Challenging Challenge Questions

Mike Just
Visiting Research Fellow
(joint work with David Aspinall)

LFCS Lab Lunch
10 March 2009

What are Challenge Questions?

- A question-answer pair
- The answer is an *authentication credential* , similar to a password
- Common questions
  - *"What is my Mother's Maiden Name?"*
  - *"What was my first pet's name?"*
  - *"What was the name of my primary school?"*
- Typically used for fallback / secondary authentication

Challenge Question Research

1. Security (Attacker's point-of-view)
   - Difficult to determine the answers to the questions?
   - Demonstration of security uses *quantitative analysis*
2. Usability (User's point-of-view)
   - Easy to choose the questions?
   - Easy to remember the answers?
   - Demonstration of usability uses *qualitative research*

- Interdisciplinary work: Security and Human Factors

- Are challenge questions Secure? Usable?

Our Approach

- Very little published research
- We wanted a systematic approach (esp. security)
- Develop security and usability models
- We gathered some experimental data
  - *HCI Class (Oct/Nov 2008)*
  - *Computer Security Class (Jan/Feb 2009)*
  - *Biology Class (Jan/Feb 2009)*
  - 170 participants submitted 500 questions

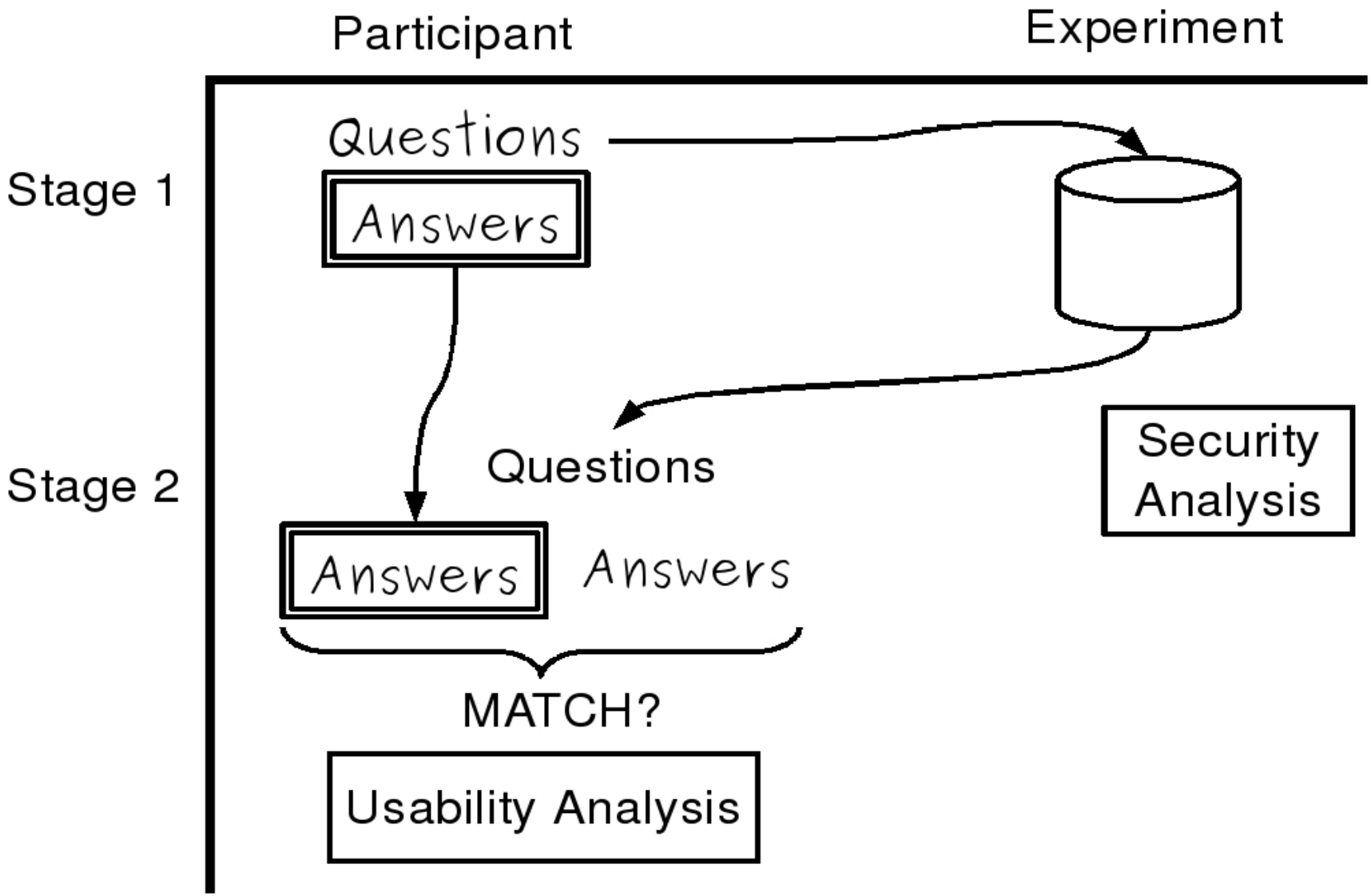
For Today's Lab Lunch Presentation

- The experimental method we used
- Our security model and analysis
  - Applied to LFCS question contributions
- Looking Ahead

