Detect and Contain

Combating Coordination Between Accounts

Robert E. Lee
Twitter: @robert_e_lee

What is Synthetic Account?

- A Synthetic Account (SA) is an account that an actor provisions/creates using names, attributes, contact methods, or other PII either not related to them or potentially not even related to an existing real human.

What is Account Takeover?

- Account Takeover (ATO) is when someone other than the authorized user successfully gains access to the user’s account

How can I detect Coordination?

- Organizations require better insights into the behavior of their users. These insights can be gleaned by analyzing ***Who* is doing *What* from *Where***.

- Fully leveraging available event data allows an organization to begin to:
  - Ease the onboarding of new customers
  - Recognize returning customers, offering a more pleasing user experience
  - **Detect and Contain unwanted users & their unwanted behavior**

Part 1 - Actors and Actions

Who is doing *What* from *Where*?

Who: Getting to know your users

- Every application that has a user base should have some notion of a Customer Identification Program (CIP).

- Common data collected during enrollment or within the use of an application include:
  - Name: Title, First, Last, Aliases
  - Contact Information: Address, Phone, Email
  - Payment information: Credit Card, Bank Account
  - State Issued ID: SSN, Driver’s License, Passport

Who: Getting to know your users

- Verification can include simple checks, such as verifying contact details (email, text message, automated voice call, mail, scanned ID card)
  - This step requires an enrollment fraudster to use contact details that they have access to
  - Verification can significantly add to a fraudster’s time and monetary costs
  - Verification can provide better nodes for Link Analysis

- External services can Validate identity by comparing the collected identity data to the public record:
  - Name <-> Address, Name <-> Phone Number, Address <-> Phone Number, Name <-> State Issued ID, etc
  - This step can reduce the number of enrollments with fictitious identities
  - You may also choose to block enrollment from certain sets of identity data, such as identities that are known to have been previously associated with account takeover, or identity theft

- Should limit how many accounts can be associated with a given node
  - 1 account per email address, 1 account per phone number, 1 account per SSN, 1 account per DDA, etc

Identity Assurance Level (IAL)

- An Identity Assurance Level is a measurement of how confident a system is that a person or organization is who they claim to be.

- Separate Identity Assurance Levels can be maintained for:

  1) A user’s legal name

  2) A user's affiliated organization(s)

     websites, press pass, drivers license, other credentials, etc

- It's important to verify Name and Organization distinctly; Many individuals should rightly be able to verify that they are Stephen King, but only one individual should be able to verify that they are the Stephen King who is associated with http://stephenking.com.

Identity Assurance Level (IAL)

- All users should be able to verify contact methods such as email and phone. Verifying one should yield a low level Identity Assurance Level (IAL). Verifying multiple should yield slightly higher IAL. These levels should be obtainable in a completely automated way.

- Once a system has a solid IAL measurement in place for its user base, it could offer its users a means to filter interactions from users who are at a lower IAL. This would allow users to filter interactions with accounts that, statistically, create more unwanted content.

- Users who have opted into filtering should have a "filtered" view where they can see all of the content that was filtered from them based on their filtration elections. They should then be able to mark something filtered as wanted if the content was desirable.

Identity Assurance Level (IAL)

- Individuals who create, take over, or control accounts at scale have an amplified voice over those with 1 account. Systems that do not allow users to obtain some IAL in an automated way, and filter interactions from users with a lower IAL, are open to manipulation at scale.

- IAL should be relatively inexpensive time/effort wise for a real person to obtain, yet quite expensive for an individual to obtain at scale for a large number of fake accounts.

What: What are my users doing?

- All of the interactions with your application are discrete events worthy of logging and analysis

- Events of particular note include:
  - Enrollment
  - Authentication
  - Profile/contact changes
  - High Risk/Sensitive transactions

What: What are my users doing?

