

SESSION ID: BR-W04

## You are what you click: Using Decoys to Identify Mobile Device Attackers

**Dr. Salvatore J. Stolfo**

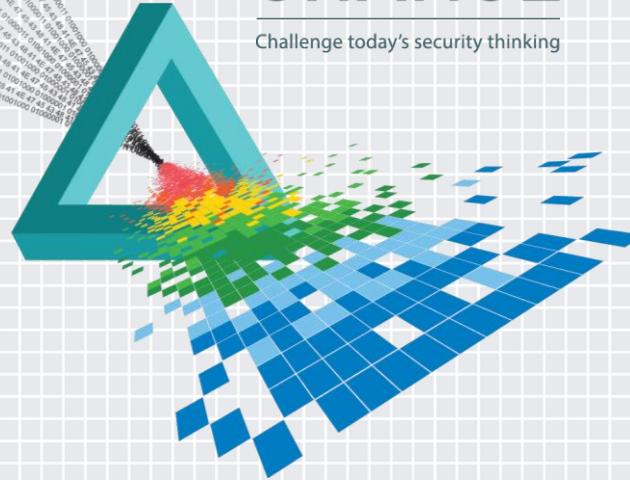
Professor  
Columbia University  
Intrusion Detection Lab  
New York, NY

**Joel Peterson**

Researcher, Columbia University &  
Software Systems Researcher  
Allure Security Technology  
New York, NY

# CHANGE

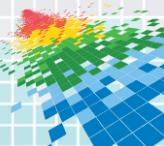
Challenge today's security thinking



# The old ways don't work....



**It's not what you know (twice), it's what you do....**



# Age of Collaboration



Dropbox



# “Anywhereization”



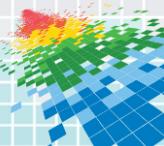
# Rise in BYOD

- ◆ **2 billion** smartphone users projected for 2015
- ◆ **61%** of employees use smartphones for work



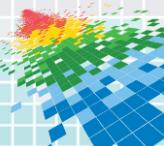
# Enterprise is borderless...and vulnerable

- ◆ Devices carry sensitive corporate data:
  - ◆ 59% of employees using BYOD haven't told their employers
- ◆ Lost or stolen:
  - ◆ 113 smart phones lost every minute
  - ◆ 1 laptop is stolen every 53 seconds
- ◆ Mobile security fails:
  - ◆ 34% of consumers fail to activate security mechanisms on their mobile devices



# Security should be designed for the people who use it

- ◆ Easy to understand controls
- ◆ Transparent to the user
- ◆ Seamless and continuous authentication



# Are you you?

- ◆ With patented machine-learning technology, RUU learns how you use your device and creates a personalized behavioral profile that continuously and seamlessly authenticates.
- ◆ If unusual behavior is detected, it's prompted to ask, "Are you you?"



# What are decoys?

## Beacons

- ◆ Enticing, believable but bogus data, documents, files, and other types of fake but realistic media
- ◆ Touch a decoy and send a beacon alert signal



Here's how it works



# What's the difference?

w9.pdf - Adobe Reader

Please fill out the following form. You can save data typed into this form.

**Request for Taxpayer Identification Number and Certification**

Form W-9  
(Rev. December 2011) Department of the Treasury Internal Revenue Service

Name (as shown on your income tax return)

Business name/deregistered entity name, if different from above

Check appropriate box for federal tax classification:  
 Individual/sole proprietor    C corporation    S corporation    Partnership    Trust/estate  
 Limited liability company. Enter the tax classification (C=C corporation, S=S corporation, P=partnership) ►  Exempt organization  
 Other (see instructions) ► Address (number, street, apt. or suite no.) Requestor's name and address (optional)  
City, state, and ZIP code  
List account number(s) here (optional)

**Part I Taxpayer Identification Number (TIN)**  
Enter your TIN in the appropriate box. The TIN provided must match the name given on the "Name" line to avoid backup withholding. For individuals, this is your social security number (SSN). However, for a resident alien, sole proprietor, or disregarded entity, see Part I instructions page 3. For other entities, enter your employer identification number (EIN). If you don't have a number, you can get a TIN on page 3.  
Note: If the account is in more than one name, see the chart on page 4 for guidelines on whose number to enter.

**Part II Certification**  
Under penalties of perjury, I certify that:  
1. The number shown on this form is my correct taxpayer identification number (or I am waiting for a number to be issued to me), and  
2. I am not subject to backup withholding because: (i) I am exempt from backup withholding, or (ii) I have not been notified by the Internal Revenue Service (IRS) that I am subject to backup withholding as a result of a failure to report all interest or dividends, or (iii) the IRS has notified me that I am no longer subject to backup withholding, and  
3. I am a U.S. citizen or other U.S. person (see below).

**Comments:** You may cross out item 3 above if you have been notified by the IRS that you are currently subject to backup withholding because you have failed to report all interest and dividends on your tax return. For real estate transactions, item 3 does not apply. For mortgage interest paid, acquisition or abandonment of secured property, cancellation or debt, contributions to an individual retirement arrangement (IRA), and generally, payments other than interest and dividends, you are not required to sign the certification, but you must provide your correct TIN. See the instructions on page 4.

**Sign Here** Signature of requester\* Date ►

**General Instructions**

Note: If a requester gives you a form other than Form W-9 to request your TIN, you must use the requester's form if it is substantially similar

w9.pdf - Adobe Reader

Please fill out the following form. You can save data typed into this form.

**Request for Taxpayer Identification Number and Certification**

Form W-9  
(Rev. December 2011) Department of the Treasury Internal Revenue Service

Name (as shown on your income tax return)

Business name/deregistered entity name, if different from above

Check appropriate box for federal tax classification:  
 Individual/sole proprietor    C Corporation    S Corporation    Partnership    Trust/estate  
 Limited liability company. Enter the tax classification (C=C corporation, S=S corporation, P=partnership) ►  Exempt organization  
 Other (see instructions) ► Address (number, street, apt. or suite no.) Requestor's name and address (optional)  
City, state, and ZIP code  
List account number(s) here (optional)

**Part I Taxpayer Identification Number (TIN)**  
Enter your TIN in the appropriate box. The TIN provided must match the name given on the "Name" line to avoid backup withholding. For individuals, this is your social security number (SSN). However, for a resident alien, sole proprietor, or disregarded entity, see Part I instructions page 3. For other entities, enter your employer identification number (EIN). If you don't have a number, you can get a TIN on page 3.  
Note: If the account is in more than one name, see the chart on page 4 for guidelines on whose number to enter.

**Part II Certification**  
Under penalties of perjury, I certify that:  
1. The number shown on this form is my correct taxpayer identification number (or I am waiting for a number to be issued to me), and  
2. I am not subject to backup withholding because: (i) I am exempt from backup withholding, or (ii) I have not been notified by the Internal Revenue Service (IRS) that I am subject to backup withholding as a result of a failure to report all interest or dividends, or (iii) the IRS has notified me that I am no longer subject to backup withholding, and  
3. I am a U.S. citizen or other U.S. person (see below).

**Comments:** You may cross out item 3 above if you have been notified by the IRS that you are currently subject to backup withholding because you have failed to report all interest and dividends on your tax return. For real estate transactions, item 3 does not apply. For mortgage interest paid, acquisition or abandonment of secured property, cancellation or debt, contributions to an individual retirement arrangement (IRA), and generally, payments other than interest and dividends, you are not required to sign the certification, but you must provide your correct TIN. See the instructions on page 4.

**Sign Here** Signature of requester\* Date ►

**General Instructions**

Note: If a requester gives you a form other than Form W-9 to request your TIN, you must use the requester's form if it is substantially similar

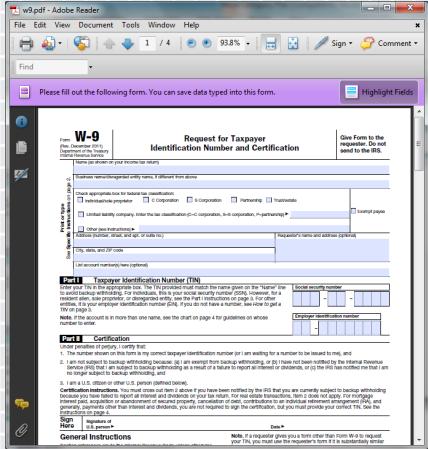
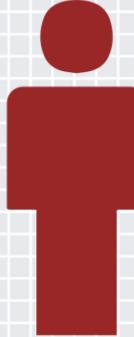
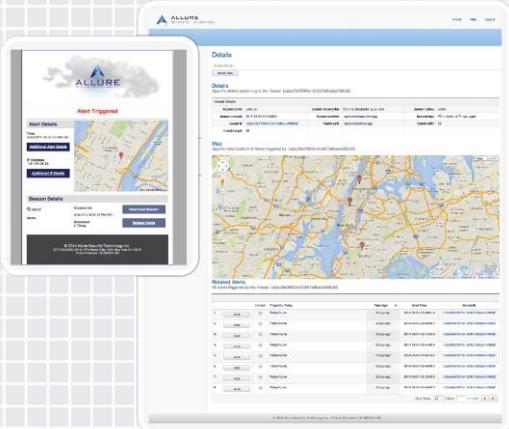


