Outline

1  Background
2  State of the Art
3  Our System
4  Locating FBSes
5  Summary
Story 1

From 95599 (Agriculture Bank of China):
We’re processing the student loan you’ve applied for, and now requiring you to transfer a deposit of ¥9900 ($1500) to the bank account XXXXXXXXX.

* Note: This is a simplified version of the actual story which involves more complex details.
Story 2

From 95566 (Bank of China): We’re processing the house mortgage for you. Please prepare ¥17,600,000 (≈ $2,600,000) ...

* Note: This is a simplified version of the actual story which involves more complex details.
GSM (Global System for Mobile Communication)

<table>
<thead>
<tr>
<th></th>
<th>Birth Year</th>
<th>User Scale</th>
<th>Speed</th>
<th>Security</th>
</tr>
</thead>
<tbody>
<tr>
<td>2G – GSM</td>
<td>1990</td>
<td>&gt; 1 billion</td>
<td>Low</td>
<td>Poor</td>
</tr>
<tr>
<td>3G – CDMA</td>
<td>2008</td>
<td>&lt; 2 billion</td>
<td>Middle</td>
<td>Middle</td>
</tr>
<tr>
<td>4G – LTE</td>
<td>2009</td>
<td>≈ 3 billion</td>
<td>High</td>
<td>Fine</td>
</tr>
</tbody>
</table>
FBS Carrier

Very high signal strength
Fake Base Station (FBS)
Which BS has the highest signal strength?

- 100 dBm
- 70 dBm
- 60 dBm
- 30 dBm

Location Update

I may have to switch my BS connection...

FBS Attack on GSM Phones
FBS Attack on GSM Phones

- 70 dBm

- 60 dBm

- 100 dBm

- 30 dBm

New Connection

GSM
FBS Can Also Impact 3G/4G Phones

3G/4G → Jamming Signal → GSM

GSM has existed for many years, so abandoning GSM also needs many years …
FBS Attack Is NOT Hypothetical

<table>
<thead>
<tr>
<th>Year</th>
<th># FBS Msgs</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013</td>
<td>&gt;&gt; 2.9 billion</td>
</tr>
<tr>
<td>2014</td>
<td>&gt;&gt; 4.2 billion</td>
</tr>
<tr>
<td>2015</td>
<td>&gt;&gt; 5.7 billion</td>
</tr>
</tbody>
</table>

N * billion
FBS Industry in China

- **Device: $400**
  - Daily income: $40

- **Device: $1000**
  - Daily income: $70

- **Device: $700**
  - Daily income: **up to $1400**
FBS Industry in China

Device: $400
Daily income: $40

Device: $1000
Daily income: $70

Device: $700
Daily income: up to $1400
2 State of the Art
Electronic Fence

Huge infrastructure costs → Poor scalability
FBS-signal Detection Car

Random walk $\rightarrow$ Limited coverage & “dull”
User Reporting

Dial 12321

Most users don’t realize the existence of FBSes
Client-side Tools

SnoopSnitch

SnoopSnitch is an Android app that collects and analyzes mobile radio data to make you aware of your mobile network security and to warn you about threats like fake base stations (IMSI catchers), user tracking and over-the-air updates. With SnoopSnitch you can use the data collected in the GSM Security Map at gsmmap.org and contribute your own data to GSM Map.

CatcherCatcher

The CatcherCatcher tool detects mobile network irregularities hinting at fake base station activity.

Do they really work in large-scale practice? …
Our System: FBS-Radar
Baidu PhoneGuard Users Opt-in

Report multiple fields of **suspicious** SMS messages

- Sender’s number is not in the recipient’s contact list
- Sender’s number is an authoritative number

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t_1 )</td>
<td>1452869570549</td>
</tr>
<tr>
<td>( s_1 )</td>
<td>-79 dBm</td>
</tr>
<tr>
<td>ID(_1)</td>
<td>460-00-39185-21492</td>
</tr>
<tr>
<td>( t_2 )</td>
<td>1452865343627</td>
</tr>
<tr>
<td>( s_2 )</td>
<td>-84 dBm</td>
</tr>
<tr>
<td>ID(_2)</td>
<td>460-00-39185-52921</td>
</tr>
<tr>
<td>( t_3 )</td>
<td>1452865278412</td>
</tr>
<tr>
<td>( s_3 )</td>
<td>-95 dBm</td>
</tr>
<tr>
<td>ID(_3)</td>
<td>460-00-39185-52112</td>
</tr>
<tr>
<td>Content of the message</td>
<td>(&lt; 140 ) characters</td>
</tr>
<tr>
<td>MAC(_1)</td>
<td>ec:26:ca:26:f6:c0</td>
</tr>
<tr>
<td>MAC(_2)</td>
<td>d0:c7:c0:aa:6a:fc</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>MAC(_n)</td>
<td>6a:3e:34:03:d8:13</td>
</tr>
</tbody>
</table>
Five Methods