- To enable Link Analysis and Behavior Monitoring capabilities, it is imperative to have complete log records:
  - ***Who***: User ID, or Personally Identifiable Information (PII)
  - did ***What***: Type of transaction
  - ***When***: Time stamp for event
  - from ***Where***: IP address, Location, Device ID
  - with what ***Result***: Success, Failure
  - and in what ***Context*** (session/role)

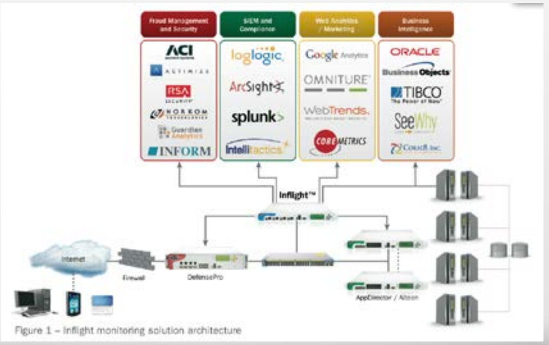


Figure 1 – Inflight monitoring solution architecture

W*here:* From where are my users accessing?

- Two common methods of determining *Where* involve the IP Address the user appears to be coming from, and the Device ID assigned to a user’s device (laptop, phone, etc).

- Location:
  - There are many services available that attempt to associate an IP Address with an approximate physical location
  - Through HTML-5 (or other methods depending on device type) it may be possible to query the device for current GPS location

W*here:* From where are my users accessing?

- There are several drawbacks to using an IP address to derive *Where*. Chiefly:
  - IP Addresses are often assigned dynamically for a short period of time
  - A single IP Address can represent many devices
    - A thousand devices behind a corporate egress point
  - A single device may cycle between multiple IP addresses; location change (home -> office), mobile device, a device proxy hopping
    - It is trivial to change which IP address a session appears to be coming from through use of proxies

For these reasons, it is desirable to supplement IP Address based identification with a more precise method. Modern Device Identification technology enables organizations to uniquely identify a device, and associate the event data with an individual device.

Where: From where are my users accessing?

- A Device ID should:
  - Accurately identify a unique device in a way that is resistant to manipulation:
    - http://samy.pl/evercookie/ -- One part Tag
    - https://panopticlick.eff.org/ -- One part Fingerprint
  - Allow for the recognition of a returning device
  - Not require active participation of the user
  - Have checks for signs of proxy use
  - Be included in the event logs
    - Once certain user behavior is observed as fraudulent, can link to other sessions from the same DeviceID
- Allow users to name their devices
  - Reference device name and location data in user communication

Part 2 – Common failures & Improvement Ideas

"*Failure is success, **if** we learn from it*"

--Malcolm Forbes

Common Methods of Account Takeover

- Enrollment Fraud - aka Synthetic Accounts (SA)
  - Attacker gets to choose their own password, answers to challenge questions, and specify contact records

- Knowledge Challenge Compromise
  - Automated brute force tools: JTR, Hydra, etc
  - Capture: Phishing, Key logging, Advanced Malware (MitM/MitB)
  - Weak "Forgotten Password" flow
  - Password reuse from previously compromised site

Common Methods of Account Takeover

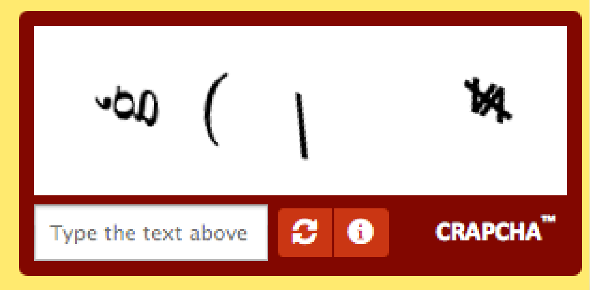
- Possession Challenge Compromise
  - Leaked seed values
  - Interception of data destined for a physical device
  - Interception of data destined for the application
  - Interception of the physical device
  - Interception of the keys contained on a physical device
- Session Hijacking
  - Cookie reuse
  - MitB

Common Methods of Account Takeover

- Biometrics Compromise
  - Without a human monitoring the process, biometric authentication is considered low assurance
  - Fingerprints are better usernames than authentication factors
    - Point: http://blog.dustinkirkland.com/2013/10/fingerprints-are-user-names-not.html - http://is.gd/ZM2wPZ
    - Counterpoint: https://blog.lookout.com/blog/2013/09/23/why-i-hacked-apples-touchid-and-still-think-it-is-awesome/ - http://is.gd/xHnVkV

Anti-Automation

- Captcha
  - Attempts to differentiate human vs computer (anti-automation protection)
    - Solves the wrong problem; Who cares if it is a human or a computer if the threshold of acceptable behavior is exceeded?
  - Provides a very [poor user experience]
  - Can be solved (defeated) by a [mechanical turk]
  - Better to have hard limits and constrain activity (No more than X count allowed in Y time)
  - Could also obfuscate/randomize form fields
    - http://www.shapesecurity.com



