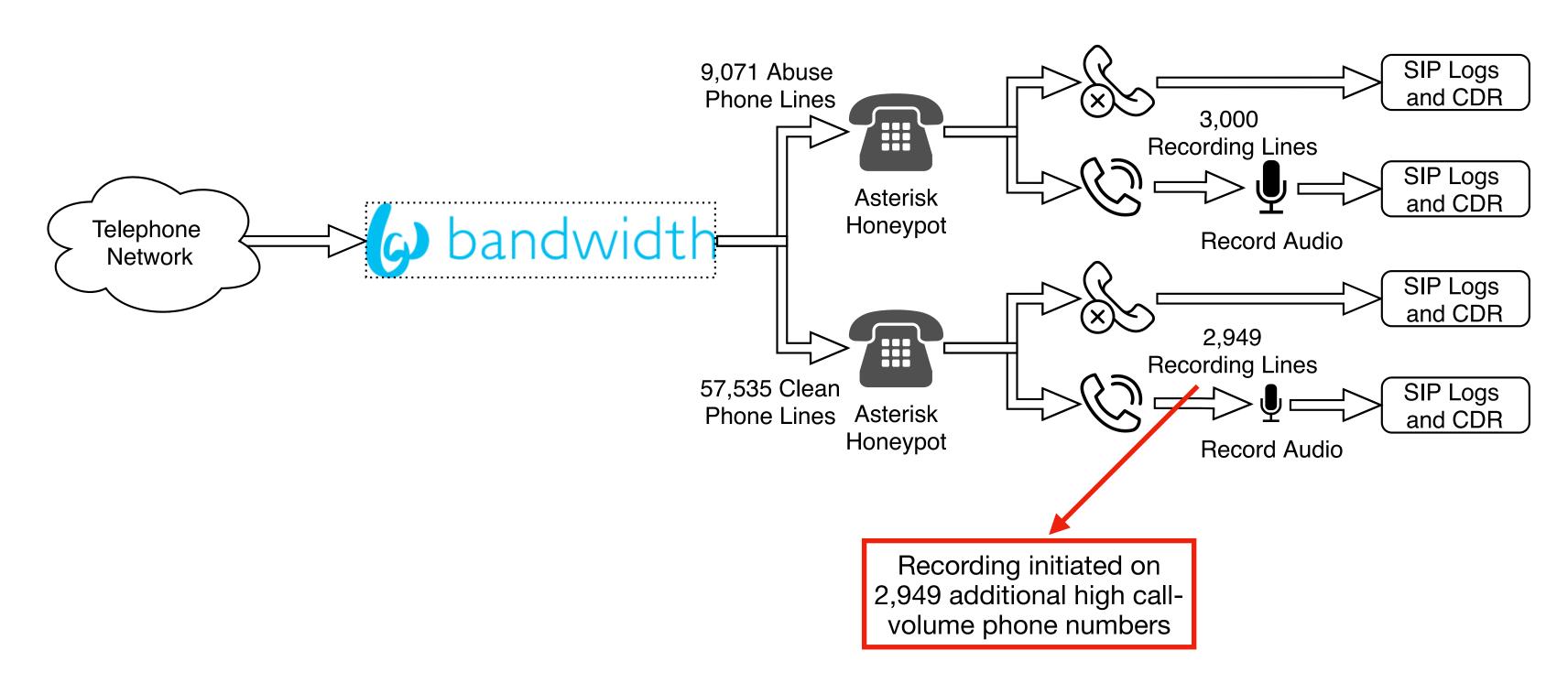
- Telephony honeypot setup to monitor unsolicited phone calls in the United States
- 3,000 randomly selected phone numbers answer phone calls while other numbers reject phone calls
- Over 57,000 clean phone lines with no history of abuse added to our honeypot
- 2,949 high call-volume phone numbers configured to answer phone calls







- Telephony honeypot setup to monitor unsolicited phone calls in the United States
- 3,000 randomly selected phone numbers answer phone calls while other numbers reject phone calls
- Over 57,000 clean phone lines with no history of abuse added to our honeypot
- 2,949 high call-volume phone numbers configured to answer phone calls





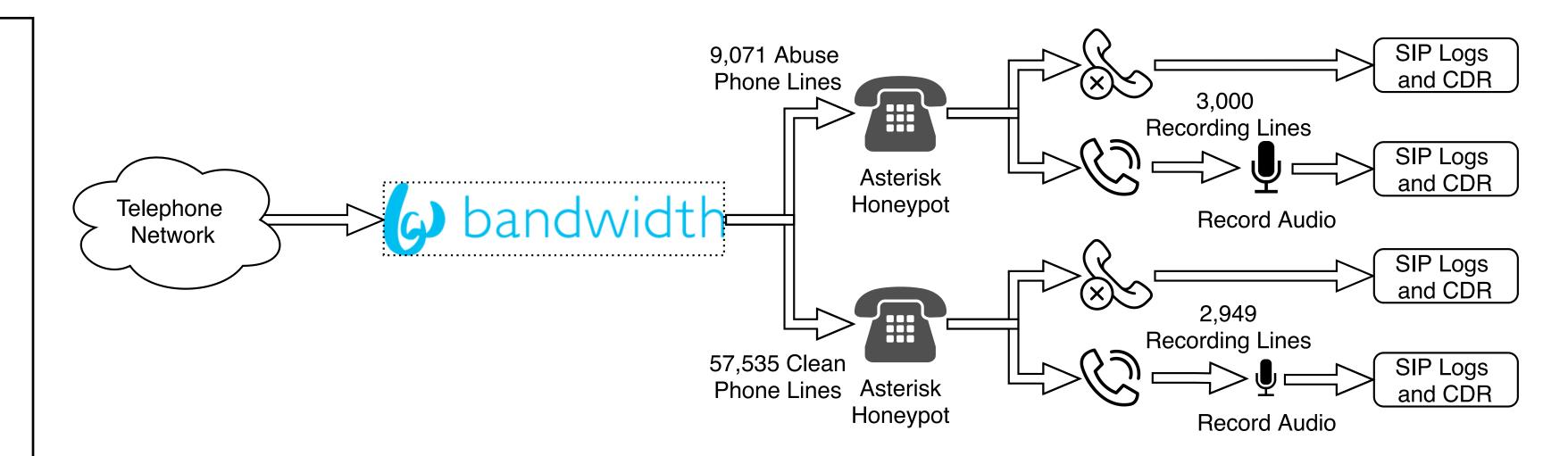


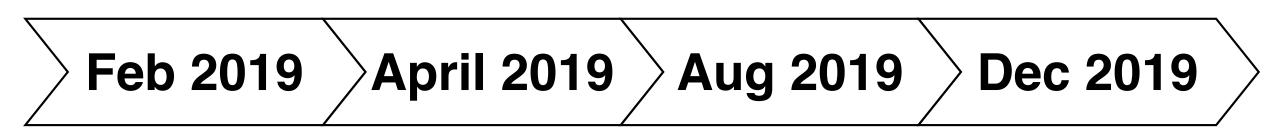
11 Months

66,000 Phone numbers

145,000 Call recordings

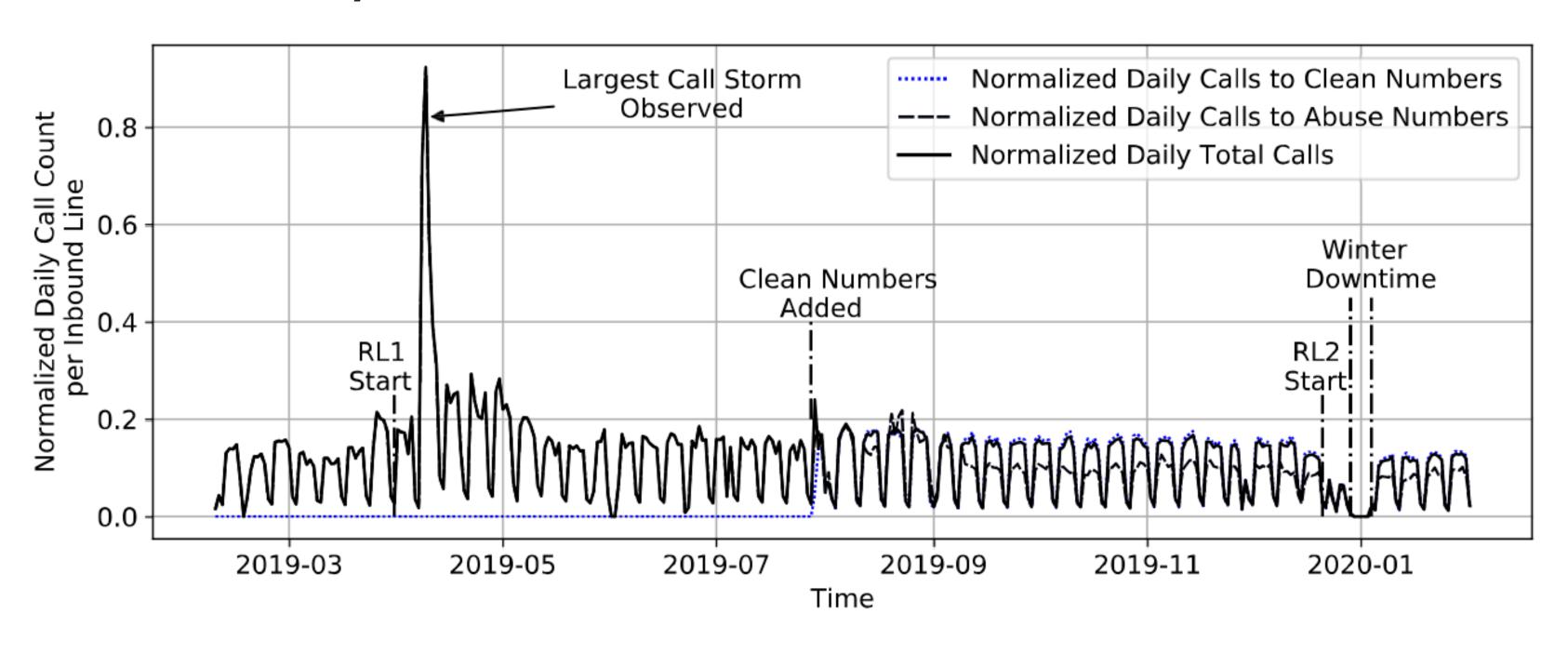
1.48 Million phone calls



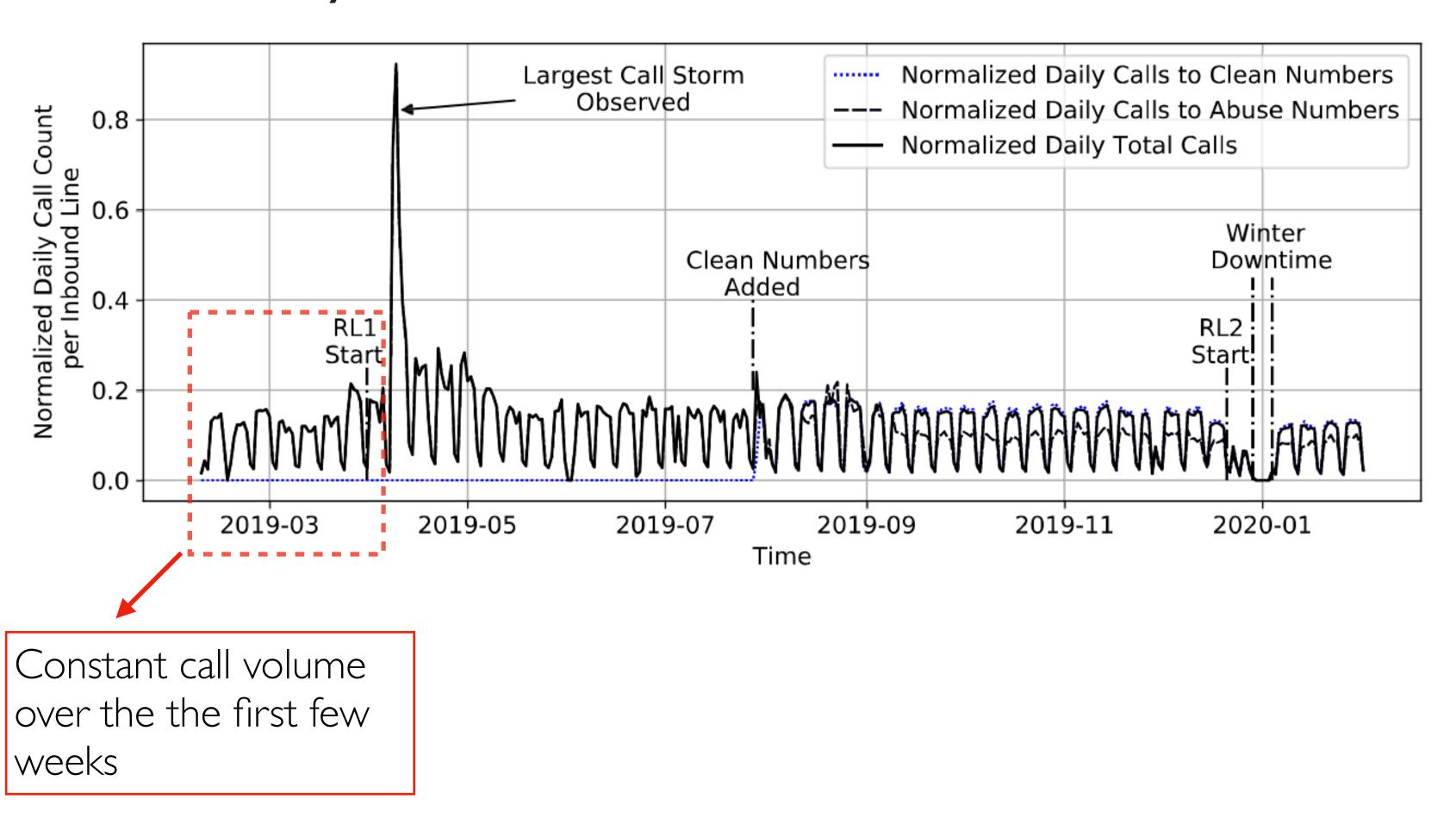


\* A detailed discussion of ethical and legal considerations of our data collection and experiments can be found in Section 3.4 of our paper