# Document is *beaconized*

Beaconized Document

Data Loss

Data Loss Alert


# Touch a decoy file, a data loss alert is emailed

**Alert Triggered**

**Alert Details**

**Time**  
1/16/2015 06:10:50 PM EST

**Additional Alert Details**

**IP Address**  
xxx.xxx.xxx.xxx

**Additional IP Details**

**Beacon Details**

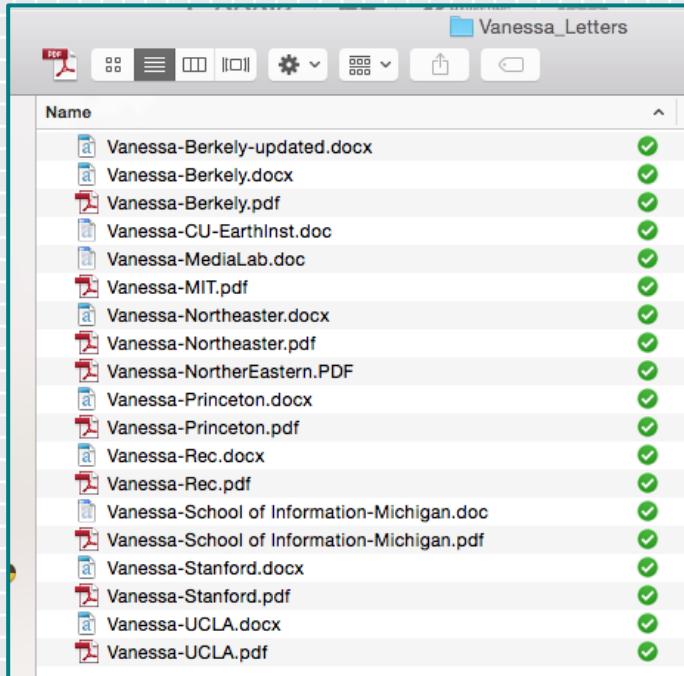
 word ebay theme      Created On  
9/12/2013 01:27:58 PM EDT      [Download Beacon](#)

Name unavailable      Accessed  
unavailable      1 Times      [Related Alerts](#)



Map data ©2015 Google

# Enticing decoy files in the cloud, too!

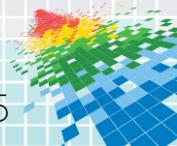


One is real, the others  
aren't – Can YOU tell?

# The Hypothesis

- ◆ We all search uniquely on our own machines....that is a user biometric captured by a behavior model computed by a machine learning algorithm.
- ◆ Decoys are a powerful tool to detect intruders who do not know the real content of a target victim's file system.

The two together detect masqueraders and provide accurate active and continuous authentication.



# (Sidebar: Decoys can also be used to detect...)

- ◆ Hackers who hijack sessions from other legitimate users
- ◆ Embedded APT actors whose malware behaves abnormally
- ◆ But, let's return our attention to Active and Continuous Authentication of users...



# RSA® Conference 2015

San Francisco | April 20-24 | Moscone Center

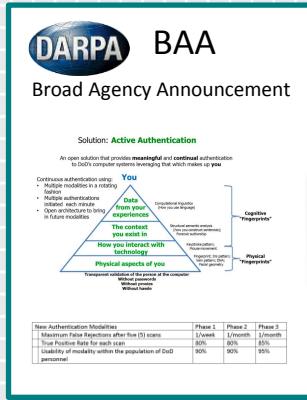


# Phase 1 DARPA Active Authentication - Desktop



#RSAC

# DARPA Phase 1 Goals

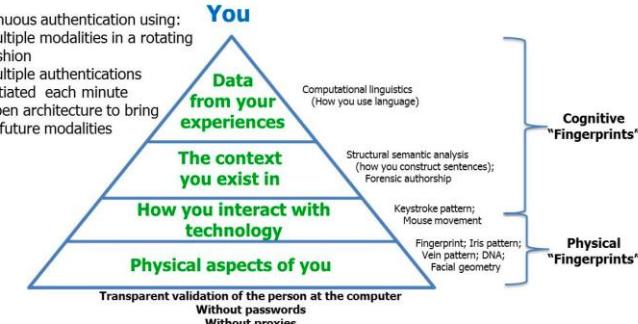


## Solution: Active Authentication

An open solution that provides **meaningful** and **continual** authentication to DoD's computer systems leveraging that which makes up **you**

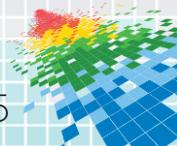
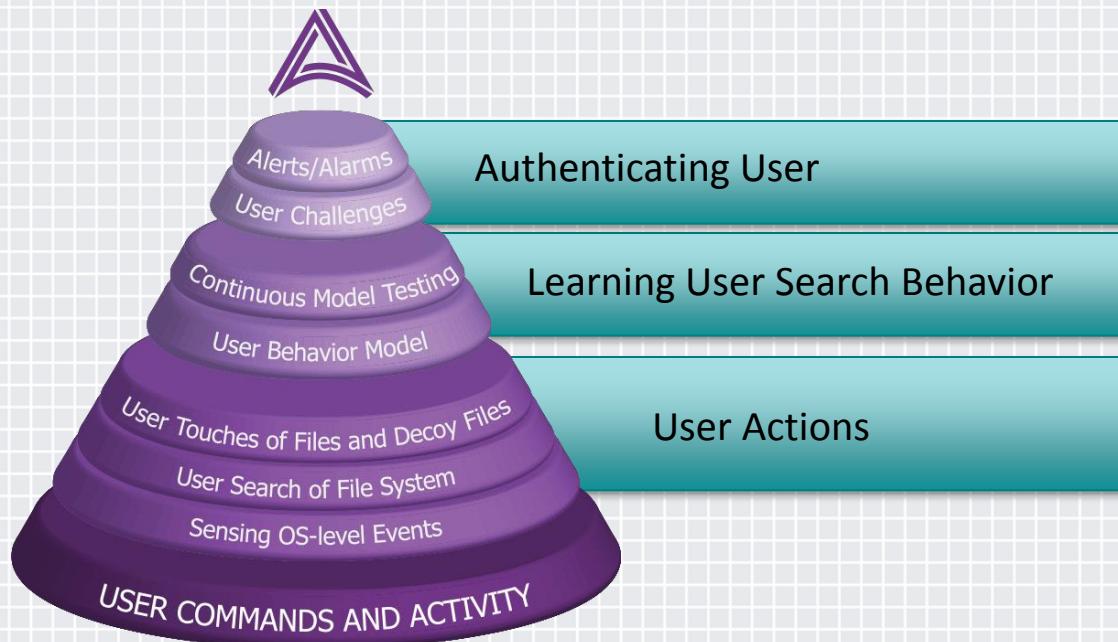
Continuous authentication using:

- Multiple modalities in a rotating fashion
- Multiple authentications initiated each minute
- Open architecture to bring in future modalities



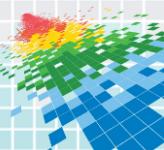
New Authentication Modalities	Phase 1	Phase 2	Phase 3
Maximum False Rejections after five (5) scans	1/week	1/month	1/month
True Positive Rate for each scan	80%	80%	85%
Usability of modality within the population of DoD personnel	90%	90%	95%

# RUU Baseline Architecture



# Initial baseline scientific user study of accuracy of modeling user behavior

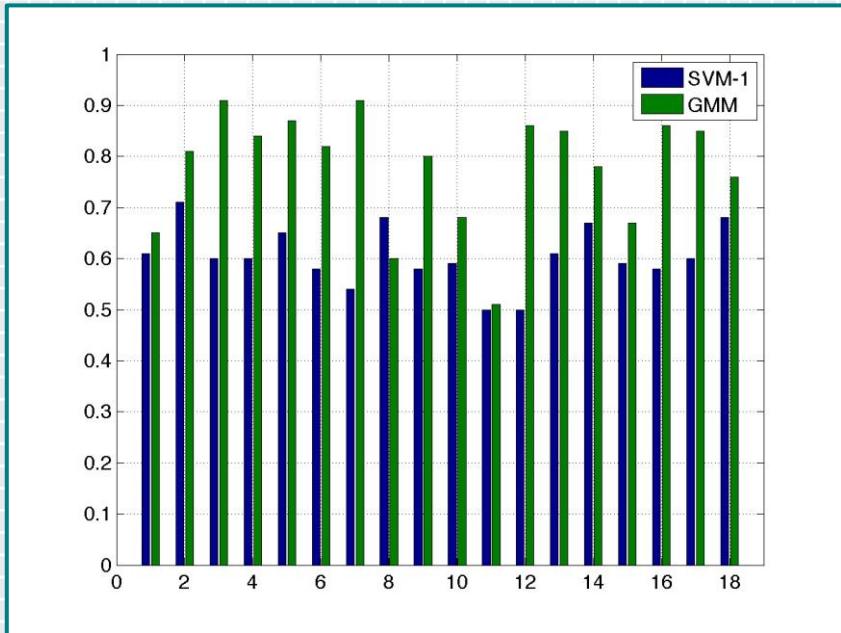
- ◆ Model baseline Volunteer Human Subject behavior; detect deviations from normal use.  
Generative model: inference, prediction, clustering, sampling, etc.
- ◆ Behavior biometrics: set of measurements on interactions between the Volunteer Human Subject and the system.
- ◆ Biometrics measurements based on OS events caused by Volunteer Human Subject action:
  - ◆ Process creation, deletion, manipulation.
  - ◆ File creation, deletion, renaming, etc.
  - ◆ Process window touches.
  - ◆ Registry key creation, manipulation, deletion.
- ◆ Four minute sliding window of measurement used.
- ◆ RUU1 dataset: 18 Volunteer Human Subjects at Columbia University, measured over the course of five weeks. Captured in 2011. RUU2 and RUU200 datasets delivered Sept 2013.



