Imperial College London

USC Viterbi School of Engineering

Characterizing Improper Input Validation Vulnerabilities

of Mobile Crowdsourcing Services

Sojhal Ismail Khan, Dominika Woszczyk, **Chengzeng You**, Soteris Demetriou, Muhammad Naveed
University of Southern California
Imperial College London

Imperial College London

USC Viterbi School of Engineering

Content

Imperial College London

USC Viterbi School of Engineering

1. Mobile crowdsourcing services (MCSs)

MCSs enable economical, rapid,and scalable data acquisition utilized for accurate information sharing.

Imperial College London

USC Viterbi School of Engineering

1. Mobile crowdsourcing services (MCSs)

MCSs enable economical, rapid,and scalable data acquisition utilized for accurate information sharing.

Imperial College London

USC Viterbi School of Engineering

1. Mobile crowdsourcing services (MCSs)

MCSs enable economical, rapid,and scalable data acquisition utilized for accurate information sharing.

Imperial College London

USC Viterbi School of Engineering

1. Mobile crowdsourcing services (MCSs)

MCSs enable economical, rapid,and scalable data acquisition utilized for accurate information sharing.

Imperial College London

USC Viterbi School of Engineering

2. Attacks on MCSs

Simon et al. [45] showed how a real-world navigation service can be fooled to make wrong predictions on traffic density, allowing an adversary to redirect traffic. **--Hard to replicate**

More systematic studies conducted by Polakis et al. [42] and Wang et al. [52] **-- Specific to the target MCS**

Imperial College London

USC Viterbi School of Engineering

3. Contributions

- **New Techniques.** Range and constraint, and semantic input exploration strategies, Simulation methods.
- **Framework for Analysis**. Feedback-driven framework, feedback monitoring mechanisms.
- **New Findings**. Previously unknown vulnerabilities for 10 high-profile MCSs.

Imperial College London

USC Viterbi School of Engineering

Content

Imperial College London

USC Viterbi School of Engineering

1. Improper Input Validation(IIV)

- **Range and Constraint Validation**. This step ensures that the input domain range is minimized to accept values meaningful to the context of the service.
- **Semantic Validation**. Semantic validation is used to validate the meanings of the inputs.

Imperial College London

USC Viterbi School of Engineering

2. Threat Model

- Adversary (A) has the access to the mobile app of the target ubiquitous crowdsourcing service.
- A can observe the traffic generated between the app and the remote service either by passive eavesdropping or active man in the middle attacks.
- A can also reverse engineer and analyze the mobile app interfacing with the service.

