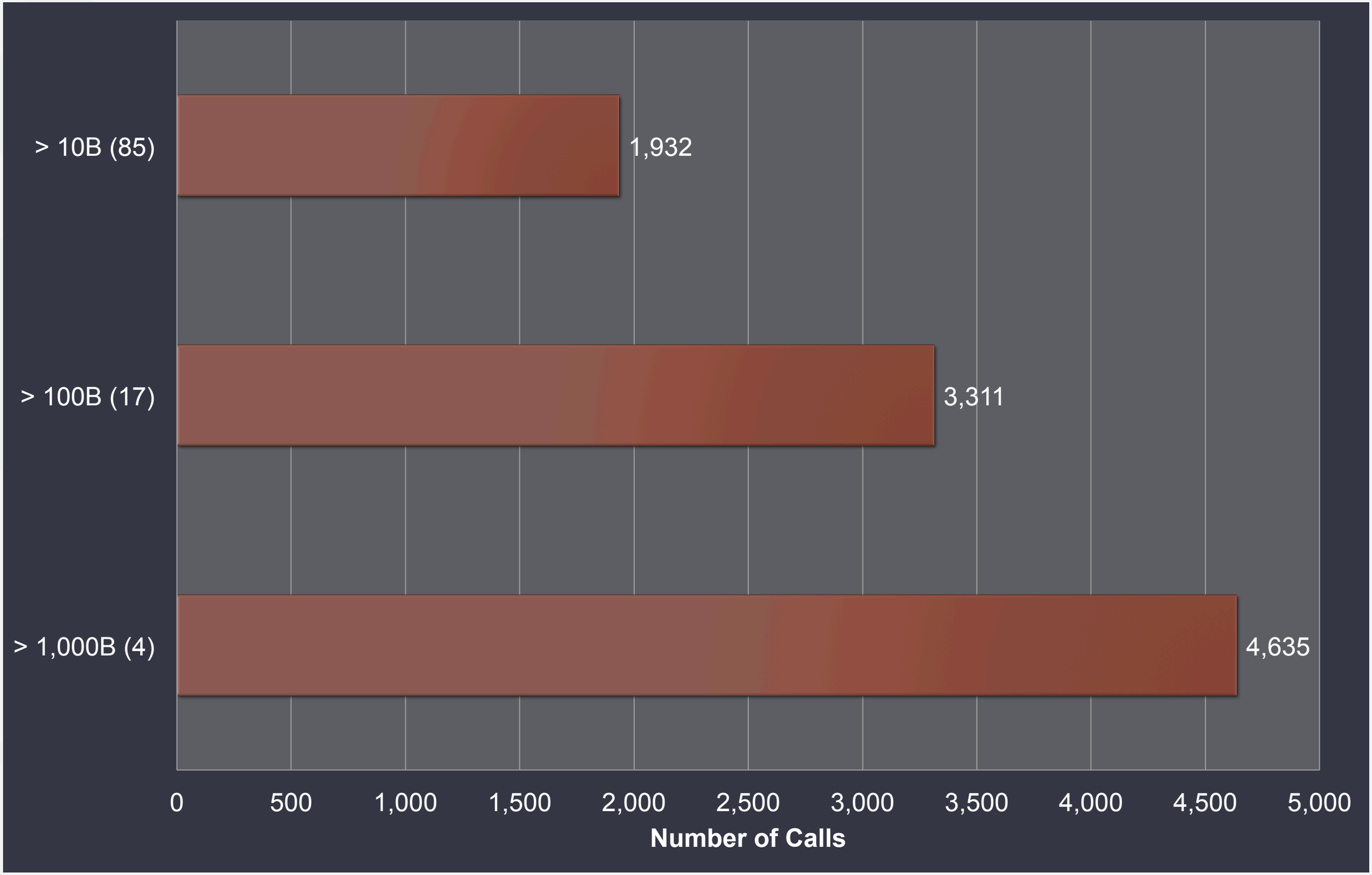
LIFECYCLE OF A PHONE FRAUDSTER: FROM ACCOUNT RECONNAISSANCE TO TAKEOVER

Vijay Balasubramaniyan, Raj Bandyopadhyay, Telvis Calhoun

7 August, 2014

ONE IN EVERY 2,901 CALLS IS FRAUD



PHONE FRAUD BY THE NUMBERS

12% – 94%

Awareness of fraud on the phone channel

5

