Challenge Question Authentication

Mike Just, Heriot-Watt University
COINS Summer School on Auth Ecosystems

31 July 2016

Outline

- Challenge Question Authentication
- Research Work
  - Approach
  - Data Collection
  - Security Evaluation
  - Usability Evaluation
- Concluding Remarks

Outline

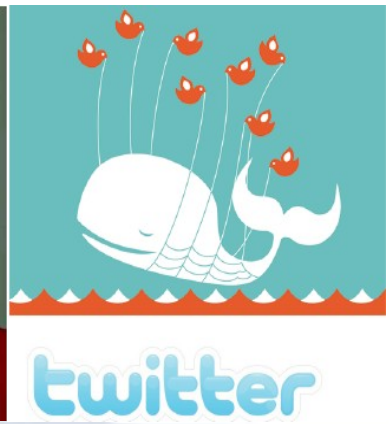
- Challenge Question Authentication
- Research Work
  - Approach
  - Data Collection
  - Security Evaluation
  - Usability Evaluation
- Concluding Remarks

Challenge Question Authentication

- What are “challenge questions?”
  - Type of *authentication credential*
  - User registers a *question* and an *answer*
  - To authenticate later, a user is posed a question(s) and must provide the answer(s)
- Used to complement passwords, or support account recovery
- As ubiquitous as passwords
- *Known* vs. *memorized*

Challenge Question Authentication (2)

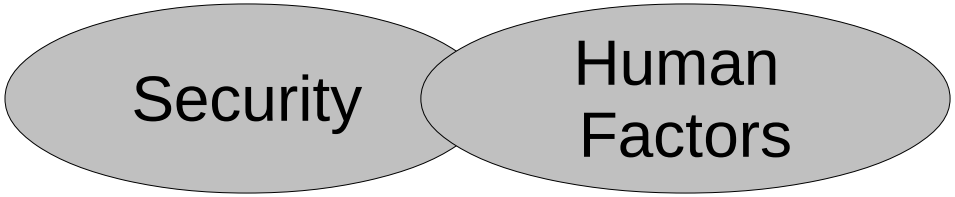
- High-profile attacks
- Are they effective?
- Can we improve?



Outline

- Challenge Question Authentication
- Research Work
  - Approach
  - Data Collection
  - Security Evaluation
  - Usability Evaluation
- Concluding Remarks

HCISec for Challenge Questions



I. Evaluate
   i. Security (Attacker's Point of View)
      - *Is it difficult to determine the answers?*
      - *Several aspects to this determination*
   ii. Usability (User's Point of View)
      - *Is it difficult to choose questions and answers?*
      - *Is it difficult to remember answers?*

Some Examples

- Consider the following examples
  - What is your mother's maiden name?
  - What is your favourite colour?
  - Who is your favourite actor?
  - What was your high school locker combination?
  - What was your first pet's name?
- Are these questions secure?
- Are these questions usable?

Security Criteria

- Guessability
  - Traditional measure in which the security level is directly proportional to the number of possible answers for a given question
- Observability
  - The security level is inversely proportional to an attacker's ability to find the answer to a given question
- "Attackers" might be strangers, acquaintances, colleagues, friends, family members

Usability Criteria

- Applicability
  - Users have sufficient information to provide a relevant answer to a question
- Memorability
  - Users can consistently recall the original answer to a question over time
- Repeatability
  - Users can consistently and accurately (syntactically) repeat the original answer to a question over time

Examples Revisited (1)

Consider the following examples

What is your mother's maiden name?

Usability HIGH **Security LOW**

What is your favourite colour?

Usability MED **Security LOW**

Who is your favourite actor?

Usability MED **Security MED**

What was your high school locker combination?

Usability LOW **Security HIGH**

What was your first pet's name?

