Digital Identity – The Ghost in the Machine

Dr.Max Kilger
Chief Social Scientist
Honeynet Project

Introduction

- **Identity**
  - Sense of identity as necessary for psychological well- being
  - Identity is a multi- dimensional construct
  - Digital technology has had a very significant impact on our sense of identity and how we construct it
    - How we see ourselves
    - How others see us
  - What might the future hold for us as ghosts in the machine

Identity as process

- **Identity as ascribed process**
  - **Identity processes begin at our very birth**
    - **Gender assignment**
    - **Given name**
- **The power of ascribed identity**
    - **What happens when you disrupt the identity process – even at birth**
- **Example – disruption of ascribed status**
  - **"Switched at birth"**

Identity as process

- **Different dimensions of identity processes**
  - **Legal**
  - **Social**
  - **Commercial**
  - **Other realms**
- **Status processes and their role in identity**
  - **Occur everyday and to everyone in this room**
  - **Involves comparing information about characteristics we possess and those that others possess**
    - **Examples include demographic variables such as age, gender, education, etc.**
    - **Other characteristics include skill levels such as expertise at coding C++**

Identity as process

- **These characteristics help identify us and our role in the social world**
  - **We gather information about others thru direct and indirect means**
    - **Direct claims made by individuals**
    - **Validation by verbal and non- verbal clues**
  - **Face to face communication**
    - **Observation of verbal and non- verbal clues in this environment very common**
    - **The large bandwidth of face to face communications**
  - **Example of face to face comms with an audience member volunteer**

Identity as process

- **The absence of verbal and non- verbal clues in the digital world disrupts the identity validation process**
  - **Common communication channel - Email**
    - **When someone named Heidi sends you an email what do you know about them?**
      - **They are female...**
        - » **How do you verify that?**
    - **Someone claims to be an expert perl programmer – how do you validate that claim?**
      - **They send you a sample of their code**
        - » **How do you know they wrote the code?**

Identity as process

- **What about digital communications by webcam, for example? Wouldn't that work?**
  - **We know from research that even with video and audio there is significantly less information to process in effective communication that results in establishing identity**
  - **If that were the case then why would I fly 10 hours to give this talk – I would just use a webcam and stay home with my bunny rabbits…**

Identity as a temporally unstable element

- **Persistence as a key attribute of identity**
- **Every element of identity has its own independent timeline**
  - **Example of an identity element with a very slow timeline - National identity card/number**
  - **Another example of a somewhat slow timeline – occupation**
  - **Example of an identity element with a very fast timeline – the number on this RSA token – it identifies me to a specific network and it changes every 60 seconds**
- **Any identity has some probability p to exist at time t and some probability q that it ceases to exist at time t+1**

Identity as a temporally unstable element

- **Online identities are a special instance of identity as temporally unstable**
  - ◆ **Passwords age for example and must be replaced**
  - ◆ **When we shop at a commercial website our session is given a session id that dies at the end of our shopping visit**
  - ◆ **People may drop one digital identity and adopt another in an online game for example**
  - ◆ **People here at the conference often use or create identities that are not linked to identities associated with them in government databases**

Identity as a temporally unstable element

- **Why might an identity cease to exist?**
    - **Individual disassociates themselves from the identity**
    - **Some entity forcible disassociates the individual from the identity**
    - **Identity expires at some given time point**
- **Devil's advocate question**
    - **Does an identity ever really cease to exist?**
        - **If some instance of it remains, either in the physical, cyber wetware world then has it really ceased to exist?**

Identity as a temporally unstable element

- **Consequences of temporally unstable identities**
  - **Persistence of identity is important for infosec purposes**
    - **Authentication and authorization often require time stable identities**
  - **Identity Management Architectures often rely upon temporally stable identities**
- **An exception to the rule – parallel identity systems that sync with some external source**
  - **RSA token one example – temporally unstable identities that change in a predictable pattern that is synced to time**

Identity as a temporally unstable element

- **Psychological consequences of temporally unstable identities**
  - Historically identity has been a more temporally stable element
    - A good example – the emphasis on "reputation" in the past
  - Digital revolution has accelerated the number and the pace at which identities emerge and expire
  - This means that we must adapt psychologically to tolerate faster and more frequent identity change in order to maintain a sense of well- being
  - Failure to adapt may result in the loss of self – the inability to develop and sustain a true sense of who one is