Avg. # of calls needed to compromise an account

$700,000

Largest financial transaction loss stopped

$0.57

Dollars lost for every call into your call center

PHONEPRINTING™

pindrop security

Call Audio

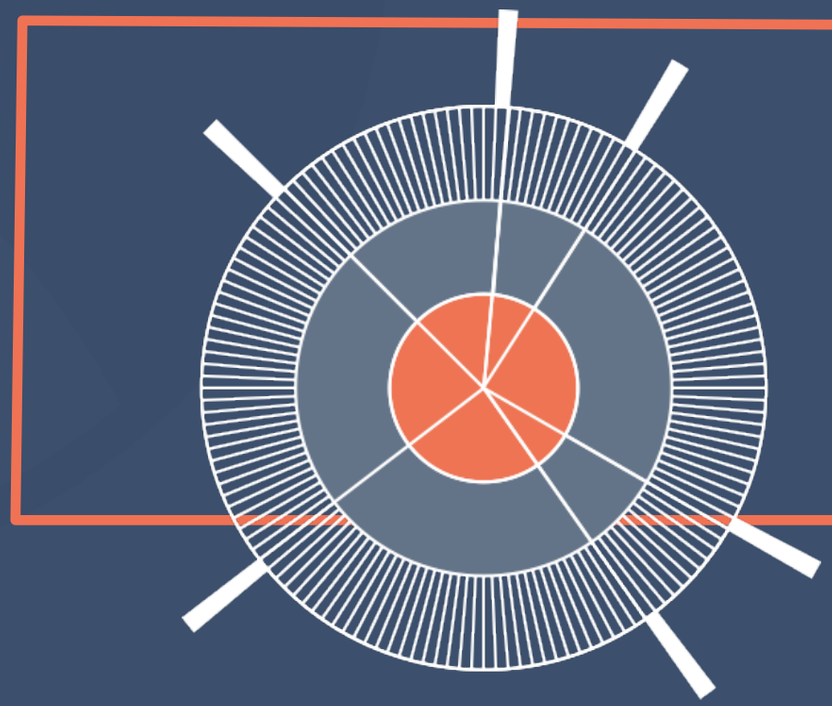
Requires 15 seconds of call audio

147 audio features in each fingerprint

LOSS
