We built a telephony honeypot with 66,000 phone numbers and uncovered over *2,600 robocalling campaigns* in a span of 11 months

Who's Calling?

*Characterizing Robocalls through Audio and Metadata Analysis*

👤 Sathvik Prasad, Elijah Bouma-Sims, Athishay Kiran Mylappan, Bradley Reaves

INTRODUCTION

- Automated and unsolicited phone calls, also called **Robocalls**, are a menace for telephone users
- We collect audio and call metadata from phone numbers we own and use that data to characterize unsolicited calls and robocalling campaigns
- We estimate the frequency of robocalls, evaluate the effects of answering unsolicited phone calls and study the operational characteristics of robocalling campaigns in the real-world

METHODOLOGY

- Using 66,606 inbound telephone lines, we deployed the largest academic telephony honeypot to collect call meta-data and call audio information
- We developed robust audio-fingerprinting based clustering techniques to identify robocalling campaigns targeting users in the phone network
- Our clustering approach is immune to caller ID spoofing or line rotation, is language agnostic, and is independent of transcription accuracy
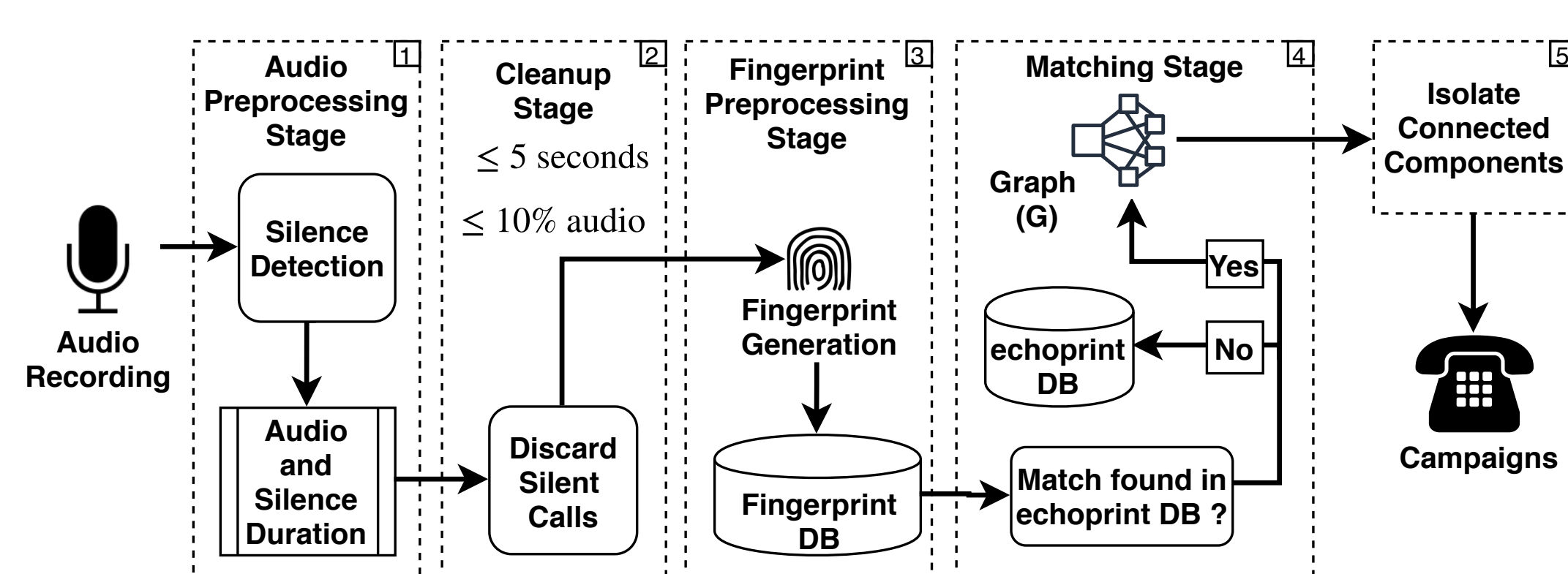