Situational identity

- **Social scientists have long observed that identity is situational**
  - ◆ **Situational identity certainly exists in the non-digital world**
    - ▪ **People hold and change roles depending upon the situation**
      - • **Example - role as concertgoer, father, air passenger**
  - ◆ **Situational identity in the digital world is much more rich and complex**
    - ▪ **Example – role as orc- battling hero in a game, member of a hacking team in the middle of an IRC chat, employee logging onto their network**

Situational identity

- **Important properties of situational identities**
  - ◆ **They are most salient when the actor is present in the situation**
  - ◆ **There is often a lack of communication between situational identities a person may experience and possess**
    - ▪ **One important reason why this is the case is that often situational identities can conflict with each other**
    - ▪ **Example – person is an info sec employee in their dayjob and is a member of a hacking group on their personal time**

Situational identity

- **Augmented Social Networks (ASN)**
  - **Concept proposed by Jordan, Hauser and Foster**
  - **ASN based on four elements;**
    - **Persistence of identity**
    - **Interoperability between online communities**
    - **Brokered relationships**
    - **Public interest matching technologies**
  - **Idea is that if you trust a person (e.g. identity) in one area you can trust them (e.g. an identity) in another area**
    - **Someone I know is a great C++ coder and so I would trust her to recommend a good book to read for fun**

Situational identity

- **Augmented Social Networks (ASN)**
  - ◆ **The breakdown in this scheme often comes in the interoperability assumption of ASN**
    - ▪ **Example –**
      - • I know someone who is an expert researcher and I trust her research judgment implicitly
      - • Jordan, Hauser and Foster suggest that I can transfer that trust to another community – in this case Ebay
      - • Would I trust her as a seller or buyer on Ebay? Not a chance – she is terrible at paying bills, has her cable shut off, forgets to make insurance payments, all sorts of financial mayhem
  - ◆ **So situational identities may often not be good candidates for schemas that attempt to transfer attributes of an entity in one situation to that same entity in another situation**

Situational identity

- **This suggests that in the digital world, large identity architectures like some federated identity systems that rely to some degree on situational identities to authenticate, validate or grant privileges or trust relationships may get themselves into trouble**
    - **That is, we know that typical info sec rules that involve "I know something"/ I have something" are simple identity functions and can be circumvented**

Identity – class versus unique identifiers

- **Unique identifier – this type of identifier ideally contains as its set only one actor (one- to- one relationship)**
  - **Example – national identity card/ number**
  - **Fingerprint**
  - **DNA**
  - **We know that even these are not perfect – but they are asymptotically unique**
- **Class identifier – this type of identifier associates an actor as belonging to a class of individuals (one to many)**
  - **Example – female, college graduate, bunny rabbit owner, User- Group- Other in unix systems**

Identity – class versus unique identifiers

- **Class and unique identifiers in the non-digital world**
    - **Examples of class identifiers – age, gender, college educated**
        - **You can attempt to alter your identity – with makeup, clothes, vocabulary**
    - **Examples of unique identifiers – passport, identity card**
- **Bandwidth is high in non- digital world and makes it difficult to alter and maintain identities**
    - **Gap between class identifiers and unique identifiers in non- digital world is large**

Identity – class versus unique identifiers

- **Class and unique identifiers in the digital world**
  - **Same examples of class identifiers – age, gender, college educated**
    - **Gender can be altered with a nickname change, age can be altered in an explicit statement, college education can be both explicitly claimed and backed by background statements**
  - **Examples of unique identifiers – bank account number, social security number in US, username/ password**
- **Bandwidth is much lower in digital world and makes it easier to alter and maintain identities**
  - **Gap between class identifiers and unique identifiers in digital world is much smaller**

Identity – class versus unique identifiers

- **Creating identities in the digital world**
  - ◆ **Use of cover information**
    - ▪ **Example – create a website with information consistent with your digital identity**
      - • Search engine crawls your website and picks up your information
      - • Search engines like Google have some inherited legitimacy that then colors your identity details
  - ◆ **Use of limited bandwidth to control/conceal telltale identity markers**
    - ▪ **In FBI Innocent Images program, 14 year old girls teach agents how to be – 14 year old girls…**

