RUHR-UNIVERSITÄT BOCHUM

Leveraging Sensor Fingerprinting for Mobile Device Authentication

Thomas Hupperich, Henry Hosseini, Thorsten Holz

HGI @ Ruhr-University Bochum

Status Quo

- Web-based fingerprinting of browsers
  - Mostly deployed for user tracking
  - No need for tracking cookies
  - Emerging technology thanks to JS & HTML5
  - Related Work: How Unique is Your Web Browser? By Eckersley,
    Cookieless Monster by Nikiforakis et al., …

- Hardware-based fingerprinting
  - Fingerprinting a system, not a browser
  - Access beyond Web context
  - Related Work: AccelPrint by Dey et al.,
    Remote physical device fingerprinting by Kohno et al., …

Approach

- Extraction of characteristic attributes (*features*) of a system
  - What attributes are characteristic?
- Combination of features as vector → Fingerprint
  - Which features are most discriminant?
- Recognize or identify a unique device
  - Machine learning classification

Why Sensors?

- Sensors are **hardware**
  - Bound to one system
- Sensors are **immutable**
  - Replacing sensors is not common
  - Tampering measurements requires system privileges
- Sensors are **characteristic**
  - Measurable hardware imperfections
- Sensors are **accessible**
  - Accelerometers & gyroscopes even via Web technology

Approach

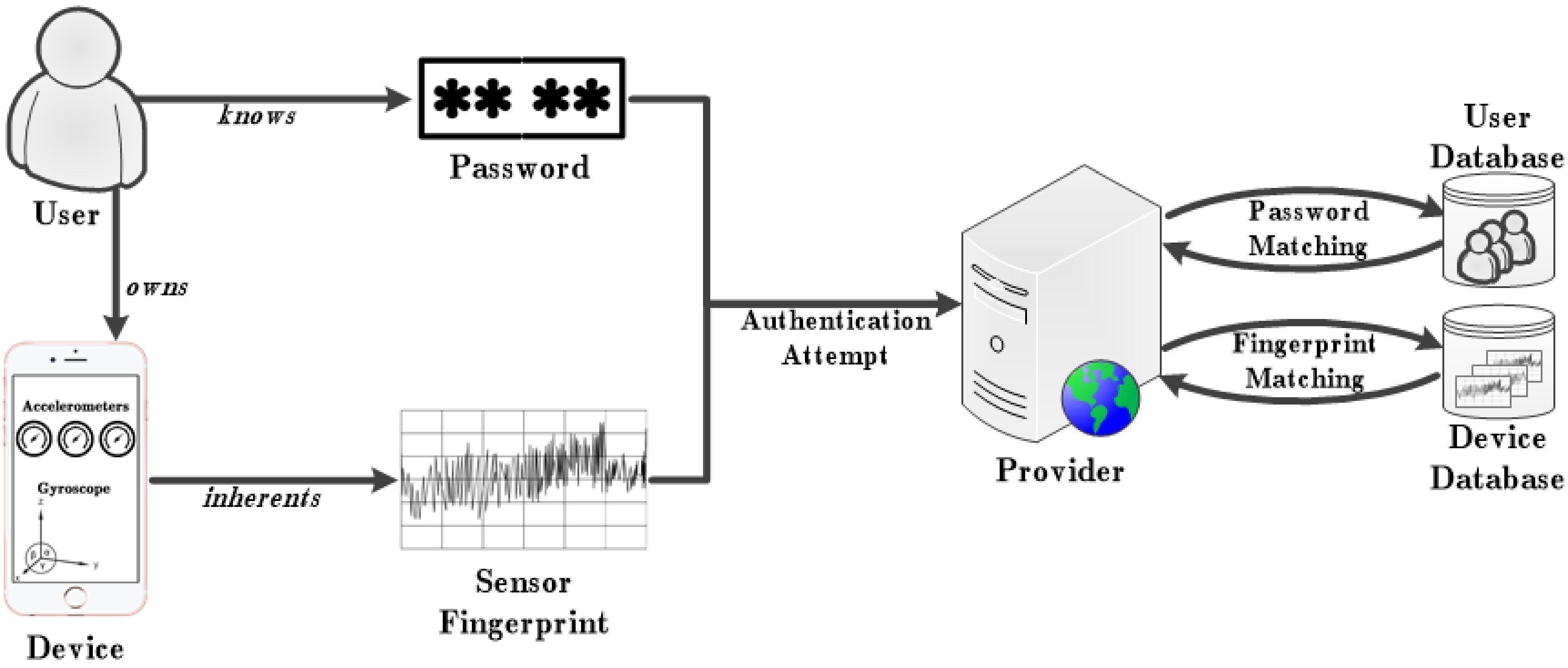
- Common use cases of fingerprinting
  - User tracking
  - Privacy breaches
  - Behavior Analysis
  - User tracking

- Is there any good purpose?
  - Fraud Detection
  - ➢ Authentication

Device Authentication

Leveraging Sensor Fingerprints

Authentication Process



- Device becomes authentication factor
- Provider is capable to verify device ownership

Fingerprinting Sensors
for Mobile Device Authentication

Data Set

- 4,989 devices
- Events and benchmarks obtained via App
- Sensor fingerprints are real-world data

Table 1. Numbers of events, benchmarks and devices per sensor type

Sensor Type	Events	Benchmarks	Devices
Acceleration	8,005,352	7,004	4,179
Magnetic Field	2,855,199	5,230	3,676
Orientation	8,047,497	6,228	4,963
Gyroscope	12,578,437	6,342	4,698
Gravity	9,061,253	5,726	4,374
Linear Acceleration	8,687,132	5,556	4,297
Rotation Vector	9,045,737	5,524	4,401

