Mobile Security
Fall 2015

Patrick Tague

#11: Mobile Sensing Benefits

[Slides c/o Dr. Jiang Zhu, slightly modified]

Class #11

- BehavioMetrics
- BehavioMetrics through mobile sensing
- Passive authentication / verification via BehavioMetrics

Why Passive Verification?

Password

A major source of security vulnerabilities. Easy to guess, reuse, forgotten, shared

Application

Different applications may have different sensitivities

Usability

Authentication too-often or sometimes too loose

Existing User Authentication

- Password management on mobile devices is either weak or unusable
  - Ex: password requirements for DHS E-file:
    - Contain from 8 to 16 characters
    - Contain at least 2 of the following 3 characters: uppercase alphabetic, lowercase alphabetic, numeric
    - Contain at least 1 special character (e.g., @, #, $, %, & *, +, =)
    - Begin and end with an alphabetic character
    - Not contain spaces, all or part of UserID, identical consecutive characters, or a recently used password
- Most users are too lazy or ignorant to use password-aid tools (Hong et al. 2009)
- Fingerprint? Gesture? Iris recognition? Face recognition? Voice recognition?

BehavioMetrics

- Derived from Behavioral Biometrics
  - Behavioral: the way a human subject behaves
  - Biometrics: technologies and methods that measure and analyzes biological characteristics of the human body
    - Finger prints, eye retina, voice patterns
- BehavioMetrics:
  - Measurable behavior to recognize or to verify identity of a human subject or subject’s certain behaviors

Mobile Sensing

- Mobile devices come with embedded sensors
  - Accelerometers, gyroscope, magnetometer
  - GPS receiver
  - WiFi, Bluetooth, NFC
  - Microphone, camera,
  - Temperature, light sensor
  - “Clock” and “Calendar”

- Connect with other sensors (e.g., EEG, EMG, GSR)
- Mobile devices are connected with the Internet
  - Upload sensor data to the cloud
  - Viewing information computing on the server side
- Users carry the device almost at all time
  - My phone “knows” where I am, what I am doing and my future activities

Mobile Sensing → BehavioMetrics

- Accelerometer
  - activity, motion, hand trembling, driving style
  - sleeping pattern
  - inferred activity level, steps made per day, estimated calorie burned
- Motion sensors, WiFi, Bluetooth
  - accurate indoor position and trace.
- GPS
  - outdoor location, geo-trace, commuting pattern
- Microphone, camera
  - From background noise: activity, type of location.
  - From voice: stress level, emotion
  - Video/audio: additional contexts
- Keyboard, taps, swipes
  - Specific tasks, user interactions, …

- *Network Factors*
- *Personal Factors*
- *Behavioral Factors*
- *Application Factors*

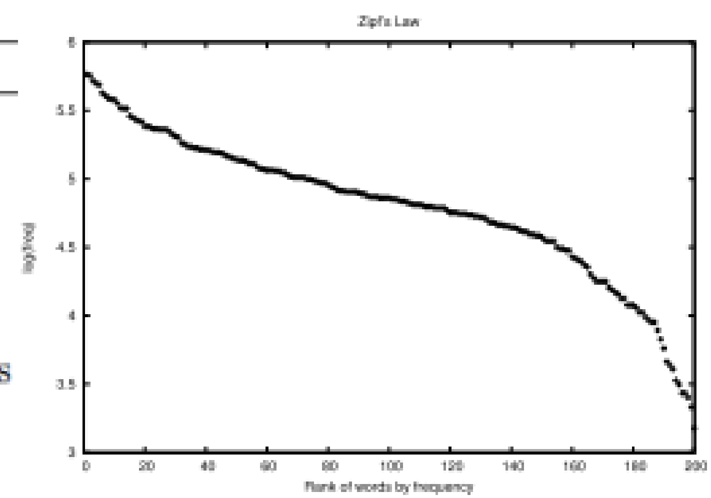
BehavioMetrics → Security

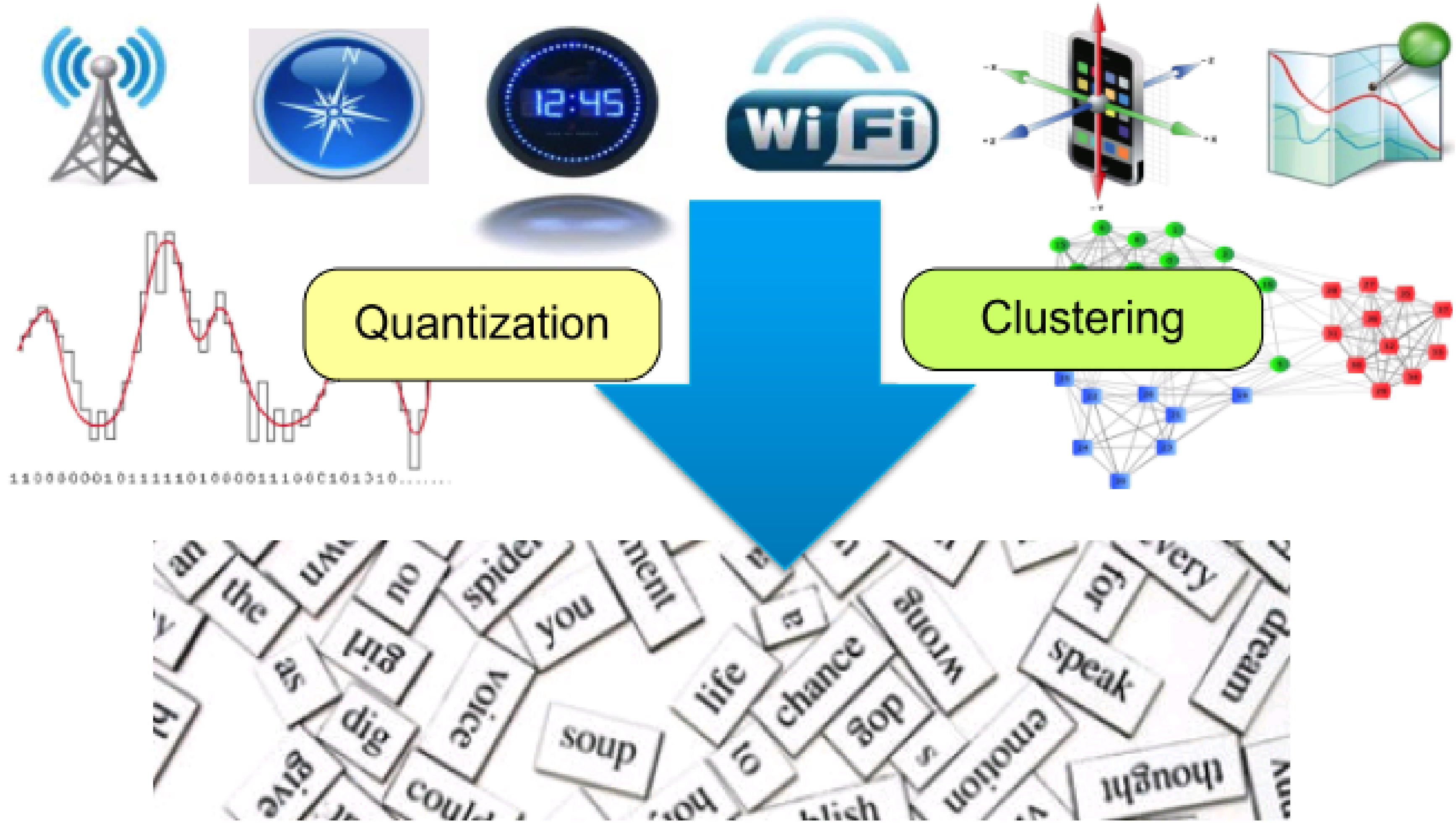
- Monitor and track user behavior on smartphones using various on-device sensors
- Convert sensory traces and other context information to personal behavior features
- Build continuous n-gram model with these features and use it for calculation of sureness scores
- Trigger various authentication schemes when certain applications are launched