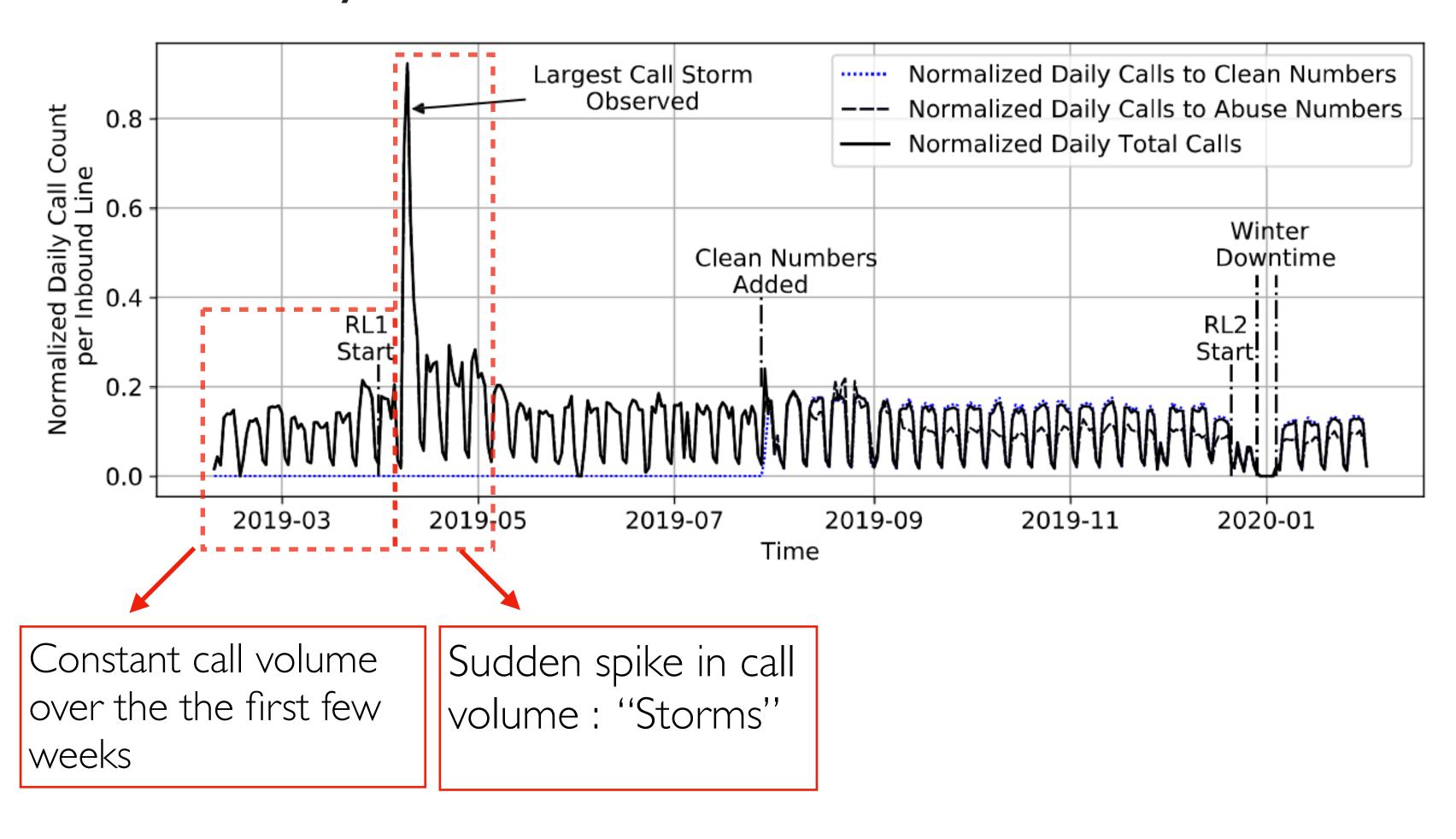




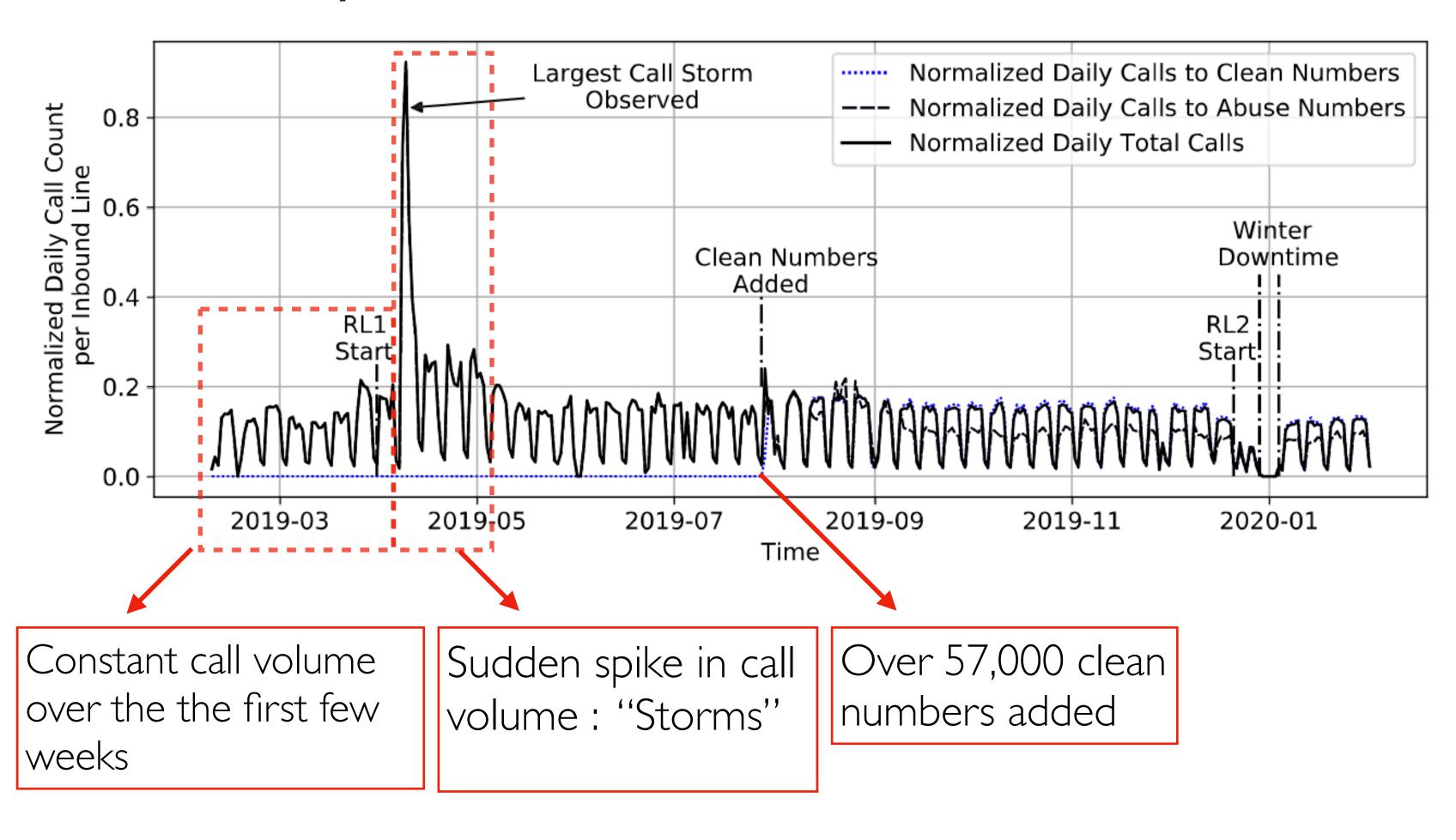




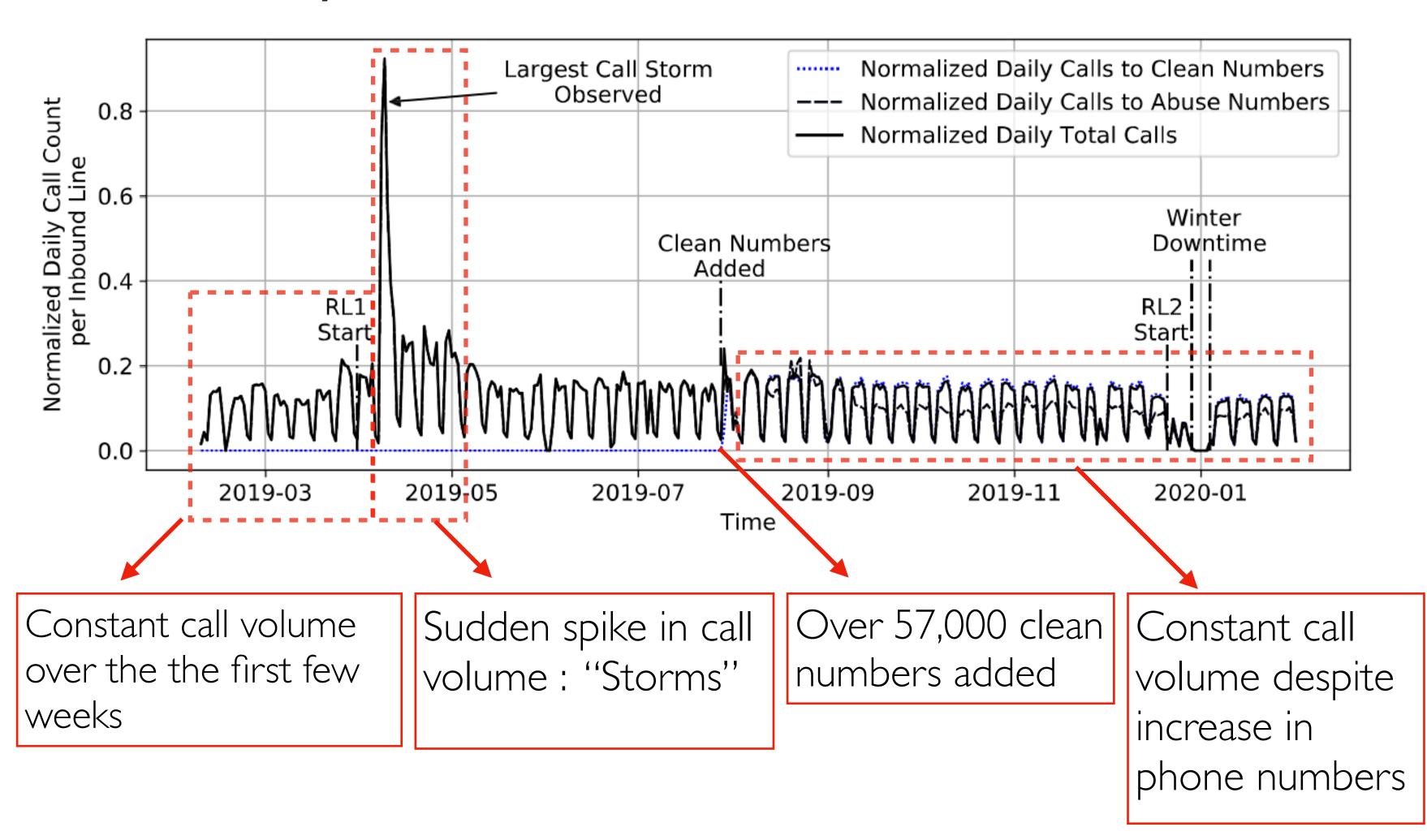




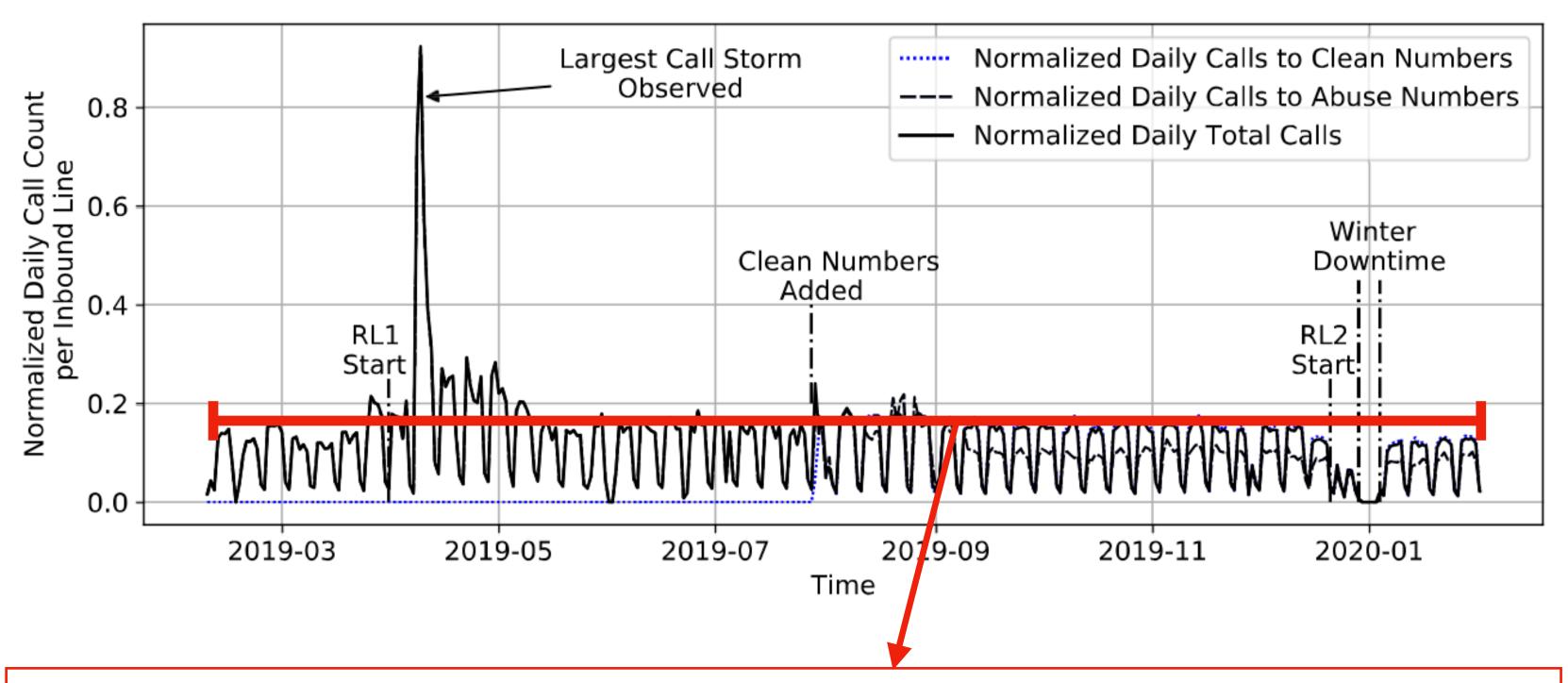








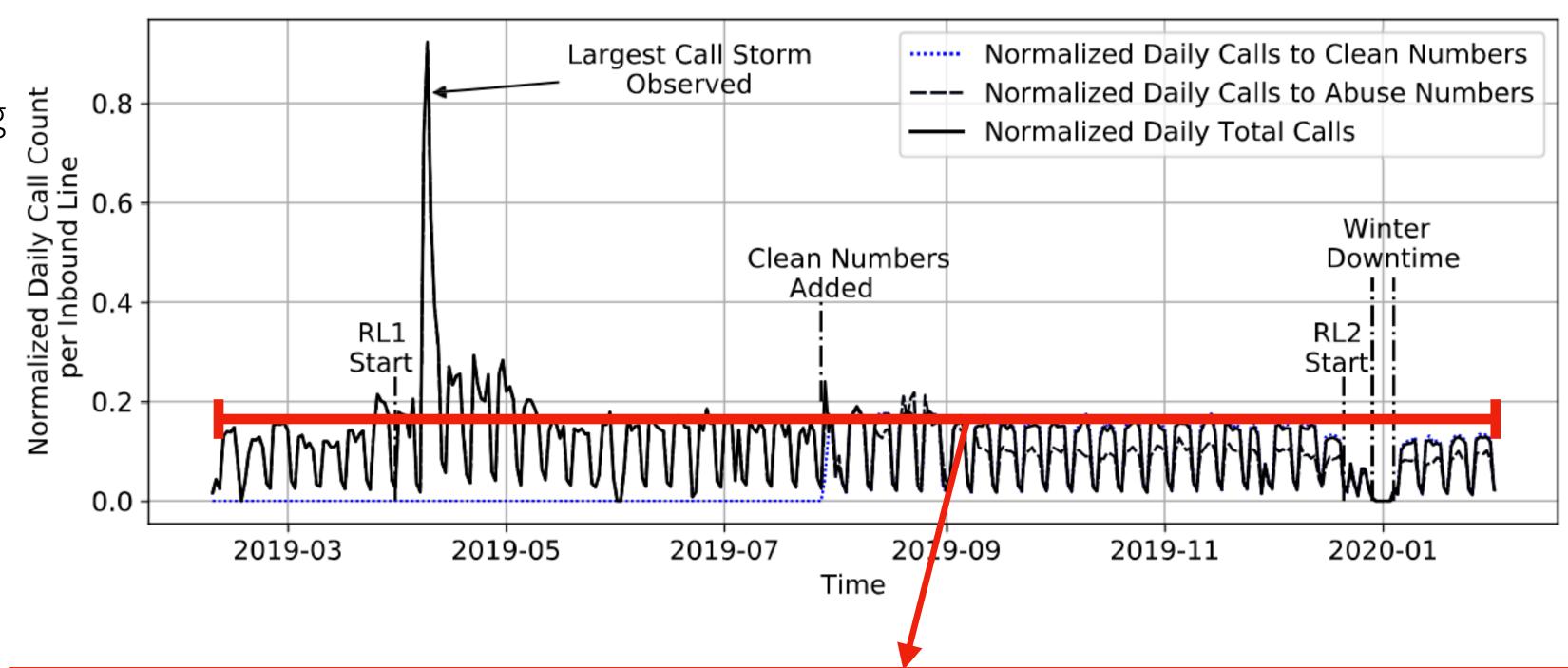




Linear model fit over weekly call volume with a slope of -9  $\times$  10 <sup>-5</sup>



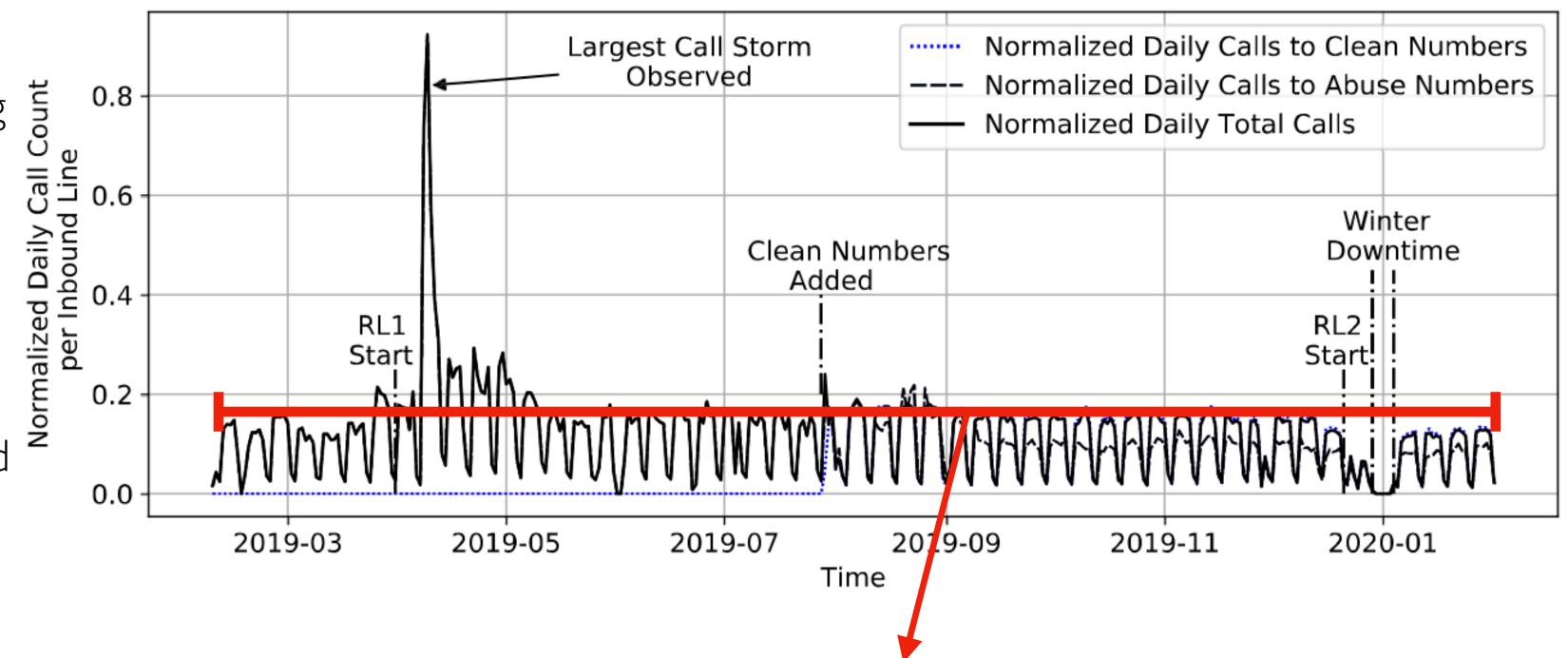
 Volume of unsolicited phone calls neither increased nor decreased during our study — it remained stationary



Linear model fit over weekly call volume with a slope of  $-9 \times 10^{-5}$ 



- Volume of unsolicited phone calls neither increased nor decreased during our study — it remained stationary