- Packet loss
- Robotization
- Dropped frames

SPECTRUM
- Quantization
- Frequency filters
- Codec artifacts

NOISE
- Clarity
- Correlation
- Signal-to-noise ratio



Phoneprint™

Phone Type Geo-Location Risk Score

FRAUD DETECTION SYSTEM

Phone Traffic

ANI

Audio

Pindrop Consortium

Reputation

Geography

Device Type

Phoneprint

Voiceprint

Analysis

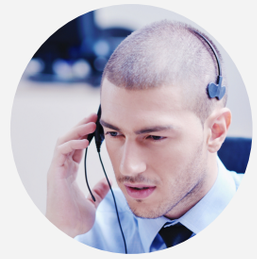
81

Risky

Scoring

API

Notification



Fraud Analyst

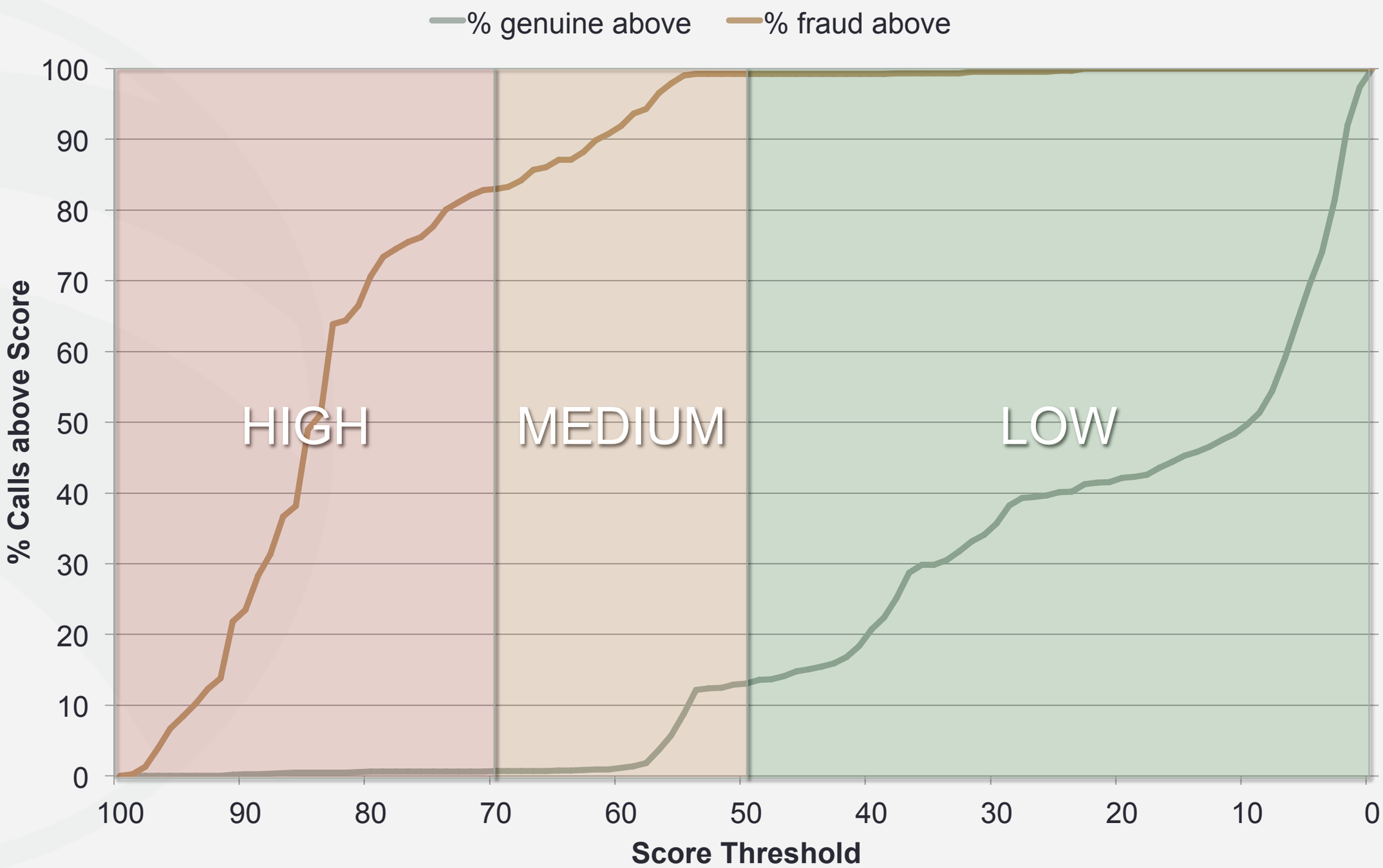
ANALYSIS AT SCALE

100 Million calls

18 Million Originating ANIs

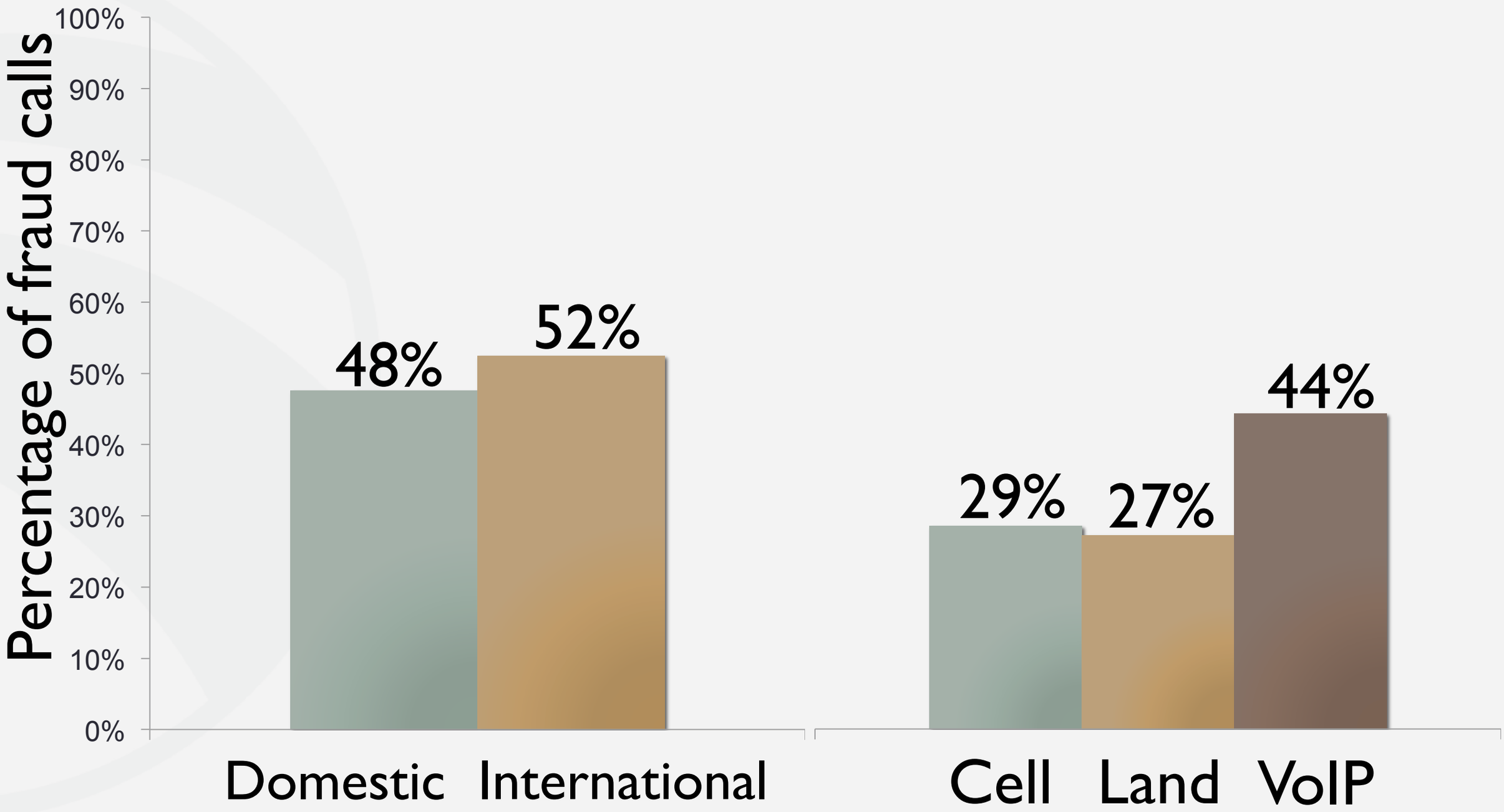
12 Million Accounts

GENUINE FROM FRAUD