"Behavioral Text"

- Human behavior/activities share some common properties with natural languages
  - Meanings are composed from meanings of building blocks
  - Exists an underlying structure (grammar)
  - Expressed as a sequence (time-series)
- Apply rich sets of Statistical NLPs to mobile sensory data

Natural Language	activity language	Example
Word	Atomic Movement	Turn upper body left
Phrase	Movement	Stand up
Sentence	Action	Climb up stairs
Paragraph	Activity	Enter building, climb up stairs and walk into office
Document	Event	Left home and ride bicycle to campus arrived at my office at 2nd floor





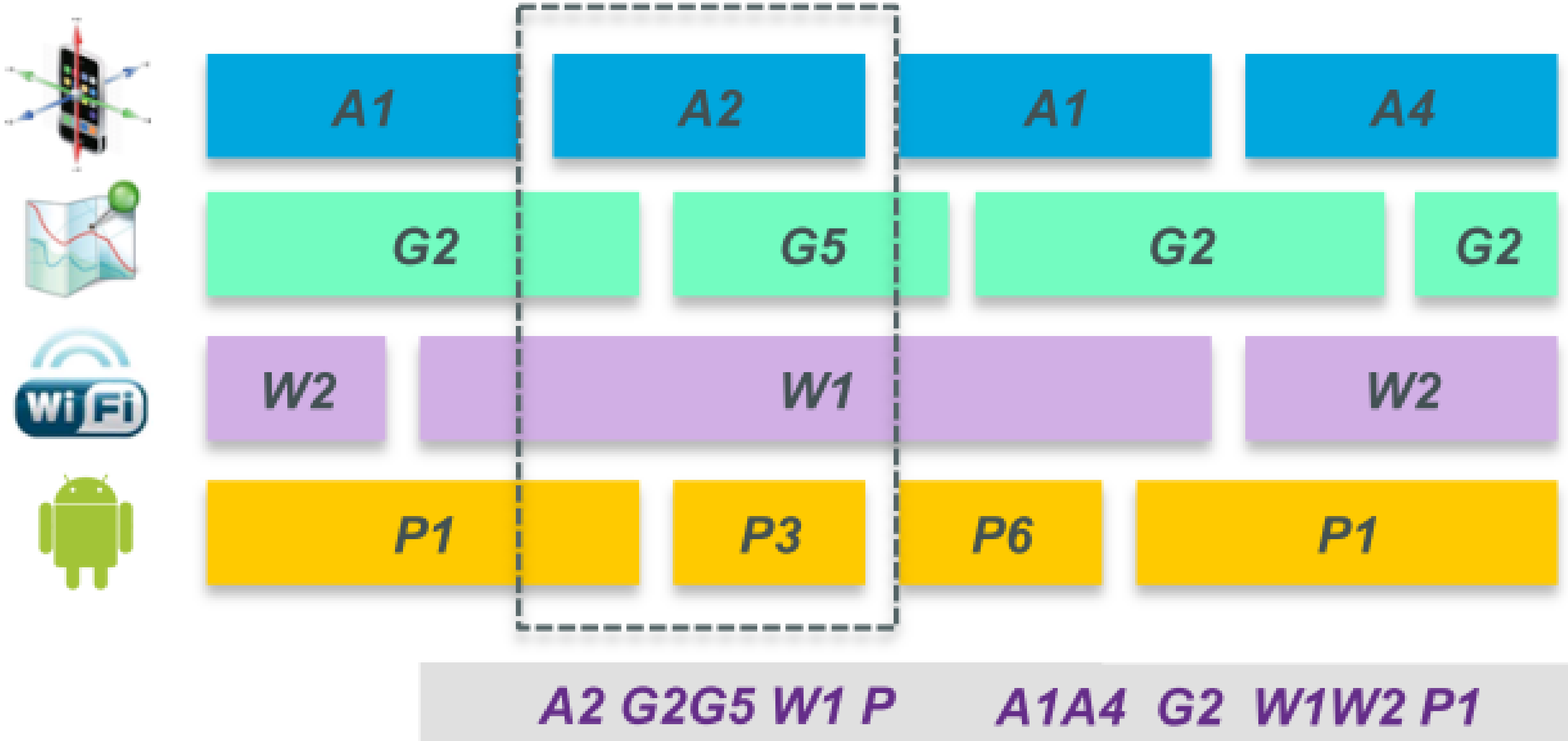
[31,271,37] [37,281,42] [37,276,47] [42,271,47] [42,266,53] [58,271,47] [53,271,47] [74,271,42] . . .

⇓

CZ DG GI FK C BI CS DC HQ BX FI FI BX FI O . . .

Sensor Data → Behavioral Text

- Convert feature vector series to label streams – dimension reduction using sliding window of specified length



Behavior ↔ Language

- Generative language model: P(English sentence | model)
  - P("President Obama signed the Bill of … "| Politics) >> P("President Obama signed the Bill of … " | Sports)
  - LM reflects the n-gram distribution of the training data: domain, genre, topics.
- With labeled behavior text data, we can train a LM for each activity type: "walking"-LM, "running"-LM and classify the activity as $i^* = \arg\max_i P(t|a_i)$

	Predicted Activity		
	walking	running	cycling
walking	95%	1%	4%
running	4%	94%	2%
cycling	2%	0%	98%

Continuous *n*-gram Model

- User activity at time *t* depends only on the last *n-1* activities
- Sequence of activities can be predicted by *n* consecutive activities in the past

$$P(l_i | l_{i-n+1}, l_{i-n+2}, \ldots, l_{i-1}) \quad \text{or} \quad P(l_i | l_{i-n+1}^{i-1})$$

- Maximum Likelihood Estimation from training data by counting:

$$P_{\mathrm{MLE}}(l_i | l_{i-n+1}^{i-1}) = \frac{C(l_{i-n+1}, \ldots, l_{i-1}, l_i)}{C(l_{i-n+1}, \ldots, l_{i-1})}$$

- MLE assign zero probability to unseen n-grams
  - Incorporate smoothing functions, discount observed grams, reserve probability for unseen grams

Classification

- Build *M* BehavioMetrics models P_0, P_1, P_2, … , P_{M-1}
  - Genders, age groups, occupations
  - Behaviors, activities, actions
  - Health and mental status
- For a new behavioral text string *L*, we calculate the probability of *L* being generated by model *m*

$$P(L, m) = P(l_1, l_2, \ldots, l_N, m) = \prod_{i=1}^{N} P_m(l_i | l_{i-n+1}^{i-1})$$

- Classification problem formulated as

$$\hat{u} = \operatorname*{argmax}_m P(L, m) = \operatorname*{argmax}_m \sum_{i=1}^{N} \log P_m(l_i | l_{i-n+1}^{i-1})$$