- We observed and characterized "Storms" a rare phenomenon of abnormally large number of unsolicited calls received within a day. We uncovered 648 storms on 223 inbound lines

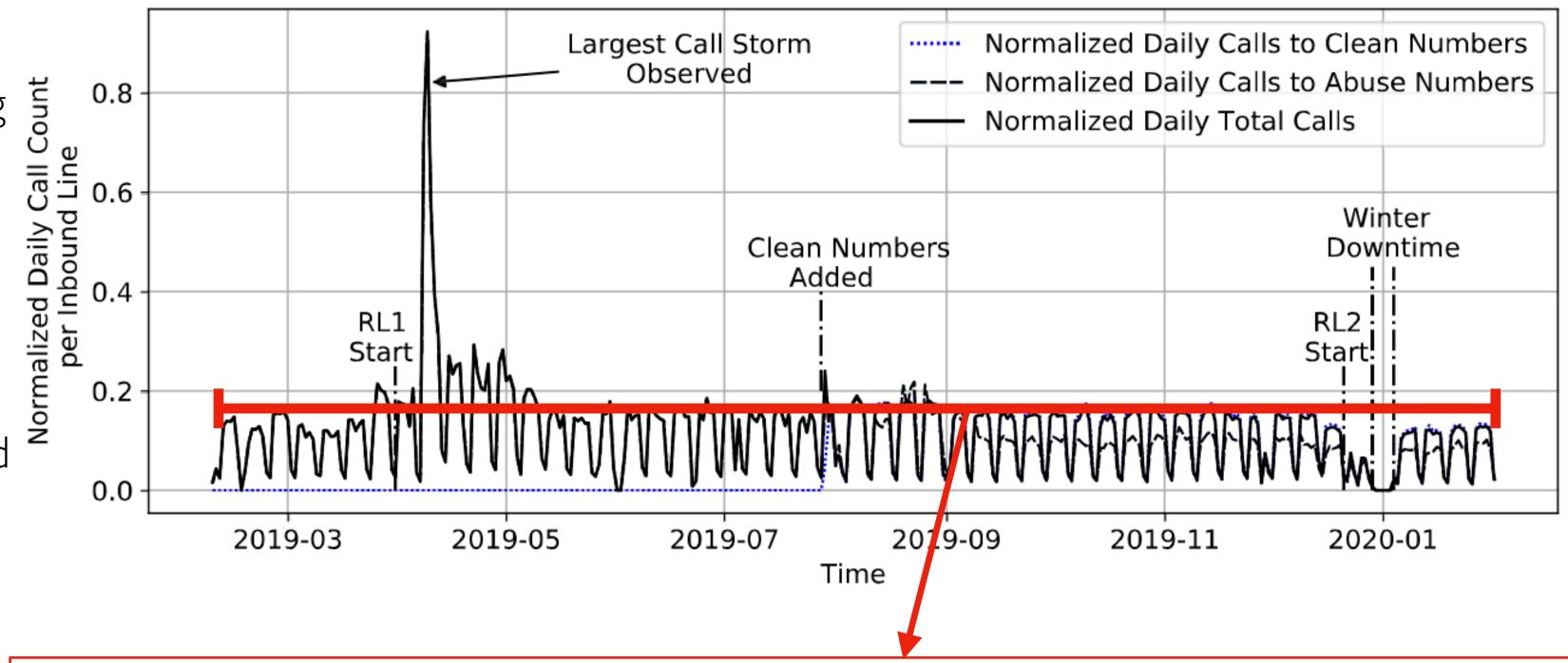


Linear model fit over weekly call volume with a slope of  $-9 \times 10^{-5}$ 



## Robocalls by the Numbers

- Volume of unsolicited phone calls neither increased nor decreased during our study — it remained stationary
- We observed and characterized "Storms" a rare phenomenon of abnormally large number of unsolicited calls received within a day. We uncovered 648 storms on 223 inbound lines
- Over 90% of these unsolicited phone calls were made during weekdays and over 80% during local working hours. We observed a strong weekly periodicity