Imperial College London

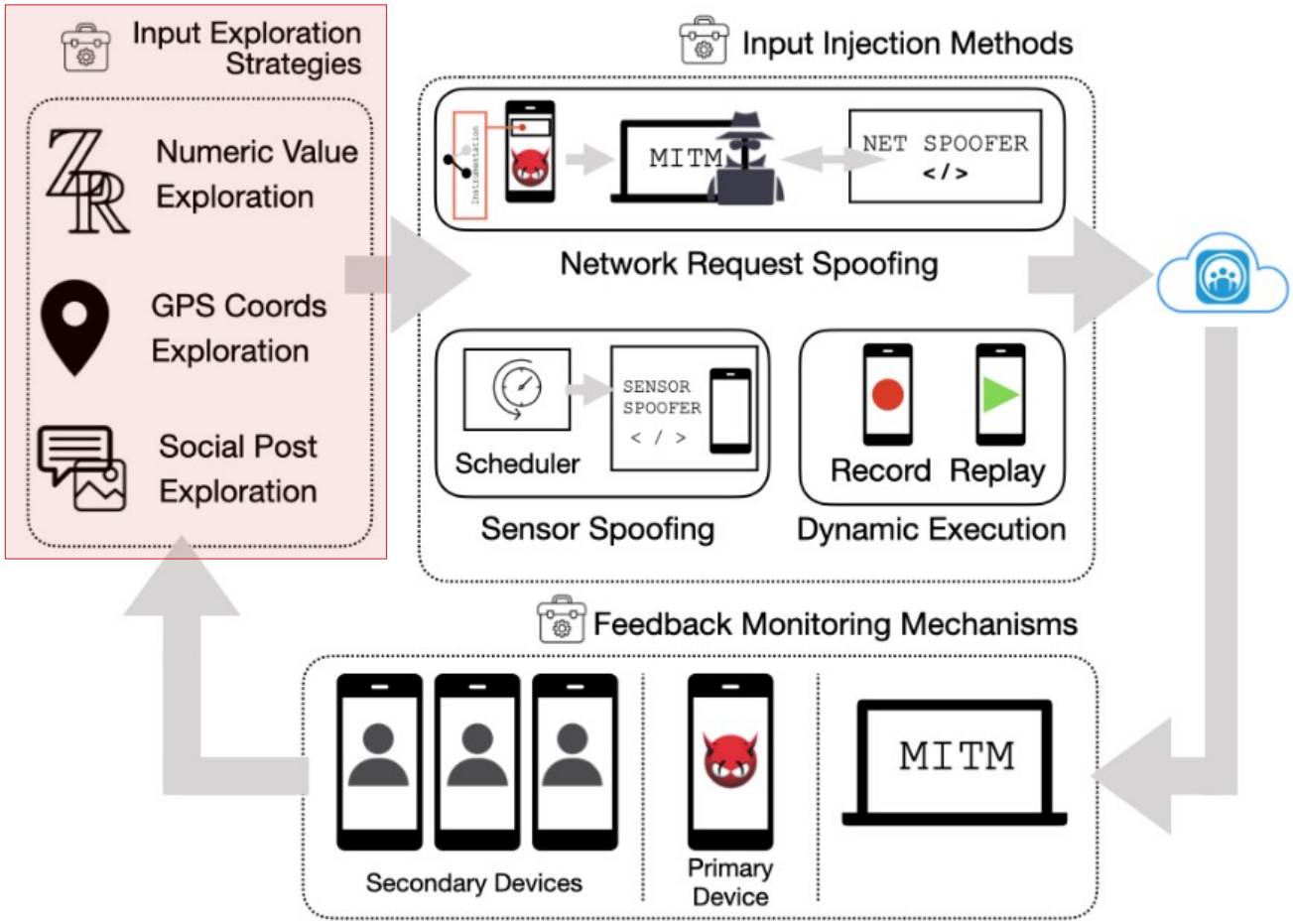
USC Viterbi School of Engineering

Content

Imperial College London

USC Viterbi School of Engineering

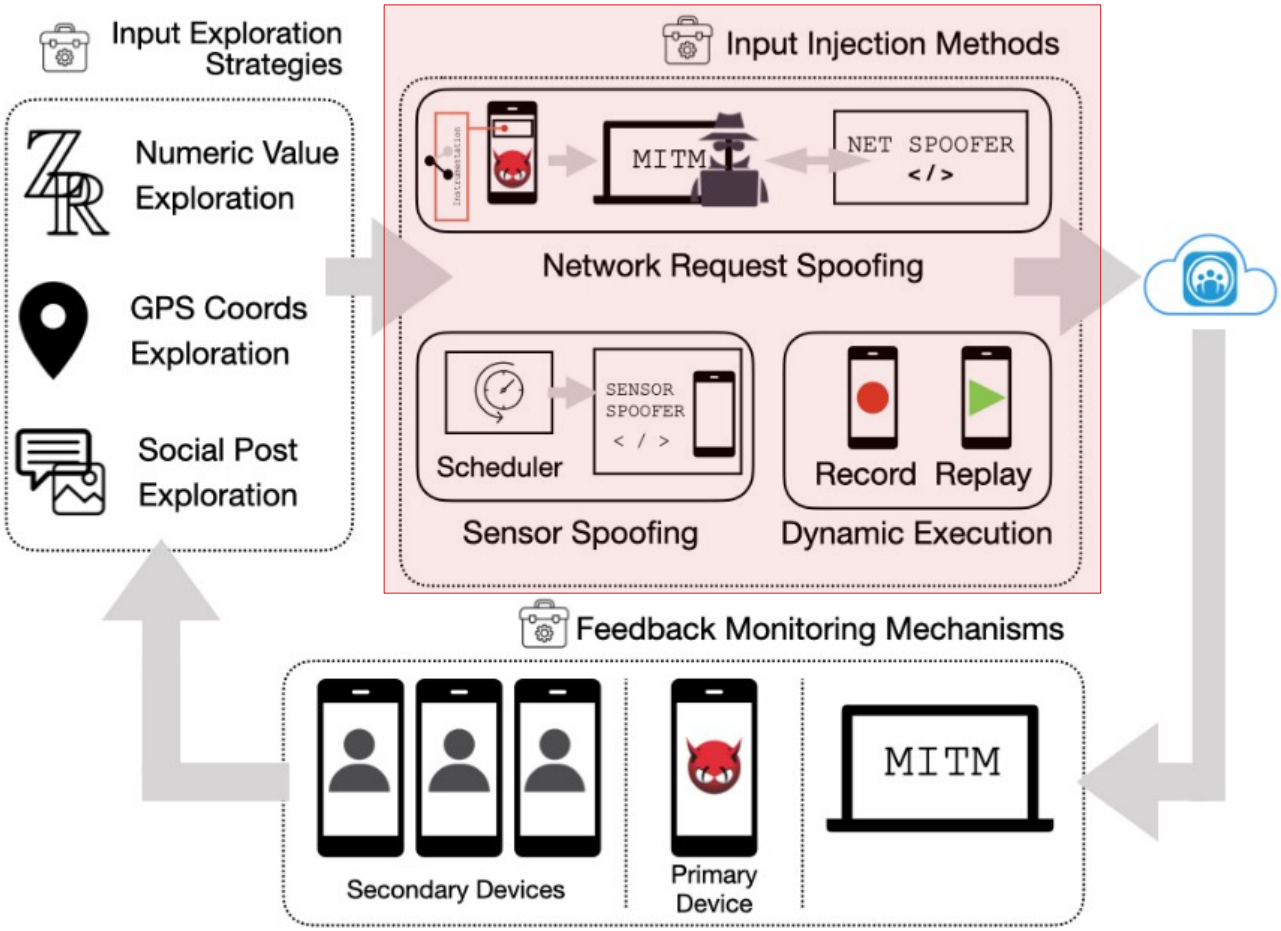
Overview



Imperial College London

USC Viterbi School of Engineering

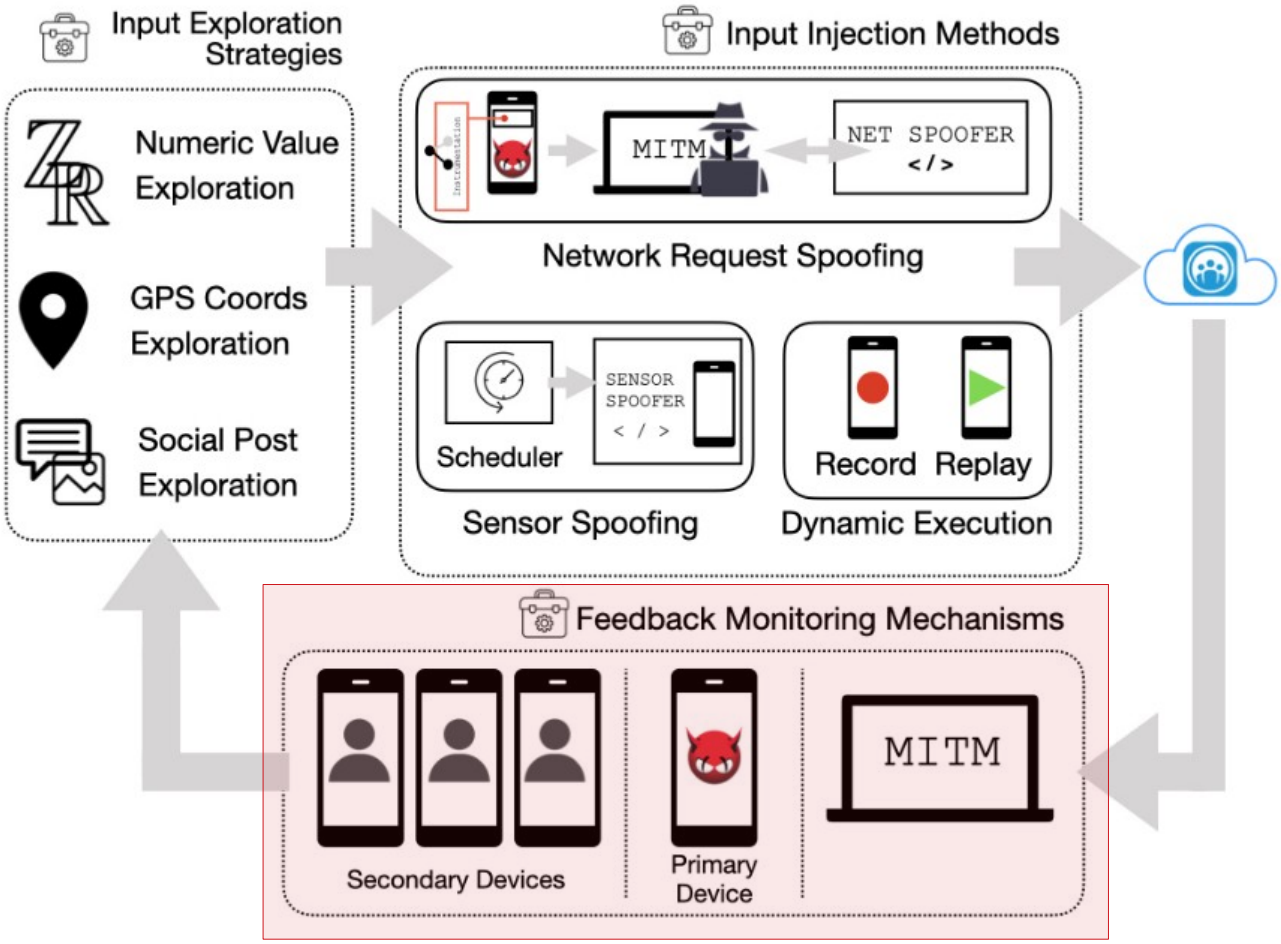
Overview



Imperial College London

USC Viterbi School of Engineering

Overview



Imperial College London

USC Viterbi School of Engineering

Content

Imperial College London

USC Viterbi School of Engineering

1. Fitness Activity Services

1.1 Settings

- **Target**: Strava
- **Injection Method**: spoofing network requests
- **Input Exploration Strategy**: numeric value exploration strategy
- **Types of Activities**: running, cycling, and swimming

The injection value is verified using a fake athlete's account.

Imperial College London

USC Viterbi School of Engineering

1. Fitness Activity Services

1.2 Results