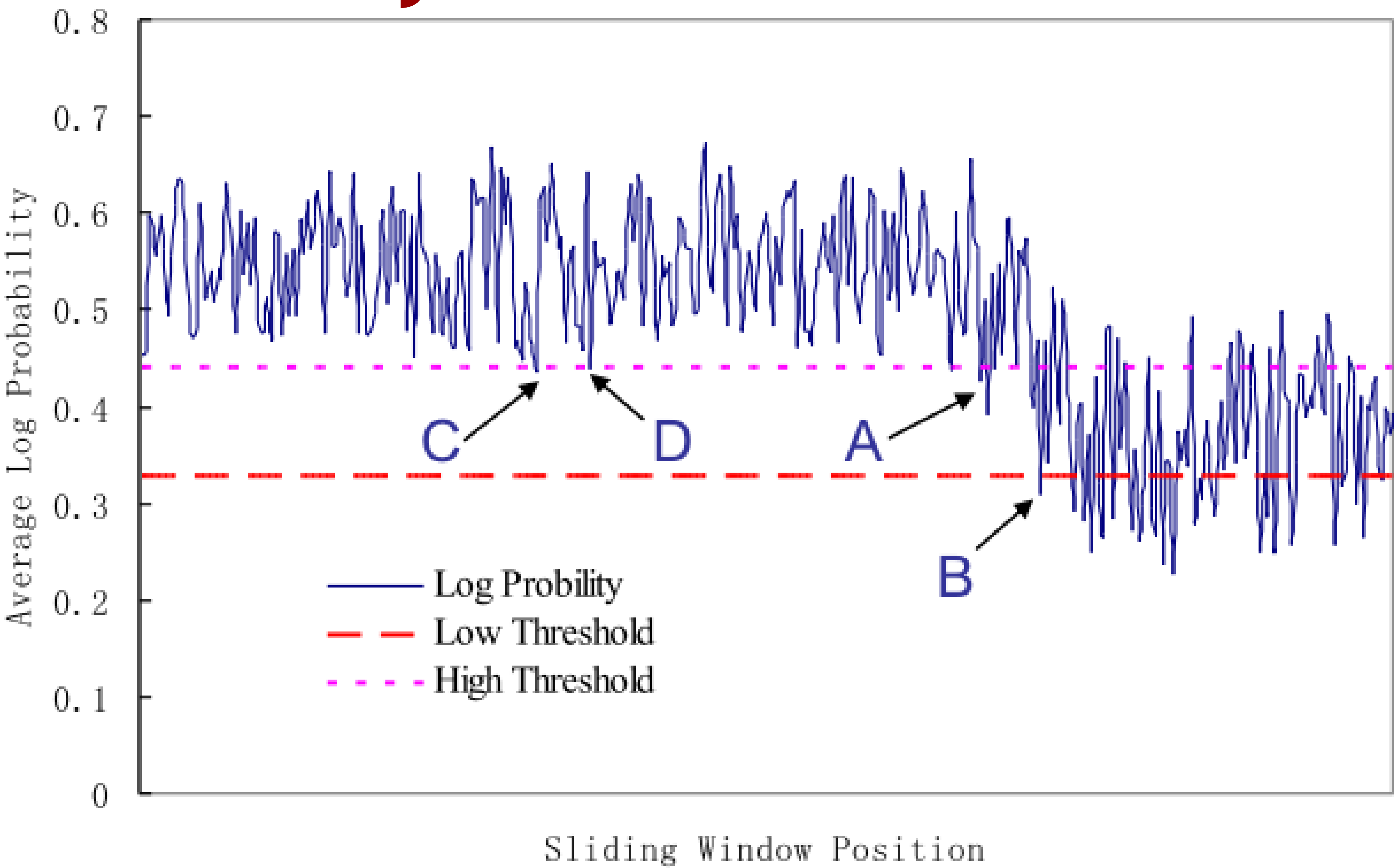
Anomaly Detection

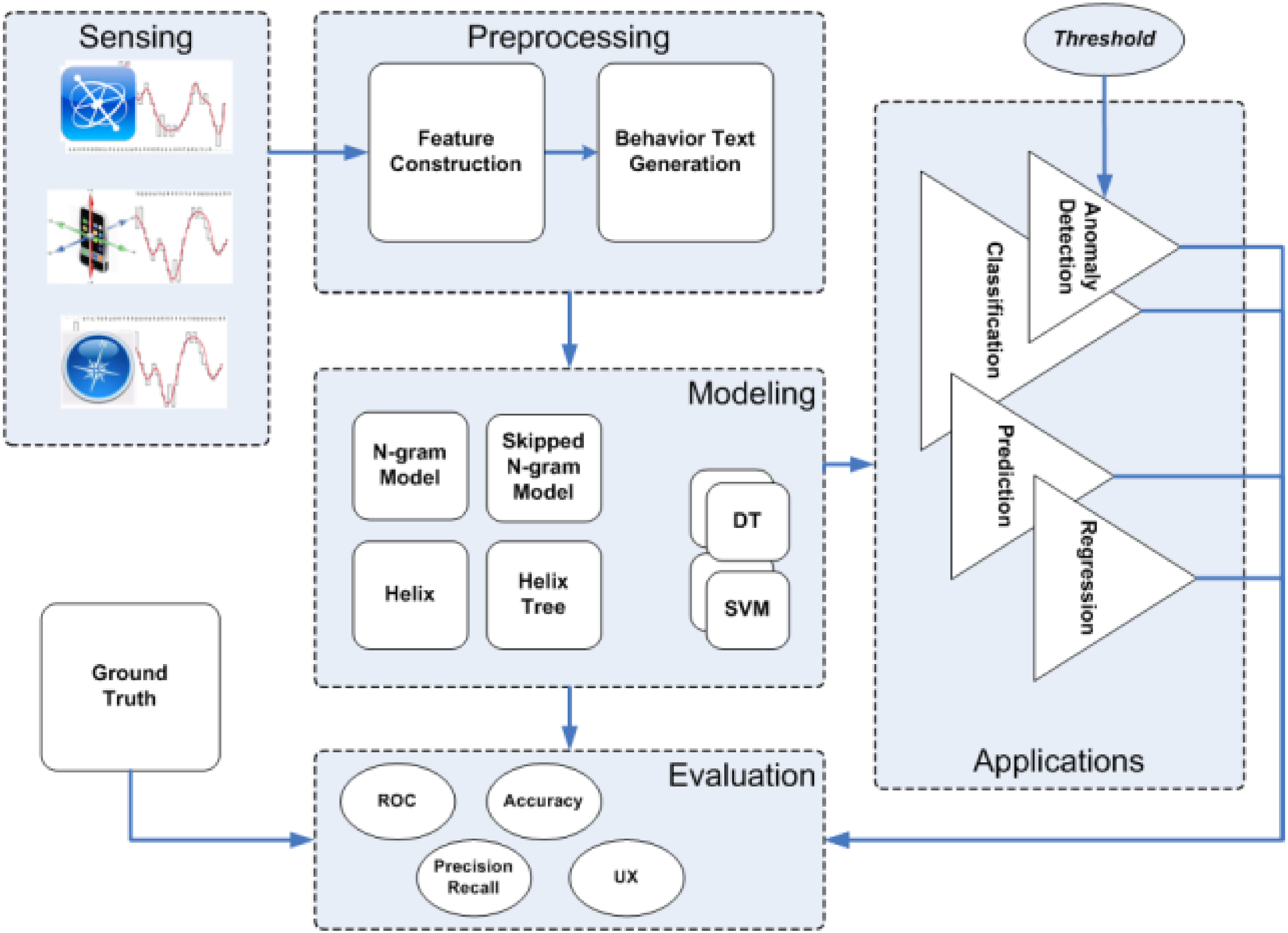
- A special binary classification problem
- Given a normal BehavioMetrics model P_n, a new behavior text sequence L, and a threshold θ, calculate the likelihood L is generated by P_n and compare with θ

$$\hat{a}(L|n,\theta) = sign[P(L,n) > \theta\ ]$$

- If the outcome is -1, flag an anomaly
- Variation caused by noise could be smoothed out statistically
- Need certain feedback to handle false positives, usually caused by unseen behaviors or sub-optimal threshold