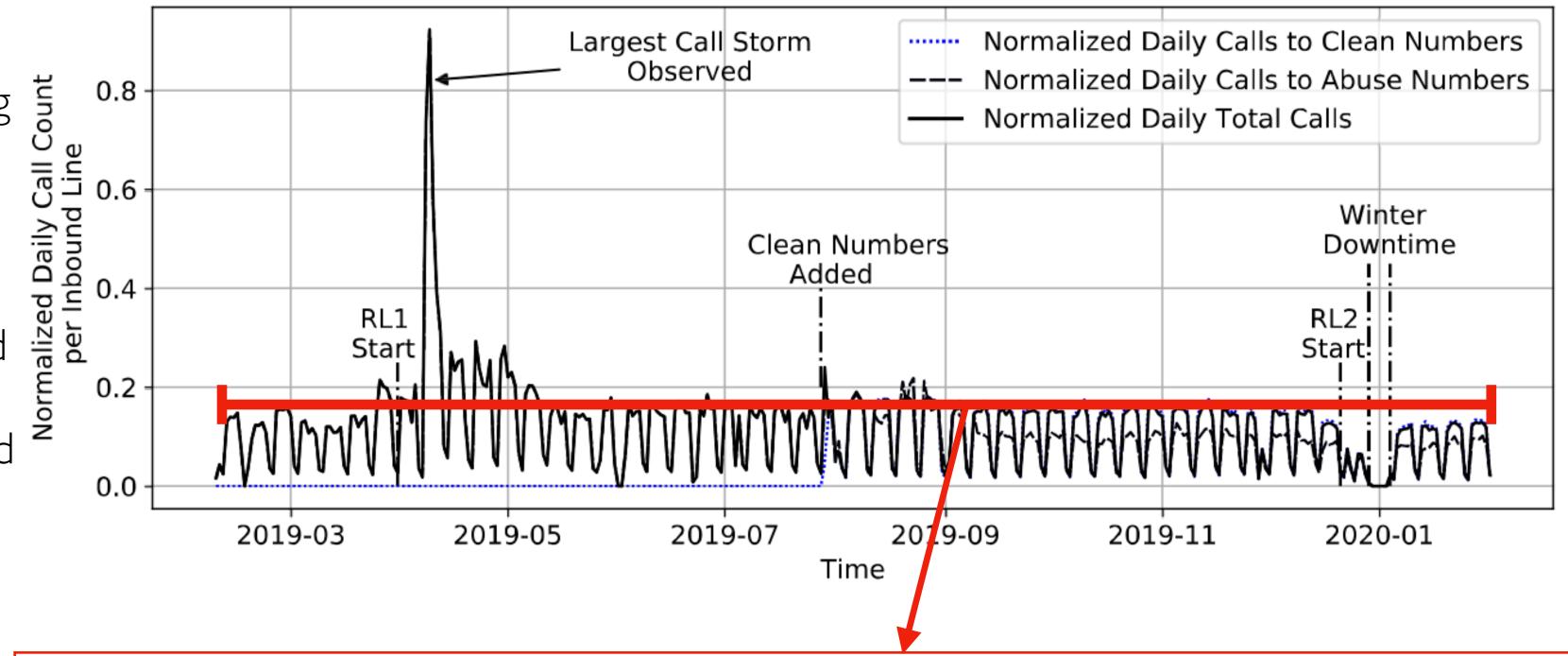


Linear model fit over weekly call volume with a slope of -9  $\times$  10 <sup>-5</sup>



## Robocalls by the Numbers

- Volume of unsolicited phone calls neither increased nor decreased during our study — it remained stationary
- We observed and characterized "Storms" a rare phenomenon of abnormally large number of unsolicited calls received within a day. We uncovered 648 storms on 223 inbound lines
- Over 90% of these unsolicited phone calls were made during weekdays and over 80% during local working hours.
   We observed a strong weekly periodicity
- Robocalls are a major problem in the United States with no signs of decline!



Linear model fit over weekly call volume with a slope of -9  $\times$  10 <sup>-5</sup>



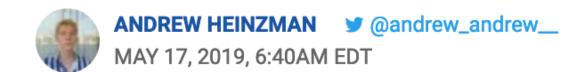








#### Tired of Robocalls? Stop Answering Your Phone





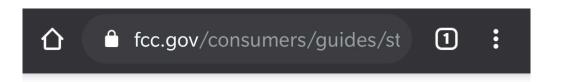




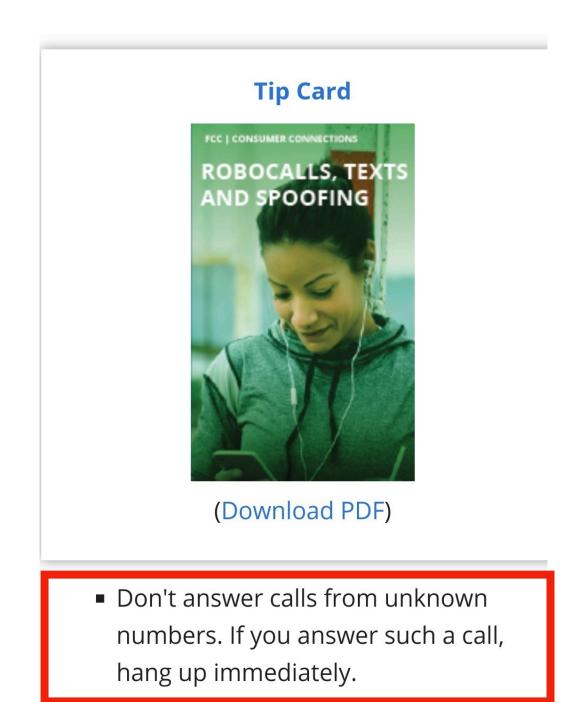


#### Tired of Robocalls? Stop Answering Your Phone





Consumer Tips to Stop Unwanted Robocalls and Avoid Phone Scams



\* How-to Geek: <a href="https://www.howtogeek.com/413873/tired-of-robocalls-stop-answering-your-phone/">https://www.howtogeek.com/413873/tired-of-robocalls-stop-answering-your-phone/</a>

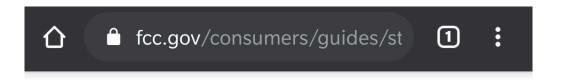




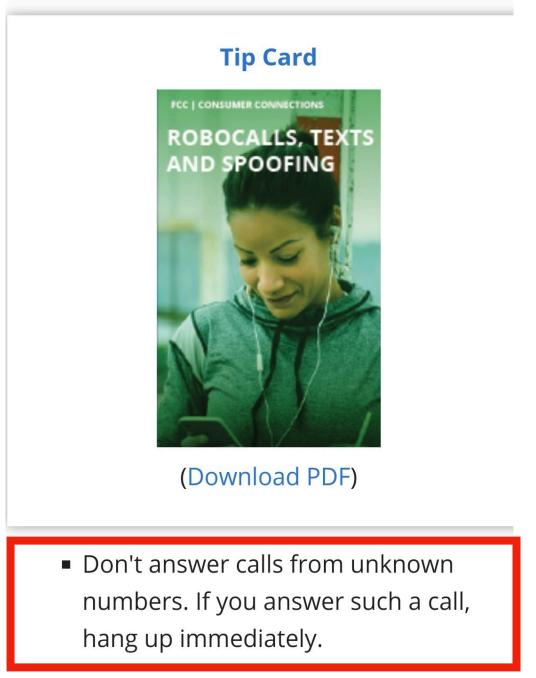
#### Tired of Robocalls? Stop Answering Your Phone



We declined every unsolicited call received on 3000 phone numbers for 6 weeks



Consumer Tips to Stop Unwanted Robocalls and Avoid Phone Scams





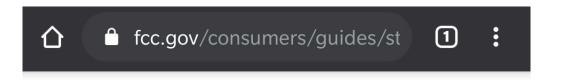
<sup>\*</sup> How-to Geek: <a href="https://www.howtogeek.com/413873/tired-of-robocalls-stop-answering-your-phone/">https://www.howtogeek.com/413873/tired-of-robocalls-stop-answering-your-phone/</a>



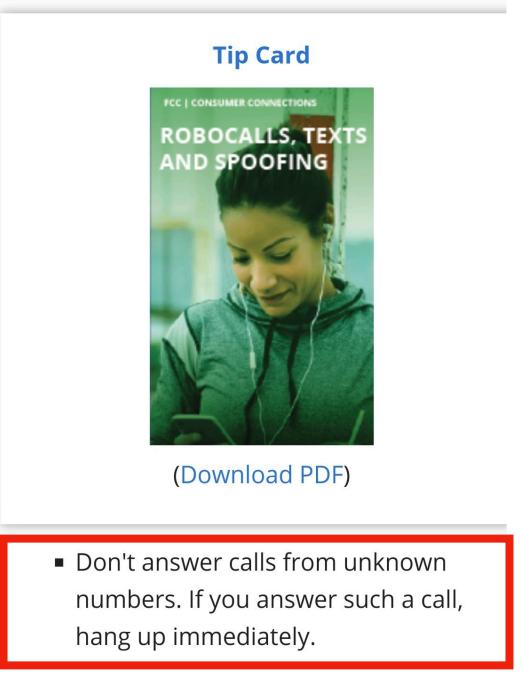
#### Tired of Robocalls? Stop Answering Your Phone



- We declined every unsolicited call received on 3000 phone numbers for 6 weeks
- Next, we answered every unsolicited call received on these numbers for 6 more weeks



Consumer Tips to Stop Unwanted Robocalls and Avoid Phone Scams





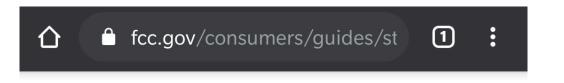
<sup>\*</sup> How-to Geek: <a href="https://www.howtogeek.com/413873/tired-of-robocalls-stop-answering-your-phone/">https://www.howtogeek.com/413873/tired-of-robocalls-stop-answering-your-phone/</a>



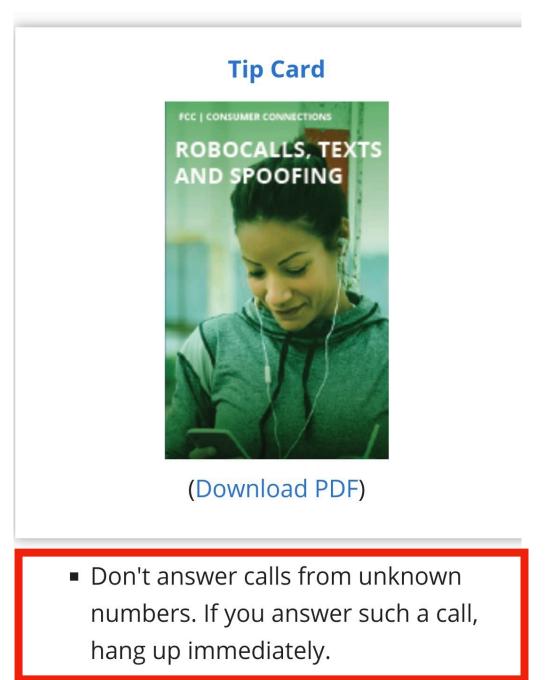
#### Tired of Robocalls? Stop Answering Your Phone



- We declined every unsolicited call received on 3000 phone numbers for 6 weeks
- Next, we answered every unsolicited call received on these numbers for 6 more weeks
- We compared the average call volumes during these two phases



Consumer Tips to Stop Unwanted Robocalls and Avoid Phone Scams



\* How-to Geek: <a href="https://www.howtogeek.com/413873/tired-of-robocalls-stop-answering-your-phone/">https://www.howtogeek.com/413873/tired-of-robocalls-stop-answering-your-phone/</a>





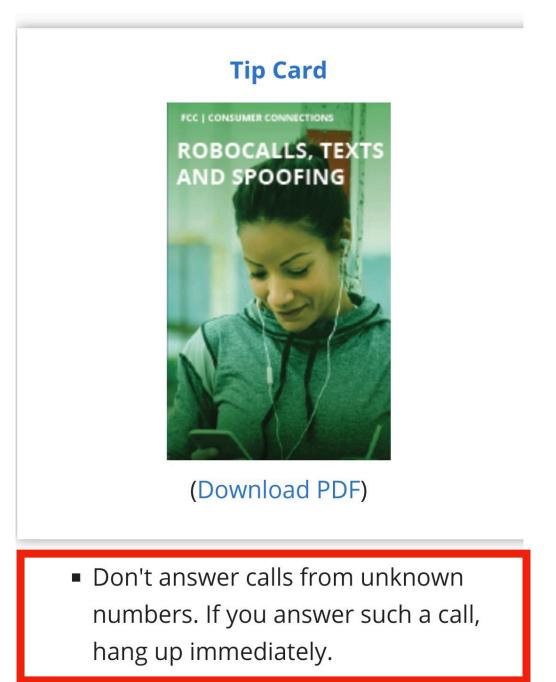
#### Tired of Robocalls? Stop Answering Your Phone



- We declined every unsolicited call received on 3000 phone numbers for 6 weeks
- Next, we answered every unsolicited call received on these numbers for 6 more weeks
- · We compared the average call volumes during these two phases
- Does answering an unsolicited phone call have a significant effect on the number of such unsolicited phone calls you receive?



Consumer Tips to Stop Unwanted Robocalls and Avoid Phone Scams



\* How-to Geek: <a href="https://www.howtogeek.com/413873/tired-of-robocalls-stop-answering-your-phone/">https://www.howtogeek.com/413873/tired-of-robocalls-stop-answering-your-phone/</a>

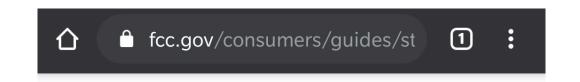


MAY 17, 2019, 6:40AM EDT

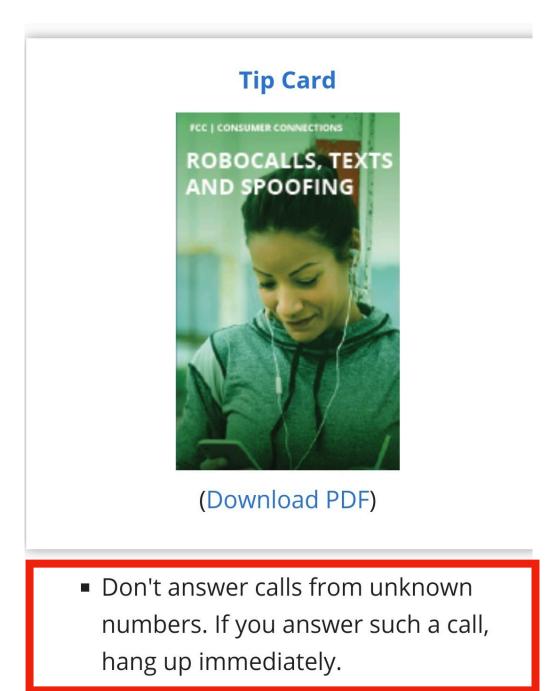
#### Will You Receive More Robocalls if You Answer them?



- We declined every unsolities all reliver on 3000 public ball reliver on 3000 public ball relivers for 6 weeks
- Next, we answered every unplicited a received on these numbers for 6 more weeks
- · We compared the average call volumes during these two phases
- Does answering an unsolicited phone call have a significant effect on the number of such unsolicited phone calls you receive?