Five Stage Audio Clustering Pipeline
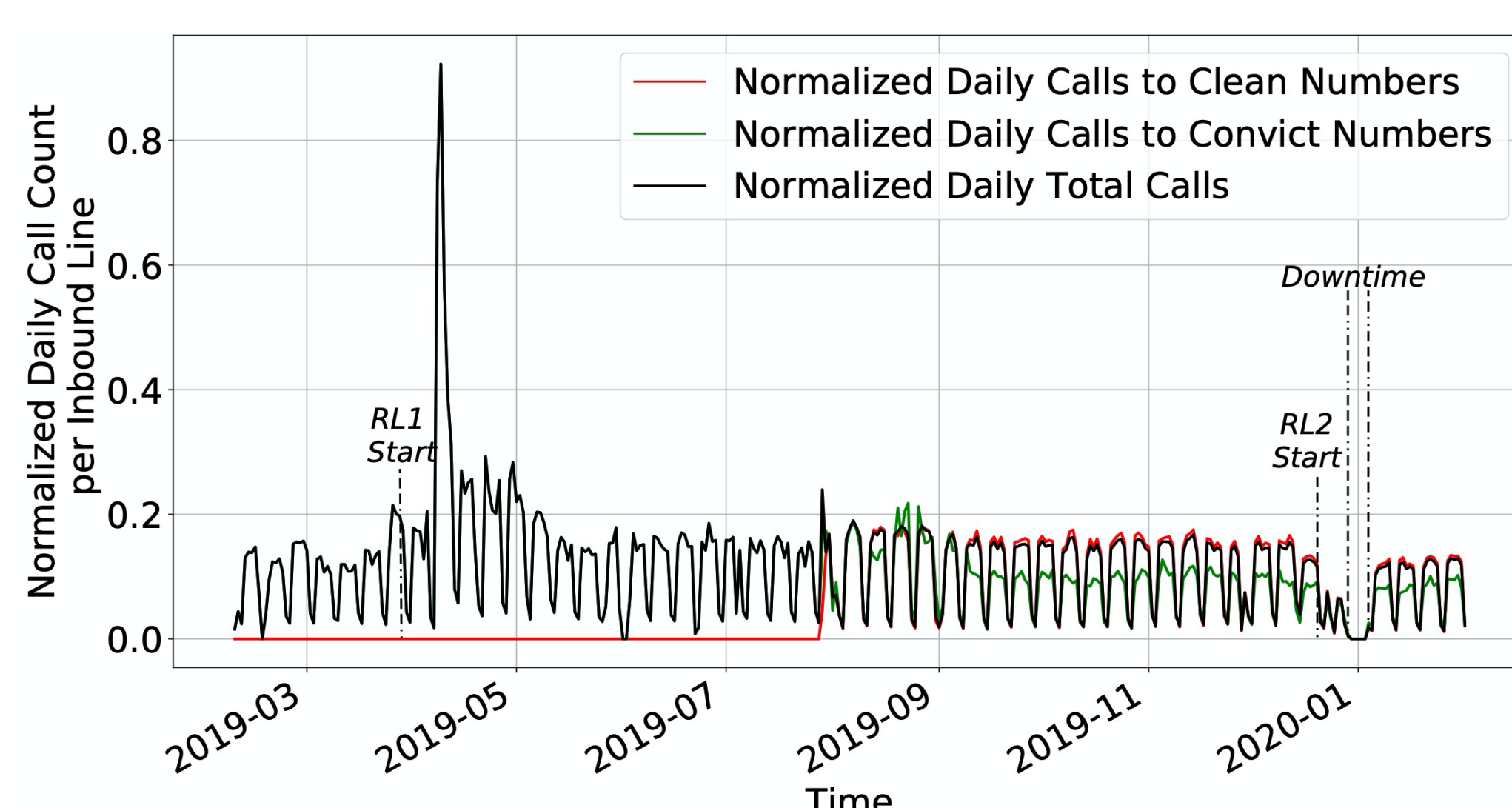
Normalized Call Volume over 11 Months

RESULTS

- We identified 2,687 robocalling campaigns
- We developed heuristics for *Wangiri* (One ring) scams and voicemail injection, and estimate that 2.91% (43,170 calls) of all unsolicited calls were attempted voicemail injections
- We observed that unsolicited calling trends were stationary throughout the duration of our study
- We uncovered two long-running large-scale SSN Fraud campaigns which targeted specific users
- We identified two campaigns which operate in Mandarin and impersonate the Chinese Consulate
- Robocalling campaigns had an average source distribution of 84.17%, indicating rampant caller ID spoofing and/or line rotation