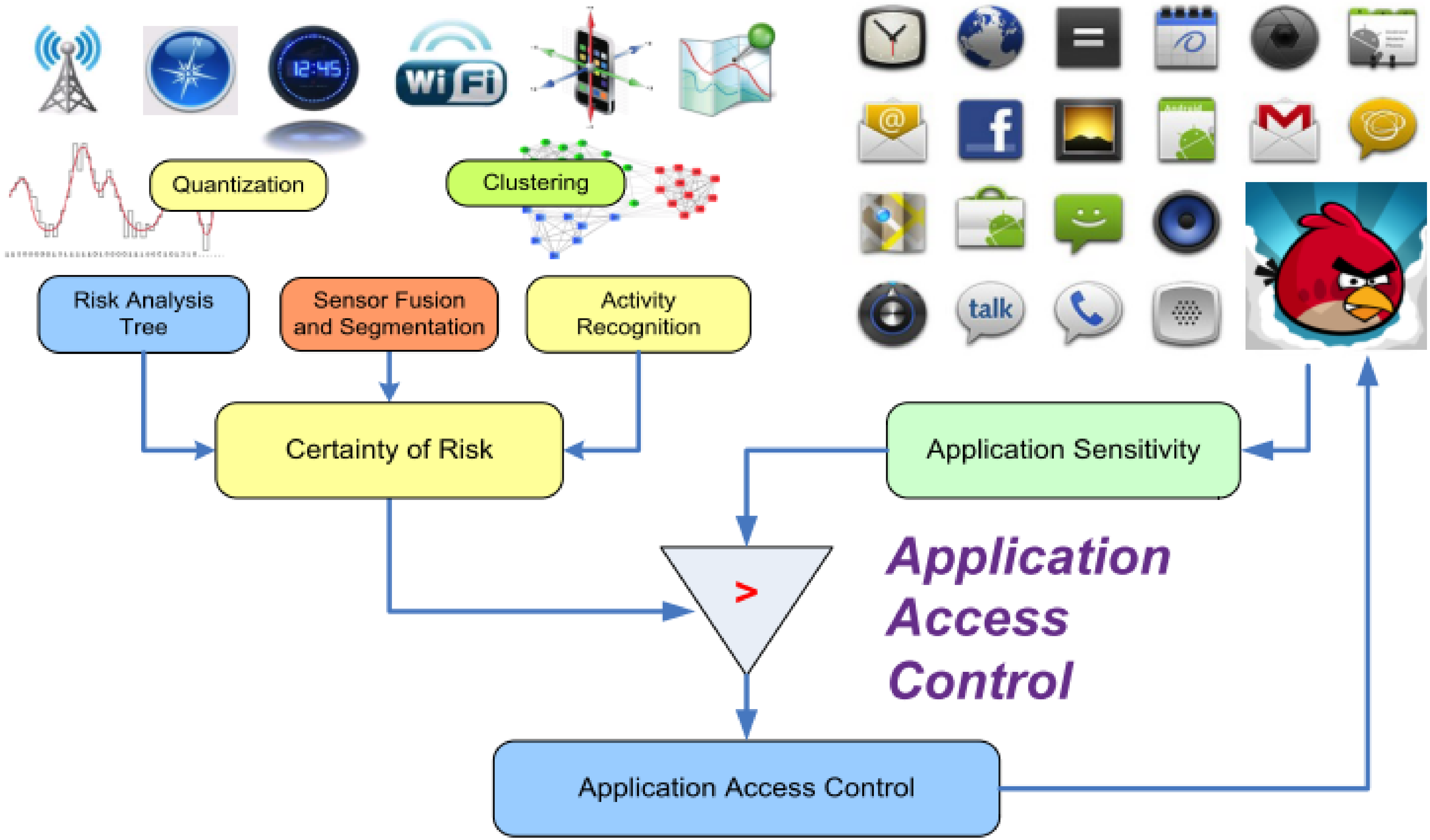
Anomaly Detection Threshold





Sensing
Preprocessing
Feature Construction
Behavior Text Generation
Threshold
Anomaly Detection
Classification
Prediction
Regression
Applications
Modeling
N-gram Model
Skipped N-gram Model
Helix
Helix Tree
DT
SVM
Ground Truth
Evaluation
ROC
Accuracy
Precision Recall
UX

SenSec – App Access Control



SenSec Certainty Scores

- SenSec uses a variety of sensor data to build user behavior models
  - Unsupervised activity segmentation and behavioral text modeling
  - Anomaly detection using risk analysis (decision) tree
  - Computes certainty score as an estimate of risk (online)

- Application access control module will decide:
  - Allow access, deny access, or ask for further authentication