Consumer Tips to Stop Unwanted Robocalls and Avoid Phone Scams



\* How-to Geek: <a href="https://www.howtogeek.com/413873/tired-of-robocalls-stop-answering-your-phone/">https://www.howtogeek.com/413873/tired-of-robocalls-stop-answering-your-phone/</a>





 Key Insight: An operation will use identical/ similar audio recording to generate large number of unsolicited phone calls



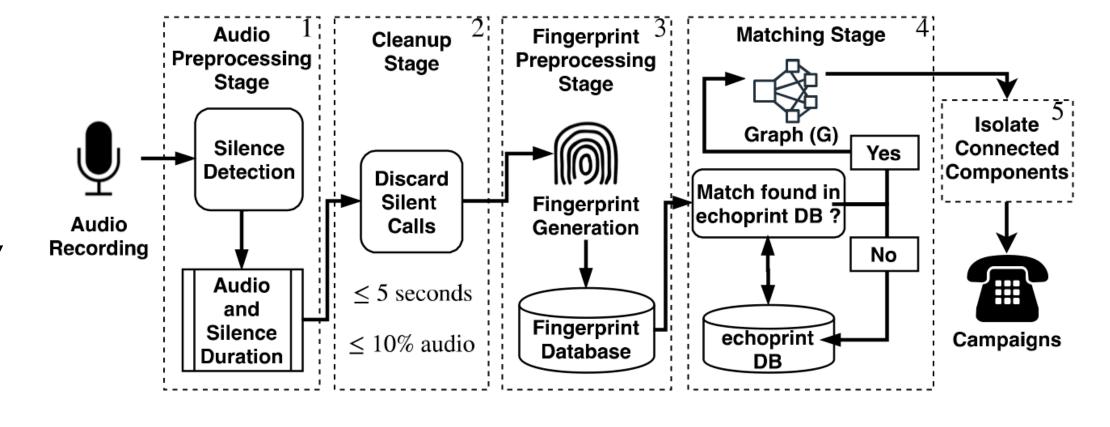
- Key Insight: An operation will use identical/ similar audio recording to generate large number of unsolicited phone calls
- Groups of calls with similar audio constitutes a robocalling campaign



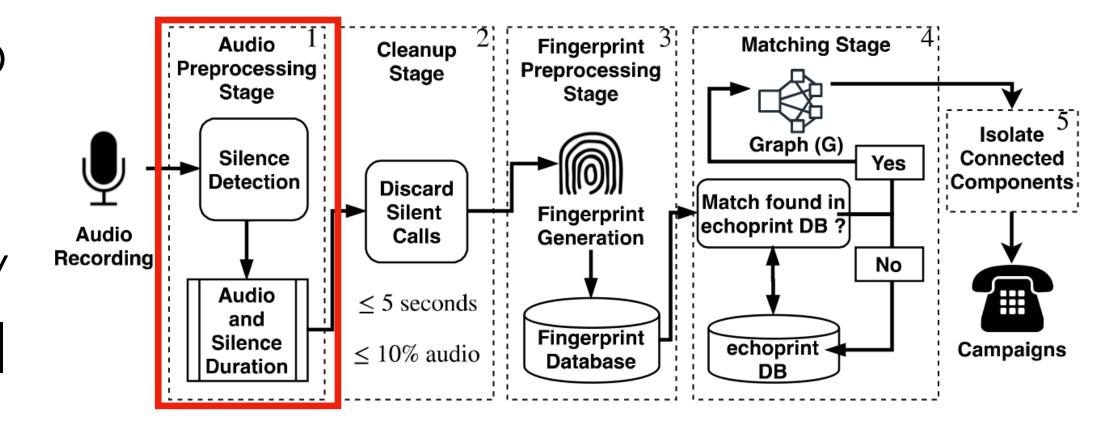
- Key Insight: An operation will use identical/ similar audio recording to generate large number of unsolicited phone calls
- Groups of calls with similar audio constitutes a robocalling campaign
- We developed a 5 stage pipeline to tightly cluster call audio by associating individual calls into broader campaigns



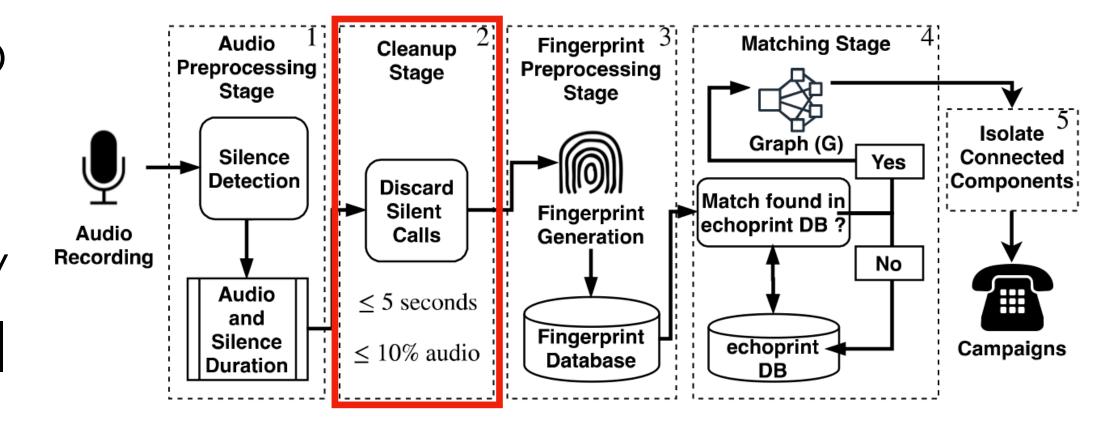
- Key Insight: An operation will use identical/ similar audio recording to generate large number of unsolicited phone calls
- Groups of calls with similar audio constitutes a robocalling campaign
- We developed a 5 stage pipeline to tightly cluster call audio by associating individual calls into broader campaigns



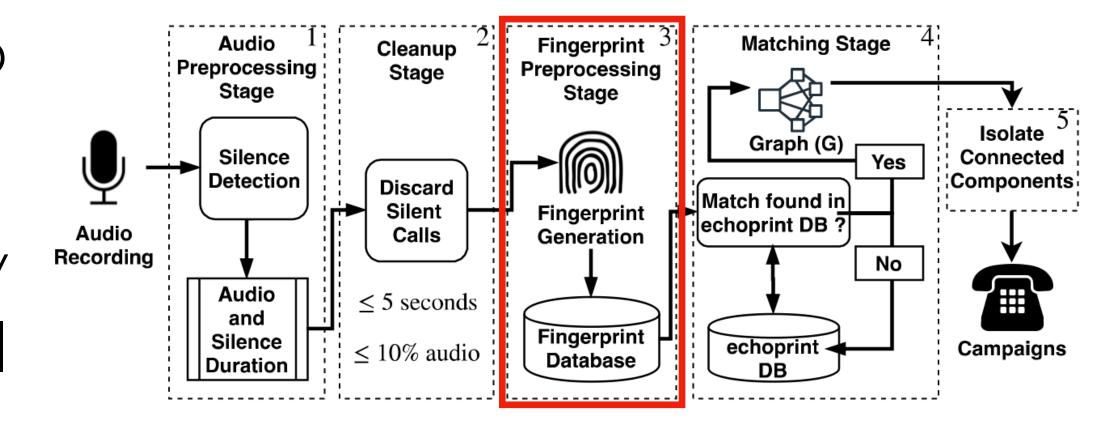
- Key Insight: An operation will use identical/ similar audio recording to generate large number of unsolicited phone calls
- Groups of calls with similar audio constitutes a robocalling campaign
- We developed a 5 stage pipeline to tightly cluster call audio by associating individual calls into broader campaigns



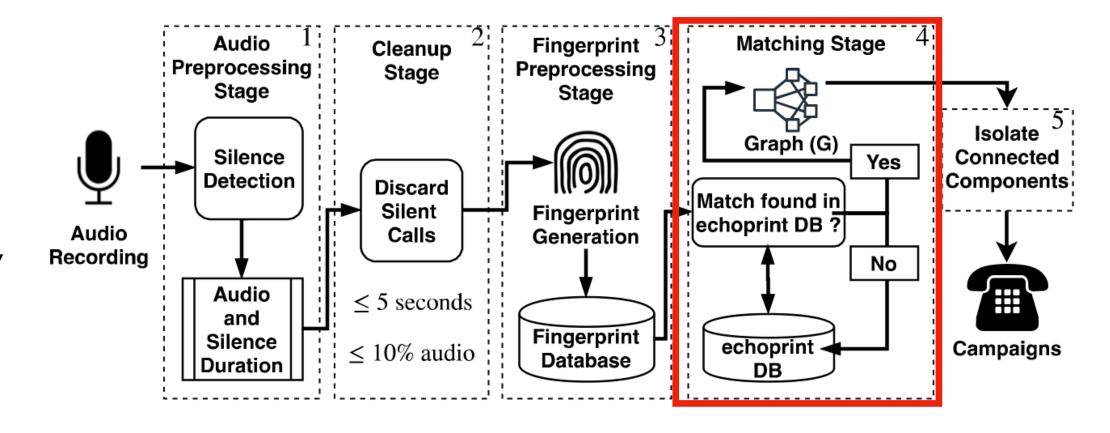
- Key Insight: An operation will use identical/ similar audio recording to generate large number of unsolicited phone calls
- Groups of calls with similar audio constitutes a robocalling campaign
- We developed a 5 stage pipeline to tightly cluster call audio by associating individual calls into broader campaigns



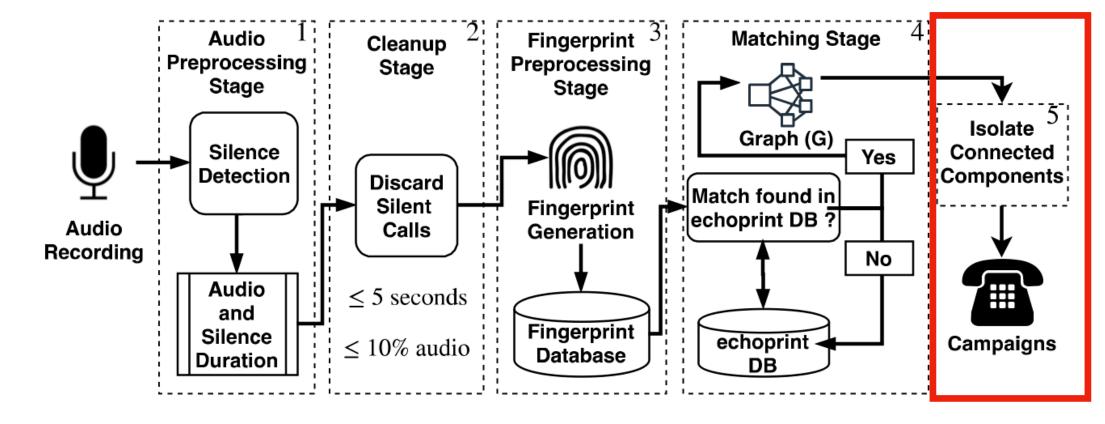
- Key Insight: An operation will use identical/ similar audio recording to generate large number of unsolicited phone calls
- Groups of calls with similar audio constitutes a robocalling campaign
- We developed a 5 stage pipeline to tightly cluster call audio by associating individual calls into broader campaigns



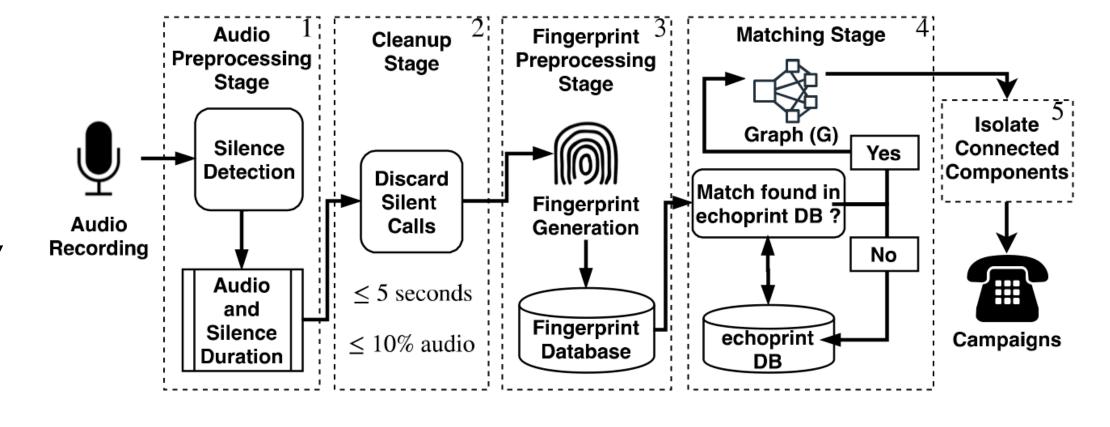
- Key Insight: An operation will use identical/ similar audio recording to generate large number of unsolicited phone calls
- Groups of calls with similar audio constitutes a robocalling campaign
- We developed a 5 stage pipeline to tightly cluster call audio by associating individual calls into broader campaigns



- Key Insight: An operation will use identical/ similar audio recording to generate large number of unsolicited phone calls
- Groups of calls with similar audio constitutes a robocalling campaign
- We developed a 5 stage pipeline to tightly cluster call audio by associating individual calls into broader campaigns