1. Signal Strength Examination
2. BS ID Syntax Checking
3. Message Content Mining
4. BS-WiFi Location Analysis
5. BS-Handover Speed Estimation

~100M users
3.1 Signal Strength Examination

-40 dBm

0.23% of user-reported suspicious SMS messages
3.2 BS ID Syntax Checking

\[
\text{BS ID} = \text{MCC} + \text{MNC} + \text{LAC} + \text{CID}
\]

- MCC: Mobile Country Code, 3 digits
- MNC: Mobile Network Code, 2 digits
- LAC: Location Area Code, 16 bits
- CID: Cell Identity, 16 bits for 2G/3G and 28 bits for 4G

0.15% of suspicious messages were sent by BSes with syntactically invalid IDs
3.3 Message Content Mining

- Bag-of-words SVM (Support Vector Machine) classifier trained on 200,000 hand-labeled SMS messages
  
  - Labelling suspicious messages;
  - Word segmentation;
  - Feature extraction;
  - Quantizing the feature vector;
  - Training the SVM model;
  - Preprocessing the test set;
  - SVM classification of the test set.

- 0.16% of suspicious messages came from authoritative phone numbers and were determined to contain fraud text content
### 3.4 BS-WiFi Location Analysis

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_1$</td>
<td>1452869570549</td>
</tr>
<tr>
<td>$s_1$</td>
<td>-79 dBm</td>
</tr>
<tr>
<td>ID$_1$</td>
<td>460-00-39185-21492</td>
</tr>
<tr>
<td>$t_2$</td>
<td>1452865343627</td>
</tr>
<tr>
<td>$s_2$</td>
<td>-84 dBm</td>
</tr>
<tr>
<td>ID$_2$</td>
<td>460-00-39185-52921</td>
</tr>
<tr>
<td>$t_3$</td>
<td>1452865278412</td>
</tr>
<tr>
<td>$s_3$</td>
<td>-95 dBm</td>
</tr>
<tr>
<td>ID$_3$</td>
<td>460-00-39185-52112</td>
</tr>
<tr>
<td>Sender’s phone number</td>
<td>+86-135-5281-9836</td>
</tr>
<tr>
<td>Content of the message</td>
<td>$\leq$ 140 characters</td>
</tr>
</tbody>
</table>

MAC$_1$ = ec:26:ca:26:f6:c0
MAC$_2$ = d0:c7:c0:aa:6a:fc
MAC$_n$ = 6a:3e:34:03:d8:13

4.1% of suspicious messages were sent by BSes that were not in their correct geolocation, i.e., they were spoofing the ID of a legitimate but distant BS.
3.4 Counterfeiting a Nearby BS ID

If I counterfeit a nearby BS ID ...

- 70 dBm

- 60 dBm

- 30 dBm

- 100 dBm

My location does not change a lot, so I needn’t switch to a new BS 😊

Current Connection

Location Update
3.5 BS-Handover Speed Estimation

- For BS-WiFi location analysis, what if the WiFi location information is not available?

\[ V_{1,2-\text{max}} = \frac{d_{1,2} + r_1 + r_2}{t_1 - t_2}, \]
\[ V_{1,2-\text{avg}} = \frac{d_{1,2}}{t_1 - t_2}, \]
\[ V_{1,2-\text{min}} = \begin{cases} \frac{d_{1,2} - r_1 - r_2}{t_1 - t_2} & \text{when } d_{1,2} > r_1 + r_2, \\ 0 & \text{when } d_{1,2} \leq r_1 + r_2, \end{cases} \]
4.5 BS-Handover Speed Estimation

threshold_{CRH} = 350 \text{ km/h},

threshold_{Highway} = 150 \text{ km/h},

>> 0.39\% of suspicious SMS messages come from FBSes
Detection Performance

- > 4.7% of suspicious messages should have come from FBSes
  - False positive rate is only 0.05% (according to user feedback), mainly due to the inaccuracy of our WiFi database

- Set-3 (by message content mining) is >98% covered by the other 4 sets
  - No need to collect the text content of users’ messages!
Arresting FBS Operators

- With the help of FBS-Radar, the police have arrested tens to hundreds of FBS operators every month.
4 Locating FBSES
Locating FBSes based on User Device Locations

- FBSes frequently move and change their IDs
  - We take both temporal and spatial locality into account

Only those FBS messages
1) using the same BS ID,
2) happening in the same time window,
and 3) located in the same spatial cluster

... can be attributed to one FBS.
Locating FBSes based on User Device Locations

- The centroid of every cluster is the estimated location of an FBS.

This location accuracy is sufficient for us to track FBSes!
Real-time Locations of FBSES

Public URL →

http://shoujiweishi.baidu.com/static/map/pseudo.html
Using extensive crowdsourced data, we evaluate five different methods for detecting FBSes in the wild, and find that FBSes can be precisely identified without sacrificing user privacy.

We present a reasonable method for locating FBSes with an acceptable accuracy.