Usability MED **Security MED**

Examples Revisited (2)

- Did you agree with the usability and security ratings on the previous page?
- Security
  - 'Observability' levels are often subjective
- Usability
  - Often depend upon context and environment, e.g. user base, user experience, guidance to users
  - Requires empirical evidence

Outline

- Challenge Question Authentication
- Research Work
  - Approach
  - Data Collection
  - Security Evaluation
  - Usability Evaluation
- Concluding Remarks

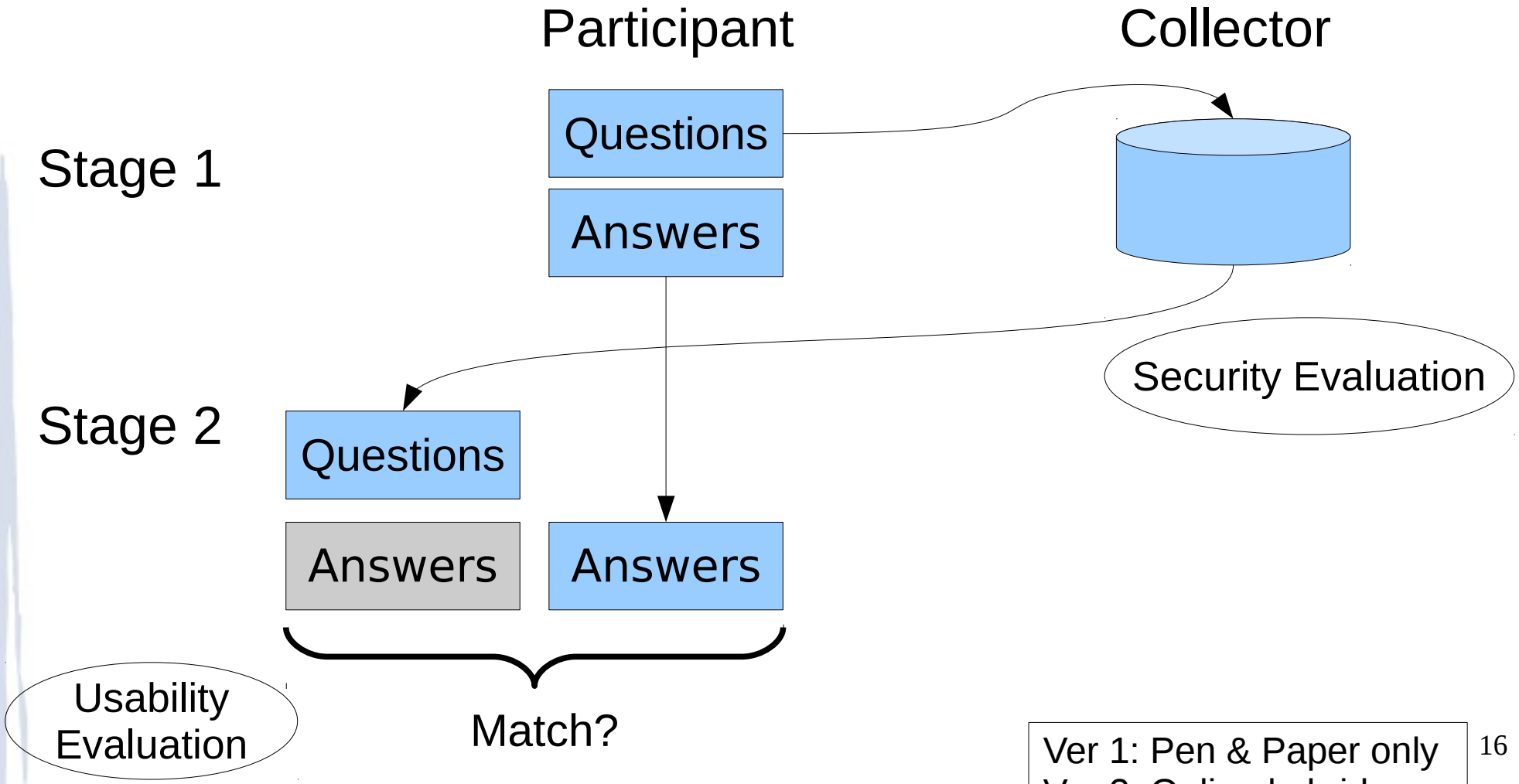
Data Collection

- Likely answer data
  - Purpose: To determine size of answer spaces
  - People, pet and place names
  - Collected source data of national statistics (and Facebook)
- Example question data
  - Purpose: To discover real user data
  - Led an experiment with 170 participants
  - Collected 500 user-chosen challenge questions

Authentication Experiment Challenges

- An ethical challenge to collect realistic data
- But often, users seem to readily submit
- Issues regarding participant behaviour
  - Equate credentials with other information?
  - Contribute *real* information?
  - Degree of freedom since *user-chosen*
- Opportunities for improved Collector behaviour
  - Challenge to ourselves: *Don't collect!*
  - Avoid having to maintain information
  - Consistent message: Keep your credentials private!