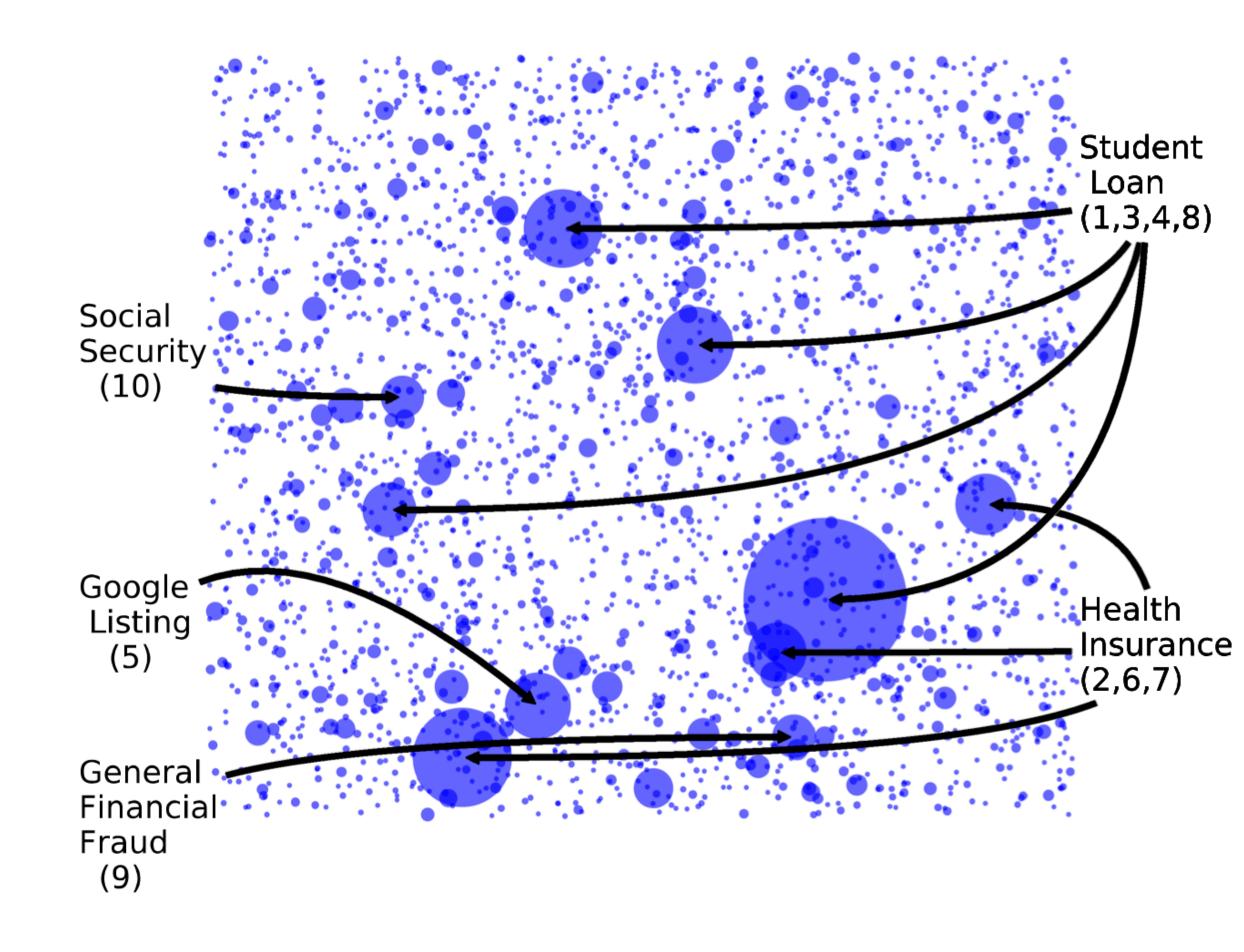


- Key Insight: An operation will use identical/ similar audio recording to generate large number of unsolicited phone calls
- Groups of calls with similar audio constitutes a robocalling campaign
- We developed a 5 stage pipeline to tightly cluster call audio by associating individual calls into broader campaigns



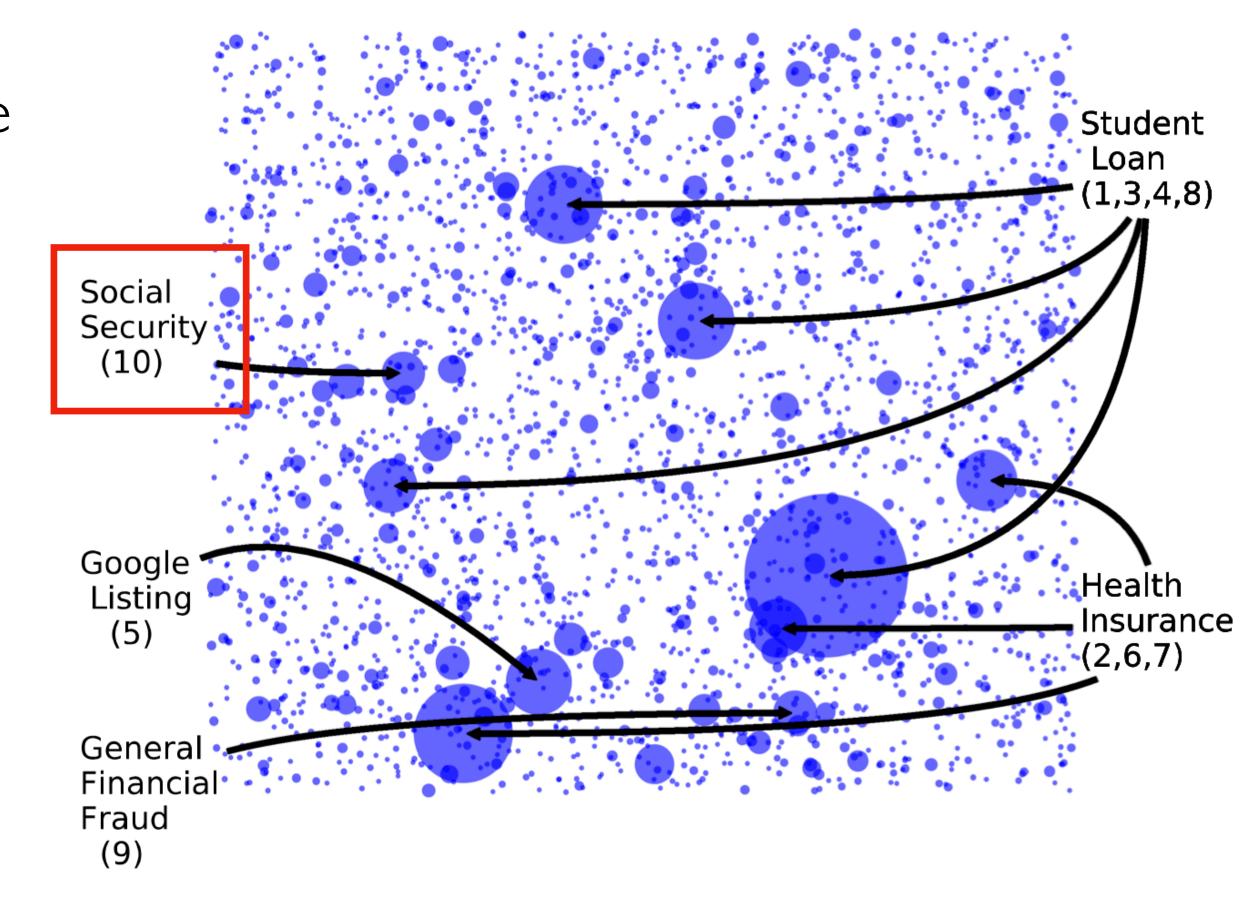


• 2,687 unique robocalling campaigns



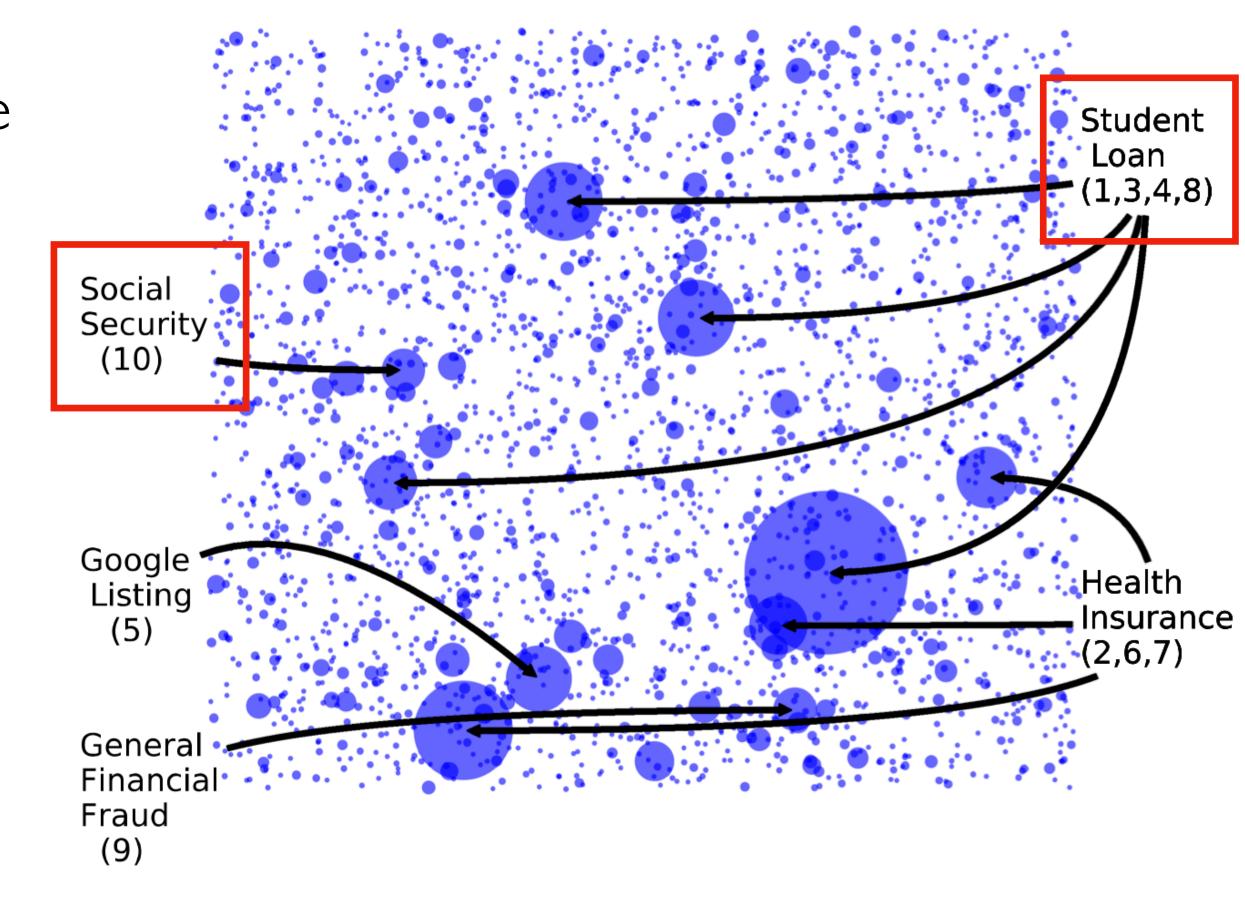


- 2,687 unique robocalling campaigns
- A fraudulent Social Security Campaign was the 10th largest campaign observed in our



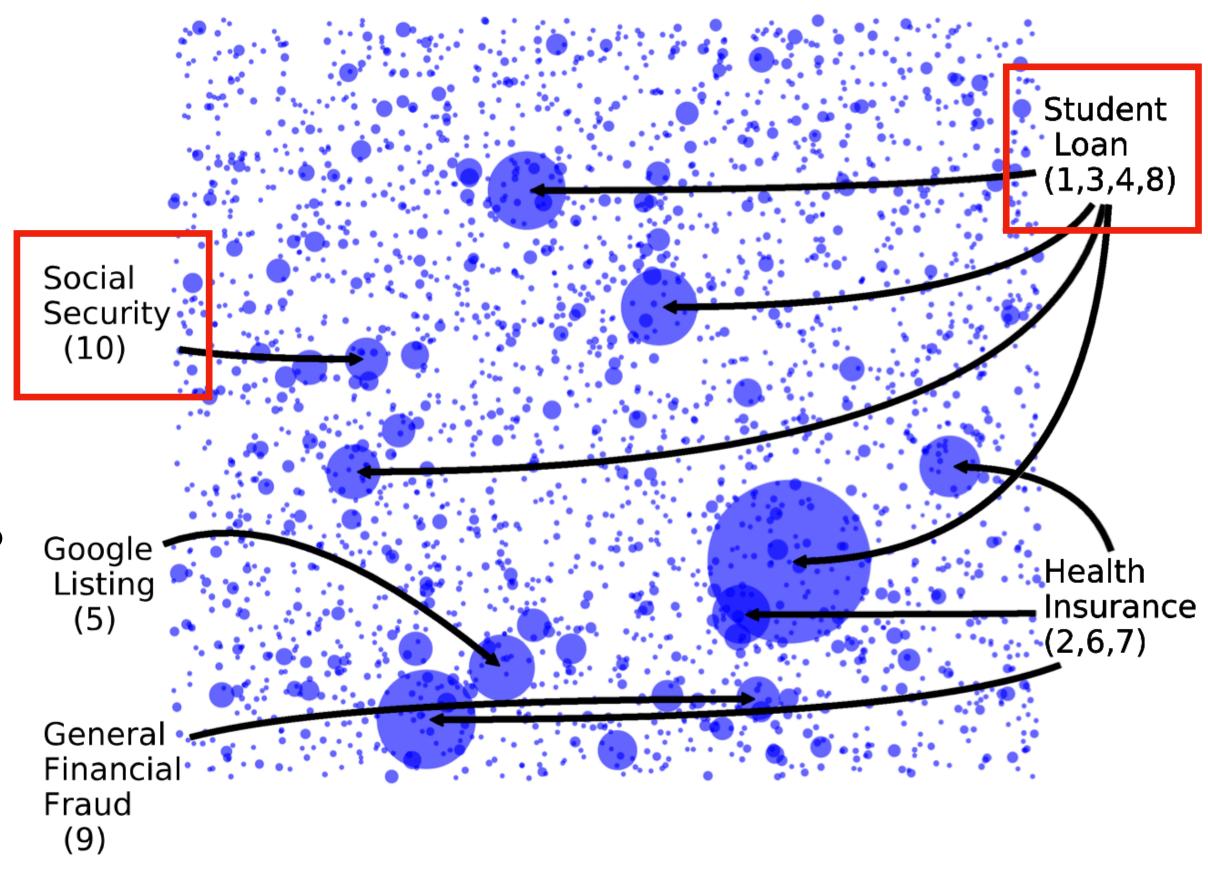


- 2,687 unique robocalling campaigns
- A fraudulent Social Security Campaign was the 10th largest campaign observed in our
- The largest campaign (a student loan robocall) had over 6,000 calls