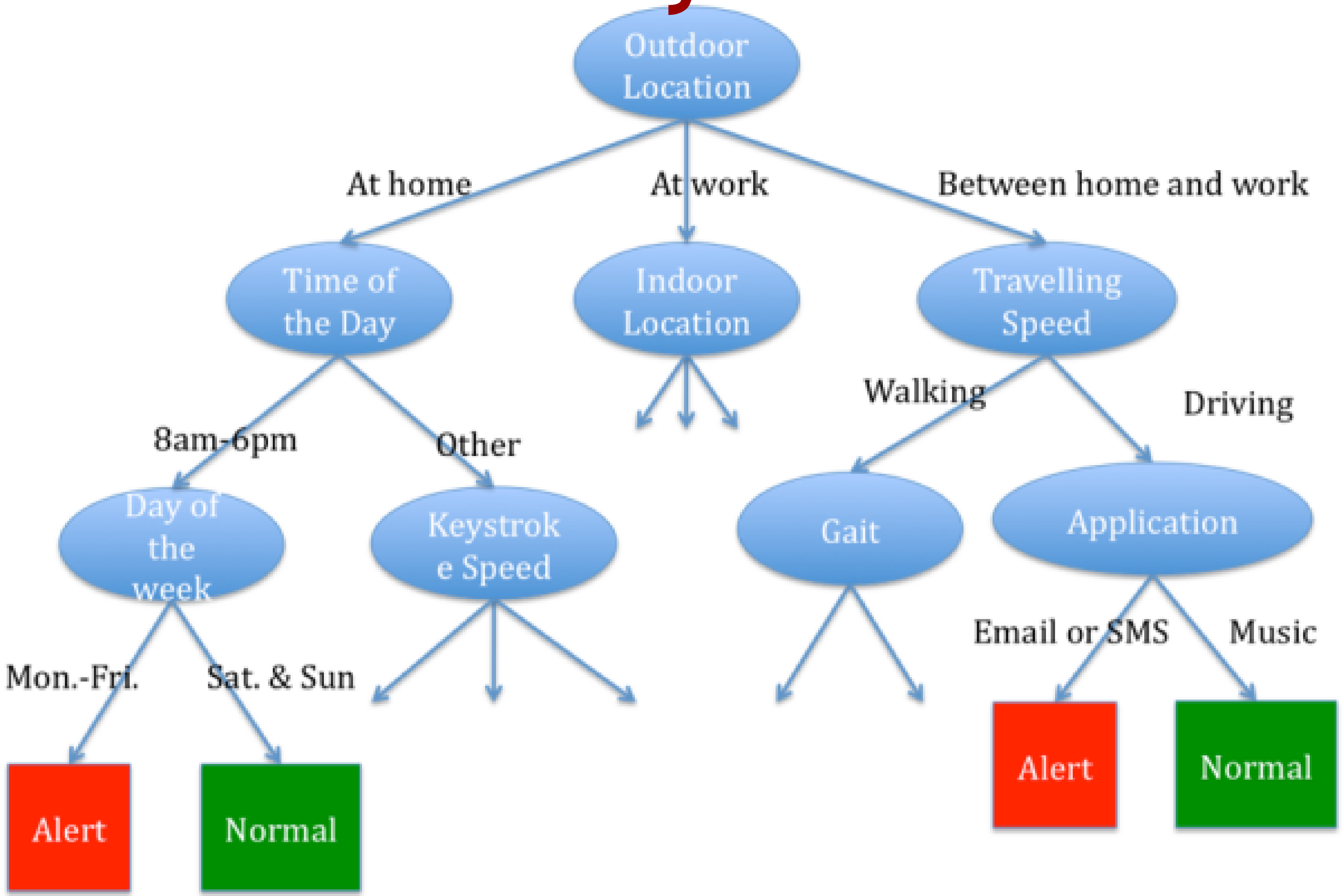
Feature Selection

- Accelerometer
  - Used to summarize acceleration stream
  - Calculated separately for each dimension
  - Meta features: total time, window size

Feature	Description	D/M
RMS	The Root-Mean-Square value	D
RMSE	The Root-Mean-Square error	D
Min	The minimum value	D
Max	The maximum value	D
AvgDeltas	The average sample-by-sample change	D
NumMax	The number of local peaks	D
NumMin	The number of local crests	D
TTP	The average time from a sample to a peak	D
TTC	The average time from a sample to a crest	D
RCR	The RMS cross rate	D
SMA	The Signal Magnitude Area	D
Total Time	The Total Time of the window	M
Window Size	The number of samples in the window	M

- GPS: location string from Google Map API and mobility path
- WiFi: SSIDs, RSSIs and path
- Applications: Bitmap of well-known applications
- Application Traffic Pattern: TCP/UDP traffic pattern vectors: [remote host, port, rate]