Authentication - Static Passwords

- Passwords are shared secrets that should only be known by the user and the authentication system
  - Passwords should be difficult to guess, yet easy to remember
  - If you use the same password on multiple sites, it's no longer a secret
  - See haveibeenpwned.com for examples of previously compromised sites

- Password policies often include:
  - An alphabet inclusion requirement (roman alphabet, numbers, special characters, case enforcement)
  - A minimum length requirement
  - An auto-lock out (hard lock, or incrementally time scaled lockout)
  - An expiration
  - A password rotation policy (can not use the same password as previous X)

Authentication - Static Passwords

- The relative strength of a policy can be calculated:
  - Alphabet requirement: roman alphabet (not case sensitive), numbers – 36 characters
  - Length: 6
  - Auto-lock out for 24 hours after 10 consecutive failed attempts (from any IP address) – rate of 10 tries/day
  - Expiration: 365 days

  - P=(10*365)/36^6, or 0.000168% of the entire password space can be exercised.
    - http://prezi.com/u1kpvimvoiwd/password-strength/

Authentication - Static Passwords

- Cap brute-force attempts
  - Lock after X consecutive failed attempts, and/or Y total failed attempts, in Z timeframe

- Don’t require the user to change their password too frequently
  - Good passwords are hard to remember

- Encourage passphrases over passwords
  - Length is crucial to increasing brute force time

Authentication - Static Passwords

- Enforce a sane entropy requirement
  - ~2.5 bits per byte seems reasonable
  - http://www.fourmilab.ch/random/

- **Enforce a dynamic wordlist check**
  - Make sure not more than N users or N% of users in the system are using the same password
  - **http://research.microsoft.com/pubs/132859/popularityiseverything.pdf**

Authentication – Challenge Questions

- Challenge questions often elicit answers that are quasi public knowledge
  - If the answer is known by more than the user and the authenticating system, then it is not valid for use in authentication

- Challenge questions are not subjected to the same complexity requirements as passwords
  - Brute forcing answers to challenge questions can often be easier than brute forcing passwords

Authentication – Challenge Questions

- Passwords and answers to Challenge Questions are both knowledge based challenges
  - An authentication system with two challenges of the same factor is still Single Factor Authentication
  - For a MFA system, consider replacing Challenge Questions with a *possession* based challenge

- Challenge Questions offer poor usability
  - Users often forget the answers

Authentication – Challenge Questions

- Static Challenge Questions are being phased out in favor of out-of-wallet/dynamic questions
  - Pro: Dynamic questions require no enrollment
  - Pro: If an attacker passively collects answers (phishing, mitm, etc), less useful for next session
  - Con: Can often be answered by anyone close to person; family members, ex-gf/spouse, close friends, someone with access to credit history
  - Con: If sourced by public record databases, by definition, not a secret, therefor invalid for authentication
  - Con: Sometimes the public record data is wrong
  - Con: Much easier to brute force – Require 2 correct answers, 4 multiple-choice answers per question – P=1/4^2, or 6.25% chance of blind guessing on 1st try