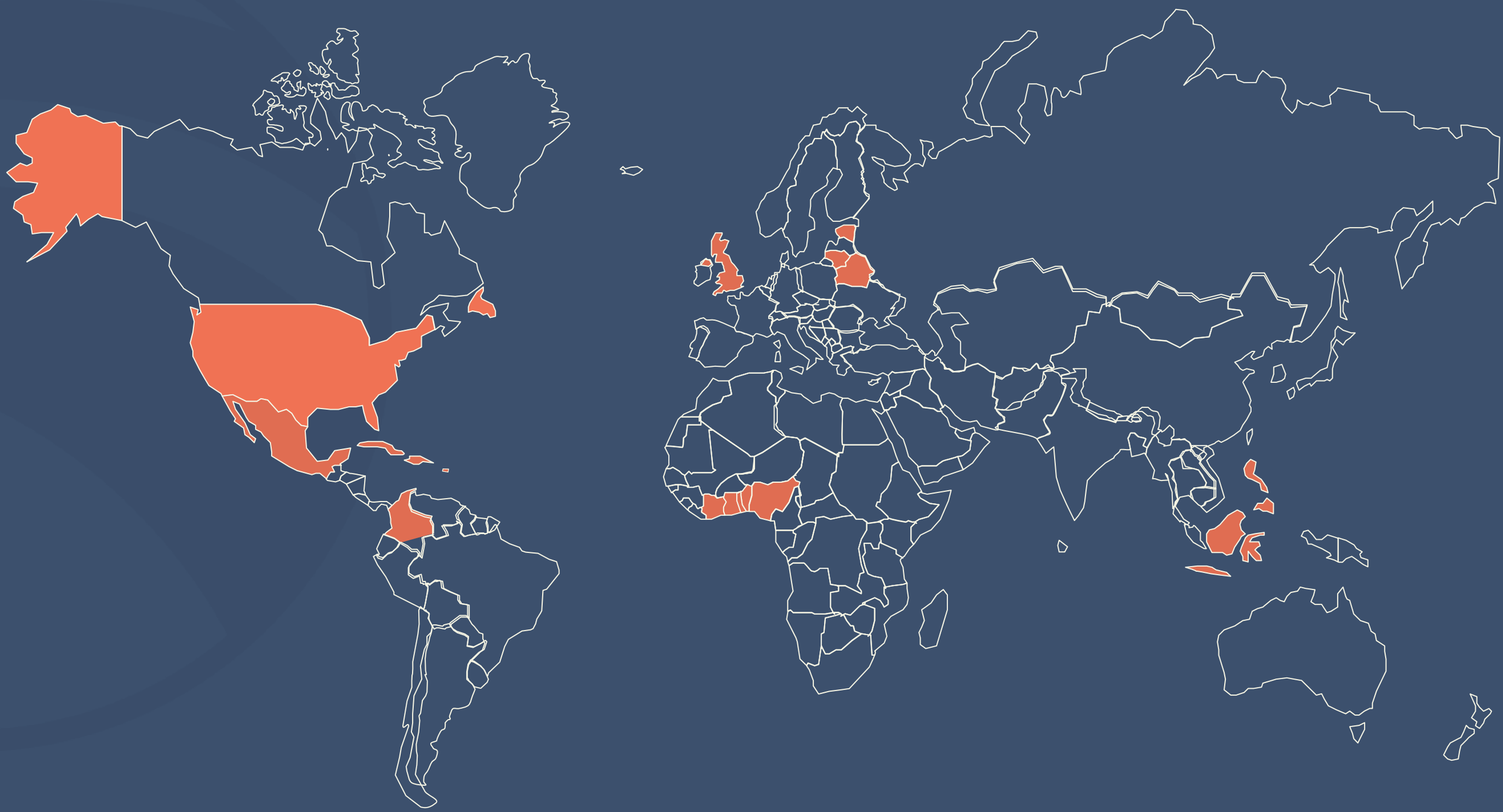


At score threshold of 60, stop 91% of fraud while passing 99% of genuine

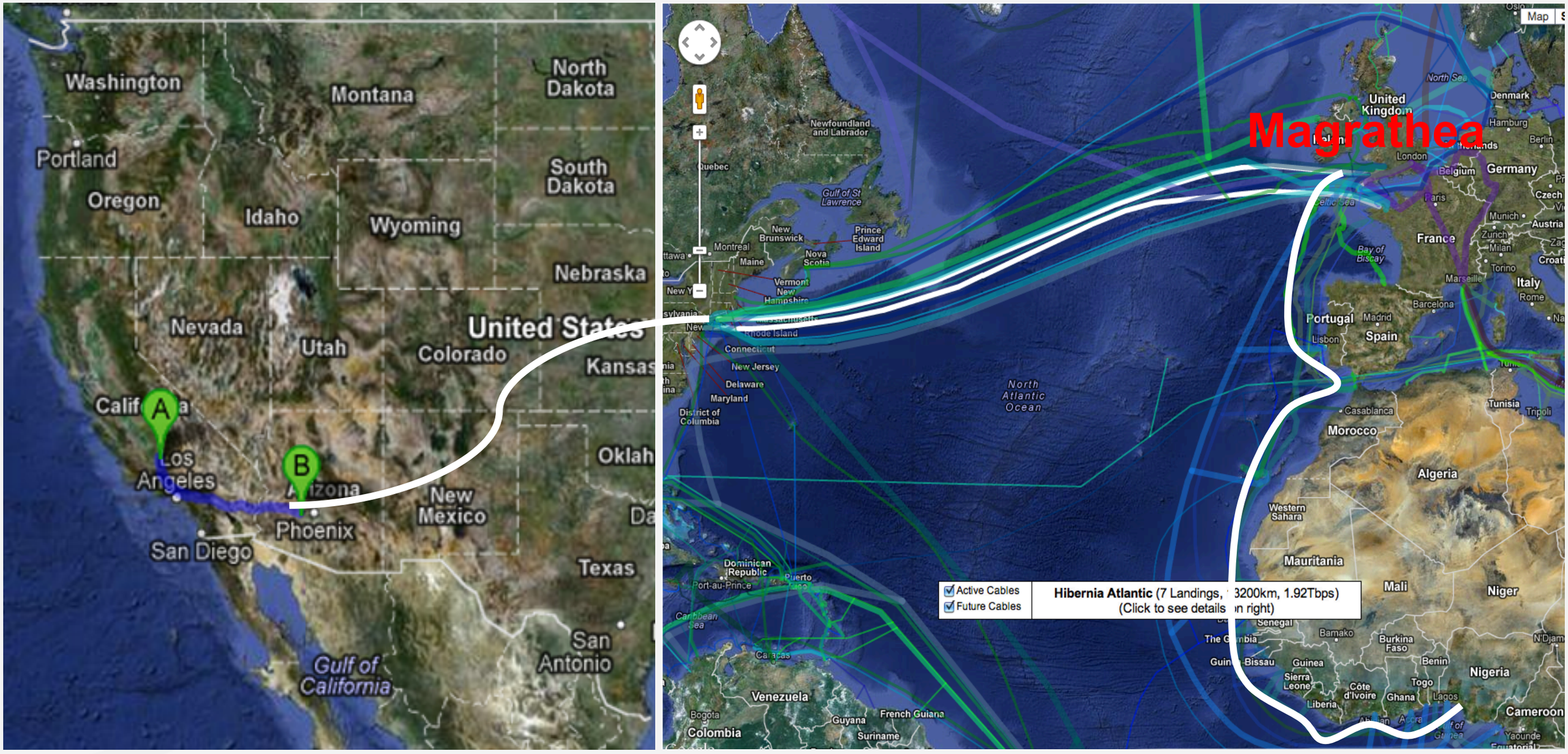
FRAUD CALL DISTRIBUTION



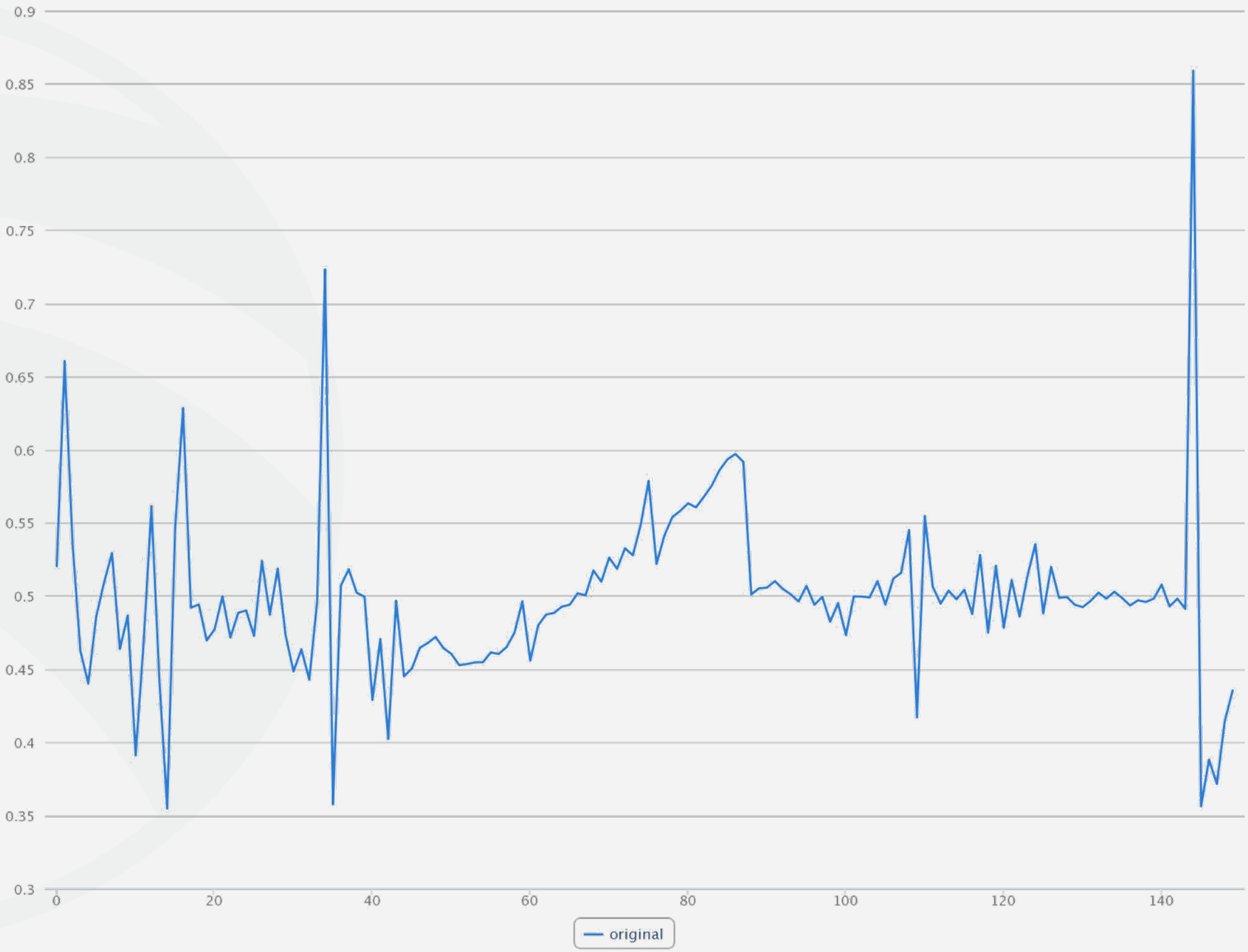
HIGH RISK COUNTRIES



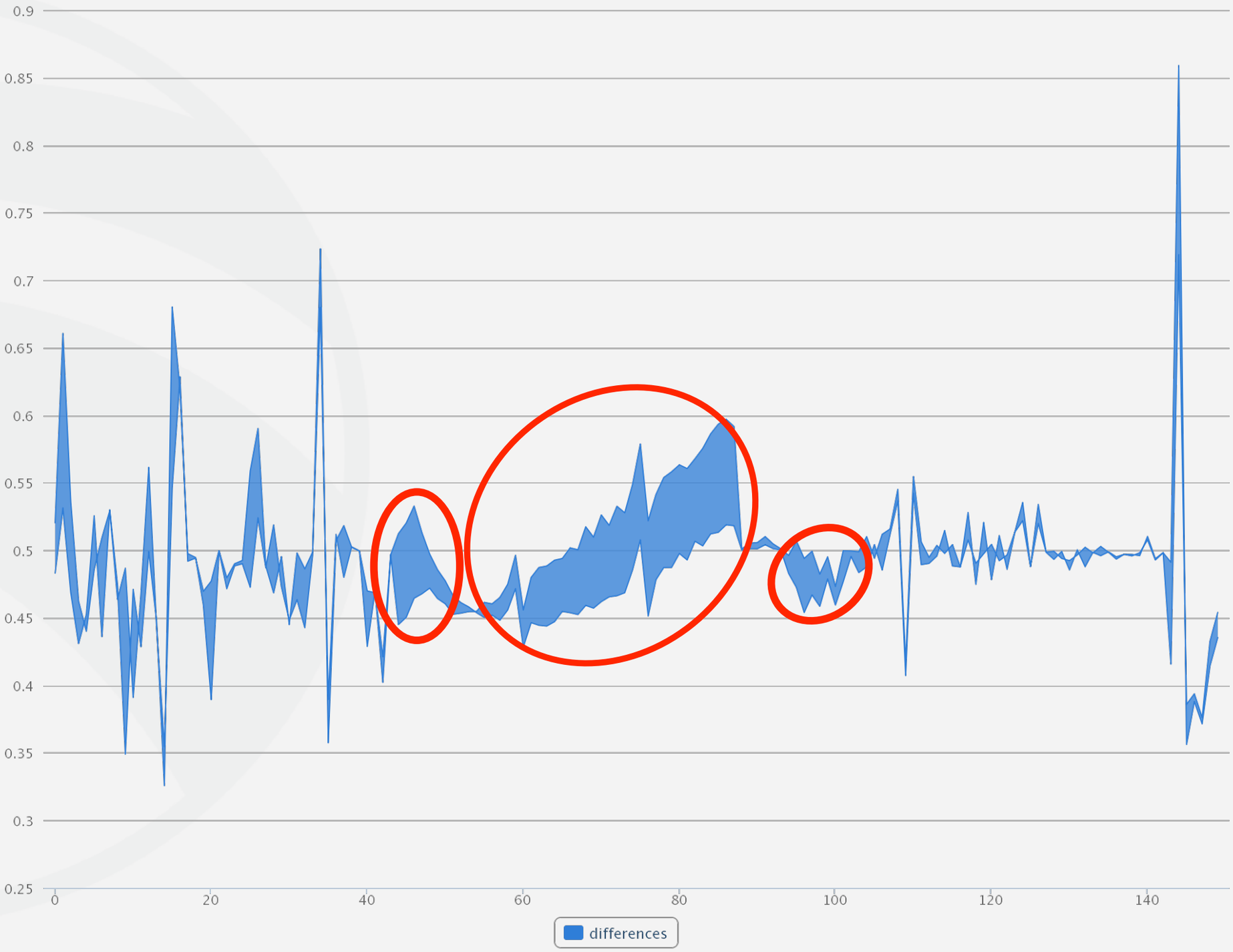
WEST AFRICA “ONE”