Risk Analysis Tree

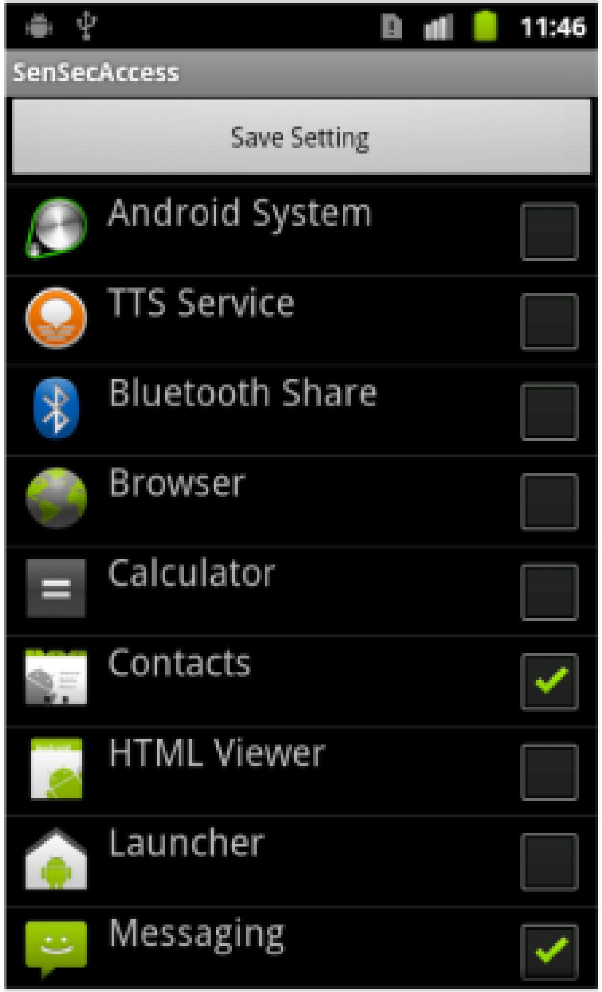


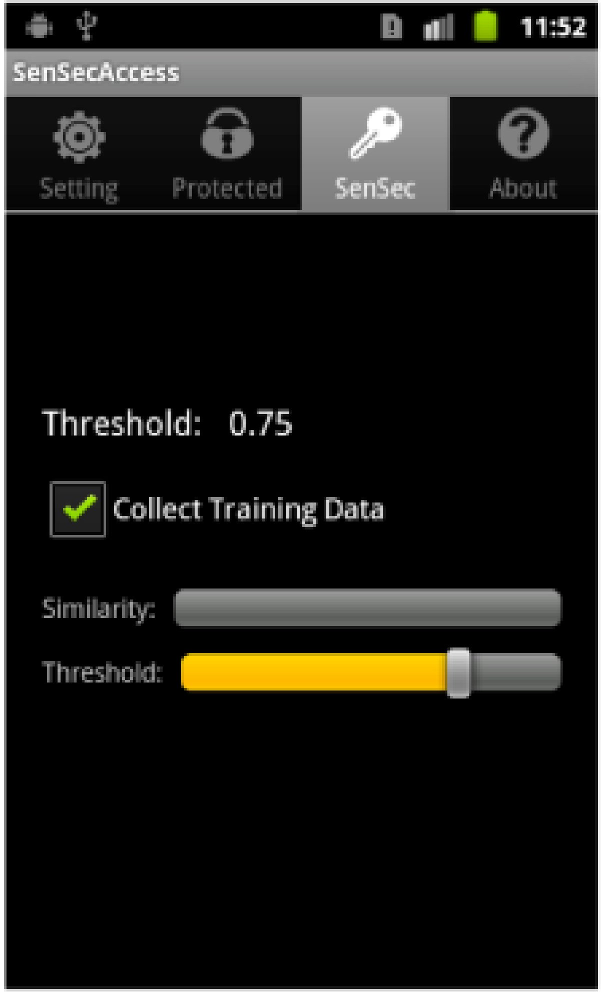
Validation / Evaluation

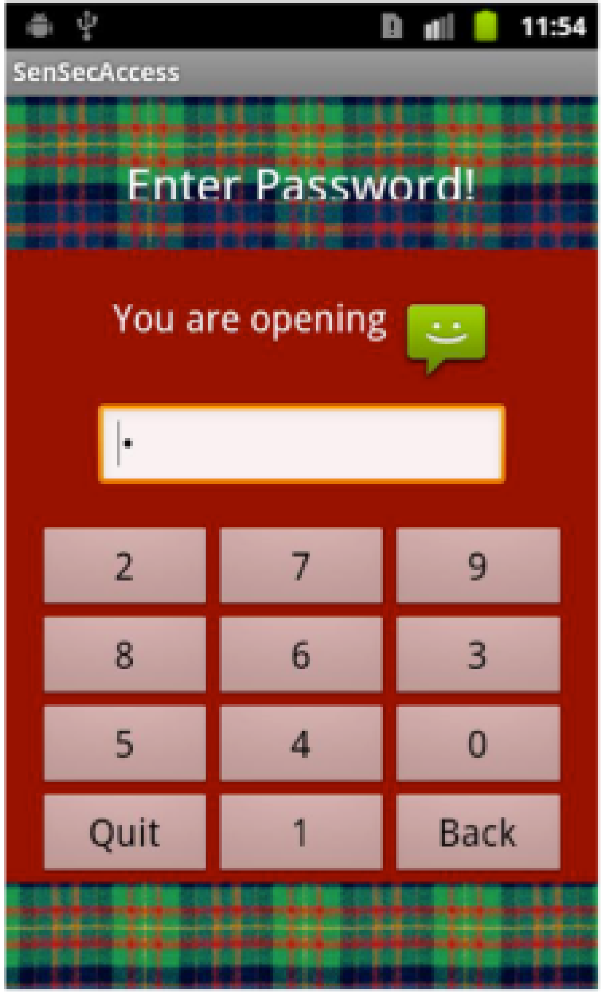
- Offline data collection (for training and testing)
  - Pick up the device from a desk
  - Unlock the device using the right slide pattern
  - Invoke Email app from the "Home Screen”
  - Some typing on the soft keyboard
  - Lock the device by pressing the "Power" button
  - Put the device back on the desk

Classification Target	No. of Classes	Accuracy
Gender	2	0.81
Age Group	3	0.79
Occupation	4	0.76
User ID	20	0.75