Authentication – Email

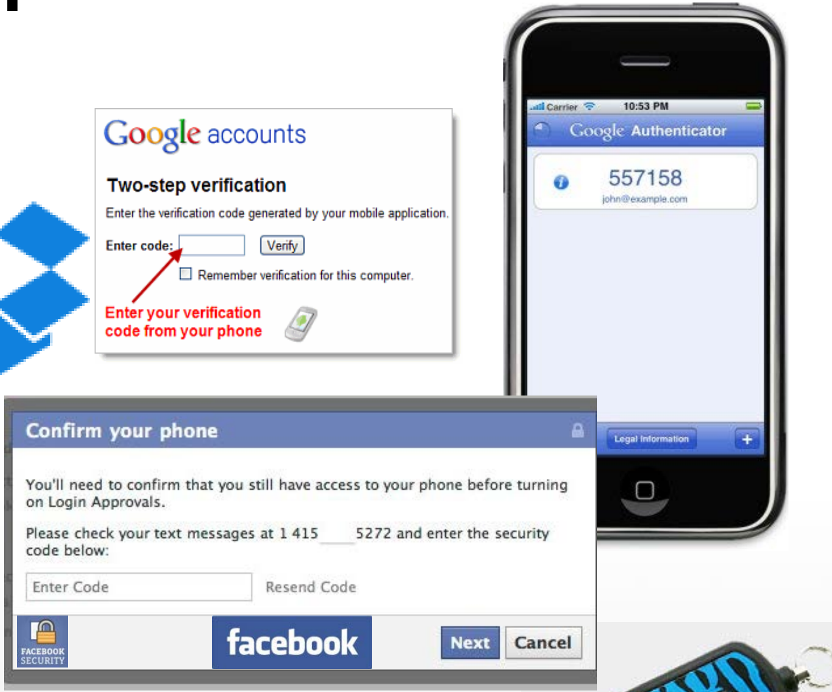
- Email is not something the user *knows*, *has*, or *is*
  - Email can therefor not be considered a factor in authentication. Still, Email can be used as part of the identity profile
  - Some major services are expiring “stale” user accounts, allowing fraudsters to open them back up under their control
- Unless otherwise provided for (SMIME, PGP, etc), email is not encrypted
  - Email is unsuitable for sending anything that requires confidentiality
  - Passwords, even One-Time Passwords require confidentiality
  - Special links that provide for privileged access require confidentiality too

Authentication - MFA

	Something you **KNOW**	Something you **HAVE**	Something you **ARE**
ATM	Pin	Card	
Application	Password, Challenge Question	Token, Phone	Retinal/Iris Scan

Users expect protection

- Examples:
  - Twitter
  - Google
  - Facebook
  - World of Warcraft
  - PayPal
  - Dropbox
  - Apple
- If free email, social media, and video games have MFA, **why doesn't your sensitive application?**



Google accounts

Two-step verification

Enter the verification code generated by your mobile application.

Enter code: [ ] Verify

☐ Remember verification for this computer.

Enter your verification code from your phone

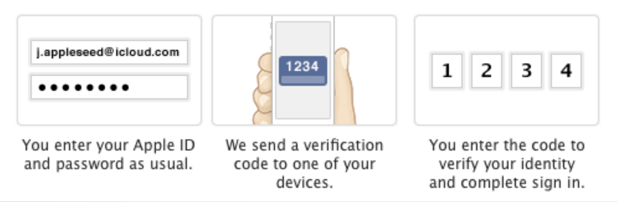
Confirm your phone

You'll need to confirm that you still have access to your phone before turning on Login Approvals.

Please check your text messages at 1 415 5272 and enter the security code below:

Enter Code

Resend Code



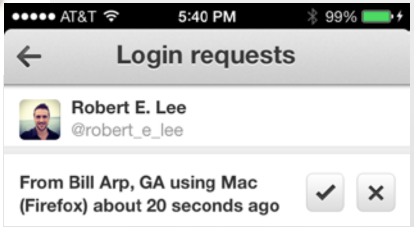
j.appleseed@icloud.com

••••••••

You enter your Apple ID and password as usual.

We send a verification code to one of your devices.

You enter the code to verify your identity and complete sign in.



Login requests

Robert E. Lee
@robert_e_lee

From Bill Arp, GA using Mac (Firefox) about 20 seconds ago ✓ ✕

Authentication – Dynamic Passwords

- A One Time Password is a dynamic knowledge challenge that is often associated with a physical object, and often time bound

- Token based OTP requires secrecy of seed value
  - https://en.wikipedia.org/wiki/SecurID

Authentication – Dynamic Passwords

- OTP's communicated to a user out-of-band (SMS, Voice) are often still collected in-band
  - Unless the user *responds* out-of-band, this is still an in-band authentication

- Typical OTP implementations rely on the application to provide authentication context to user
  - In a MitM/MitB/Phishing scenario, user can be socially engineered into giving an attacker their OTP credential

Authentication – Dynamic Passwords

- While OTP's do not represent something the user has, they may provide higher Authentication Assurance than static passwords

- As an industry, we should consider replacing static passwords with dynamic passwords

Authentication – Cookies

- Cookies are digital artifacts that can be copied and reused
  - If two or more people can simultaneously “have” the item, it is invalid for *possession* based authentication

- FFIEC definition of Complex Device Identification
  - Have the value in the authentication cookie change each session
  - Restrict use of cookie to systems with identical fingerprint