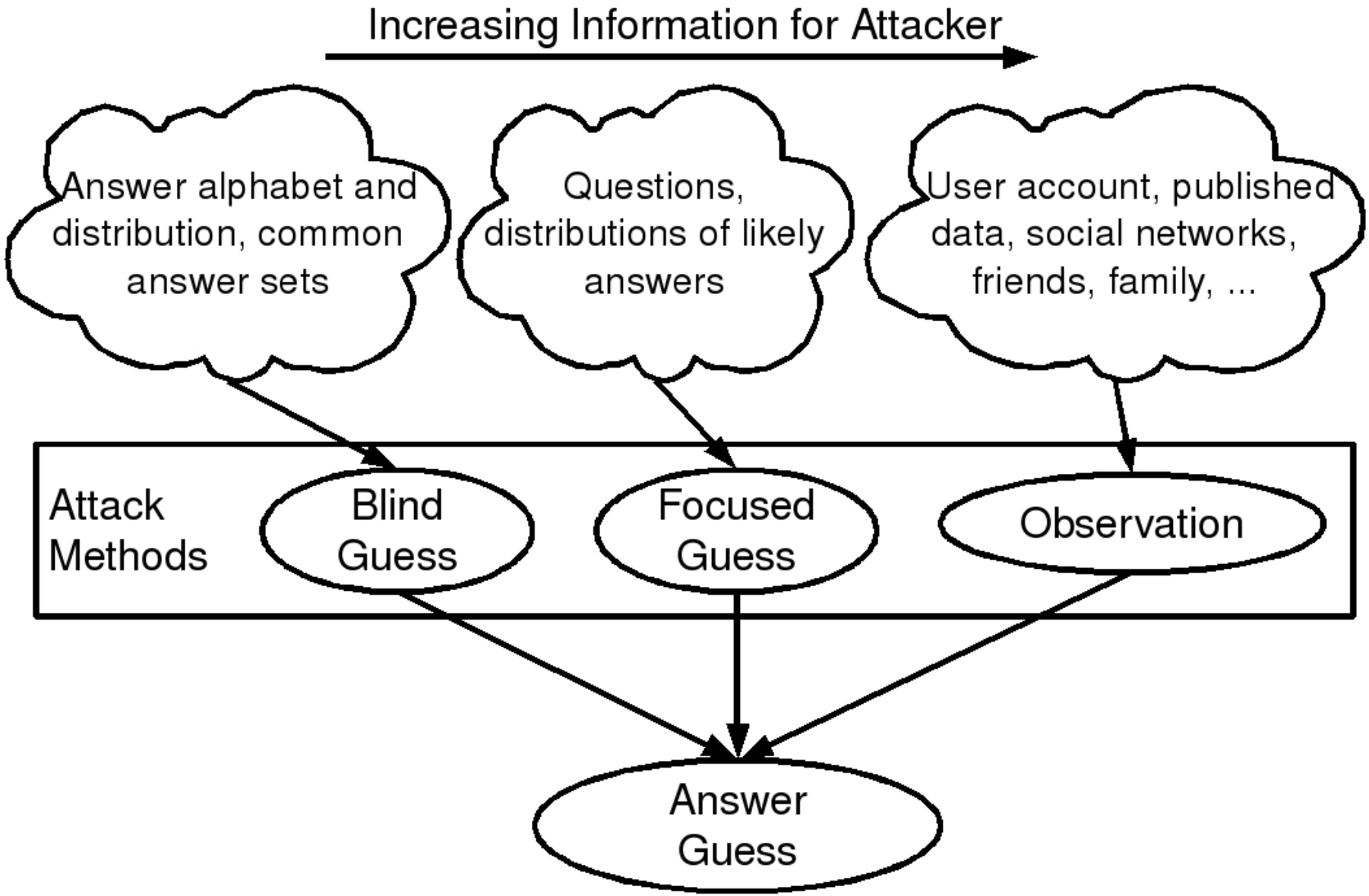
Challenge Question Experiment

- Collecting authentication data can be tricky
  - Users consistently told to not reveal information
  - Effective analysis requires this information
  - Ethically, we could ask for authentication information
  - But, if users participate, will they give real information?

Challenge Question Experiment (2)



Security Model and Analysis



LFCS Challenge Questions - Blind Guess

- Blind Guess - Brute force attack
- Security Levels based on equivalence to passwords
  - 6-character alphabetic password (2^34)
  - 8-character alphanumeric password (2^48)
- Security Levels
  - Low (2^34) Medium (2^48) High
- Answer entropy: 2.3 bits (first 8 chars), then 1.5 bits
- Blind Guess (16 LFCS participants - 48 questions)
  - Average answer length = 8.96 chars (median = 7)
  - By Question: 41 Low, 6 Medium, 1 High
- Blind Guess - By User
  - Q1: 11 Low, 5 Medium, 0 High
  - Q1, Q2: 6 Low, 3 Medium, 7 High
  - Q1, Q2, Q3: 1 Low, 3 Medium, 12 High

LFCS Challenge Questions - Focused Guess

- Attacker knows the Challenge Questions
- Security Levels same as for Blind Guess
- Answer types and space
  - Personal Names **44%** 10^4 - 10^5
  - Place 15% 10^2 - 10^5
  - Name 13% 10^3 - 10^7
  - Number **8%** 10^1 - 10^4
  - Time/Date 2% 10^2 - 10^5
  - Ambiguous **19%** 10^8 - 10^15
- Focused Guess by Question
  - 39 Low, 9 Medium, 0 High
- Focused Attack by User
  - Q1- 13 Low, 3 Medium, 0 High
  - Q1, Q2 - 4 Low, 8 Medium, 4 High
  - Q1, Q2, Q3 - 0 Low, 3 Medium, 13 High

LFCS Challenge Questions - Observation

