Cryptanalysis of a Cognitive Authentication Scheme

Philippe Golle, PARC

David Wagner, UC Berkeley

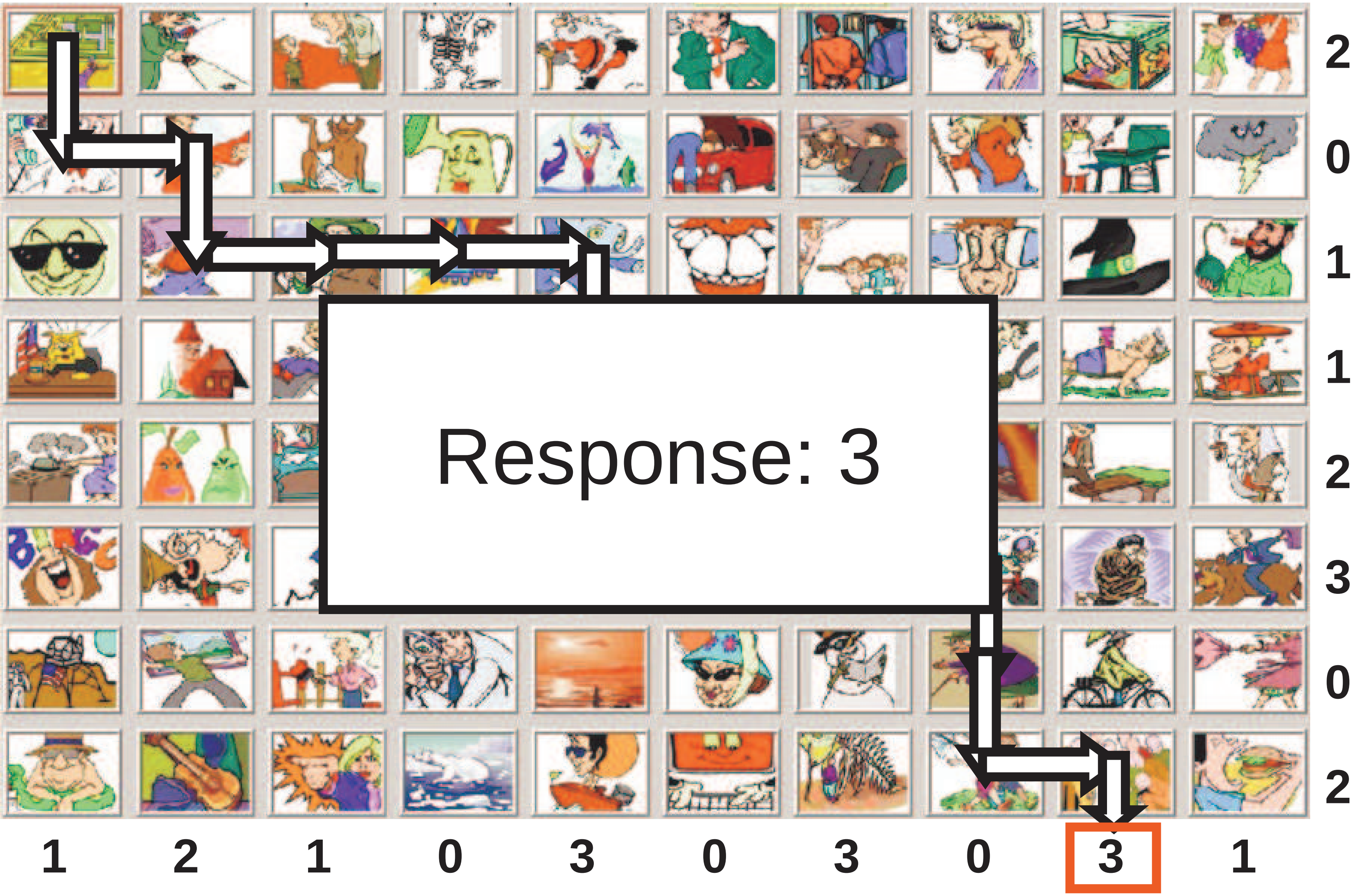
Problem Statement

- How can I log into my bank without keyloggers/eavesdroppers stealing my credentials?

A recent proposal [Weinshall]

- Server has a set of 80 images
- My secret is a subset of 30 images I recognize
- Protocol performs 10 rounds of challenge-response authentication
  - Server asks question about the shared secret
  - Human responds

A Round of Challenge/Response



Response: 3

2
0
1
1
2
3
0
2

1 2 1 0 3 0 3 0 3 1

Cryptanalysis

- Associate a boolean variable x_i to each image
  - 80 boolean variables $x_1, \ldots, x_{80}$
- For each known challenge-response pair, write a SAT formula expressing that $x_1, \ldots, x_{80}$ are consistent with this pair
- Apply an off-the-shelf SAT solver

- Result: Reveals the secret after observing 10 authentications and 7 seconds of CPU time

Parting Thoughts

- Advice to cryptanalysts:
  For schemes that have small circuits, try applying a SAT solver

- More details: eprint.iacr.org/2006/258/