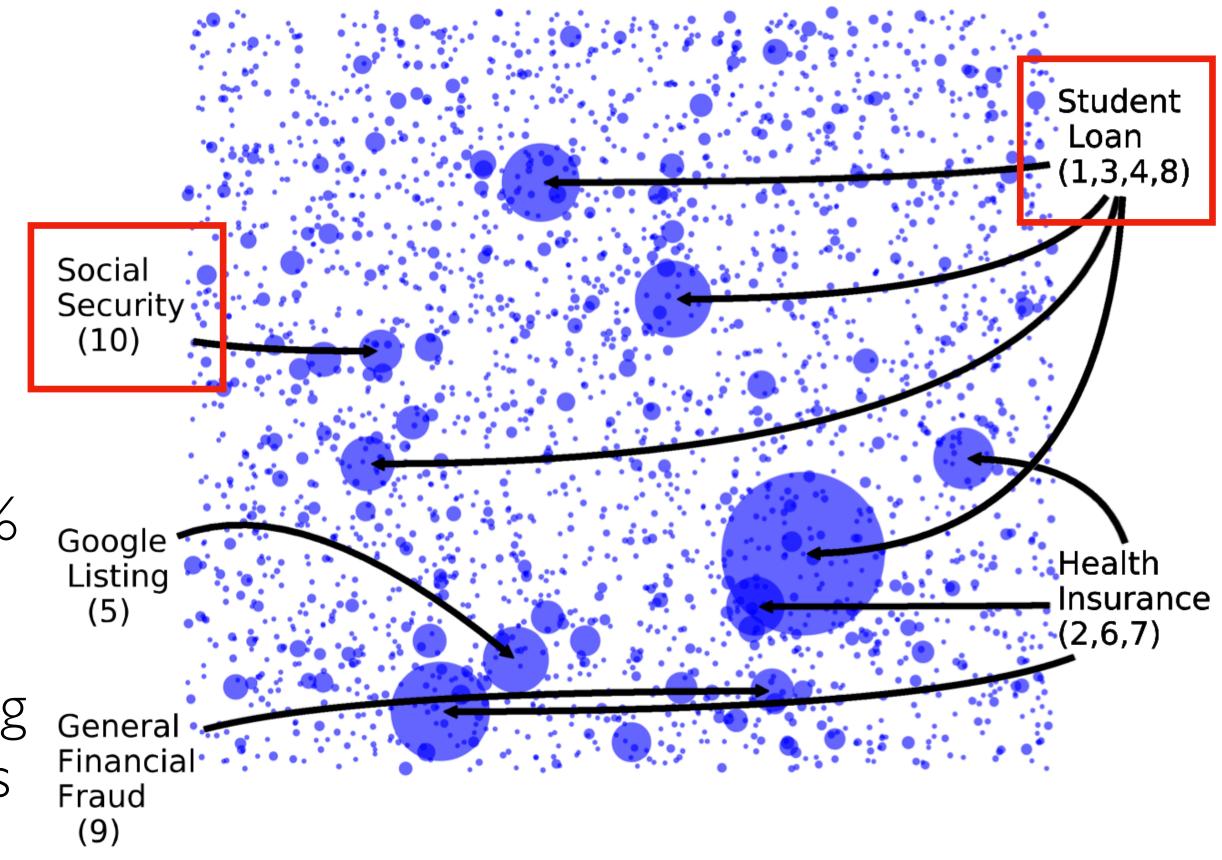


- 2,687 unique robocalling campaigns
- A fraudulent Social Security Campaign was the 10th largest campaign observed in our
- The largest campaign (a student loan robocall) had over 6,000 calls
- We received only a few calls from most campaigns over 11 months (<27 calls from 95% campaigns)





- 2,687 unique robocalling campaigns
- A fraudulent Social Security Campaign was the 10th largest campaign observed in our
- The largest campaign (a student loan robocall) had over 6,000 calls
- We received only a few calls from most campaigns over 11 months (<27 calls from 95% campaigns)
- Campaign identification through audio clustering allowed us to study campaign specific behaviors which were previously impossible to measure





#### NC STATE UNIVERSITY

What strategies do robocalls use to entice their victims?



#### **NC STATE** UNIVERSITY

## What strategies do robocalls use to entice their victims?



#### **NC STATE** UNIVERSITY

## What strategies do robocalls use to entice their victims?

**Caller ID Spoofing**: Robocalling campaigns regularly spoof caller ID or rotate between calling numbers

• Trivial strategy to evade allow-list/deny-list based robocall mitigation techniques



- Trivial strategy to evade allow-list/deny-list based robocall mitigation techniques
- How rapidly does a campaign change caller ID? Source Distribution =  $\frac{\text{(Number of unique caller ID used by the campaign)}}{\text{(Size of the campaign as seen in our honeypot)}}$



- Trivial strategy to evade allow-list/deny-list based robocall mitigation techniques
- How rapidly does a campaign change caller ID? Source Distribution =  $\frac{\text{(Number of unique caller ID used by the campaign)}}{\text{(Size of the campaign as seen in our honeypot)}}$
- Average Source Distribution of 84.17%



- Trivial strategy to evade allow-list/deny-list based robocall mitigation techniques
- How rapidly does a campaign change caller ID? Source Distribution =  $\frac{\text{(Number of unique caller ID used by the campaign)}}{\text{(Size of the campaign as seen in our honeypot)}}$
- Average Source Distribution of 84.17%
- The largest campaign had 99.93% Source Distribution: almost every call among the 6,055 calls had a different caller ID



**Caller ID Spoofing**: Robocalling campaigns regularly spoof caller ID or rotate between calling numbers

- Trivial strategy to evade allow-list/deny-list based robocall mitigation techniques
- How rapidly does a campaign change caller ID? Source Distribution =  $\frac{\text{(Number of unique caller ID used by the campaign)}}{\text{(Size of the campaign as seen in our honeypot)}}$
- Average Source Distribution of 84.17%
- The largest campaign had 99.93% Source Distribution: almost every call among the 6,055 calls had a different caller ID

**Neighbor Spoofing**: Some campaigns use "Neighbor Spoofing" — a technique to match the area code and exchange code (first 6 digits) of the caller ID with the called number



**Caller ID Spoofing**: Robocalling campaigns regularly spoof caller ID or rotate between calling numbers

- Trivial strategy to evade allow-list/deny-list based robocall mitigation techniques
- How rapidly does a campaign change caller ID? Source Distribution =  $\frac{\text{(Number of unique caller ID used by the campaign)}}{\text{(Size of the campaign as seen in our honeypot)}}$
- Average Source Distribution of 84.17%
- The largest campaign had 99.93% Source Distribution: almost every call among the 6,055 calls had a different caller ID

**Neighbor Spoofing**: Some campaigns use "Neighbor Spoofing" — a technique to match the area code and exchange code (first 6 digits) of the caller ID with the called number

• 77 campaigns used neighbor spoofing techniques



**Caller ID Spoofing**: Robocalling campaigns regularly spoof caller ID or rotate between calling numbers

- Trivial strategy to evade allow-list/deny-list based robocall mitigation techniques
- How rapidly does a campaign change caller ID? Source Distribution =  $\frac{\text{(Number of unique caller ID used by the campaign)}}{\text{(Size of the campaign as seen in our honeypot)}}$
- Average Source Distribution of 84.17%
- The largest campaign had 99.93% Source Distribution: almost every call among the 6,055 calls had a different caller ID

**Neighbor Spoofing**: Some campaigns use "Neighbor Spoofing" — a technique to match the area code and exchange code (first 6 digits) of the caller ID with the called number

- 77 campaigns used neighbor spoofing techniques
- 14 campaigns used neighbor spoofing for every call they generated



#### **NC STATE** UNIVERSITY

What strategies do robocalls use to entice their victims?



#### **NC STATE UNIVERSITY**

### What strategies do robocalls use to entice their victims?

**Targeted Operation**: Robocalling campaigns are highly targeted and few top campaigns targeted specific numbers



### What strategies do robocalls use to entice their victims?

**Targeted Operation**: Robocalling campaigns are highly targeted and few top campaigns targeted specific numbers

• How targeted are the robocalls from a campaign? Spread =  $\frac{\text{(Number of Unique Destination Numbers)}}{\text{(Size of the campaign as seen in our honeypot)}}$ 



### What strategies do robocalls use to entice their victims?

**Targeted Operation**: Robocalling campaigns are highly targeted and few top campaigns targeted specific numbers

- How targeted are the robocalls from a campaign? Spread =  $\frac{\text{(Number of Unique Destination Numbers)}}{\text{(Size of the campaign as seen in our honeypot)}}$
- We observed an average Spread of 78.30%



### What strategies do robocalls use to entice their victims?

**Targeted Operation**: Robocalling campaigns are highly targeted and few top campaigns targeted specific numbers

- How targeted are the robocalls from a campaign? Spread =  $\frac{\text{(Number of Unique Destination Numbers)}}{\text{(Size of the campaign as seen in our honeypot)}}$
- We observed an average Spread of 78.30%
- The largest campaign had 19.60% Spread: about 5 calls were made from this campaign to the same number on average

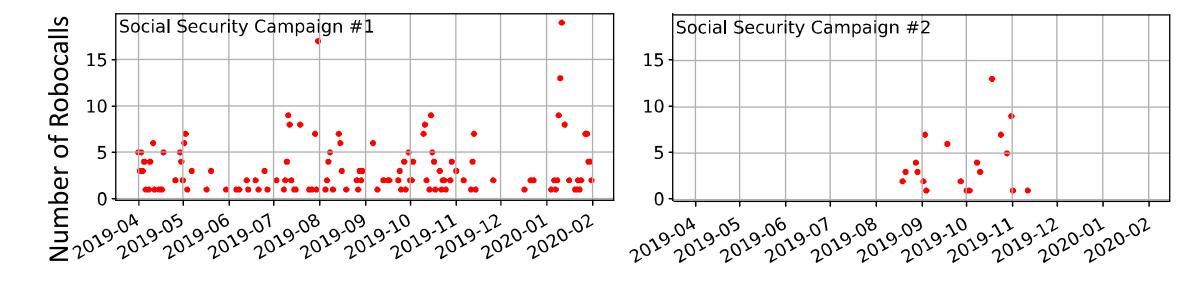


#### **NC STATE UNIVERSITY**

Adversarial Campaigns Target the Vulnerable Population

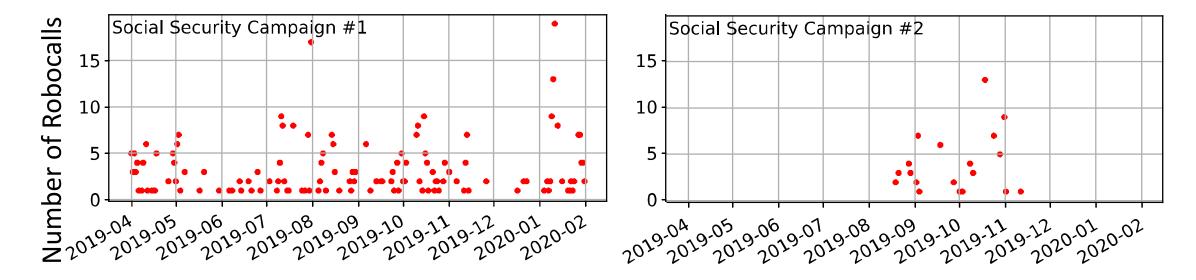


Social Security Fraud Campaigns





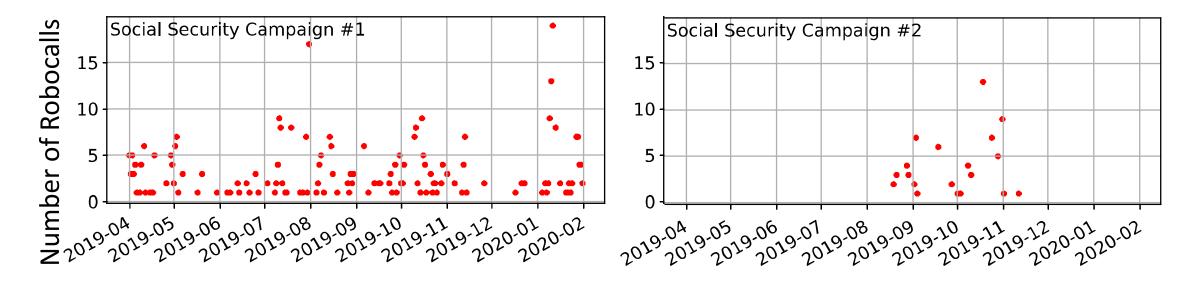
Social Security Fraud Campaigns



 We uncovered two distinct long-running Social Security Fraud campaigns where the first campaign operated almost throughout the year