Authentication – Forgotten Password

- Forgotten Password features often utilize weaker authentication controls than normal authentication flow
  - In many MFA implementations, Forgotten Password relies solely on a possession challenge to reset the knowledge credential, rendering the MFA solution in actuality, Single Factor Authentication
  - Especially worrisome if your “possession” challenge is mistakenly using Email

- Resetting a password is a high risk transaction
  - Protect it 1st with a valid *possession* challenge
  - Supplement with alternate *knowledge* challenge

Assume Compromised System

Web Browser

Trusted Path (OOB)

User Interaction – Alerting

- Users should be notified when important changes are made

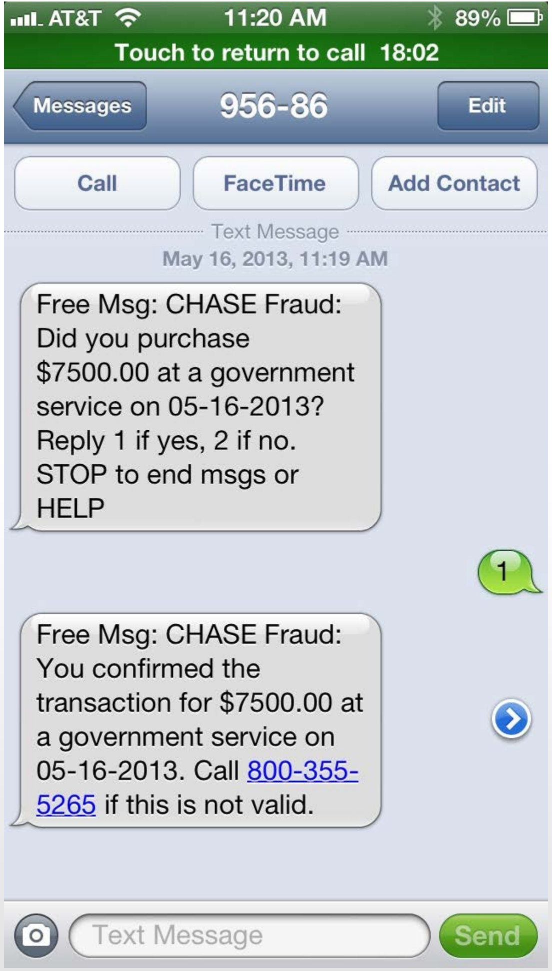
- If alerted of a sensitive change they did not make, users will naturally contact you

Transaction Intent Verification

Blind Authentication	Context Aware Authentication
• User starts to log into shopping site with Username and Password • User is prompted for an OTP • User Enters OTP into web browser to complete authentication	• User starts to log into shopping site with Username and Password • User receives a message OOB showing an authentication attempt, complete with the username, device name, IP address, location, and date/time of attempt. • User responds OOB “approve” to complete authentication
• User initiates a purchase for $10 worth of paper towels • User prompted for their 16-digit CC number and CVV • User enters CC, CVV and seemingly completes purchase • Days go by • User discovers their account was used to purchase a TV and shipped it to an unknown address • Possible lawsuit filed; Bad PR	• User initiates a purchase for $10 worth of paper towels • User receives a message OOB showing a purchase attempt for a TV • User denies the transaction and reports the transaction as fraud • Company helps user reset account credentials and clean infected device

Armed with contextual awareness, users can make intelligent authentication decisions. Responding OOB reduces chance of giving authentication credential to attacker.

Transaction Intent Verification (Two-way OOB + Context), SMS Example



May 16, 2013, 11:19 AM

Free Msg: CHASE Fraud: Did you purchase $7500.00 at a government service on 05-16-2013? Reply 1 if yes, 2 if no. STOP to end msgs or HELP

1

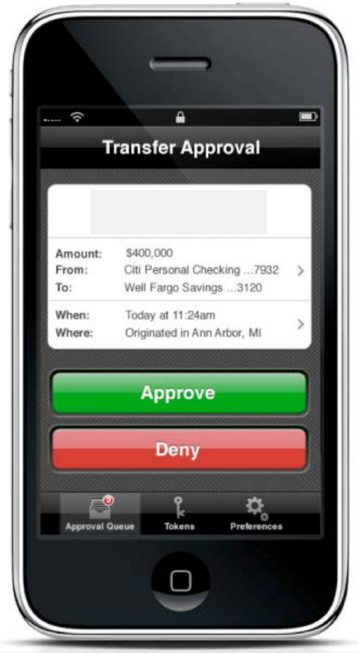
Free Msg: CHASE Fraud: You confirmed the transaction for $7500.00 at a government service on 05-16-2013. Call 800-355-5265 if this is not valid.