# Fisher Linear Discriminant Analysis

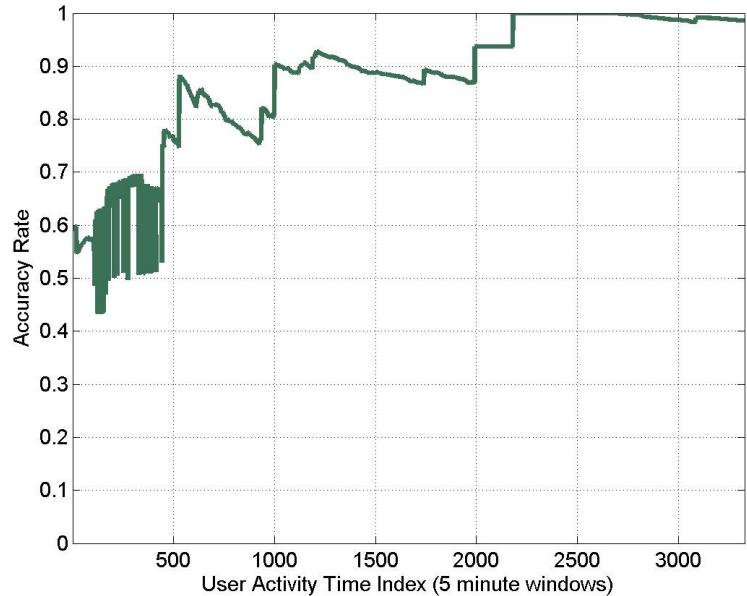
Feature	FLD Score
Number of unique processes	0.0359
Number of delete key actions	0.0018
Number of processes created	0.0015
Number of files touched	0.0013
Number of registry flush key actions	0.0012
Number of user touches	0.0011
Number of registry key queries	0.0011
Number of registry value queries	0.0010
Number of processes destroyed	0.0010
Number of open key actions	0.0010
Number of manual search actions	0.0009

# Accuracy Improvements ... choose wisely



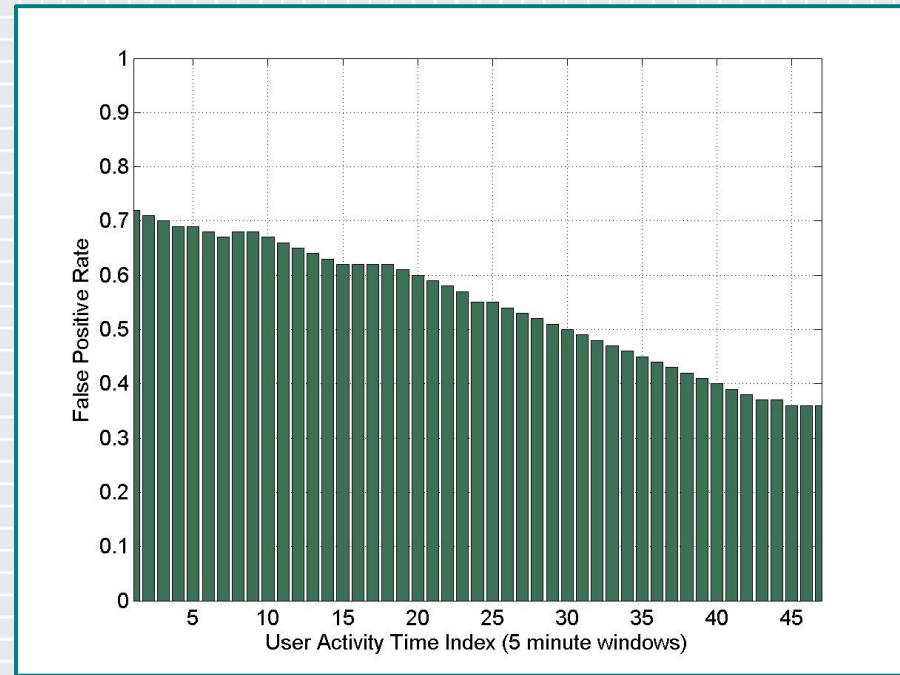
Accuracy over the initial RUU dataset. GMM model with Fisher features, improved accuracy and faster.

# True Positive Rate increases with training



Accuracy improves over time.  
As more data is observed, the  
accuracy of the user's model  
improves. And...

# False Positive Rate Decays, too...



Maintaining and improving model performance over time is an important goal.  
Continuous learning methods work well.

# RSA® Conference 2015

San Francisco | April 20-24 | Moscone Center

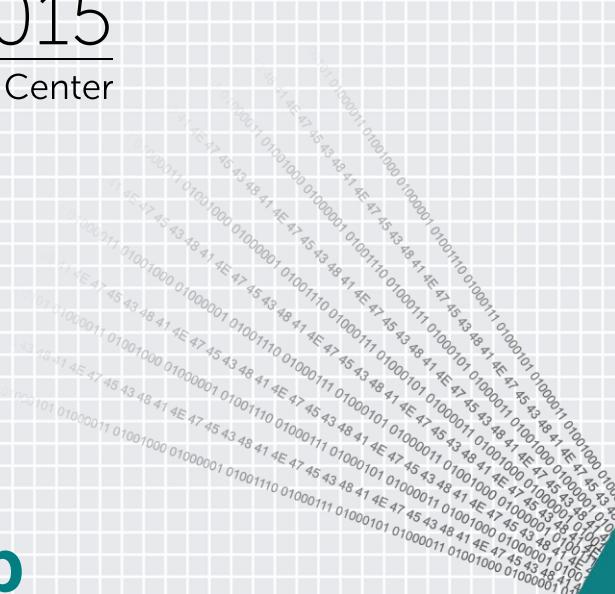


# Phase 2 – Desktop

## Sensor Improvements

# Automatic Decoy Placement

# Larger formal user study to detect masqueraders



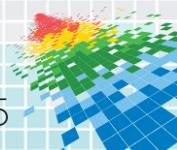
# RUU Host Sensors

## Phase 1

- ◆ Volunteer Human Subject data acquisition uploaded to server for analytics and performance bundled with Decoy Document Distribution
  - ◆ Identify most discriminating features
  - ◆ Measure decoy touch behavior

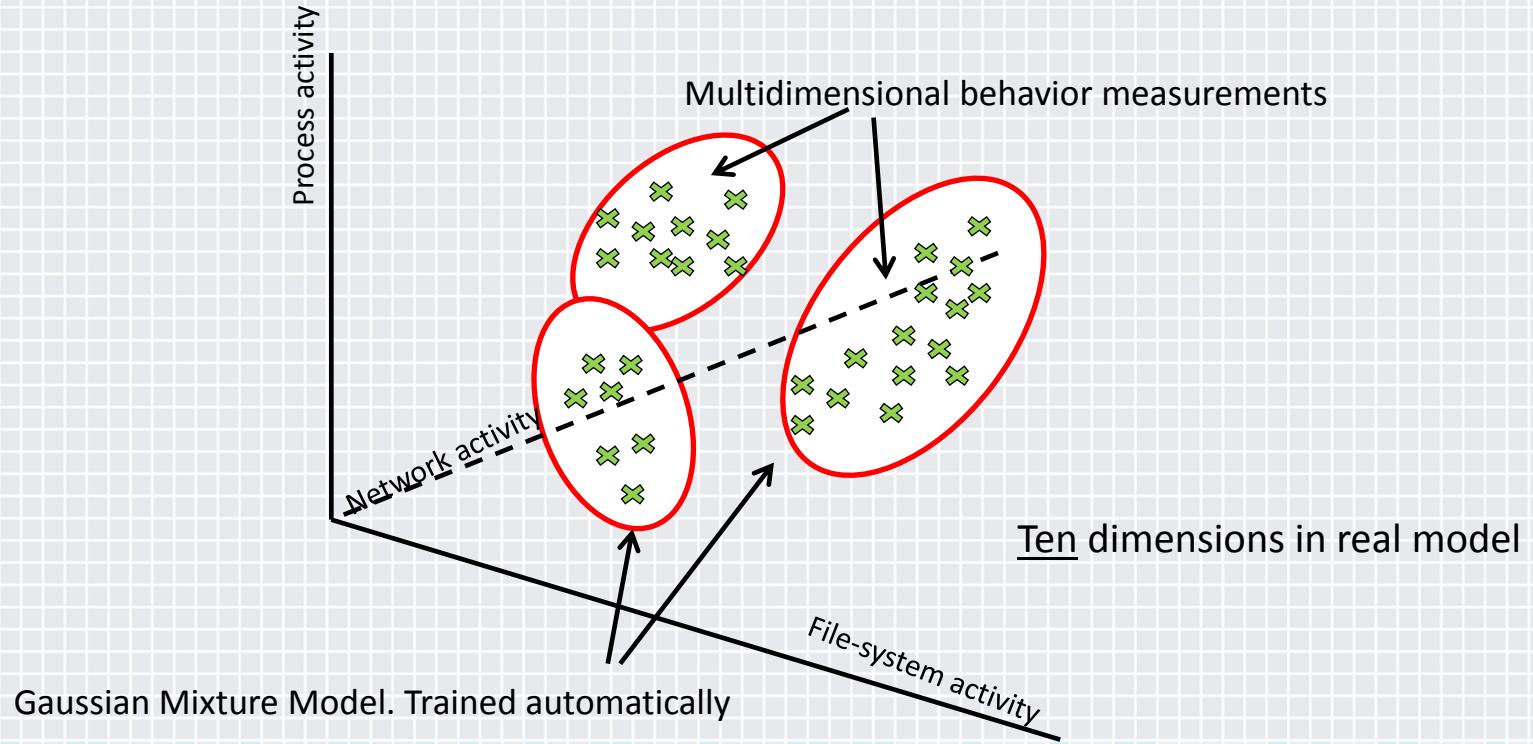
## Phase 2

- ◆ Volunteer Human Subject data acquisition on local host for automatic analysis and active authentication with mitigation strategy, also bundled with Decoy Document Distribution
  - ◆ Continuous learning
  - ◆ Automate Decoy Placement
  - ◆ Self-measurement of performance
  - ◆ Re-authentication strategies