Identity – deterministic versus probabilistic identifiers

- **Deterministic identifiers**
  - Characteristics or pieces of information that are known about yourself by others with a high level of confidence
    - Your gender, age, status as a jeep owner
  - These characteristics are asymptotically deterministic because they can be verified and validated by corroborating clues
- **Probabilistic identifiers**
  - Characteristics or pieces of information that are know with some probability much less than limit(p=1.0)
    - Often involve modifiers or levels of a certain attribute or characteristic

Identity – deterministic versus probabilistic identifiers

- **Probabilistic identifiers (continued)**
  - ◆ **Example – "expert C++ programmer"**
    - ▪ **Issue – identifier is relational to others in the universe**
    - ▪ **Issue – more difficult to verify and validate in a lower bandwidth environment like the web**
  - ◆ **These probabilistic identifiers often cause significant status conflicts in the online communities because they are difficult to validate**
    - ▪ **This is one reason why there is a lot of conflict in the hacking community**
    - ▪ **Conferences like this one help reduce these conflicts by facilitating face- to- face interaction and validation**

Identity – deterministic versus probabilistic identifiers

- **Another meaning…**
  - Commercial, governmental and military entities create databases and digital identities within them
  - These probabilistic identifiers may be either deterministic or probabilistic
  - Here probabilistic can be of a different nature
    - Many times information about an individual is incomplete or missing
      - Missing or incomplete because of logistical issue
      - Missing or incomplete because the concept or variable is a complex, abstract identity

Identity – deterministic versus probabilistic identifiers

- **Another meaning…**
  - **What happens when information is missing or incomplete?**
    - **It is often modeled – statistical models often used…**
    - **Many orgs like credit bureaus may model your income or other characteristics like education or ethnic origin**
  - **Other kinds of identities may be modeled as well**
    - **Probability that the individual is a terrorist for example**
    - **Which census block groups are the most likely ones that need to be searched for a criminal or terrorist**

Sources of Identity – Self or Other

- **Classic social interactionist thoery**
  - **George Herbet Mead**
    - **Concept of the "I" and the "me"**
      - **"me" is the social self – identity that arises from the way in which others interact with you**
      - **"I" is the novel or individualistic way in which the individual reacts to the social world**
  - **Thus our psychological sense of self and identity is created both from ourselves as well as our perception of how others see us**

Sources of Identity – Self or Other

- **Computers and machines as social identities**
  - ◆ **In Mead's theory of social interaction, one of the main definitions of a social actor is an entity that exchanges meaningful social symbols**
  - ◆ **In our interaction with computers we exchange meaningful social symbols with the computer, so is the computer a social actor and thus has a social/ digital identity?**
  - ◆ **Some evidence suggests that the answer is yes**
    - ▪ **Notice the way in which people – both tech and non-tech – anthropomorphize computers and computing machinery**
      - • **With names**
      - • **With moods and personality**
      - • **Sometimes with a sense that they are sentient entity**

Future of the digital individual

- **Digital identity fragmentation**
  - **Our sense of identity is likely to become more and more fragmented as we and others build additional digital identities for ourselves. Will we be able to adapt to this without some psychological penalties? Will we just become "ghosts in the machine?"**

- **Does social theory about the sense of identity and self have implications for technical arenas such as Identity Management schemas?**
  - **Will these systems progress beyond the usual "I have something" , "I know something", etc.**
  - **Does a more psychological understanding of identity contribute to identity schemes such as Cameron's "7 Laws of Identity" ?**

Future of the digital individual

- **Can we outwit/ outrun entities that attempt to identify us without our knowledge?**
  - ◆ **My sense is that digital technology not only gives others the ability to efficiently construct multiple identities for us but also at the same time allows individuals to develop multiple identities for themselves that they can create and shed faster than other- controlled identity schemas can collect and assemble them.**

Future of the digital individual

- **Thanks for listening**
- **I'm Dr. Max Kilger – at least I think I'm Dr. Max Kilger – uh- oh – I'd better go check…**