- Maximum boundary for the duration: **31,622,400 sec = 8784 hours = 1 year**
- Maximum distance: 50,000 km (**running around the Earth 1.25 times**)
- The maximum accumulated distance: 4,294,967,295 (**2^{32}**) meters

Imperial College London

USC Viterbi School of Engineering

2. Pricing Services

2.1 Settings

- **Target**: Basket Savings
- **Injection Method**: spoofing network requests
- **Input Exploration Strategy**: numeric value exploration

The injected value is verified using a secondary passive device.

Imperial College London

USC Viterbi School of Engineering

2. Pricing Services

2.2 Results

Table 1: Basket: Trader Joe's & Amazon Prime(*)

Product	Value	Min	Max	*Value	*Min	*Max
Apples	0.49	0.05	2.0	1.58	0.16	4.0
Bananas	0.19	0.09	2.0	0.55	0.06	2.0
Strawberries	0.99	0.09	2.0	5.0	2.21	8.3
Eggs	1.99	0.2	4.0	2.12	0.21	6.0
Chicken Breasts	2.69	0.27	6.0	3.25	0.33	8.0
Organic whole Milk	3.49	0.35	8.0	3.76	0.38	8.0

Table 2: Basket: Milk on Trader Joe's

Product	Gallons	Value	Min	Max