# RUU Sensor Identity Engine – 10 Dimensions works well

## Learns User Search Behavior and OS-level Behavior Modeling

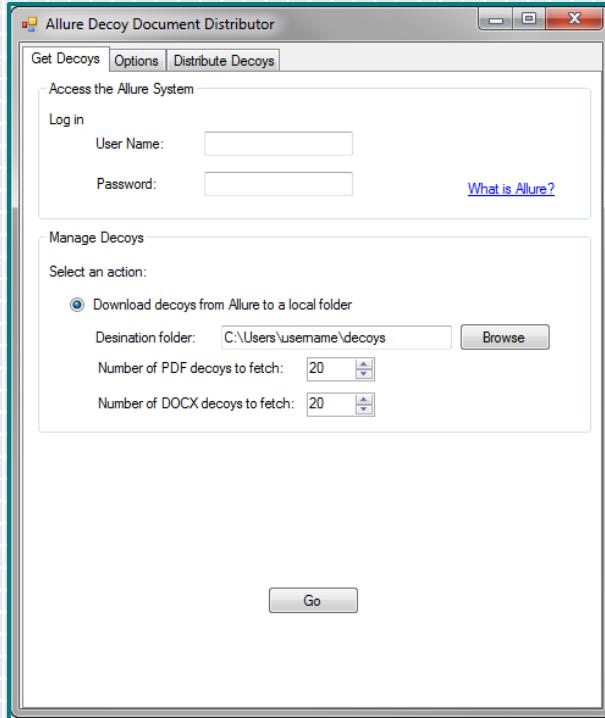


# RUU Decoy Distribution

- ◆ How to deploy decoys in scale throughout an organization?
- ◆ Manual placement
  - ◆ Tedium
  - ◆ Requires survey of Volunteer Human Subject habits
- ◆ Alternative approach
  - ◆ Distribute via an automated application
  - ◆ Decoy Document Distributor (DDT)



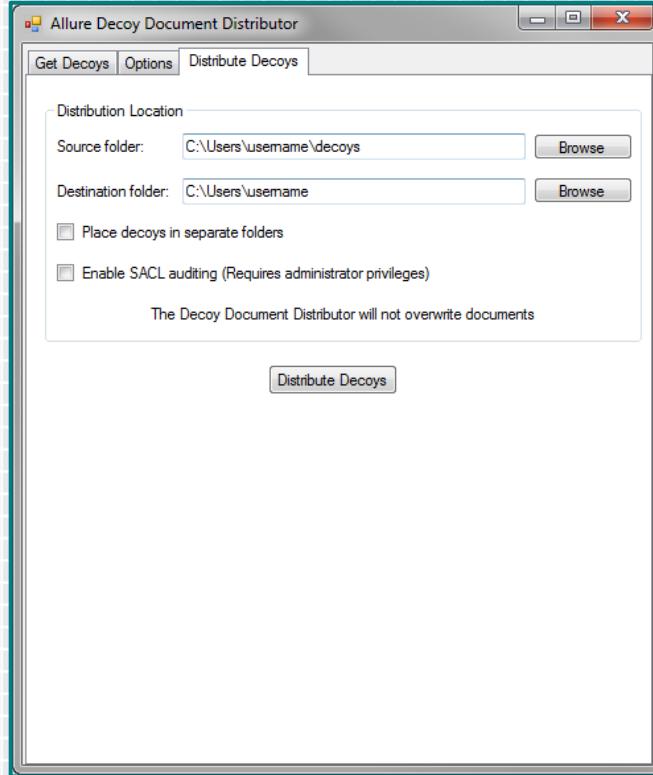
# Decoy Document Distributor (DDT)



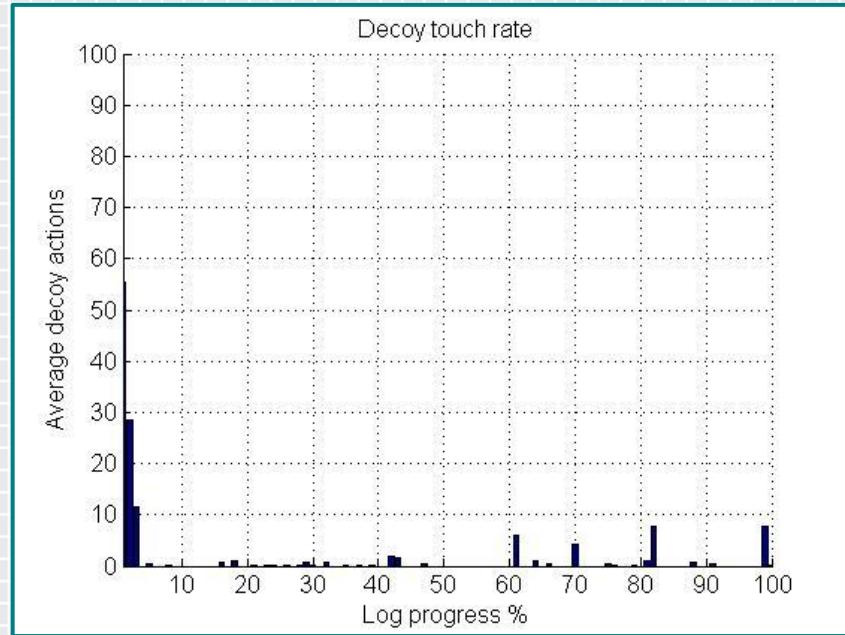
- ◆ Fetches decoys from server

# DDT Analyzes User's file system

- ◆ Automatic deployment of decoys to strategic file locations



# RUU Average Decoy Touch Rate of real user

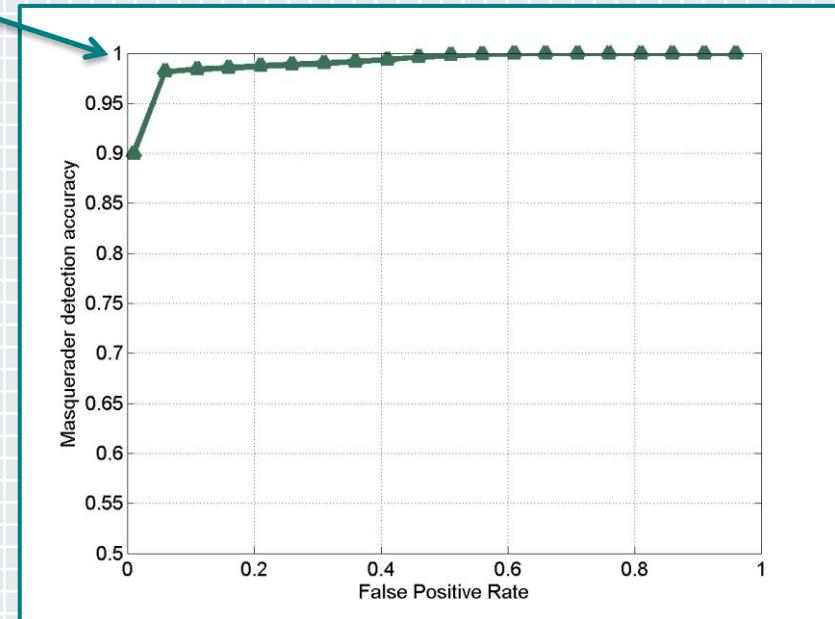


- ◆ Most decoy touches are caused by initial deployment.
- ◆ Curiosity decays rapidly!