Experiment



Ver 1: Pen & Paper only
Ver 2: Online hybrid

Outline

- Challenge Question Authentication
- Research Work
  - Approach
  - Data Collection
  - Security Evaluation
  - Usability Evaluation
- Concluding Remarks

Security Model

- Security analysis had been very *ad hoc*

Blind Guess

- Attacker has no additional information
- Attack success ↔ Answer length

Focused Guess

- Attacker knows the challenge questions
- Attack success ↔ Size of answer space

Observation

- Attacker knows the user
- Attack success ↔ Availability of information

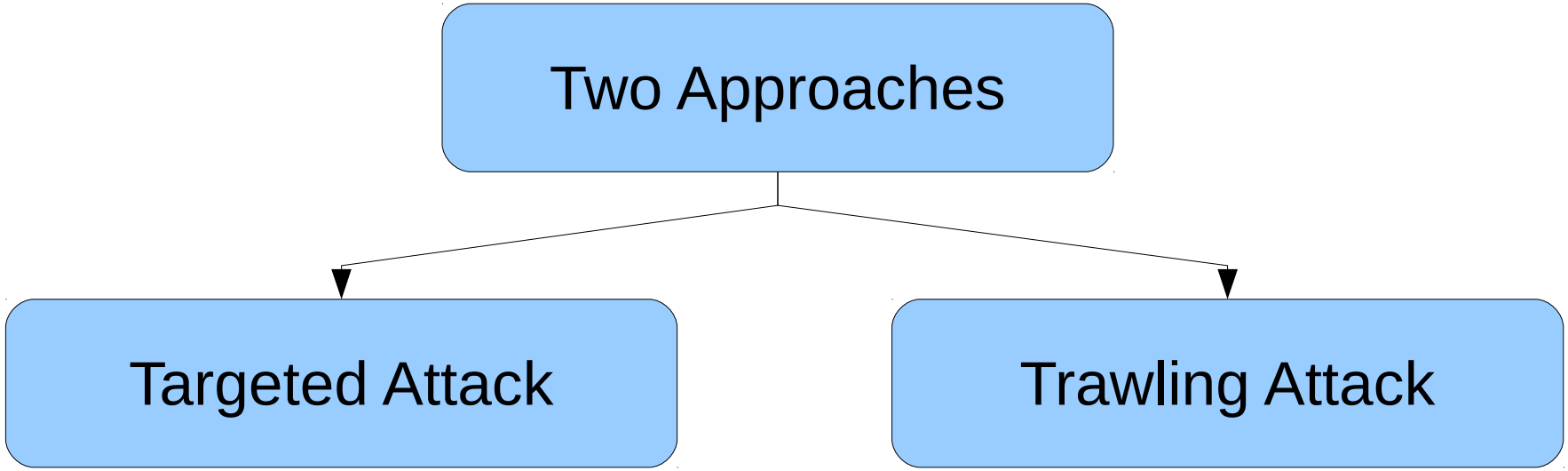
Security Model – Levels

- Security Levels are a baseline against which we evaluate we evaluate challenge questions
- Blind Guess and Focused Guess
  - Low: < 6-character alphabetic password (2^{34})
  - Medium: < 8-character alphanumeric password (2^{48})
  - High: ≥ 2^{48}
- Observation
  - Low: Answer publicly available
  - Medium: Answer not public, but known to Friends & Family
  - High: Neither

Security Evaluation – Blind

- Evaluating answers (with only the length)
  - *Assumption*: Alphabet of 26 lowercase letters
  - Entropy: 1.5 bits/char, but 2.3 for short text [Shannon]
  - *Answer entropy*: 2.3 bits (1st 8 chars), then 1.5 [NIST]
- Results by question (180)
  - Average answer length: 7.5 characters
  - Low (174) – Medium (4) – High (2)
- Results by user (60)
  - Q1: Low (59) – Medium (1) – High (0)
  - Q1,Q2: Low (38) – Medium (13) – High (9)
  - Q1,Q2,Q3: Low (5) – Medium (19) – High (36)

Security Evaluation – Focused



- Evaluated **experiment data** from rough estimates of answer space size
- Targeted attack against **specific user**

- Evaluated **source data** and measured likelihood of attack success
- Trawling attack succeeds for **any user**