TAKEAWAYS

- **We find no evidence that answering unsolicited phone calls increases the number of unsolicited phone calls received by a user in North America**
- Our work presents new tools that telephone service providers can adopt to better target operational robocalling campaigns
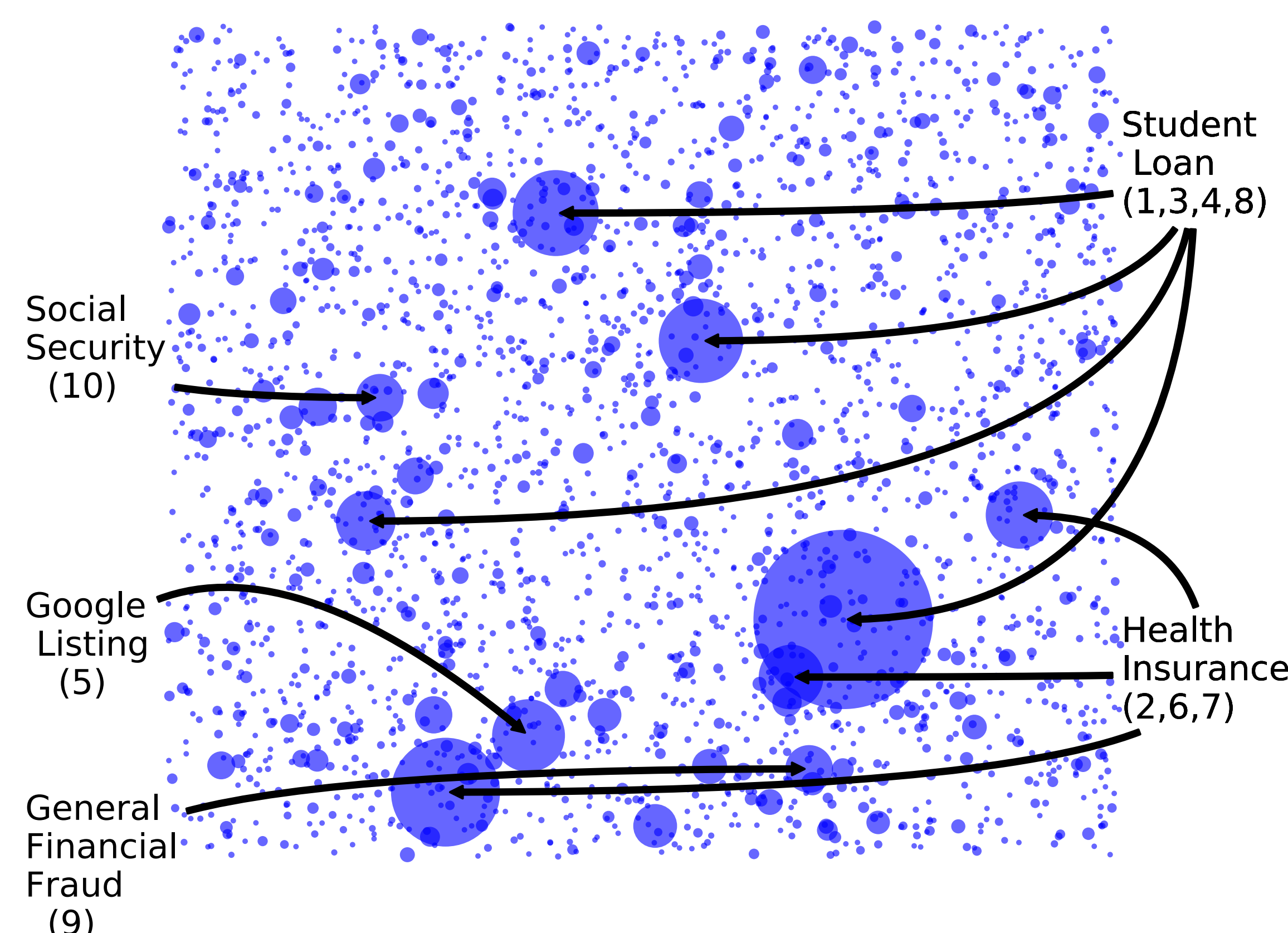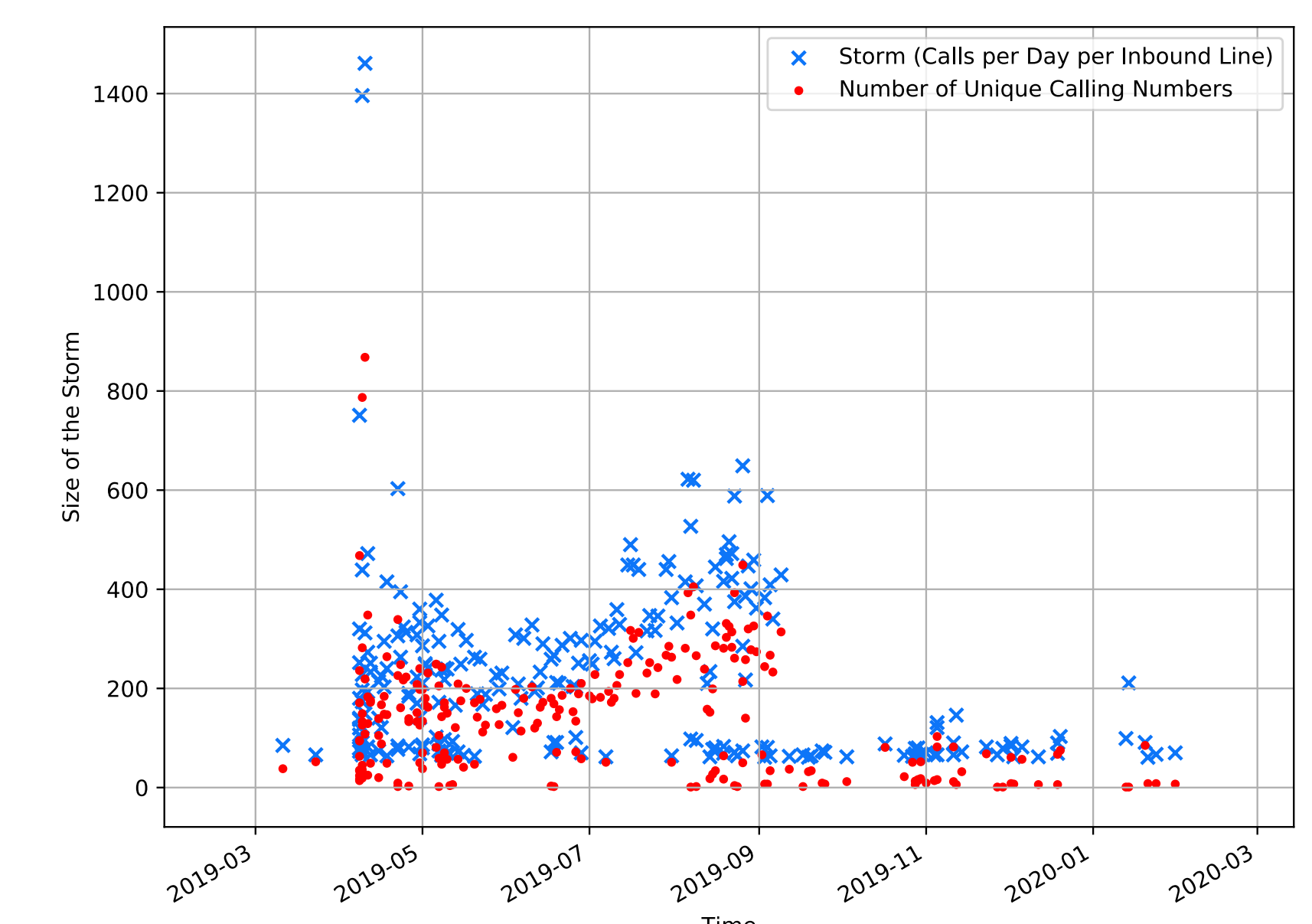