# Masquerader Detection Accuracy with user models and decoys: Average ROC

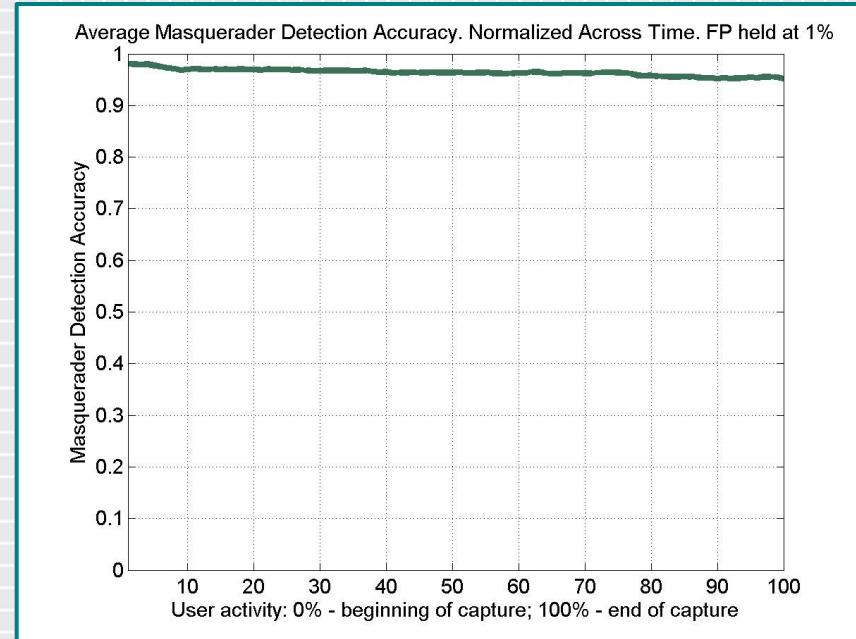
RUU models vs. *masquerader* data. Influential factors:  
masqueraders used “smash and grab.” (They didn’t play games.)

Optimal

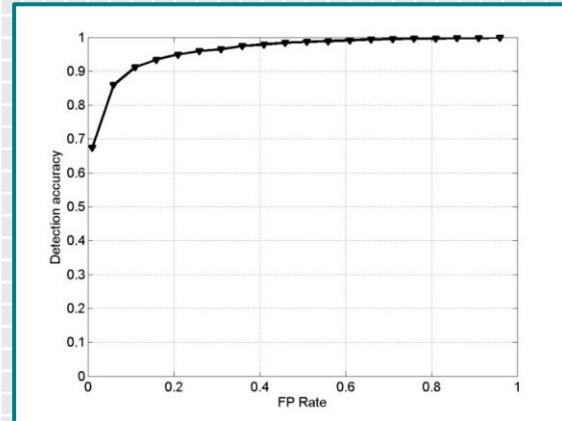


# Accuracy of detecting masqueraders over time is consistently high

Human subject activities are scaled as a percentage of capture progress (0%-100%). Average performance across all users.



# Accuracy translated to detection latency – users emit observables at different rates



## Experiment

- ◆ Overall Average Attacker Detection Across All Users
- ◆ 160 Users
- ◆ 1 week average capture period

Time until detection (TTD) given evaluation frequency for a 40-hour work week.

Frequency	Total Samples	FP Req.	Acc.	Evals	TTD
1m	2400	0.042%	49.55%	5	5m
2m	1200	0.083%	50.29%	5	10m
3m	800	0.125%	51.46%	5	15m
4m	600	0.167%	53.11%	4	16m
5m	480	0.208%	54.00%	4	20m

- ◆ Evaluation interval: 3 minutes
- ◆ Active authentication corresponds to Bernoulli trial: Probability that masquerader evades detection in 5 consecutive evaluations is less than 5%.
- ◆ Detection within 15 minutes with 95% confidence

## Experiment Results

- ◆ 95% detection accuracy at 1% false positive rate
- ◆ Constraint: 1 FP per 40 hour work week
  - ◆ Fifteen minutes until detection with 95% confidence

# Discussion – user model alone works, too

## Masquerader ID and number of decoy touches by masquerader

- ◆ Masqueraders had higher than normal volumes of activity; exhibited “smash and grab” behavior
- ◆ 10 decoys were distributed randomly on the test environment
- ◆ Nearly every masquerader touched several decoys, didn’t matter where they were placed
- ◆ Some touched no decoys, but were still detected

ID	1	2	3	4	5	6	7	9	10	12	13	14	15	16	18	19	20	21	22	23
#	7	0	4	3	5	8	10	5	7	3	2	3	5	4	6	5	14	0	3	10

Masquerader detection even without decoy touches!

# What about mitigation?

What do you do when you detect a masquerader?

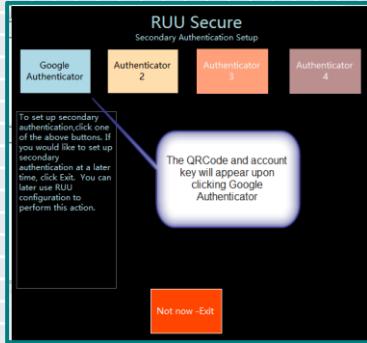
- ◆ De-authenticate and challenge the user to re-authenticate
  - ◆ This also provides an opportunity to update and improve the user model, ground truth is revealed
- ◆ Several possible re-authentication strategies, here's one...



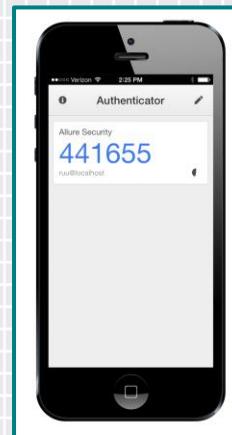
# RUU Secondary Authentication: When Desktop Locks

## Secondary Authentication

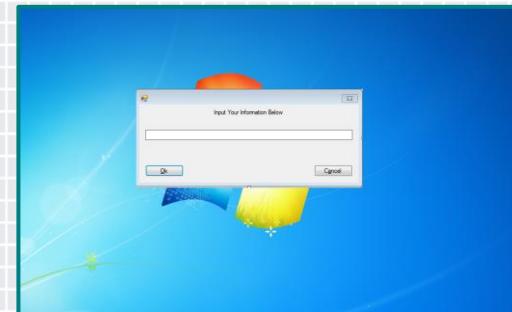
- ◆ Time-based One-time Password Algorithm for secondary authentication (RFC 6238)



When installing RUU the user is prompted to enable secondary authentication

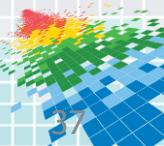
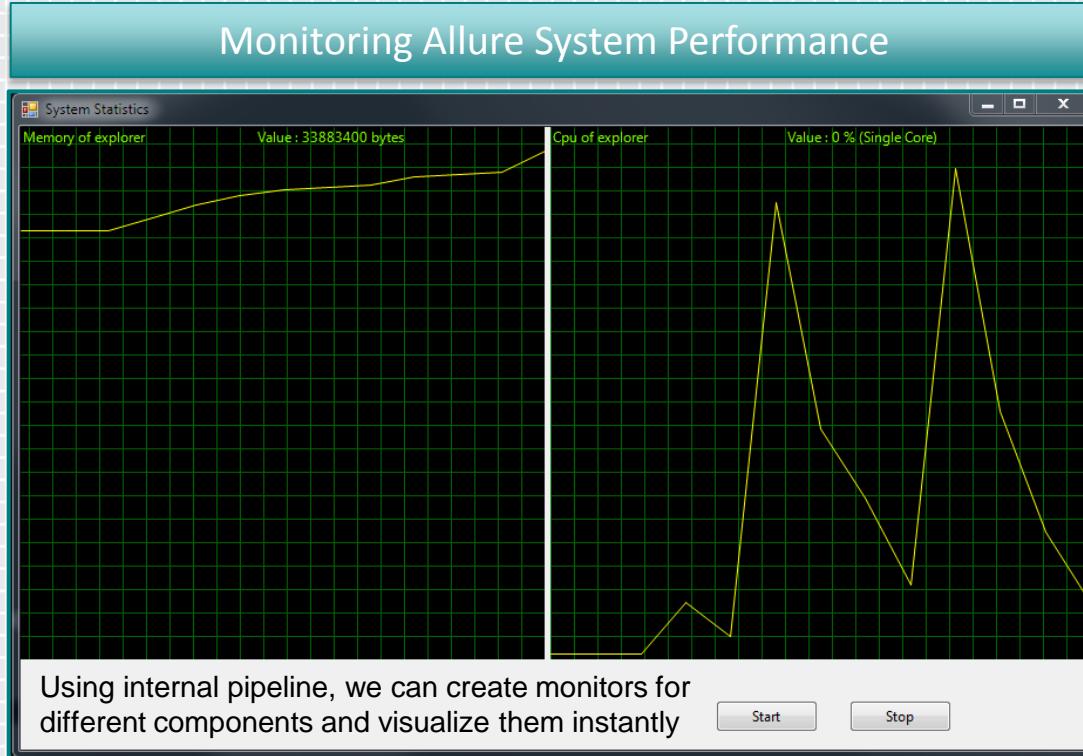


Google Authentication running on an iPhone as the authentication agent



When RUU locks and the user re-authenticates the secondary authentication is requested