Whole Milk 1	0.5	1.29	0.13	4.0
Whole Milk 2	0.5	2.29	0.23	6.0
Organic Whole Milk 1	0.5	2.99	0.30	6.0
Organic Whole Milk 2	1	5.69	1.71	10.58
Homogenized Whole Milk	1	5.99	1.80	6.59

Imperial College London

USC Viterbi School of Engineering

3. Transportation Services

3.1 Settings

- **Target**: Transit
- **Injection Method**: dynamic execution + sensor spoofing
- **Platform**: Genymotion non-root emulators
- **Scheduler**: emulate the speed of movement of the adversarial device

An injection is considered successful when it can affect the expectation of the bus arrival time on the observer device.

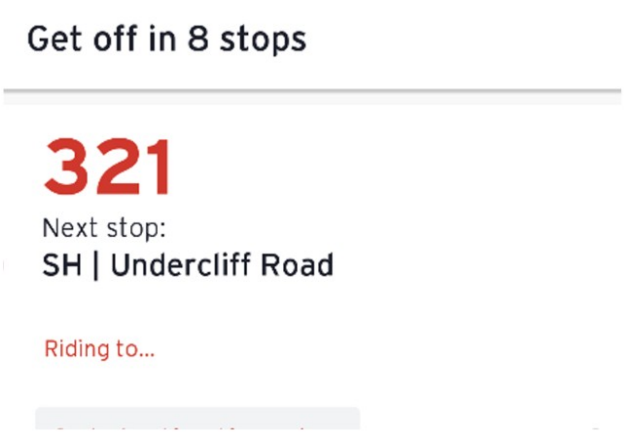
Imperial College London

USC Viterbi School of Engineering

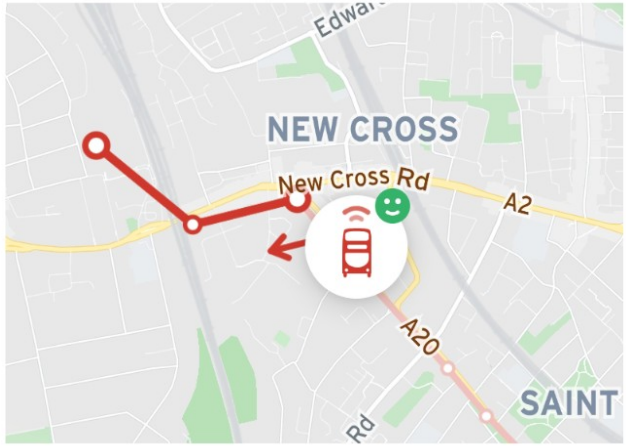
3. Transportation Services

3.2 Linear value exploration

- Fake speed values from 0 to 1000km/h with a step size of 10 km/h.
- **97%** fake movements succeeded in fooling Transit that the adversarial emulator is actively riding a fake bus.