Top 10 Unsolicited Phone Campaigns

Data Collection Setup
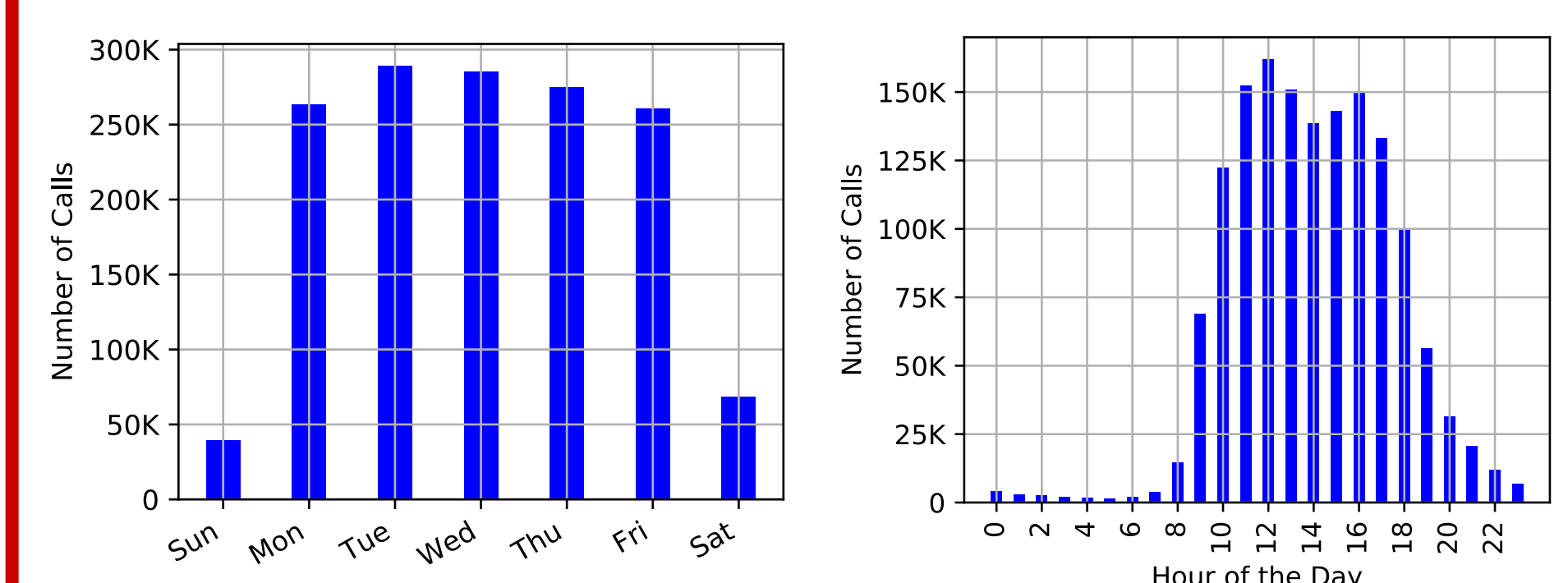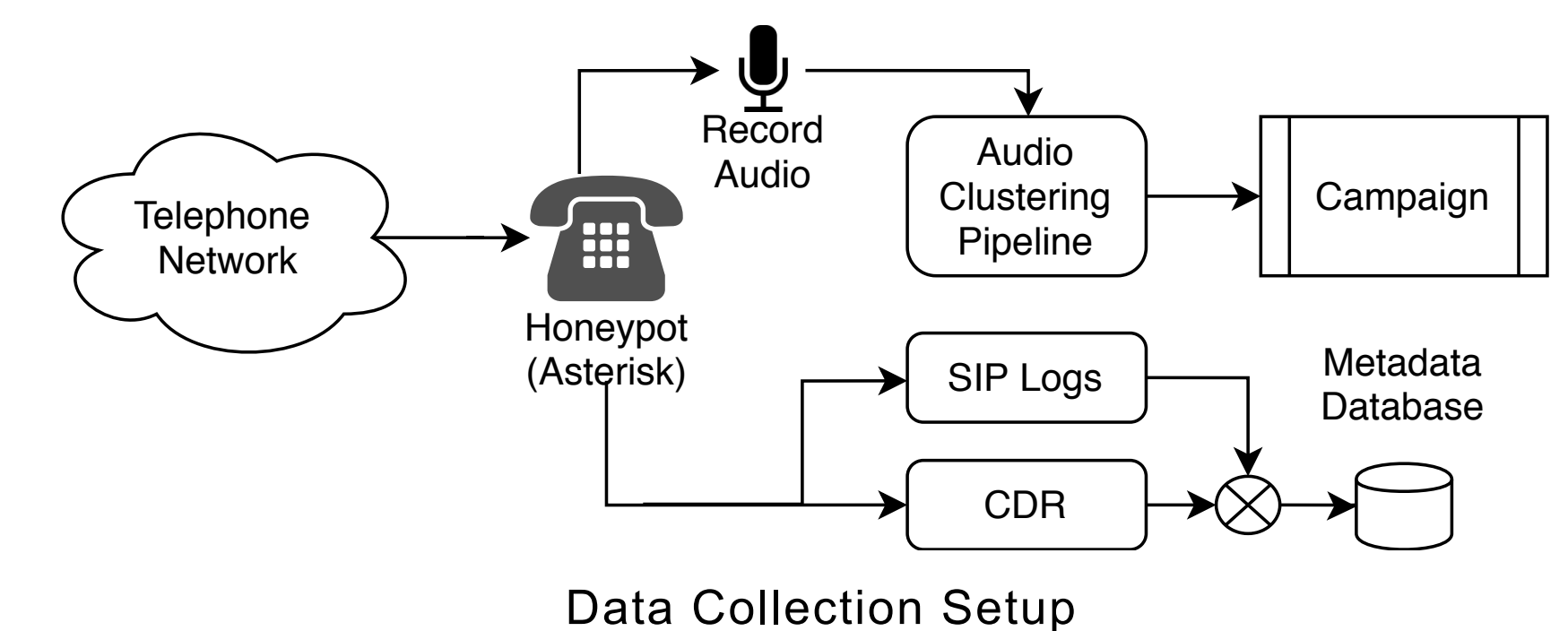
Robocalls are more frequent during working hours (EST 9 AM-5 PM) and weekdays
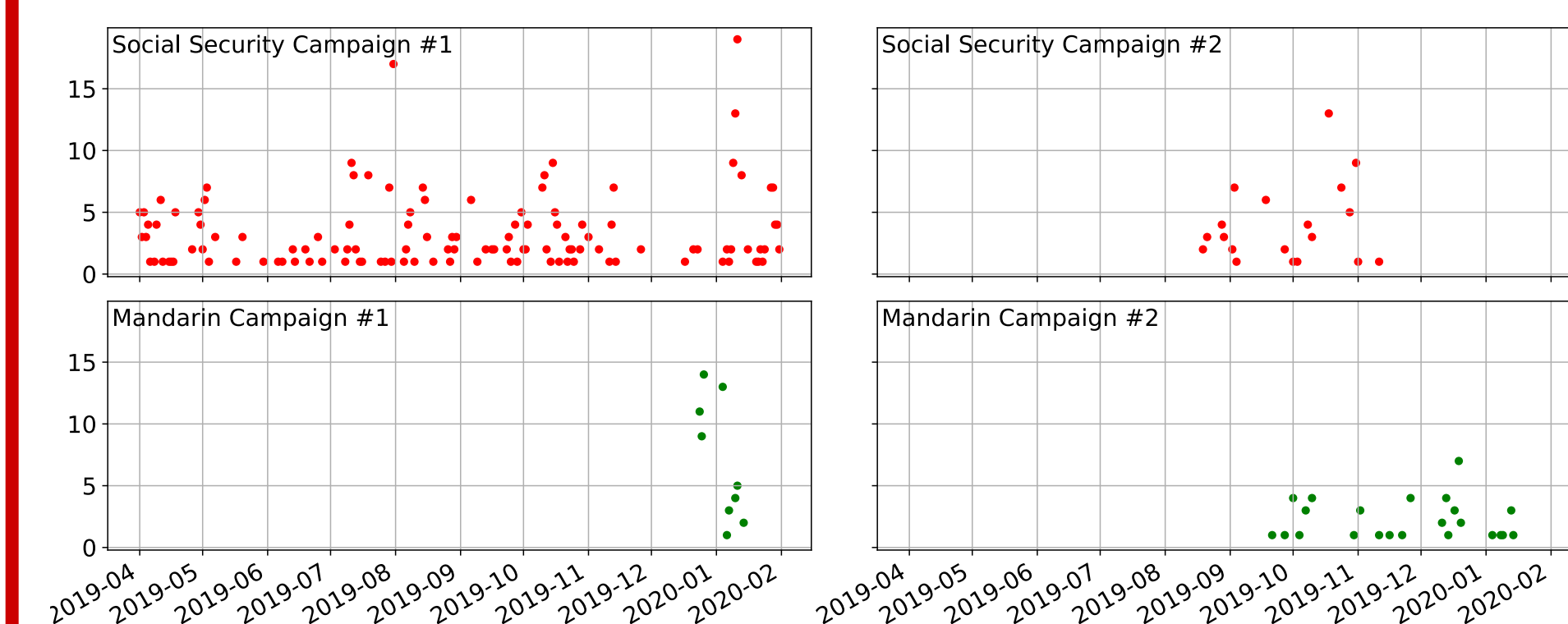
Fraudulent robocalling campaigns are active throughout the year

Bursts of high call-volume events of unsolicited phone calls, also called "Storms", are more frequent than previously reported

Number of Calls Received	1,400,000+
Number of Calls Recorded	146,000+
Duration of Study	11 Months
Number of Inbound Telephone Lines	66,606

NC STATE UNIVERSITY

NSF

WSPR