# Monitoring and displaying RUU Sensor Performance: System Monitoring in scale for BYOB Management



# RSA® Conference 2015

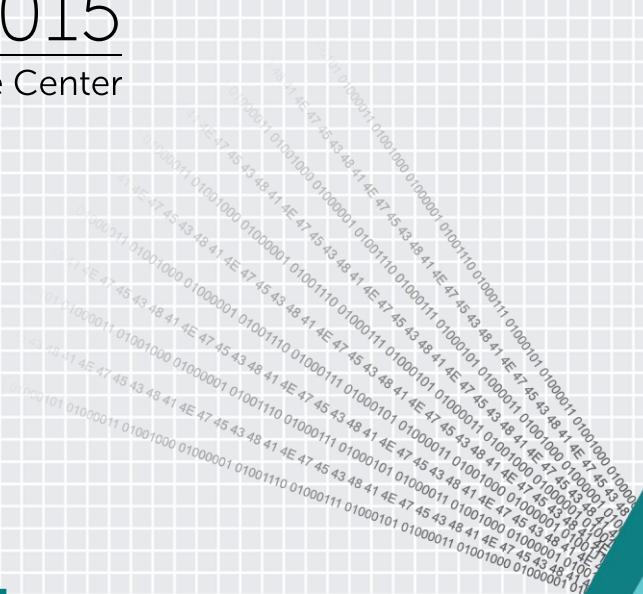
San Francisco | April 20-24 | Moscone Center



# mRUU – Mobile Phones

## Decoy Apps

## Decoy Clouds



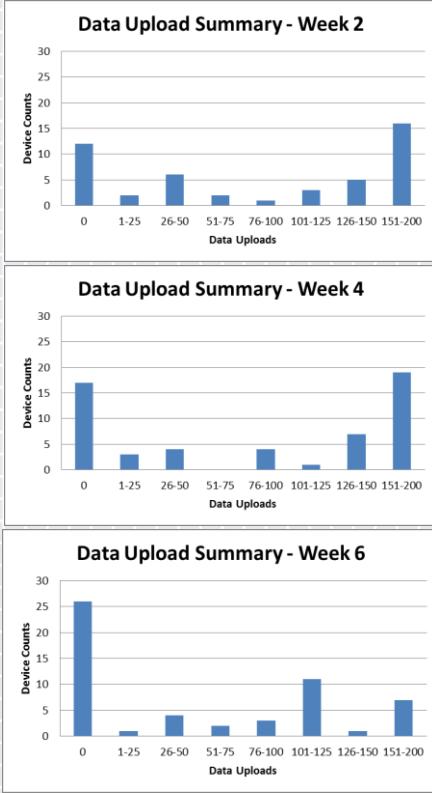
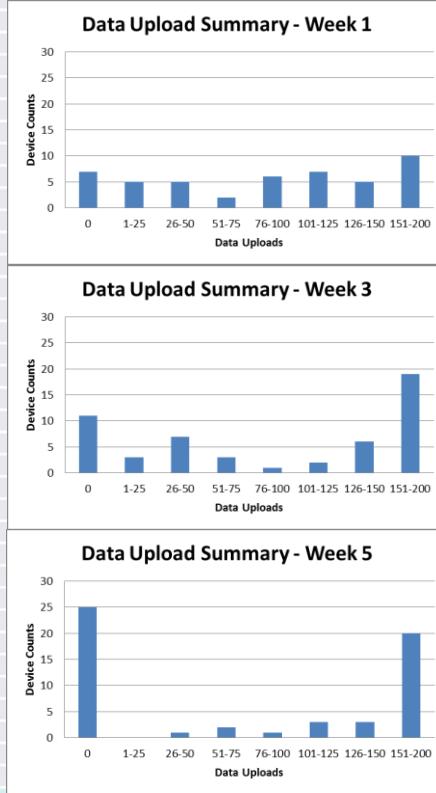
# mRUU Study

## IRB-Approved User Studies

- ◆ January 2014: Pilot study
  - ◆ Preliminary Activity Collector
  - ◆ Users gathered from Accenture and Columbia University
  - ◆ Used to inform modeling approach
- ◆ July-August 2014: Full scale user study with 53 Accenture users
  - ◆ Fully developed activity collector
    - ◆ More efficient
    - ◆ Collection of auxiliary activity data
  - ◆ Used for final Identity Engine design and accuracy analysis



# mRUU Study Results

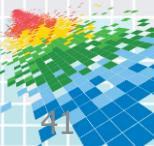
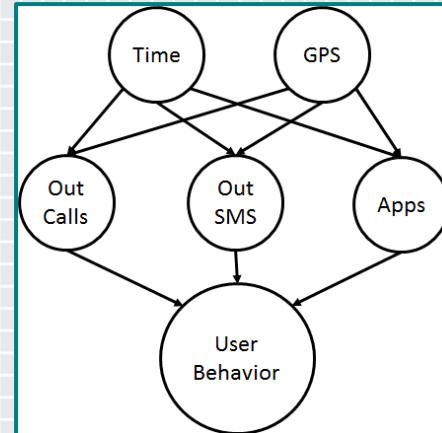


Participant Upload  
Distributions

# mRUU Update

## Sensor Updates

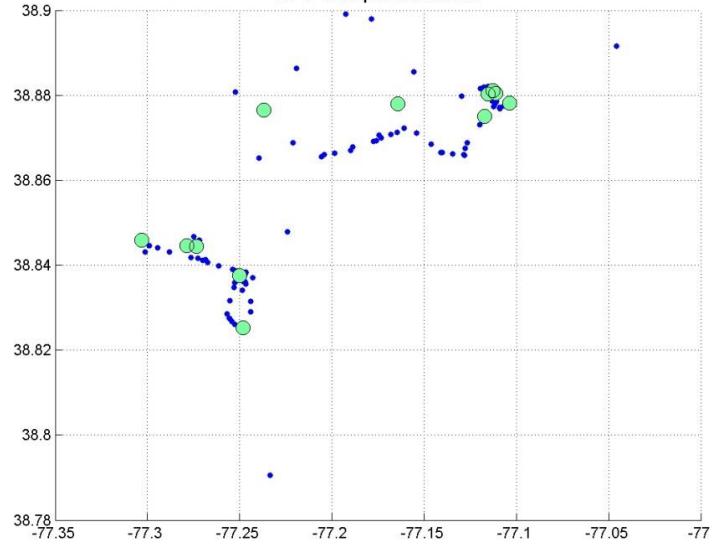
- ◆ Implemented Identity Engine using adapted modeling technique which incorporates:
  - ◆ Activity hotspots
  - ◆ Temporal information
  - ◆ Location information



# mRUU Location Based Modeling of User App Behavior

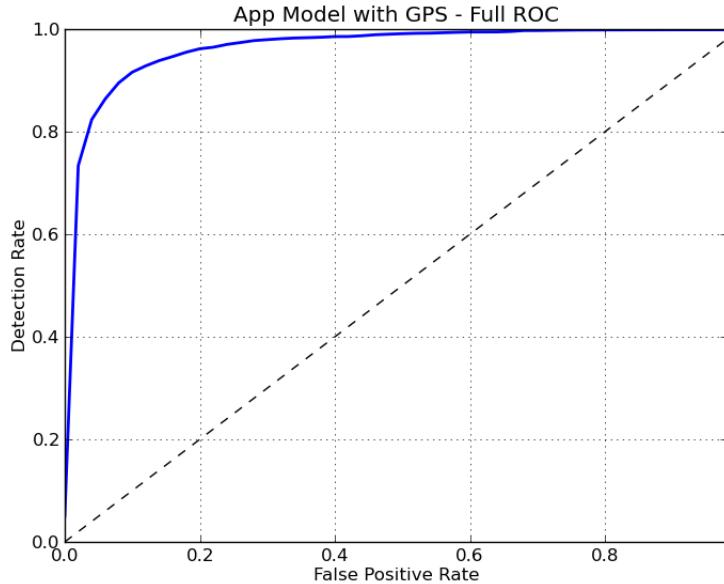
## Location Based Sub-Modeling

GPS Hotspot Detection



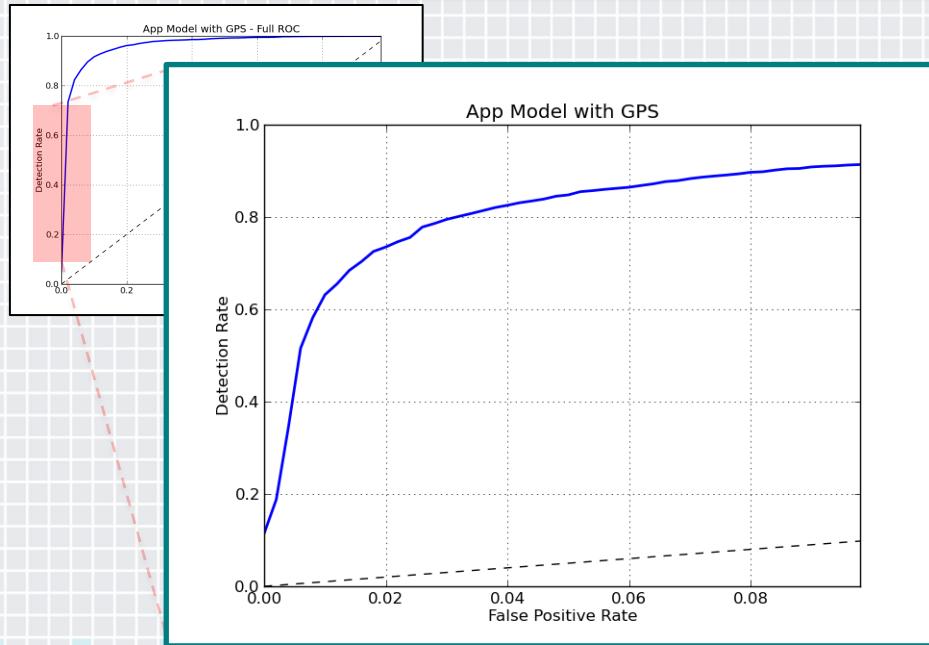
# Accurate Modeling of user app behavior