Authentication – Text/Voice

- Text/Voice based authentication can be intercepted
  - Some platforms (blackberry, android, etc) allow for running unsigned code and have been targeted for special SMS interception malware: https://www.google.com/#q=zitmo
  - NIST's opinion on SMS: https://threatpost.com/nist-recommends-sms-two-factor-authentication-deprecation/119507/
  - Calls can be forwarded http://www.snopes.com/inboxer/scams/forward.asp

Authentication – Trusted Path

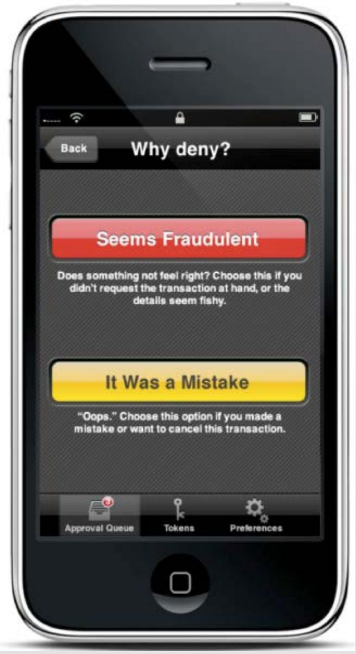
- Ideal to use native applications with signed, encrypted, two-way messaging
  - Better Security
  - Better Usability



Transfer Approval

Amount: $400,000
From: Citi Personal Checking ...7932
To: Well Fargo Savings ...3120

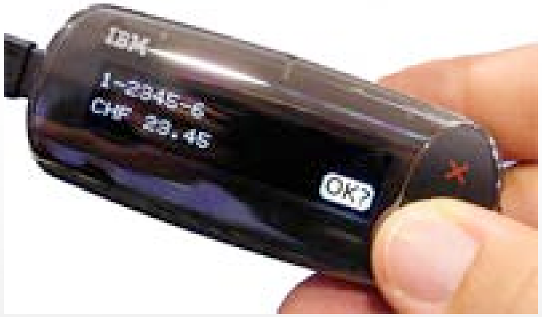
When: Today at 11:24am
Where: Originated in Ann Arbor, MI



Why deny?

Does something not feel right? Choose this if you didn't request the transaction at hand, or the details seem fishy.

"Oops." Choose this option if you made a mistake or want to cancel this transaction.



Example Vendors:
DuoSecurity
IBM

Link Analysis

- Many organizations are not performing thorough root cause analysis or link analysis for confirmed security incidents
  - Unaware of what control(s) failed; stay blind of what to fix
  - Unaware of true scope of incident; unable to reach out to other affected users, business partners

Part 3 – Putting it all together

“*The trick to forgetting the big picture is to look at everything close-up*”

--Chuck Palahniuk

So what?

- Too little security == lack of customer trust
- Too much security == customer frustration

Our Challenge

- Differentiate legitimate users from attackers
  - Low friction for users
  - High friction for attackers

Detect and Contain: Building Blocks

- Ensure log completeness
  - *Who* did *What*, *When*, from *Where*, *Result*, *Context*

- Identify high risk transactions in your application
  - Enrollment
  - Authentication
  - Profile/contact changes
  - High Risk/Sensitive transactions

- Determine appropriate Identity Assurance and Authentication Assurance Requirements to perform transactions
  - Deploy appropriate controls for risk tolerance
  - Interesting implications for SSO/Federation

Detect and Contain: Building Blocks

- Set up authentication checkpoints within your application to protect high risk transactions
  - May be too onerous to challenge users every time at checkpoints
  - Can scale back challenges based on thresholds (challenge on money movement over X amount)

- Ensure healthy incident response program
  - Don’t just track incident count; where possible, identify which control failed, and how it failed
  - If incident is associated with a vulnerability, track incident count against known vulnerability
  - Perform Link Analysis to see which other sessions/actors are related to incident

Detect and Contain: Finishing Touches

- Deploy a real-time risk scoring system:
  - Automatic learning from the event data stream
    - Multiple models: User vs Self and User vs App Population
  - Should take all pieces of Identity, and reputation of identity into consideration
  - Should also incorporate learnings from manually tracked incidents
  - **Tries to answer question: Is this individual behaving as they normally do, or is their behavior anomalous?**