- Attacker tries to obtain or observe the answer
- Security Levels defined qualitatively
  - Low - Answer publicly available and known to F&F
  - Medium - Answer known to F&F
  - High - Neither
- We assigned security levels by
  - Subjective analysis
  - Participant input (provided us an upper bound only)
- Observation by Question
  - 18 Low, 21 Medium, 9 High
- Observation by User
  - Q1 - 4 Low, 9 Medium, 3 High
  - Q1, Q2 - 3 Low, 8 Medium, 5 High
  - Q1, Q2, Q3 - 3 Low, 7 Medium, 6 High

LFCS Challenge Questions - Overall

- Simultaneous application of all attack methods
- Overall security rating is a 3-tuple
  - (Blind Guess, Focused Guess, Observation)
- Results (out of 16 participants)
  - All Low - 0 participant
  - All High - 5 participants (31%)
  - No Lows - 12 participants (75%)

Looking Ahead

- Longer-term usability studies
  - Current results showed memorability or repeatability issues for about 8% of the questions
- Independence of questions/answers
- Do users answer questions honestly?
- Dynamic answer assessment / User training
  - 'Authentication Aardvark'
- What security levels are tolerable?
  - Is 'Medium' or 'Low' sufficient against an Observation

Further Information

- Knowledge-Based Authentication Project
  - http://homepages.inf.ed.ac.uk/mjust/KBA.html
- Questions
  - Email: mike.just@ed.ac.uk
- Note
  - This version of the presentation has been shortened to remove specific questions provided by LFCS members. If you're an LFCS member and would like to discuss the original presentation, contact Mike at the above email.