FRAUDSTER PROFILE



VOICE DISTORTION



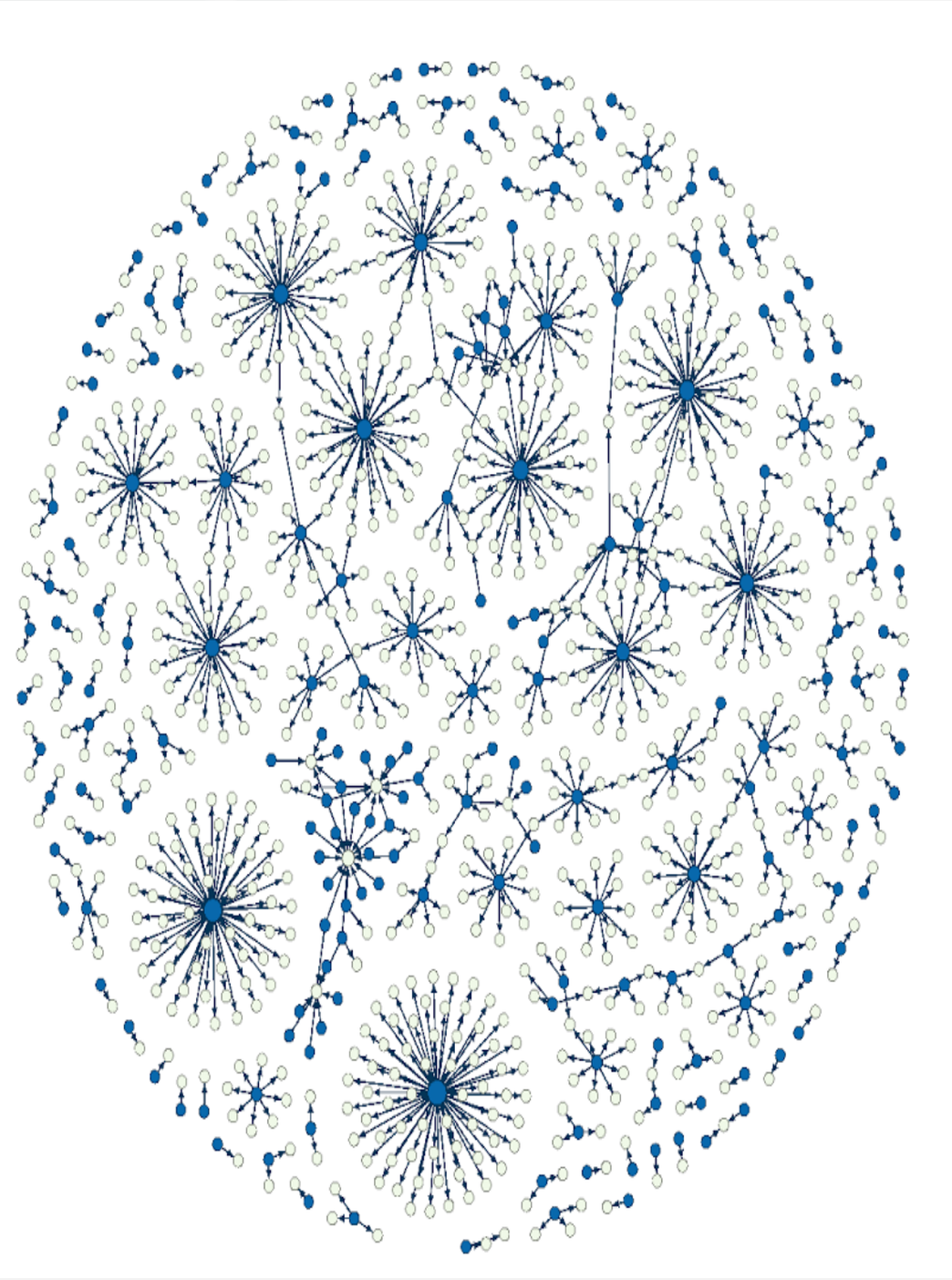
pindrop security

CDR ANALYSIS

CHARACTERISTICS OF CDR DATA

- Basic metadata about calls
  - Source and destination ANIs
  - Start and end timestamps of call
- Advanced application-specific metadata
  - IVR call flow information
  - Account numbers and other user information
  - Tower & base station information (for cell networks)

GRAPH REPRESENTATION OF CDRS



- Directed graph: ANIs as nodes, edges from source to target
- Edges annotated with timing and other information
- Can include other elements as nodes e.g. account numbers

FEATURES FOR CDR ANALYSIS

- Reputation features
  - Carrier, device type, prepaid status of source ANI
  - Complaints against source ANI
- Velocity (graph) features
  - Number of ANIs & accounts targeted by source ANI
  - Frequency and duration of calls from ANI
  - Application-specific features e.g. ANI scanning, number of authentication attempts

FEATURES FOR CDR ANALYSIS

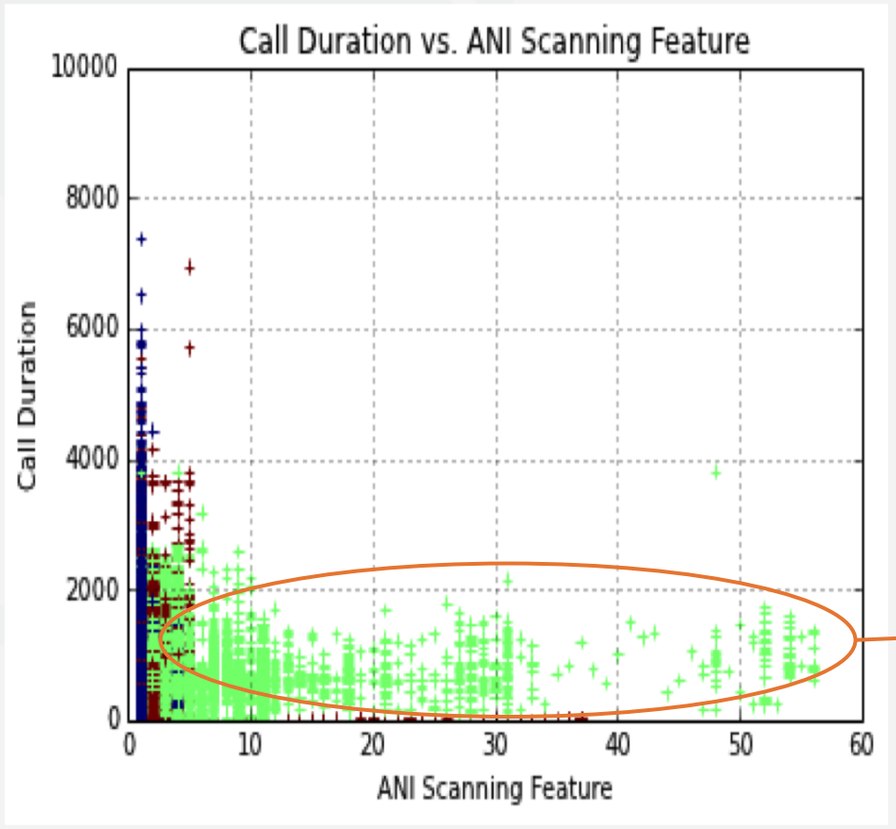
- Behavior features from IVR
  - Call flow sequence in the IVR e.g. [Account Entry, PIN Entry, Balance Check]
  - Use call flow sequences from single or multiple calls
  - Break up sequences into short chunks
  - Represent chunks as fixed-length, numeric vectors
  - Select top K features using feature selection techniques e.g. chi-squared

CASE STUDY 1: CALLING CARD TELCO

- Premium rate services fraud
  - Fraudsters using stolen calling cards to call fake ‘premium’ numbers abroad
  - Use of automated robots to discover valid customer ANIs (ANI scanning) and dial out using those ANIs
- Our CDR analysis approach
  - Create features based on graph analysis, duration of calls, and interval between subsequent calls
  - Create a custom feature to identify scanning

WE DETECT ANI SCANNING

- Detect over 80% of premium rate fraud, up to 10 days before actual fraud calls
- ANI scanning feature detects 50% of fraud



Premium Rate Fraud

CASE STUDY 2: BENEFITS PROVIDER

- Fraudulent claims in state benefits
  - Fraudsters suspected of performing reconnaissance over IVR to find valid info
  - Use of valid account info for account takeovers
- Our approach
  - Combine reputation, velocity and behavior features
  - Train model over labeled set of calls
  - Use model to score incoming calls