- Integrate application checkpoints with risk scoring system (Risk Based Adaptive Authentication)
  - Challenge based on risk score, and risk tolerance
  - Delivers the best blend between security and usability
  - **Automatic User Experience changes**

Detect and Contain: Finishing Touches

- Work with Data Scientists
  - You can only write pattern matching rules once you know what you’re looking for
  - To stay ahead of the incident curve, invest in people and technology (Big Data) that can help you identify new trends

Base Services

- **Device Identification Service**
  - **Input**: Device Tag and/or Fingerprint data
  - **Output**: Universal Device Identifier, Confidence Level
  - **Scope**: Web enabled devices, smart phone applications, and desktop applications

- **User Identification Service**
  - **Input**: PII
  - **Output**: Universal Identity Identifier, Confidence Level

- **Reputation Service(s)**
  - **Input**: Device ID, IP Address, PII Elements
  - **Output**: Reputation Score

- **Risk Score Service**
  - **Input**: IP Address, Device ID, User ID or PII, Reputation Scores, Transaction details
  - **Output**: Risk Score, recommendation (Ie: allow, deny, stepped authentication)
  - Enables RBA

Types of Assurance

- Authentication Assurance Level (AAL)
  - Assurance that returning user is person who created account
  - Can be calculated on a per user/per session basis
  - Function of what authentication challenges a user has passed, and when (level to decay over time)

- Identification Assurance Level (IAL)
  - Assurance that user is identity they claim to be
  - Can be calculated on a per user basis
  - Function of what identity challenges a user has passed, and when (level to decay over time)

Static AAL

- Authentication types, communication mechanisms, and device type should be modeled together when specifying AAL values

Authentication Type	Communication Mechanism	Device Type	AAL
OTP	Email	n/a	5
OTP	SMS	n/a	10
OTP	Token	RSA	12
TIV	Email	n/a	15
TIV	SMS	n/a	20
TIV	Push	n/a	25
TIV	NFC	YubiKey X	30

Dynamic AAL

- Incident Response rates should feed back into the AAL calculations
  - For every confirmed case of ATO in your application, Customer Service should note/log what authentication controls were compromised by the attacker to perform the operation
  - When a given authentication type, over a specific communications mechanism, or on a specific device type fails to differentiate the real user from the attacker, the measured AAL for challenges of that combination should be adjusted

Dynamic AAL

- Measure the failure rate divided by the total times that combination has been fired

Authentication Type	Communication Mechanism	Device Type	Failure Count/Challenge Count in past 90 days
OTP	Email	n/a	341/40566
OTP	SMS	n/a	209/50823
OTP	Token	RSA	178/59799
TIV	Email	n/a	285/54383
TIV	SMS	n/a	144/65462
TIV	Push	n/a	85/74684
TIV	NFC	YubiKey X	38/91212

Risk Based Authentication

- Authentication Assurance Requirement (AAR)
  - – The policy specified minimum AAL a user/session must have to be authorized to perform an operation
  - – Enforced by a reference monitor/policy engine

- Authentication Assurance Delta (AAD)
  - – Difference between AAR and AAL

Risk Based Authentication

- Risk Score
  - A dynamically calculated measurement of perceived risk for a given operation
  - Can start off with a numeric range to represent risk, such as 0-1000
    - Can then assign bounded ranges for different degrees of risk such as
      - Low: 0-299
      - Medium: 300-699
      - High: 700-1000
  - Different types of operations should be modeled differently
    - Ie, SignIn operations will differ from purchase decisions or configuration changes
    - Inexpensive starting point for risk calculation can be threshold based

Risk Based Authentication

- Reference Monitor Example
  - Low friction for legitimate user
  - High friction for attacker

Operation	Data Classification	Risk Score	AAR
Write	Critical	Low	20
Write	Critical	Medium	30
Write	Critical	High	Disallowed

Risk Based Authentication

- Authentication Orchestration
  - Knows what discrete authentication types, communications mechanisms, and device types are available for a given user
    - Challenge Types: Password, Challenge Question, OTP, TIV
    - Communications Mechanisms: Email, SMS, Push
    - Device Types: Ex - Yubikey X
  - Can take user through least friction challenges leveraging available challenge types, communications mechanisms, and device types to achieve AAR