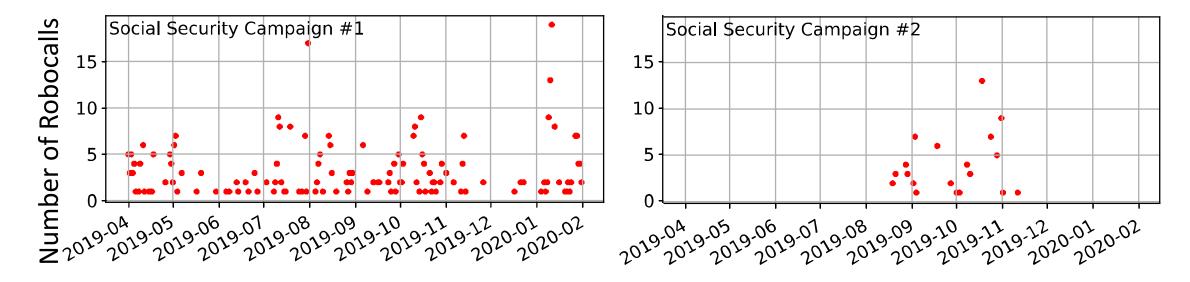
#### Social Security Fraud Campaigns



- We uncovered two distinct long-running Social Security Fraud campaigns where the first campaign operated almost throughout the year
- Both these Social Security campaigns impersonated a government agency, frequently used toll-free numbers as caller ID and targeted specific phone numbers



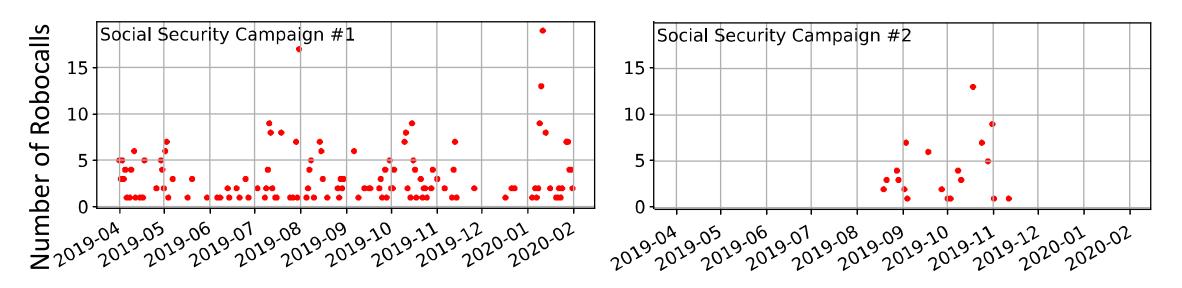
#### Social Security Fraud Campaigns



- We uncovered two distinct long-running Social Security Fraud campaigns where the first campaign operated almost throughout the year
- Both these Social Security campaigns impersonated a government agency, frequently used toll-free numbers as caller ID and targeted specific phone numbers
- Since Social Security is relevant year-round, such scams allow adversaries to target their victims not just during tax seasons

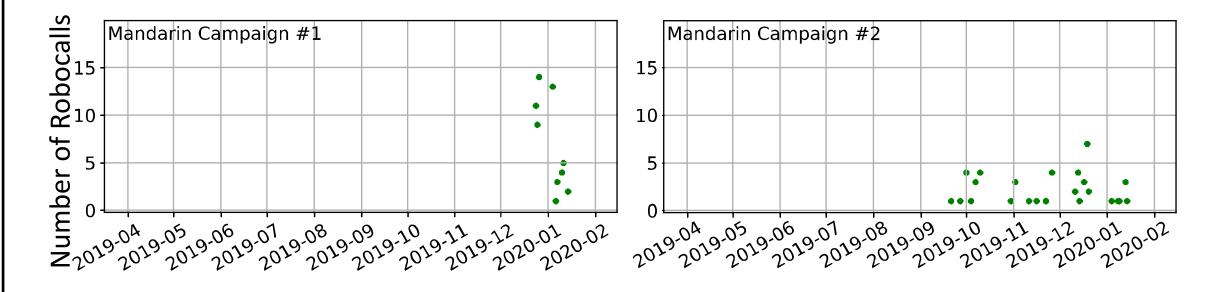


#### Social Security Fraud Campaigns



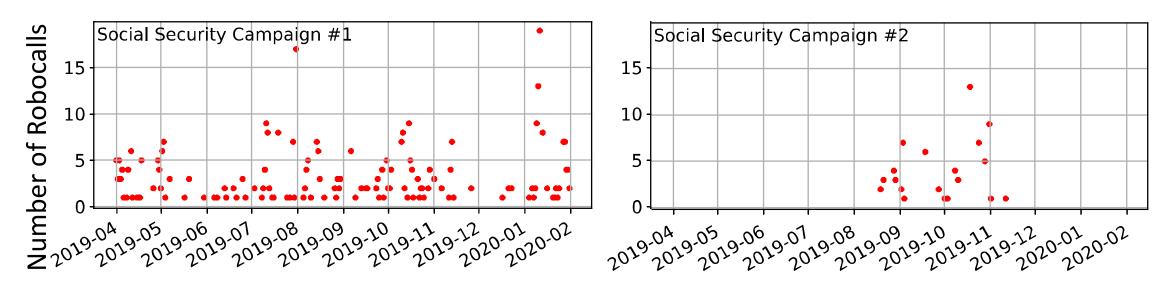
- We uncovered two distinct long-running Social Security Fraud campaigns where the first campaign operated almost throughout the year
- Both these Social Security campaigns impersonated a government agency, frequently used toll-free numbers as caller ID and targeted specific phone numbers
- Since Social Security is relevant year-round, such scams allow adversaries to target their victims not just during tax seasons

#### Fraudulent Mandarin Campaign



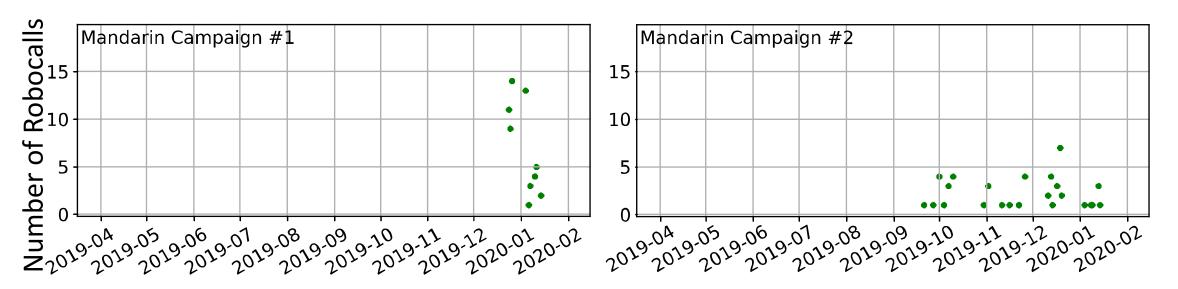


### Social Security Fraud Campaigns



- We uncovered two distinct long-running Social Security Fraud campaigns where the first campaign operated almost throughout the year
- Both these Social Security campaigns impersonated a government agency, frequently used toll-free numbers as caller ID and targeted specific phone numbers
- Since Social Security is relevant year-round, such scams allow adversaries to target their victims not just during tax seasons

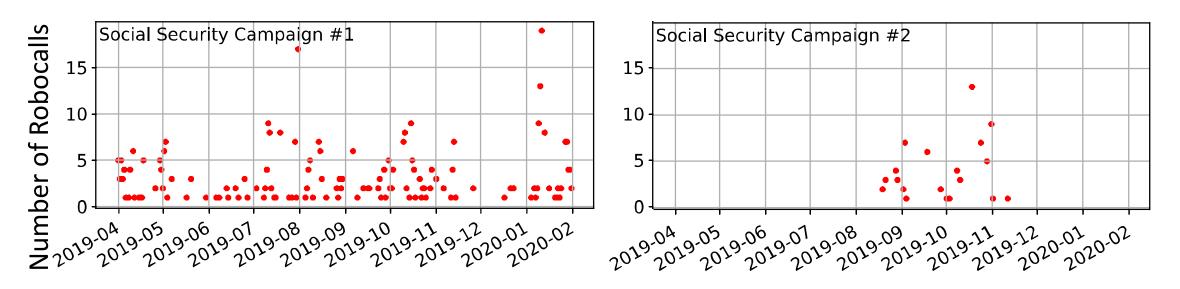
#### Fraudulent Mandarin Campaign



 Our language-independent campaign identification pipeline uncovered two Chinese robocalling campaigns impersonating the Chinese Consulate and targeting the Mandarin speaking population in the United States

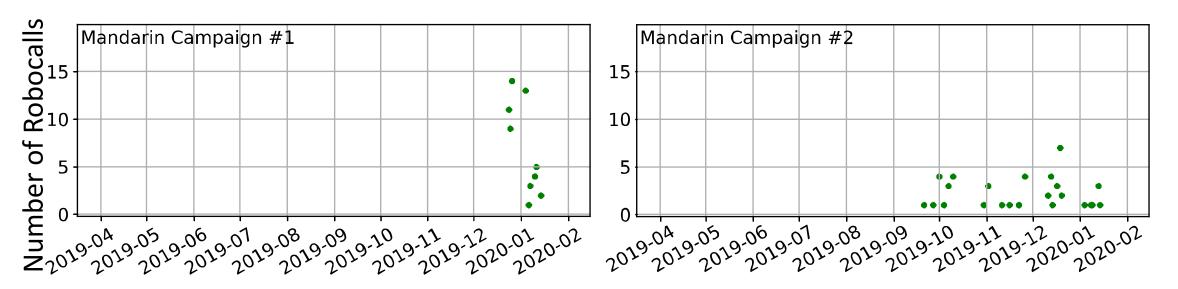


### Social Security Fraud Campaigns



- We uncovered two distinct long-running Social Security Fraud campaigns where the first campaign operated almost throughout the year
- Both these Social Security campaigns impersonated a government agency, frequently used toll-free numbers as caller ID and targeted specific phone numbers
- Since Social Security is relevant year-round, such scams allow adversaries to target their victims not just during tax seasons

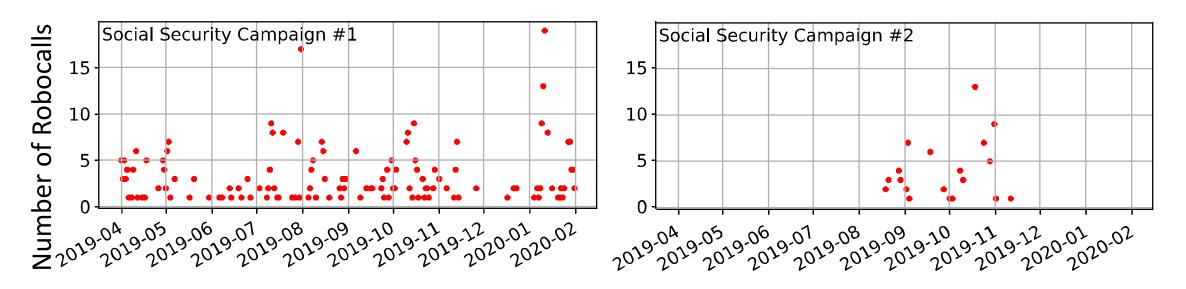
#### Fraudulent Mandarin Campaign



- Our language-independent campaign identification pipeline uncovered two Chinese robocalling campaigns impersonating the Chinese Consulate and targeting the Mandarin speaking population in the United States
- These campaigns threatened their victims and was designed to defraud recent immigrants and non-native English speakers

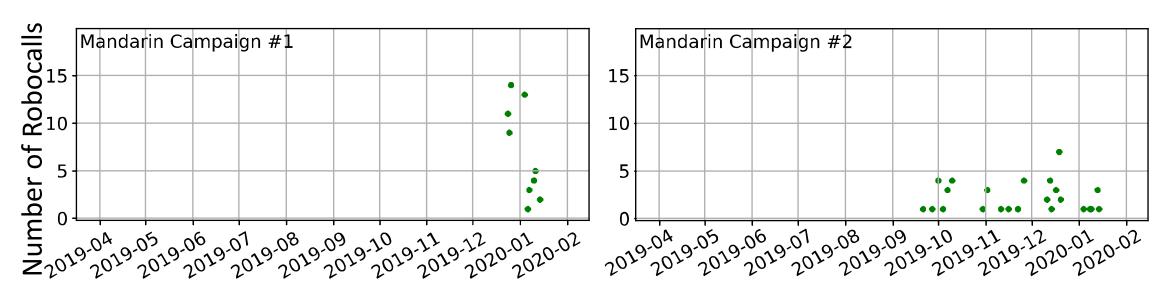


### Social Security Fraud Campaigns



- We uncovered two distinct long-running Social Security Fraud campaigns where the first campaign operated almost throughout the year
- Both these Social Security campaigns impersonated a government agency, frequently used toll-free numbers as caller ID and targeted specific phone numbers
- Since Social Security is relevant year-round, such scams allow adversaries to target their victims not just during tax seasons

#### Fraudulent Mandarin Campaign



- Our language-independent campaign identification pipeline uncovered two Chinese robocalling campaigns impersonating the Chinese Consulate and targeting the Mandarin speaking population in the United States
- These campaigns threatened their victims and was designed to defraud recent immigrants and non-native English speakers
- Adversaries impersonate government agencies to gain the victim's attention and boost their credibility



### But wait, there's more!

- · Heuristics to detect and measure voicemail spam and Wangiri scams
- How unsolicited phone calls abuse CNAM (Caller ID Name)
- How callers use "\*67" to mask their true caller ID and remain anonymous
- How to systematically characterize "Storms" high call volume events
- Discussion on language-agnostic and noise-resistant audio clustering pipeline
- An in-depth discussion of ethical and legal consideration of our study
- · How can our techniques help service providers combat fraudulent robocalls



### Key Takeaways

- Robocalling is a major security issue in the United State. We desperately need effective solutions to combat robocalls and protect phone users
- Even though we suggest users to be cautious when answering calls from unknown numbers, answering a robocall does not necessarily increase the number of robocalls
- Illegal robocalls are highly targeted. They regularly spoof caller ID or rotate between numbers and perform neighbor spoofing to entice victims
- Fraudulent robocalls can operate for long durations without being detected. They are designed to target vulnerable population

### Acknowledgements

• We would like to thank Bandwidth Inc. and Nomorobo for their support

• Link to the paper: <a href="https://robocallobservatory.org">https://robocallobservatory.org</a>

Contact: Sathvik Prasad (snprasad@ncsu.edu)