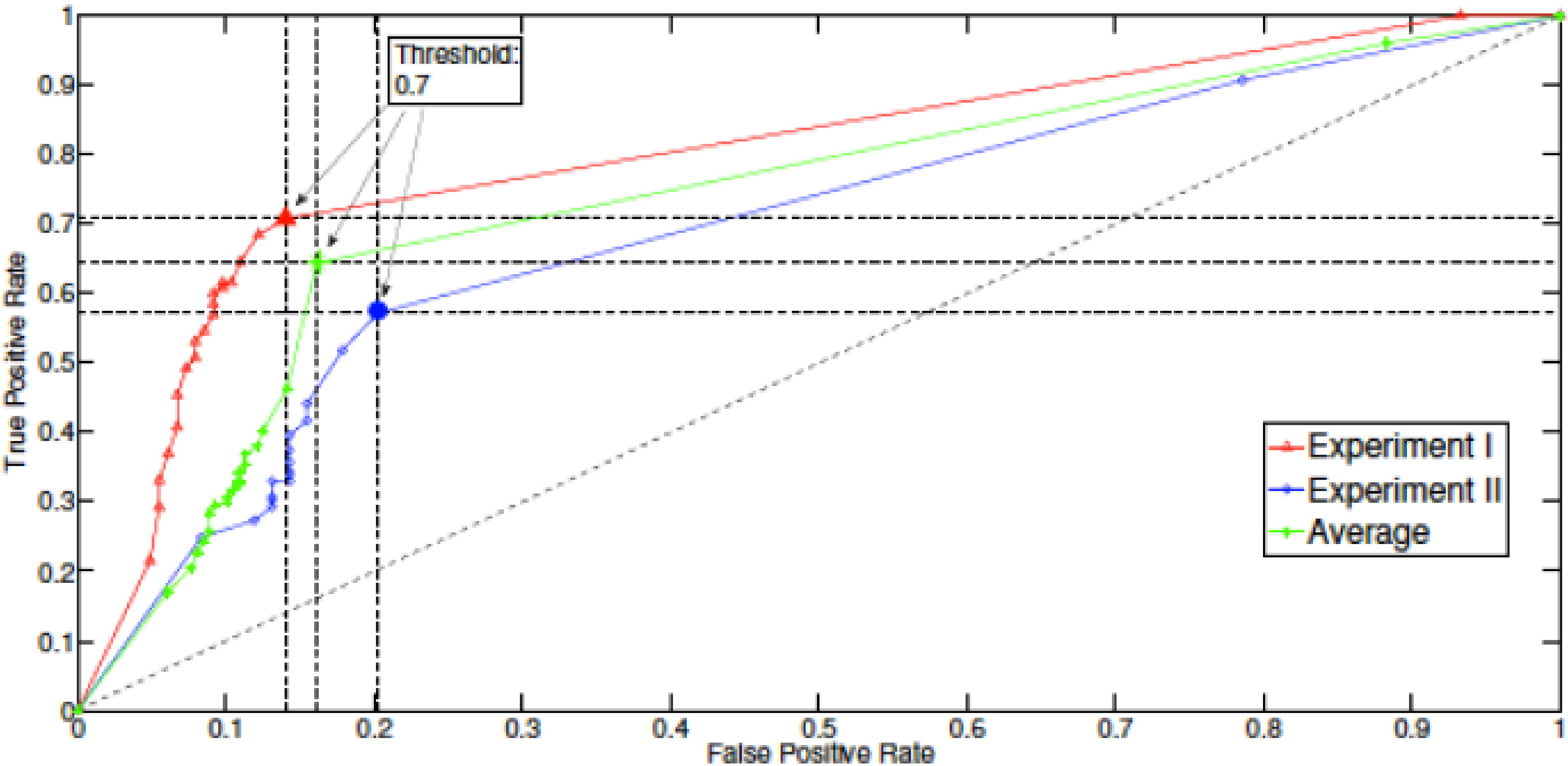
SenSec App v1.0







Experiment: Detecting Theft

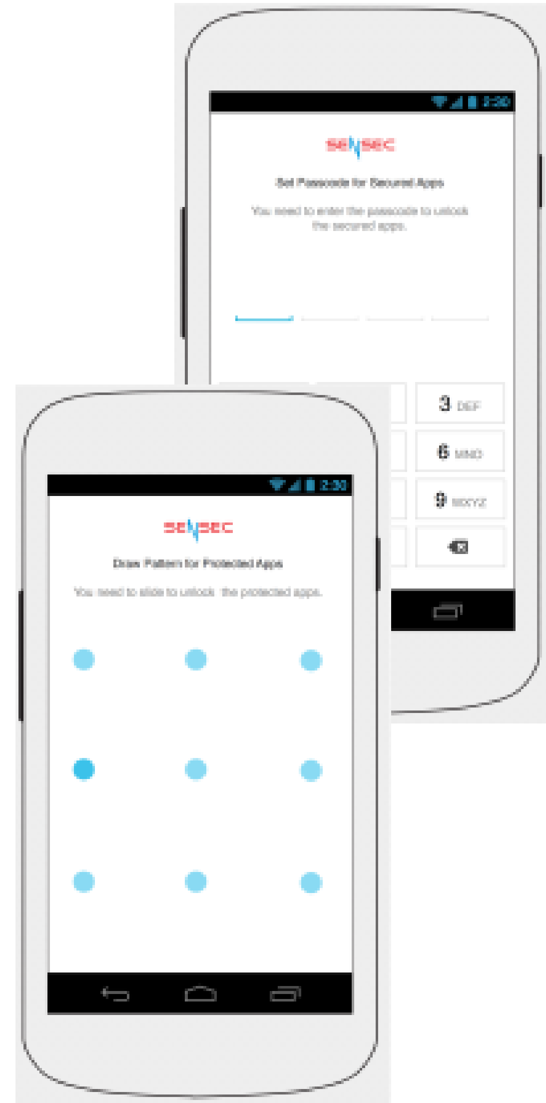


- 71.3% true-positive rate, 13.1% false positive

Practical Issues w/ v1.0

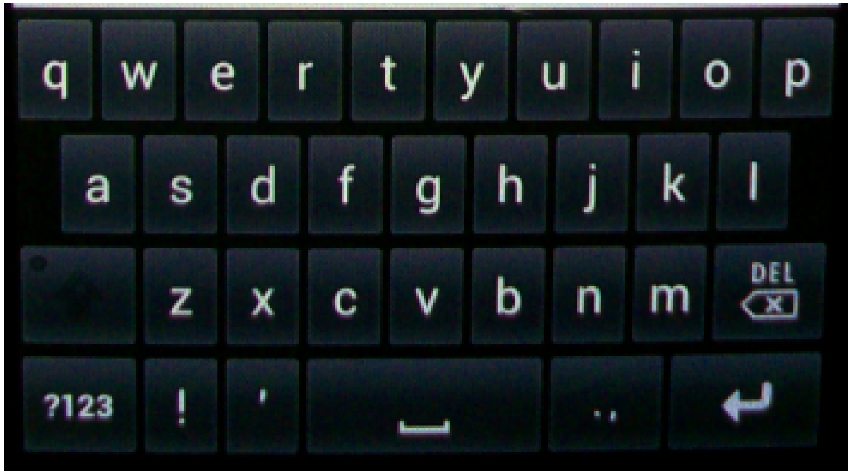
- Alpha test Jun 2012, Google Play Store release Oct 2012
  - False Positive: 13% FPR still annoying users sometimes
- Possible Solutions
  - Use adaptive modeling
  - Adding the trace data shortly before a false positive to the training data and update the model
  - Change passcode validation to sliding pattern
  - A false positive will grant a “free ride” for a configurable duration
  - Assumption: just authenticated user should control the device for a given period of time
  - “Free Ride” period will end immediately if abrupt context change is detected.

Evolution to SenSec v2.0



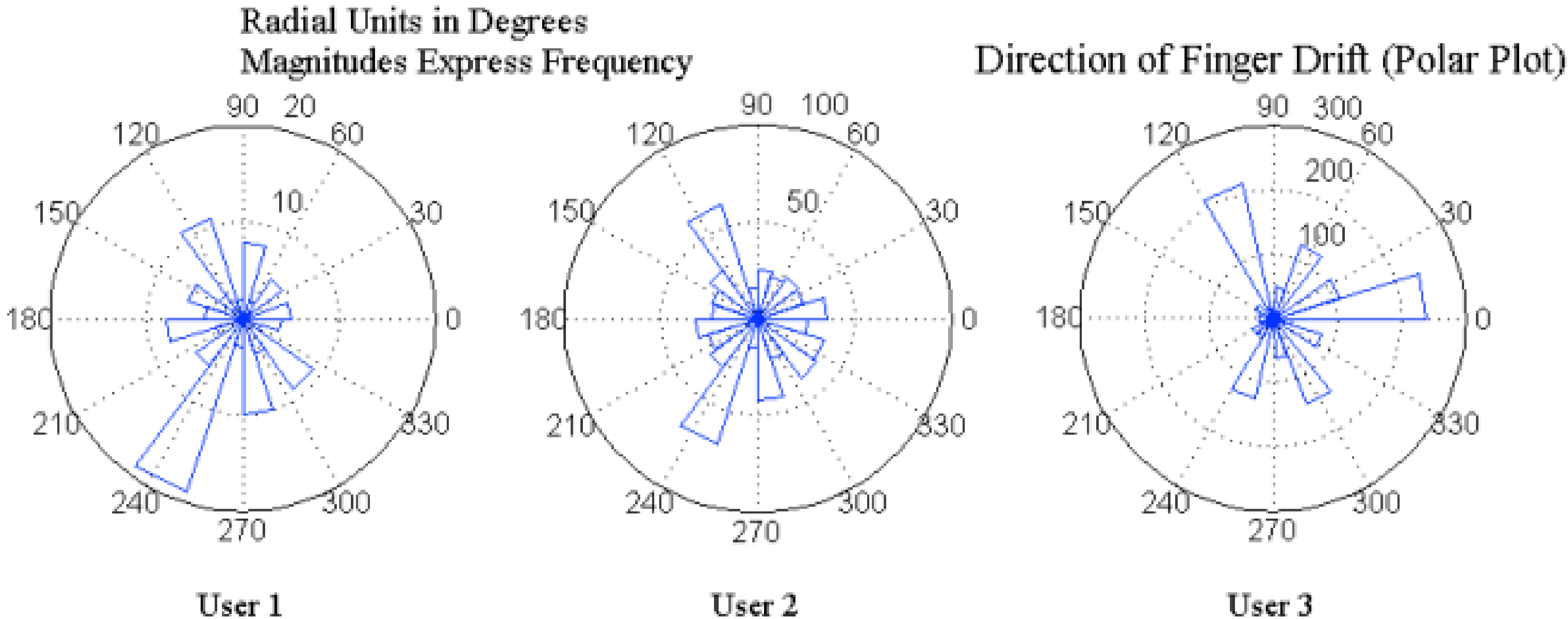
+ Soft Keyboard Interaction

- Hypothesis: the micro-behavior a user interacts with the soft keyboard reflects his/her cognitive and physical characteristics
  - Cognitive fingerprints: typing rhythms, correction rate, delay between keys, duration at each key…
  - Physical characteristics: area of pressure, amount of pressure, position of contact, shift …