Modeling where you use Apps is very accurate



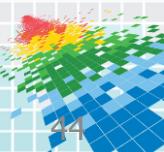
# mRUU Study Results- classification accuracy with no FP

## Application Usage Model Accuracy



- ◆ Behavior eval every 2 min
- ◆ 4 hours total = 120 Evals/day
- ◆ Goal: 1 FP/day = 0.00833

False-Positives per day	Percent of Foreign behavior identified
1	62%
2	70%
3	78%
4	80%

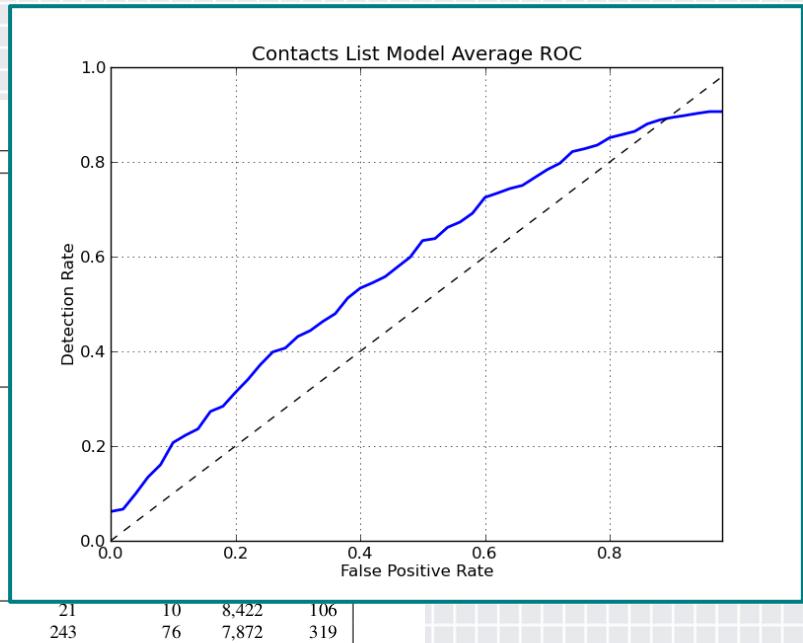


# mRUU Study Results

## Contact List Accuracy

**When contacts list accessed  
Ineffective – too few samples**

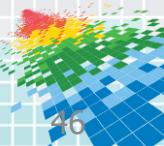
User	Days	Hours	Apps	Contact List	GPS	Phone
1	6 days	11:20:00	2,718,911	3,728	2,051	12
2	22 days	20:12:00	3,490,217	143	7,940	107
3	24 days	0:41:00	15,425,996	2,861	13,938	151
4	3 days	3:19:00	3,767,563	83	1,870	10
5	27 days	17:44:00	9,415,586	412	16,235	249
6	20 days	5:49:00	3,142,314	10	3,506	0
7	27 days	12:00:00	5,628,189	2,353	7,082	368
8	10 days	11:06:00	9,311,562	119	6,255	53
9	90 days	0:07:00	9,793,582	16,261	8,840	197
10	46 days	21:27:00	14,548,709	717	8,965	176
11	14 days	21:13:00	4,659,372	72	2,871	7
12	27 days	13:28:00	35,406,045	131	16,170	27
13	26 days	22:09:00	27,127,850	1,081	15,252	369
14	16 days	21:31:00	7,335,354	863	7,829	109
15	24 days	19:32:00	5,216,493	77	12,157	0
16	24 days	1:54:00	17,703,599	4,189	13,290	265
17	3 days	10:38:00	2,739,994	103	2,029	26
18	19 days	21:54:00	12,941,415	318	8,095	34
19	8 days	21:22:00	8,623,558	131	5,239	49
20	24 days	6:08:00	9,884,105	518	14,189	376
21	27 days	13:41:00	14,385,298	99	15,899	7
22	26 days	6:33:00	23,098,123	835	14,694	480



# mRUU User Study

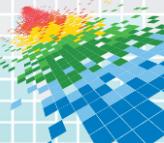
## Conclusions

- ◆ Users' mobile application usage habits can successfully be used to derive behavioral biometric identifiers
- ◆ The discriminative power of application usage patterns can be augmented using temporal and geographic information
- ◆ Additional usage data (eg contacts, etc.) provides poor discriminative measurements



# Introducing Decoy Apps and Decoy Clouds

- ◆ The mobile RUU app automatically creates decoy apps from unused apps or downloads strategic decoy apps
- ◆ Masqueraders are herded to pre-positioned decoy file system and decoy cloud services when they fail to re-authenticate



# Under the Hood

- ◆ Bad Behavior or Touching of decoy apps
  - de-authenticates the user
    - ◆ Locks the device
    - ◆ Captures a picture of the current user and records background ambient sound
    - ◆ Sends an alert out of band to the user
    - ◆ Re-authenticates by a second factor
      - ◆ Failure: Load Decoy Clouds and Decoy file system
      - ◆ Capture data on attacker



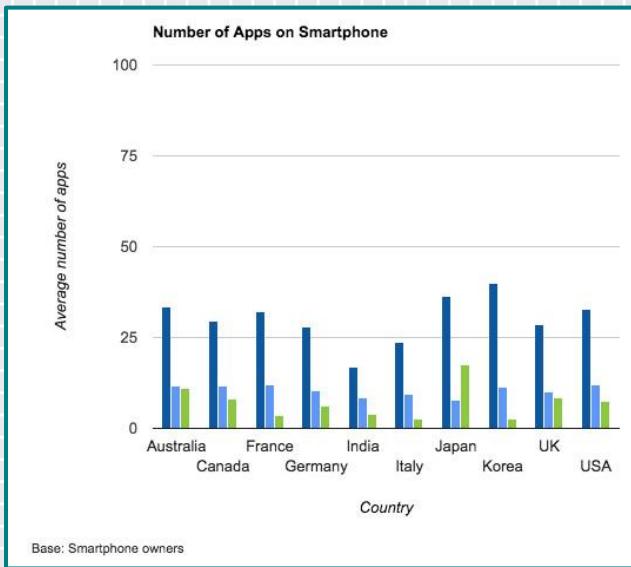
# Decoy Apps are intuitive

- ◆ Authentic looking apps that hold fake but enticing information to the adversary
- ◆ An attacker does not know what is a Decoy App and what is a Real App
- ◆ They are simple to use
- ◆ They are simple to understand
- ◆ They do not increase resource use



# Bloatware is turned into a Security Feature

Numerous Unused Apps available as decoys



- Number of Apps on Smartphone
- Number of actively used Apps (last 30 days)
- Number of paid Apps

# Onboard unused apps become decoys or strategic decoy apps are installed



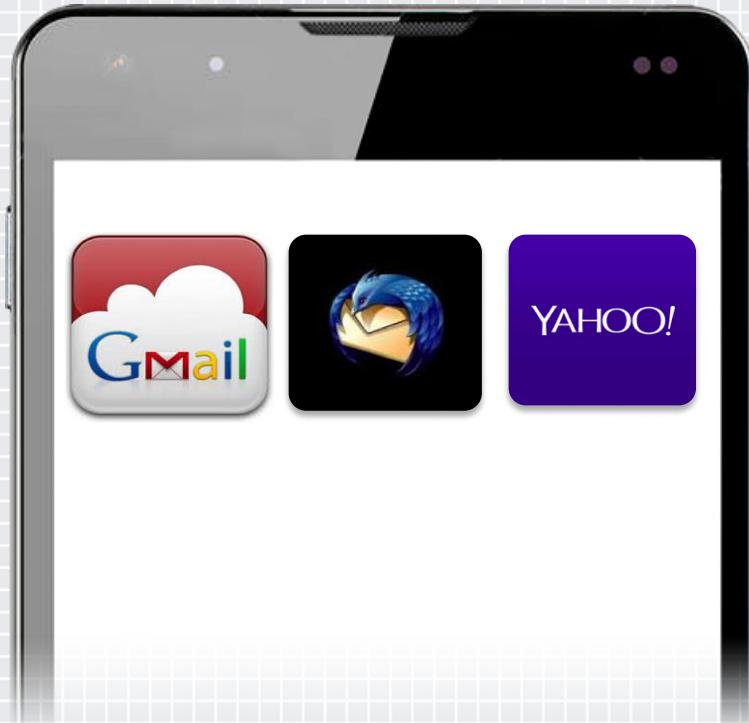
# Who do you really bank with?

One is real, the others  
aren't – Can YOU tell?