WE FIND IVR RECONNAISSANCE

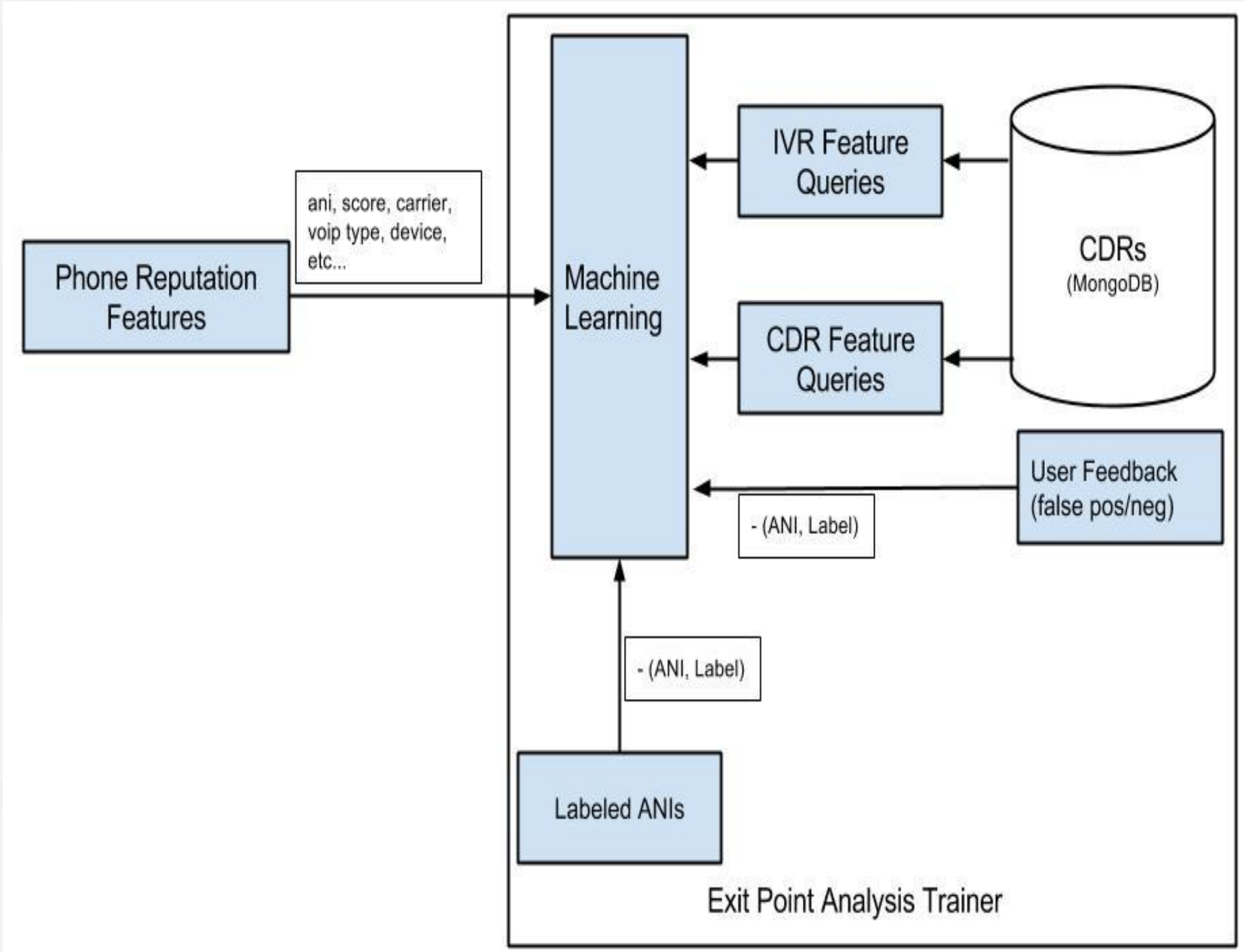
- Suspicious activity on 46% of accounts up to 2 months before fraud
- Specific instances of reconnaissance in IVR

ACEStart	ACEStart
Language Menu	ENGLISH
Balance PAN - Begin	
PAN Entry	**********8888
PAN Entry	Error
PAN Entry	Timeout
PAN Entry	Timeout
PAN Retry Menu	TryAgain
PAN Entry	**********8839
PAN Entry	Timeout
PAN Entry	Timeout
PAN Retry Menu	TryAgain
PAN Entry	**********3994
PAN Entry	Business Logic
ACEStart	ACEStart
Language Menu	ENGLISH
Balance PAN - Begin	
PAN Entry	**********8883
PAN Entry	Timeout
PAN Entry	Timeout
PAN Retry Menu	Error
PAN Retry Menu	TryAgain
PAN Entry	**********8883
PAN Entry	Timeout
PAN Entry	Timeout
PAN Retry Menu	Hang Up

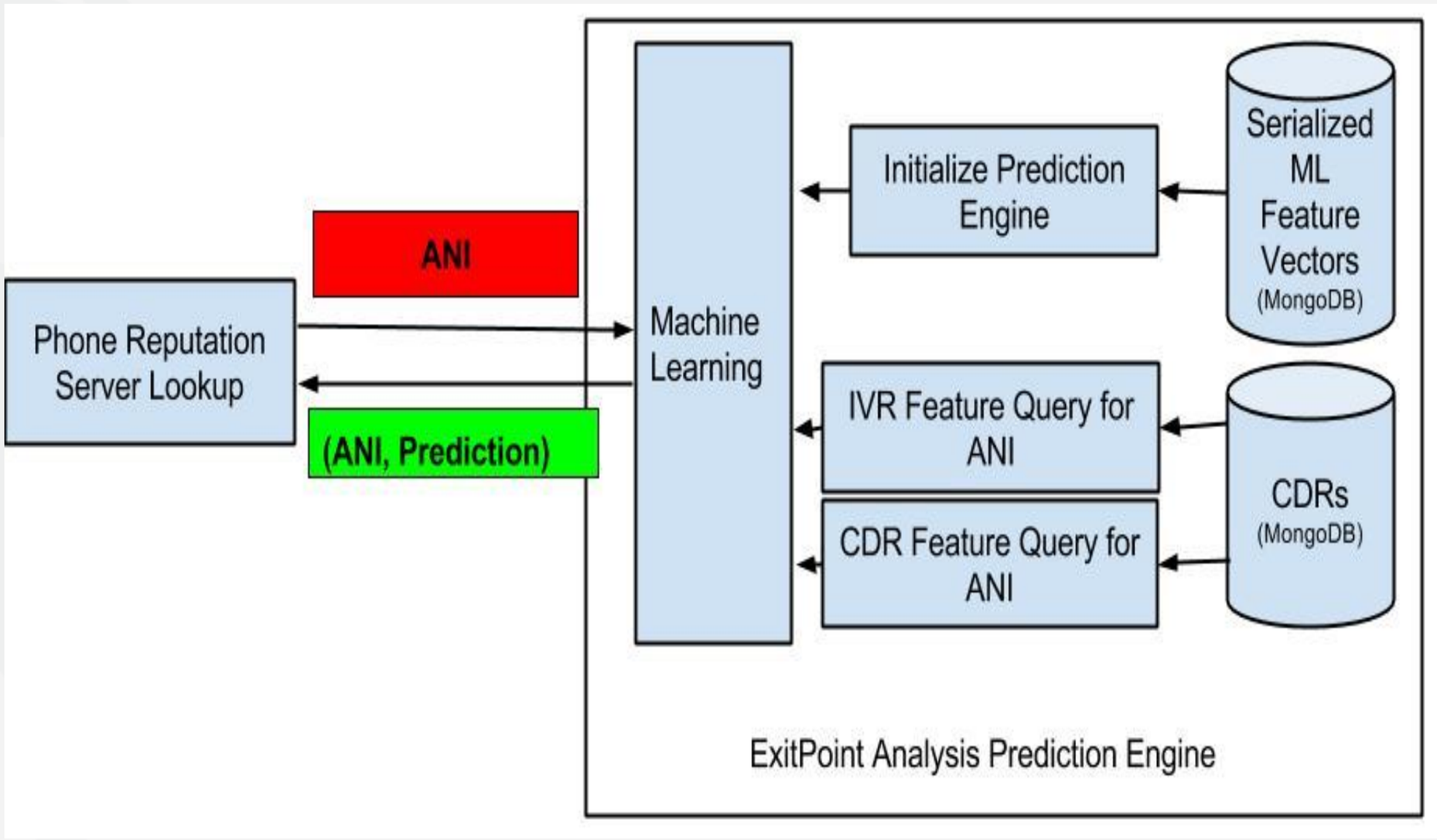
ACEStart	ACEStart
Language Menu	ENGLISH
Balance PAN - Begin	
PAN Entry	**********8839
PAN Entry	Error
PAN Entry	**********8399
PAN Entry	Timeout
PAN Entry	Timeout
PAN Retry Menu	Hang Up
ACEStart	ACEStart
Language Menu	ENGLISH
Balance PAN - Begin	
PAN Entry	**********9491
PIN Entry	****
PIN Entry	Balance Playback
ACEStart	ACEStart
Language Menu	Error
ACEStart	ACEStart
Language Menu	ENGLISH
Balance PAN - Begin	
PAN Entry	**********3629
PIN Entry	****
PIN Entry	Balance Playback