FBS-Radar is currently in use by ~100M people. It protects users from millions of malicious messages from FBSes every day, and has helped the authorities arrest numerous FBS operators every month.
Backup slides
FBS Attack: Passive vs. Active

Passive: IMSI-catcher

Active: Push spam/fraud SMS messages with spoofed phone numbers

<table>
<thead>
<tr>
<th>Year</th>
<th># FBS Msgs</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013</td>
<td>&gt;&gt; 2.9 billion</td>
</tr>
<tr>
<td>2014</td>
<td>&gt;&gt; 4.2 billion</td>
</tr>
<tr>
<td>2015</td>
<td>&gt;&gt; 5.7 billion</td>
</tr>
</tbody>
</table>

Rarely reported in China, but sometimes reported in the US.
Ground Truth

- Our ONLY ground truth comes from users’ feedback

We think this message comes from an FBS. What do you think?

- Yes: 99.95%
- No: 0.05%

Manual double-check
Why not use GPS?

- Most people turn GPS off in most time to save battery, so we have to ask users for GPS privilege

 Locattion accuracy increases by 20%?

 User scale decreases by 20%? for harassment...
Localizing User Devices based on WiFi Information

- The centroid of the **dominant** cluster is the estimated location of the user device.
Spam and Fraud SMS Messages

**Spam (Ads)**

“We are selling excellent, cheap goods and food from Jul. to Aug. 2016. Visit our shops at the People’s Square as soon as possible!” --- sent from a (usually not well-known) mart or grocery.

“We provide very cheap and legal invoices that can help you quickly make a big fortune. Don’t hesitate, dial us via the phone number: 010-61881234!” --- sent from a (usually not well-known) company.

**Fraud**

“Dear user, you are lucky to be the winner of this month’s big award! You will be offered 10-GB FREE 4G traffic by clicking on this URL: http://www.10086award.com.” --- sent from 10086 (China Mobile).

“Dear customer, you have failed to pay for this year’s management fee of 100 dollars. If you do not pay for it before Jul. 30th, you will face a fine of 500 dollars. You should pay it by transferring money to the following bank account: …” --- sent from 95533 (Bank of China).
FBS-Radar: 4-fold Design Goals

- Detect as many FBSes as possible with very few false positives, without specialized hardware.
- Automatically filter spam/fraud FBS messages from user devices with a high precision.
- Provide actionable intelligence about geolocations of FBSes to aid law enforcement agencies.
- Use minimal resources on client side, minimize collection of sensitive data, and not require root.
FBS-Radar & Baidu PhoneGuard

Crowdsourced data from ~100 Million Users
Database and List

BS ID $\rightarrow$ <lat, lon, radius, tag>

BS-location Database

WiFi MAC $\rightarrow$ <lat, lon, tag>

WiFi-location Database

User Report $\rightarrow$ Content-free Analysis

Message Logs

Content-free Analysis $\rightarrow$ Content Analysis

SVM Machine Learning Cluster

$\approx$ 1500 phone numbers

Authoritative Phone Number List
FBS-Radar: Timeline

- **2014.01-07**
  Design & Implementation

- **2014.08**
  Online released

- **2015**
  ~17.5 million suspicious SMS messages reported per day

- **2016**
  ~32 million suspicious SMS messages reported per day
Informed Consent from Users

3.1 Baidu PhoneGuard provides smartphone users with comprehensive security protections, such as smartphone acceleration, garbage cleaning, app management, anti-spam, anti-phishing, traffic monitoring, anti-virus, secure marketing, and so forth. To achieve specific functions and services, Baidu PhoneGuard may collect the following data from users:

......

3.1.4 The anti-spam service (offered in part by FBS-Radar) can help users detect spam/fraud SMS messages. During the detection, Baidu PhoneGuard may need to scan the users’ contact lists. In order to intelligently identify and filter spam/fraud SMS messages, Baidu PhoneGuard may need to analyze the features of SMS messages sent from suspicious phone numbers. If an SMS message is determined to be spam/fraud, it will be quarantined on the user device.

Baidu PhoneGuard never collects data without user authorization. Meanwhile, users can opt-out of the anti-spam service in the app settings at any time. Furthermore, users can report wrongly determined spam/fraud SMS messages (to our team) through the app. The report will be uploaded to our servers, and this action is only performed after your active operation.

The cloud-side anti-spam function is able to further increase the detection precision and recall, as well as decrease the false positive rate. Data of suspicious spam/fraud SMS messages is encrypted and uploaded to the Baidu Cloud Security Center for analysis. These data is securely stored on Baidu servers, and will never be leaked out (unless required by law enforcement agencies). In addition, these data is never used to match with users' personal information.

3.2 This app does not contain malicious code aiming to undermine data on user devices or acquire private data of users.

......
Opt-in Options for Users

- Detection Rules (for FBS-Radar)
- Intelligent Detection ON/OFF
- Cloud-side Detection ON/OFF
- Content Detection of Suspicious SMS Messages ON/OFF
- Suspicious Voice Call & SMS Message Detection ON/OFF
- Contacts’ Voice Call & SMS Message Detection ON/OFF