Adversarial side



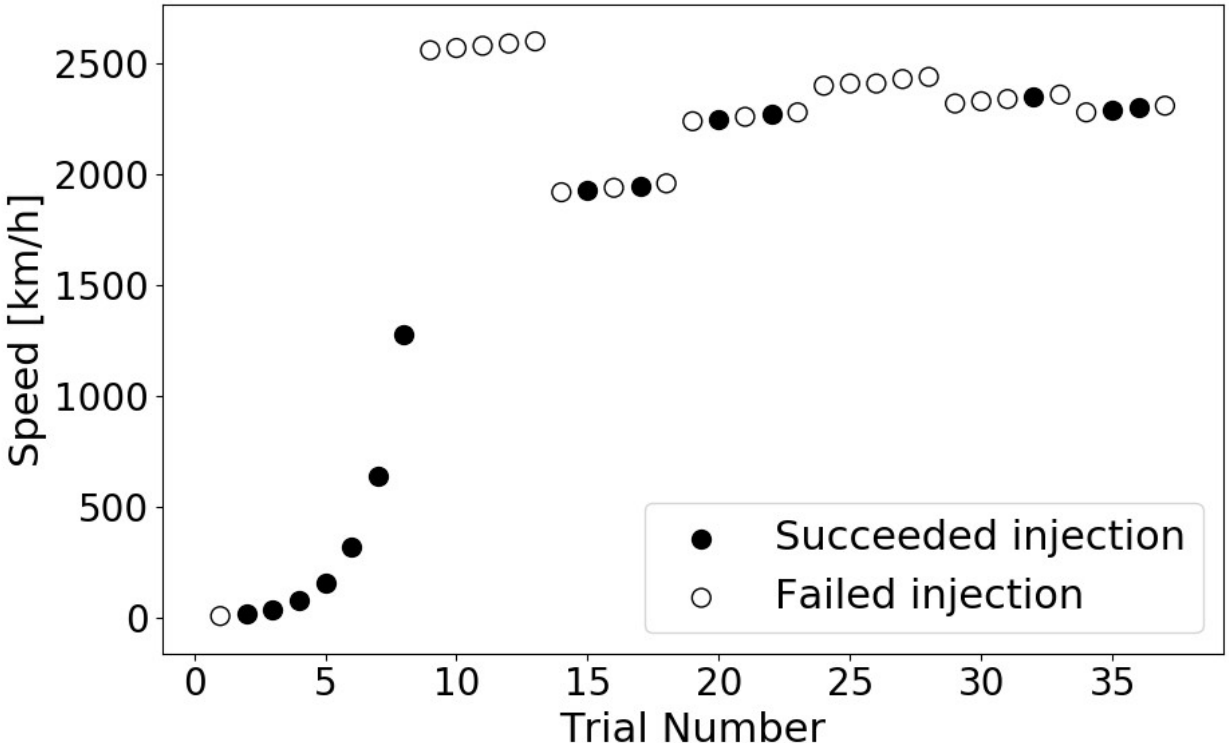
Victim side

Imperial College London

USC Viterbi School of Engineering

3. Transportation Services

3.3 Supersonic speeds



Speeds up to 2350km/h are possible

4. Location-based Services

4.1 Settings

- **Target**: Police Detector
- **Injection Method**: spoofing network requests
- **Input Exploration Strategy**: GPS coordinates exploration

A reverse engineered API is used for observing the success or failure of the injection