- ANI: 209-532-XXXX
- 7 consecutive calls in 1 hour
- Sequence of invalid Account number (PAN) entry attempts, followed by successful PAN entry and PIN entry.

ML TRAINING AT SCALE



REAL-TIME PREDICTION ARCHITECTURE



MONGODB: LESSONS LEARNED

- Bulk Ingest
  - Use Journaled Write-Concern for Inserts.
    Acknowledged Write-Concern for Updates.
- Query
  - Use Aggregations API + Indexes to generate IVR and CDR features.
- Prediction
  - Store feature vectors as Binary BSON objects.

CONCLUSION

- Account takeover – acoustical anomalies
  - > 80% TDR, < 2% FPR
  - 52% coming from international locations
- Account reconnaissance – CDR analysis
  - 46% detected 2 months before attack
- Detect pre-crime, zero day attacks and repeat attacks on the phone channel by complete understanding of lifecycle of a fraudster

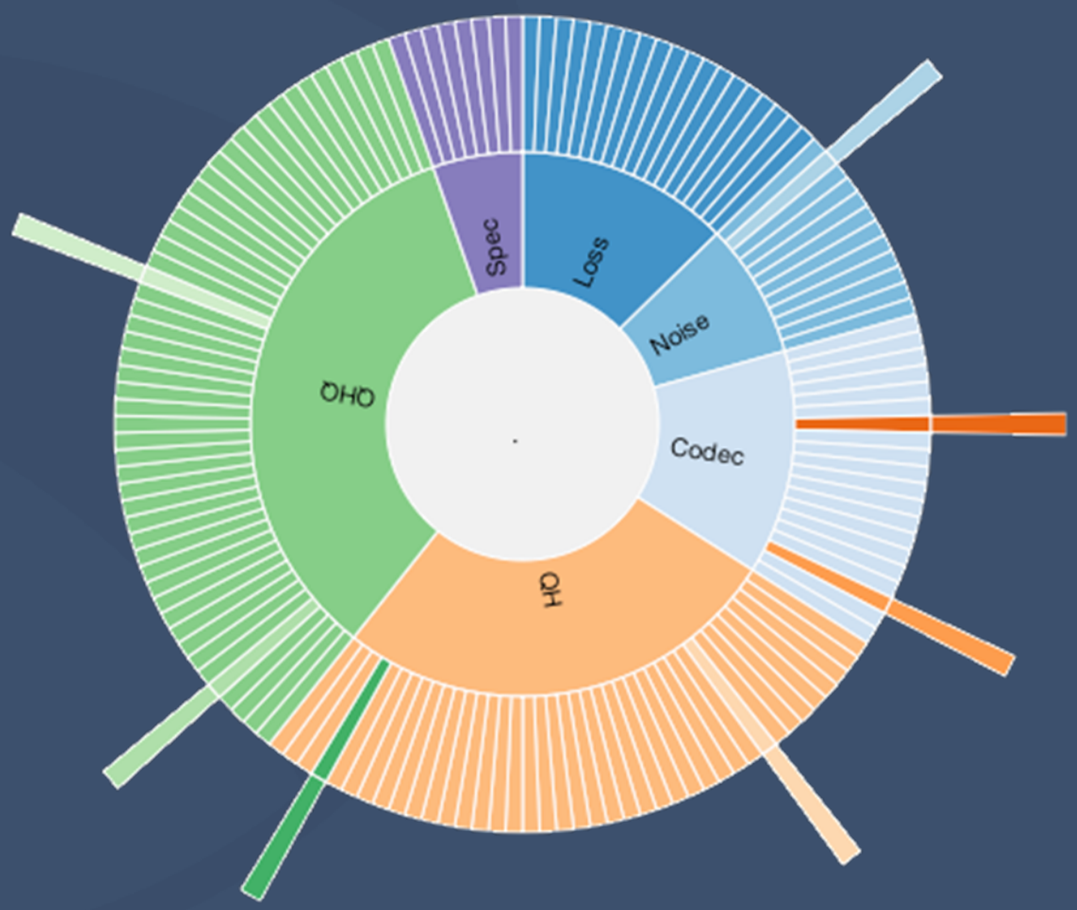
PINDROP SECURITY

vijay@pindropsecurity.com
raj@pindropsecurity.com
tcalhoun@pindropsecurity.com

PINDROP PHONEPRINT

pindrop security

Good



Fraud