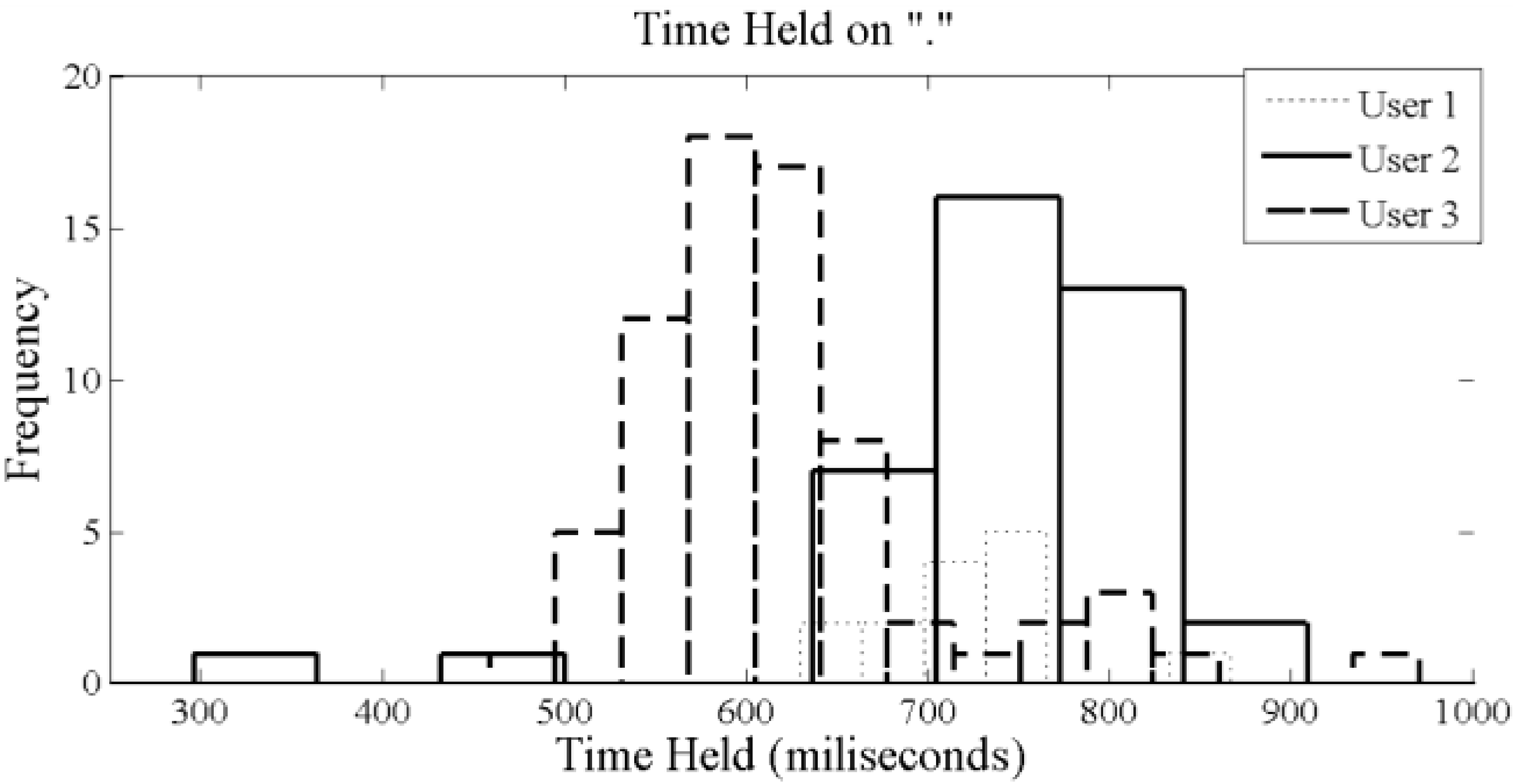


Drift

- When pressing a key, the lifting-up position drifts away from the touch-down position



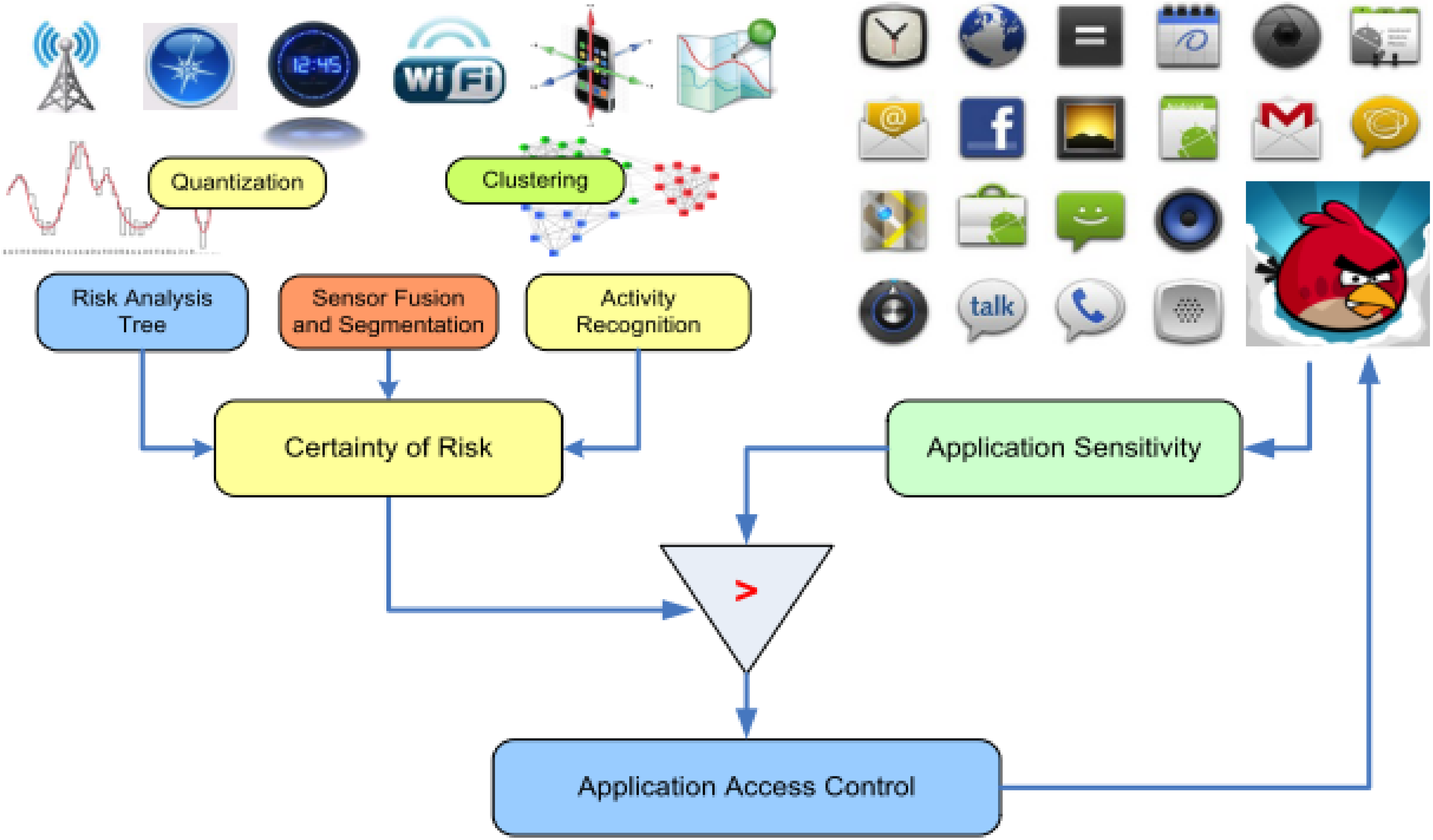
Key Holding Time



Passive Auth via SoftKey

- Discriminative model can identify a user at 99% accuracy with just one keypress:
  - When all users' behavior is known
  - Models trained over 4000 keys each from 4 users

- Generative model to detect unauthorized use from an unknown user
  - Only the authorized user's behavior is known
  - After 15 key presses: detection rate is 86% (14% false negative) and only 2.2% false positive



- Experiments to discover anomaly usage with ~80% accuracy with only days of training data

Some Open Questions

- Extended data set for feature construction
  - TCP/UDP traffic, sound, ambient light, battery, etc.
- Data and Modeling
  - Gain more insights into the data, features and factorized relationships among various sensors
  - Try other classification methods and compare results
- Enhanced security of SenSec components
  - Integration with Android security framework and other applications
- Privacy as expectation (Liu et al., 2012)
  - Data management, usage, sharing, trust, etc.
- Energy efficiency

Oct 29:

Mobile Malware

Nov 3:

NO CLASS