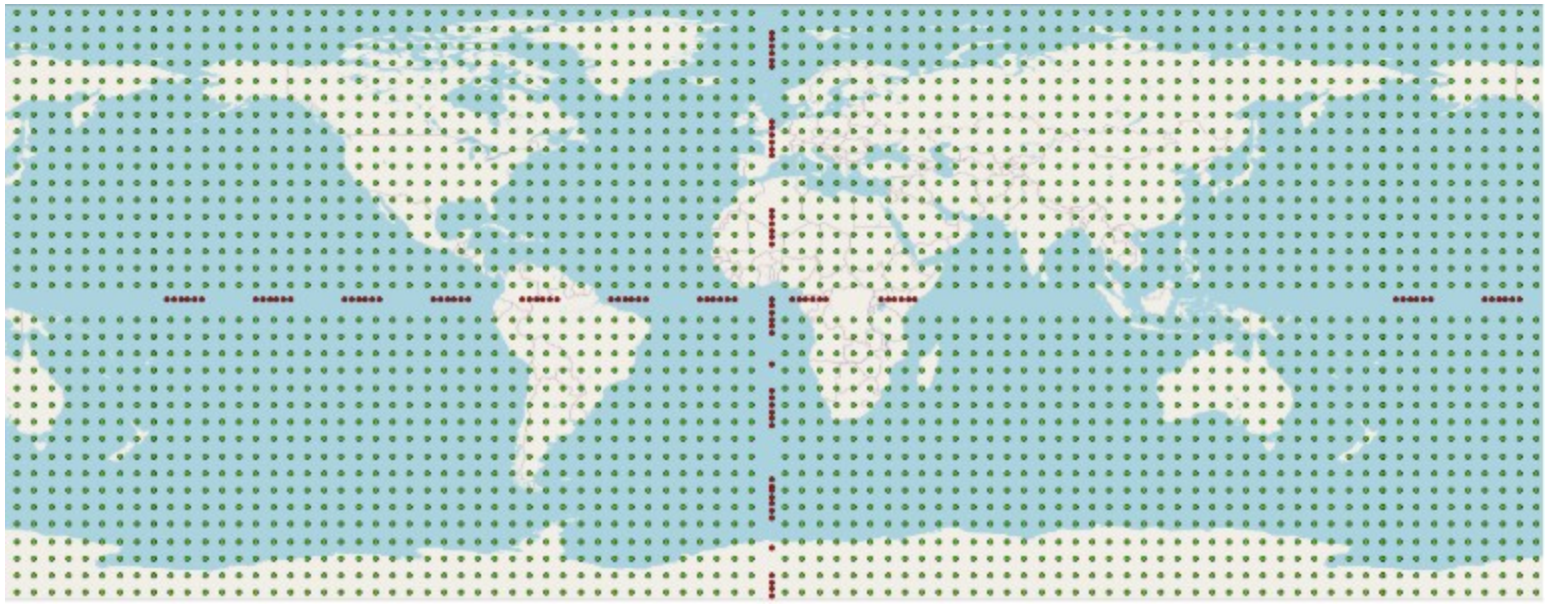
Imperial College London

USC Viterbi School of Engineering

4. Location-based Services

4.2 Results

- Successful injections for the entire 2D range of points.
- Point of interests (POIs) can be inserted with a precision of up to 5 decimal places but no two POIs can be closer than 0.002.



Imperial College London

USC Viterbi School of Engineering

5. Safety Services

5.1 Settings

- **Target**: Neighbors by Ring
- **Injection Method**: dynamic execution
- **Input Exploration Strategy**: post generation

If the post appears on the devices, it is marked as accepted

Imperial College London

USC Viterbi School of Engineering

5. Safety Services

5.2 Results

Fake Injection Success Rates

Strategy	Image	Crime	Safety	Lost Pet	Unexpected Activity	Total
RSG	n/a	n/a	n/a	n/a	n/a	0/100
SGP	n/a	9/25	6/25	10/25	1/25	23/100
SGA	n/a	22/25	19/25	16/25	9/25	66/100
SGP	Irrelevant	9/25	6/25	10/25	1/25	23/100
SGA	Irrelevant	22/25	**20/25**	16/25	9/25	**67/100**
SGP	Relevant	9/25	6/25	**16/25**	**2/25**	**33/100**
SGA	Relevant	**23/25**	**20/25**	**25/25**	**12/25**	**80/100**

RSG: random sentence generation
SGP: sentence generation with pre-trained GPT-2
SGA: sentence generation with adapted GPT-2

Imperial College London

USC Viterbi School of Engineering

5. Safety Services

5.2 Results

Fake Injection Success Rates

Strategy	Image	Crime	Safety	Lost Pet	Unexpected Activity	Total
RSG	n/a	n/a	n/a	n/a	n/a	0/100
SGP	n/a	9/25	6/25	10/25	1/25	23/100
SGA	n/a	22/25	19/25	16/25	9/25	66/100
SGP	Irrelevant	9/25	6/25	10/25	1/25	23/100
SGA	Irrelevant	22/25	**20/25**	16/25	9/25	**67/100**
SGP	Relevant	9/25	6/25	**16/25**	**2/25**	**33/100**
SGA	Relevant	**23/25**	**20/25**	**25/25**	**12/25**	**80/100**

RSG: random sentence generation
SGP: sentence generation with pre-trained GPT-2
SGA: sentence generation with adapted GPT-2