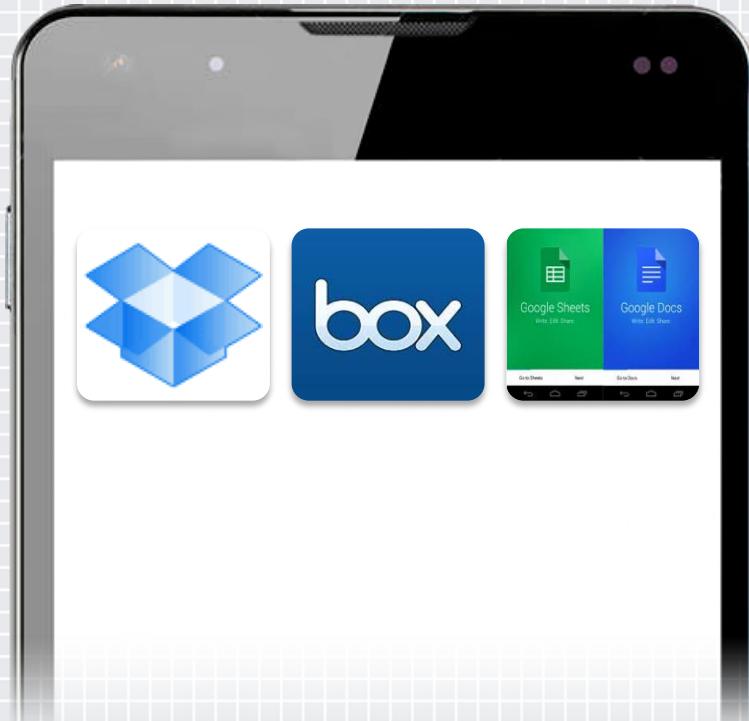
# Which is your real email client?

One is real, the others  
aren't – Can YOU tell?



# Who is Your Cloud Provider?

One is real, the others  
aren't – Can YOU tell?



# Which is your real corporate VPN?

One is real, the others  
aren't – Can YOU tell?



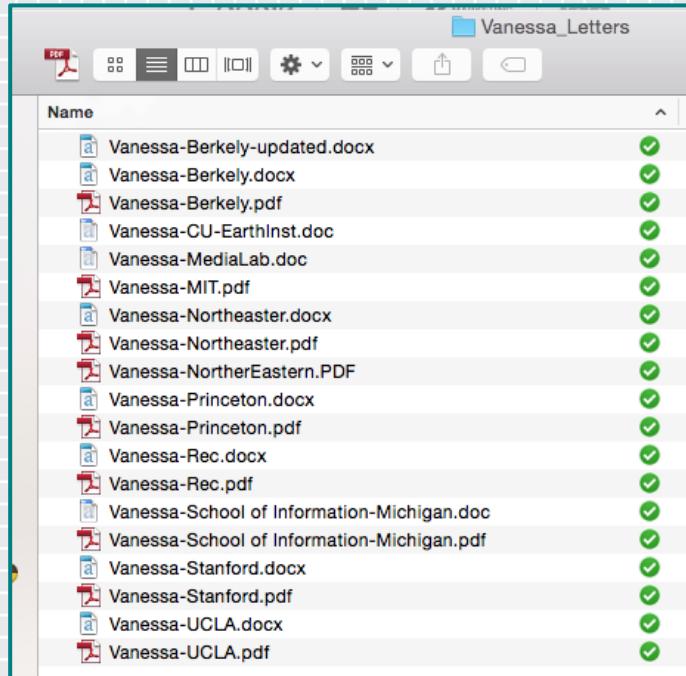
# Which is your real Facebook?

One is real, the others  
aren't – Can YOU tell?

Note: 2-D Passcode!



# Recall, enticing decoy files in the cloud, too!



One is real, the others  
aren't – Can YOU tell?

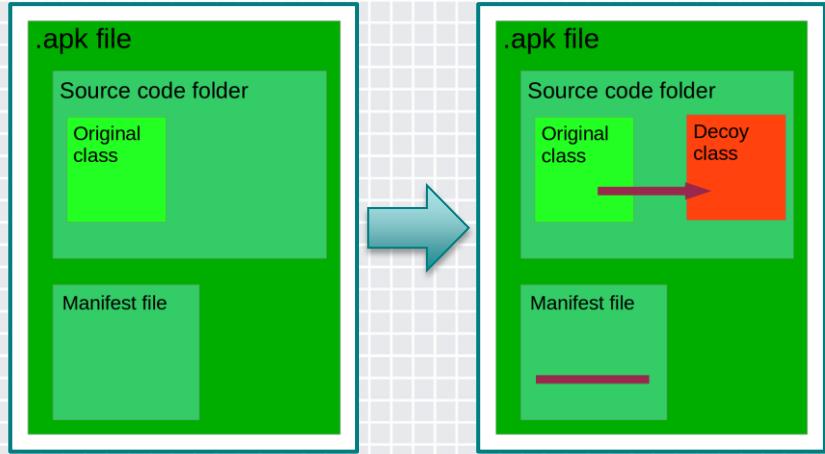


# How do we do it?



# Any app is easily converted into a decoy

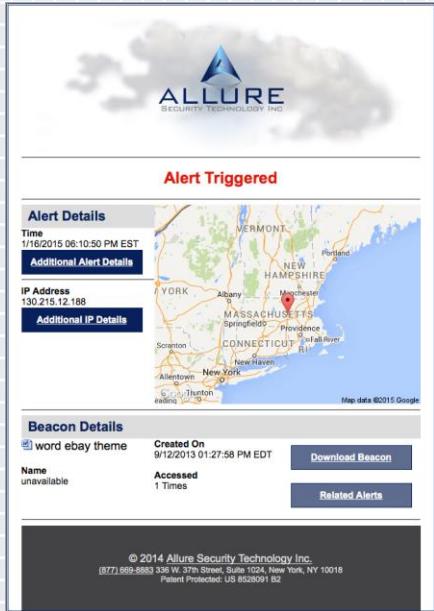
1. Unpack and disassemble .apk file.
2. Copy decoy functionality classes into assembly code folder
3. Insert code into original classes to run decoy functionality
4. Insert permissions for decoy functionality into manifest file
5. Reassemble and repack .apk file.



# Touch a decoy app, the phone locks and alerts



# ...includes location, picture & recording



**ALLURE**  
SECURITY TECHNOLOGY INC.

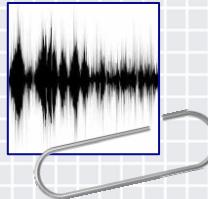
**Alert Triggered**

**Alert Details**  
Time  
1/16/2015 06:10:50 PM EST  
**Additional Alert Details**

IP Address  
130.215.12.188  
**Additional IP Details**

**Beacon Details**  
word ebay theme  
Name  
unavailable  
Created On  
9/12/2013 01:27:58 PM EDT  
Download Beacon  
Accessed  
1 Times  
Related Alerts

© 2014 Allure Security Technology Inc.  
(877) 859-8883 338 W. 37th Street, Suite 1024, New York, NY 10018  
Patent Protected: US 8528091 B2



# Sample Decoy App email alert

From: [rapd.cn@gmail.com](mailto:rapd.cn@gmail.com)  
Subject: Beacon Activated  
Date: June 16, 2014 at 11:27:37 AM EDT  
To: [sal@alluresecurity.com](mailto:sal@alluresecurity.com)

Somebody at /172.18.0.215 has accessed your beaconized application.  
Open attachments for more details.



# RSA® Conference 2015

San Francisco | April 20-24 | Moscone Center



Have we offered sufficient gifts to the demo gods?

# Demonstration

enter



# Alternative Unlock strategy, challenge the user, the phone knows your most recent behavior

#RSAC



With whom did you last chat?

- John Public
- Jane Doe
- Bill Jones
- None of the Above



# Answer wrong, try once more



What city did you last visit?

- Philadelphia
- New York
- Menlo Park
- None of the Above



# Answer wrong again...



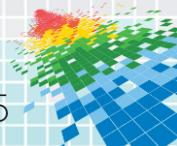
## Configurable mitigation strategies

- Brick the phone
- Upload tracking data
- Unlock and Load a decoy file system
- Alert Security Personnel



# DARPA Sponsorship

- ◆ DARPA ADAMS – Anomaly Detection at Multiple Scales
  - ◆ Insider threat
- ◆ DARPA Active Authentication
  - ◆ Masquerader/Impersonator threat
- ◆ \$10 Million of research support, transitioned from Columbia University IDS Lab to Allure Security Technology



# The Research Team



Sal Stolfo



Malek Ben Salem



Jon Voris



Yingbo Song



Joel Peterson



Shlomo Hershkop

# Apply What You Have Learned Today

- ◆ Next week you should:
  - ◆ Review corporate security policy for BYOD
  - ◆ Identify the number of employee phones stolen or compromised
  - ◆ Measure how many employees have no security controls on their devices
- ◆ In the first three months following this presentation you should:
  - ◆ Measure employee mobile access to critical corporate infrastructure
  - ◆ Evaluate corporate access and authentication controls
  - ◆ Explore a deployment strategy for advanced mobile authentication
- ◆ Within six months you should:
  - ◆ Identify and deploy solutions to protect employee mobile devices



# Thank you...



- ◆ Resources and contact
  - ◆ [www.cs.columbia.edu/ids](http://www.cs.columbia.edu/ids)
  - ◆ [www.alluresecurity.com](http://www.alluresecurity.com)

<http://www.channelpronetwork.com/article/Mobile-Device-Security-Startling-Statistics-on-Data-Loss-and-Data-Brea>

<http://transition.fcc.gov/bureaus/oet/tac/tacdocs/meeting12414/TAC-MDTP-Report-v1.0-FINAL-TAC-version.pdf>

<http://mashable.com/2012/11/08/smartphone-theft-city/>