Security Evaluation – Focused (2)

Targeted Attack

- Analysis of user-chosen questions and answers

Q Type	Freq	Space Est.
Proper Name	50%	10^4 – 10^5
Place	20%	10^2 – 10^5
Name	18%	10^3 – 10^7
Number	3%	10^1 – 10^4
Time/Date	3%	10^2 – 10^5
Ambiguous	6%	10^8 – 10^{15}

- Results by question (180)
  - Low (167) – Medium (0) – High (13)
- Results by user (60)
  - Q1: Low (58) – Medium (0) – High (2)
  - Q1,Q2: Low (46) – Medium (11) – High (3)
  - Q1,Q2,Q3: Low (5) – Medium (28) – High (27)

Security Evaluation – Focused (3)

Trawling Attack

- Security can be measured from likely answer data
- E.g., the distribution of surnames can be used for many questions, such as “What was your mother’s maiden name?”
- Data has shown that single questions are relatively insecure.
  - For example, US statistical data (2000) reveals only 150K surnames, 1.2K male first names and 4K female first names (1990).
- Three surnames from South Korea are used by 15% of users
- Pet names are harder to guess than first names

Security Evaluation – Focused (4)

Trawling Attack

- Analysis of national statistic data for people, pet & place names
- Shannon Entropy a poor estimate in this case
- We adapted other measures to better approximate the guesswork required of a trawling attacker
- With 3 guesses at each of multiple accounts, success rates increase greatly (e.g., success every 80 accounts)
- Observations
  - Pet names more difficult than (US) forenames
  - South Korea: Kim, Lee, Park → 50% of surnames
  - Knowing ethnicity can double attack effiency

Security Evaluation – Observation Guess

- Targeted Observation
  - Subjectively based upon an estimate of the availability of a particular answer
  - Querying the user as to the answer availability (but not accepting a user over-estimate)
  - Assessed empirically, by having other users pose as attackers to guess answers
- Empirically, answers are highly susceptible to guesses by family, friends, and even acquaintances
- Biggest threat to challenge question security

Security Evaluation – Observation Guess

- Recall criteria: Answer public? Answer known to F&F?
- Evaluating answers
  i. Subjective assessment
  ii. Participant input (upper bound only)
  iii. (Can also assess with *real attackers* – not done here)
- Results by question (180)
  - Low (124) – Medium (54) – High (2)
- Results by user (60)
  - Did not “sum” for multiple questions (used max)
  - Low (24) – Medium (34) – High (2)

Security Evaluation - Overall

- Overall rating: (Blind, Focused, Observation)
- Results from experimental data (60)
  - All Low (1)
  - All High (0)
  - No Lows (31 or 50%)
  - (H,H,M) or (M,H,M) (15 or 25%)
  - (H,H,M) (11 or 20%)
- Not all attack dependencies yet explored

Outline

- Challenge Question Authentication
- Research Work
  - Approach
  - Data Collection
  - Security Evaluation
  - Usability Evaluation
- Concluding Remarks

Usability Model

- Applicability
  - Sufficient information to register an answer?
  - E.g., “What was my first pet's name?”
- Memorability
  - Recall original answer over time?
- Repeatability
  - Precisely repeat original answer over time?
  - Syntactic: Correct spelling
  - Semantic: Changes over time, e.g., Favourites

Usability Evaluation

- Despite using “already known” answers, memorability & repeatability results are weak
- Results of 10% - 25% of failed authentication
  - Both for admin. and user chosen questions
  - Even for young participants (and memories)
- Possible reasons
  - Syntactic: Difficulty with precise recall
  - Semantic: Answers change over time
  - False answers

Evaluation – Summary

- Significant issues with the security and usability of challenge questions
- Key observations
  - Multiple questions improve security
  - Need novel approach to mitigating Observation
  - Improving usability is a big challenge
  - Current solutions are terribly boring

Outline

- Challenge Question Authentication
- Research Work
  - Approach
  - Data Collection
  - Security Evaluation
  - Usability Evaluation
- Concluding Remarks

Concluding Remarks

- Current challenge question solutions have numerous security and usability challenges
- Some remaining potential for authentication using personal knowledge
- Yet, longer term solutions are likely elsewhere (hardware, biometrics, “Someone you know”)
- Secure HCI is a useful interdisciplinary approach to traditional security problems