Feature Set

- Calculated features over all sensor events:

Time Domain	Mean	Standard Deviation	Average Deviation	Skewness
	Kurtosis	Root Mean Square	Lowest Value	Highest Value
Frquency Domain	Spectral Standard Deviation	Spectral Centroid	Spectral Skewness	Spectral Kurtosis
	Spectral Crest	Irregularity-K	Irregularity-J	Smoothness
	Flatness			

Classifier

- Set of commonly used classifiers
- Designed to handle (mostly) numeric values

- k-Nearest Neighbor (k-NN)
- Support Vector Machines (SVM)
- Bagging Tree (BT)
- Random Forest (RF)
- Extra Trees (ET)

Experiments Set-Up

- Data once grouped by device *models* and once grouped by single *devices*
  - Model: Recognition of a specific model, e.g., „Nexus 5"
  - Device: Recognition of a specific mobile phone, e.g., „Henry's Nexus 5"

- For both data sets the ***R****aw events* as well as the ***F****eature Set* are compared
  - R = Raw Measurements (no feature calculation)
  - F = Feature Set (extraction of characterisitc attributes)

- All classifiers are used and compared

- Single-Sensor Experiment: One specific sensor is taken into account
- Multi-Sensor Experiments: Recognition by *groups* of sensors

Evaluation

Leveraging Sensor Fingerprints for Mobile Device Authentication

Single-Sensor Experiments

Sensor	Identifier	Data	Classifier	Average Precision
Acceleration	**Device**	**F**	**ET**	**78.2300**
	Device	R	k-NN	62.6781
	Model	F	ET	69.3900
	Model	**R**	**BT**	**76.4570**
Magnetic Field	Device	F	ET	78.0100
	Device	**R**	**RF**	**96.3808**
	Model	F	ET	57.9100
	Model	**R**	**ET**	**96.4232**
Orientation	Device	F	ET	75.2400
	Device	**R**	**k-NN**	**98.2033**
	Model	F	ET	58.7400
	Model	**R**	**k-NN**	**98.1090**
Gyroscope	**Device**	**F**	**BT**	**49.4400**
	Device	R	k-NN	41.4460
	Model	**F**	**BT**	**50.5000**
	Model	R	k-NN	45.1595

Sensor	Identifier	Data	Classifier	Average Precision
Gravity	Device	F	ET	60.9500
	Device	**R**	**k-NN**	**82.9912**
	Model	**F**	**ET**	**54.7200**
	Model	R	k-NN	9.9967
Lin. Acceleration	**Device**	**F**	**BT**	**58.9200**
	Device	R	k-NN	18.8124
	Model	**F**	**BT**	**48.3500**
	Model	R	k-NN	10.1388
Rotation Vector	Device	F	ET	70.7200
	Device	**R**	**k-NN**	**99.8063**
	Model	F	ET	55.5700
	Model	**R**	**k-NN**	**99.8216**

R = raw data, F = features
k-NN = k-NearestNeighbor, BT = BaggingTree,
ET = ExtraTrees, RF = RandomForest
showing only best performing classifiers
bold rows show maximum precision rate

Evaluation

Single-Sensor Experiments

- The use of these (widely used) mathematical features is questionable
  - Using raw data may also yield a high recognition precision
- Recognition of device models is about as hard as recognizing single devices
- Acceleration sensors and gyroscopes are realatively bad for recognition

Evaluation

Leveraging Sensor Fingerprints for Mobile Device Authentication

Multi-Sensor Experiments

- Sensors grouped

Sensors	Identifier	Data	Classifier	Average Precision
Accelerometers	**Device**	**F**	**BT**	**92.4782**
	Device	R	BT	88.6941
	Model	**F**	**ET**	**91.5432**
	Model	R	BT	89.6469
Accelerometers & Gyroscope	Device	F	ET	88.5019
	Device	**R**	**BT**	**88.8444**
	Model	F	ET	92.3950
	Model	**R**	**RF**	**95.0076**
All Available Sensors	Device	F	ET	98.6026
	Device	**R**	**ET**	**99.9806**
	Model	F	ET	98.1615
	Model	**R**	**RF**	**99.9950**
No Accelerometers	Device	F	RF	97.2484
	Device	**R**	**ET**	**99.9922**
	Model	F	ET	97.4589
	Model	**R**	**ET**	**99.9821**
No Accelerometers & No Gyroscope	Device	F	RF	94.6407
	Device	**R**	**RF**	**99.9848**
	Model	F	RF	96.0450
	Model	**R**	**ET**	**99.9671**

R = raw data, F = features
k-NN = k-NearestNeighbor, BT = BaggingTree,
ET = ExtraTrees, RF = RandomForest
showing only best performing classifiers
bold rows show maximum precision rate

Multi-Sensor Experiments

- Generally very high recognition precision
- Higher recognition precision if data from several sensors is combined
- Raw sensor reading more effective than feature set
- Accelerometed-based recognition may rely on feature set, but:
- Accelerometers and gyroscopoe have almost no effect
  - Recognizing devices by other sensors is more effective

Conclusion

Leveraging Sensor Fingerprints for Mobile Device Authentication

- Obtained sensor data of almost 5,000 mobile devices
- Tested raw sensor readings against well-established feature set
- Classification tests with five different classifiers

- The feature set is suitable for accelerometes *only*
  - Computational effort for calculating features can be saved

- Recognition of devices and models is best when sensors are combined
  - Up to 99.98% for single devices and 99.995% for models

- Hardware-based device fingerprinting with sensor data
  - is feasible and
  - a valid method for device authentication (when based on multiple sensors)

RUHR-UNIVERSITÄT BOCHUM

! Thank You for Your Attention !

Leveraging Sensor Fingerprinting
for Mobile Device Authentication