Imperial College London

USC Viterbi School of Engineering

5. Safety Services

5.2 Results

Fake Injection Success Rates

Strategy	Image	Crime	Safety	Lost Pet	Unexpected Activity	Total
RSG	n/a	n/a	n/a	n/a	n/a	0/100
SGP	n/a	9/25	6/25	10/25	1/25	23/100
SGA	n/a	22/25	19/25	16/25	9/25	66/100
SGP	Irrelevant	9/25	6/25	10/25	1/25	23/100
SGA	Irrelevant	22/25	**20/25**	16/25	9/25	**67/100**
SGP	Relevant	9/25	6/25	**16/25**	**2/25**	**33/100**
SGA	Relevant	**23/25**	**20/25**	**25/25**	**12/25**	**80/100**

RSG: random sentence generation
SGP: sentence generation with pre-trained GPT-2
SGA: sentence generation with adapted GPT-2

Imperial College London

USC Viterbi School of Engineering

5. Safety Services

5.2 Results

Fake Injection Success Rates

Strategy	Image	Crime	Safety	Lost Pet	Unexpected Activity	Total
RSG	n/a	n/a	n/a	n/a	n/a	0/100
SGP	n/a	9/25	6/25	10/25	1/25	23/100
SGA	n/a	22/25	19/25	16/25	9/25	66/100
SGP	Irrelevant	9/25	6/25	10/25	1/25	23/100
SGA	Irrelevant	22/25	**20/25**	16/25	9/25	**67/100**
SGP	Relevant	9/25	6/25	**16/25**	**2/25**	**33/100**
SGA	Relevant	**23/25**	**20/25**	**25/25**	**12/25**	**80/100**

RSG: random sentence generation
SGP: sentence generation with pre-trained GPT-2
SGA: sentence generation with adapted GPT-2

Imperial College London

USC Viterbi School of Engineering

Content

Imperial College London

USC Viterbi School of Engineering

1. countermeasures

1.1 Input validation

Input validation can be a great addition in our defense arsenal which can minimize the adversary's incentive.

- Easy to implement
- Can be immediately deployed with a software update on the server side
- Do not assume any capabilities on the participants' devices

Imperial College London

USC Viterbi School of Engineering

1. countermeasures

1.2 Results

App Domain	App Type	Example Countermeasure	Function	Reduction
Strava	Fitness	Restrict running distance (d) to be at most the world record	$0 > d \leq 350m \pm e_1$	98.65%
Map My Run	Fitness	Restrict running distance (d) to be at most the world record	$0 > d \leq 350m \pm e_1$	99.58%
Fitbit	Fitness	Restrict running distance (d) to be at most the world record	$0 > d \leq 350m \pm e_1$	99.58%
Transit	Transportation	Enforce bus speed (v) according to highway code—70mph in UK Motorways	$0 > v \leq 70mph \pm e_2$	94.25%
Basket Savings	Pricing	Use auxiliary data sources to verify price	|aux_price - reported_price| < threshold	Varies
Police Detector	Location	Restrict distance ($d()$) between inserted location ($loc(i)$) and the nearest road segment to be within 10m	$d(loc(i), near(loc(i))) \leq 10m.$	99.89%
NbR	Safety	Use metrics based on user reputation	reputation(user) > threshold	Varies

Imperial College London

USC Viterbi School of Engineering

1. countermeasures

1.2 Results

App Domain	App Type	Example Countermeasure	Function	Reduction
Strava	Fitness	Restrict running distance (d) to be at most the world record	$0 > d \leq 350m \pm e_1$	98.65%
Map My Run	Fitness	Restrict running distance (d) to be at most the world record	$0 > d \leq 350m \pm e_1$	99.58%
Fitbit	Fitness	Restrict running distance (d) to be at most the world record	$0 > d \leq 350m \pm e_1$	99.58%
Transit	Transportation	Enforce bus speed (v) according to highway code—70mph in UK Motorways	$0 > v \leq 70mph \pm e_2$	94.25%
Basket Savings	Pricing	Use auxiliary data sources to verify price	|aux_price - reported_price| < threshold	Varies
Police Detector	Location	Restrict distance ($d()$) between inserted location ($loc(i)$) and the nearest road segment to be within 10m	$d(loc(i), near(loc(i))) \leq 10m$.	99.89%
NbR	Safety	Use metrics based on user reputation	reputation(user) > threshold	Varies

Imperial College London

USC Viterbi School of Engineering

1. countermeasures

1.2 Results

App Domain	App Type	Example Countermeasure	Function	Reduction
Strava	Fitness	Restrict running distance (d) to be at most the world record	$0 > d \leq 350m \pm e_1$	98.65%
Map My Run	Fitness	Restrict running distance (d) to be at most the world record	$0 > d \leq 350m \pm e_1$	99.58%
Fitbit	Fitness	Restrict running distance (d) to be at most the world record	$0 > d \leq 350m \pm e_1$	99.58%
Transit	Transportation	Enforce bus speed (v) according to highway code—70mph in UK Motorways	$0 > v \leq 70mph \pm e_2$	94.25%
Basket Savings	Pricing	Use auxiliary data sources to verify price	|aux_price - reported_price| < threshold	Varies
Police Detector	Location	Restrict distance ($d()$) between inserted location ($loc(i)$) and the nearest road segment to be within 10m	$d(loc(i), near(loc(i))) \leq 10m.$	99.89%
NbR	Safety	Use metrics based on user reputation	reputation(user) > threshold	Varies

Imperial College London

USC Viterbi School of Engineering

1. countermeasures

1.2 Results

App Domain	App Type	Example Countermeasure	Function	Reduction
Strava	Fitness	Restrict running distance (d) to be at most the world record	$0 > d \leq 350m \pm e_1$	98.65%
Map My Run	Fitness	Restrict running distance (d) to be at most the world record	$0 > d \leq 350m \pm e_1$	99.58%
Fitbit	Fitness	Restrict running distance (d) to be at most the world record	$0 > d \leq 350m \pm e_1$	99.58%
Transit	Transportation	Enforce bus speed (v) according to highway code—70mph in UK Motorways	$0 > v \leq 70mph \pm e_2$	94.25%
Basket Savings	Pricing	Use auxiliary data sources to verify price	|aux_price - reported_price| < threshold	Varies
Police Detector	Location	Restrict distance ($d()$) between inserted location ($loc(i)$) and the nearest road segment to be within 10m	$d(loc(i), near(loc(i))) \leq 10m$.	99.89%
NbR	Safety	Use metrics based on user reputation	reputation(user) > threshold	Varies

Imperial College London

USC Viterbi School of Engineering

1. countermeasures

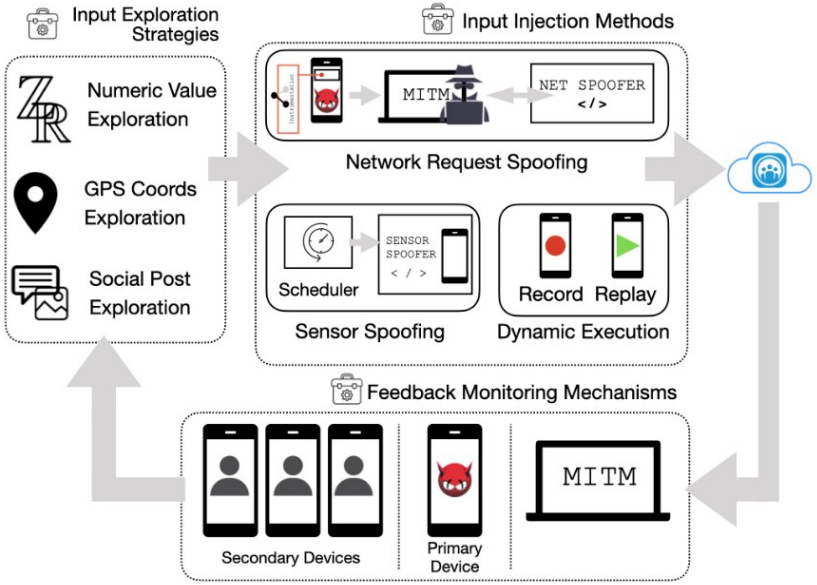
1.2 Results

App Domain	App Type	Example Countermeasure	Function	Reduction
Strava	Fitness	Restrict running distance (d) to be at most the world record	$0 > d \leq 350m \pm e_1$	98.65%
Map My Run	Fitness	Restrict running distance (d) to be at most the world record	$0 > d \leq 350m \pm e_1$	99.58%
Fitbit	Fitness	Restrict running distance (d) to be at most the world record	$0 > d \leq 350m \pm e_1$	99.58%
Transit	Transportation	Enforce bus speed (v) according to highway code—70mph in UK Motorways	$0 > v \leq 70mph \pm e_2$	94.25%
Basket Savings	Pricing	Use auxiliary data sources to verify price	|aux_price - reported_price| < threshold	Varies
Police Detector	Location	Restrict distance ($d()$) between inserted location ($loc(i)$) and the nearest road segment to be within 10m	$d(loc(i), near(loc(i))) \leq 10m.$	99.89%
NbR	Safety	Use metrics based on user reputation	reputation(user) > threshold	Varies

Imperial College London

USC Viterbi School of Engineering

Thank you



chengzeng.you19@imperial.ac.uk