Exploring Extrinsic Motivation for Better Security:
A Usability Study of Scoring-Enhanced Device Pairing

Alexander Gallego
alex@yieldmo.com
YieldMo, Inc.

Nitesh Saxena
saxena@cis.uab.edu
University of Alabama at
Birmingham
Department of Computer and
Information Science

Jonathan Voris
jvoris@cs.columbia.edu
Columbia University
Department of Computer
Science

COLUMBIA | ENGINEERING
The Fu Foundation School of Engineering and Applied Science

Motivation

- Mobile platform security risks:
  - Eavesdropping
  - Man-in-the-middle attacks
- Examples
  - Bluetooth headset eavesdropping
  - "Evil twin" access point in a cyber cafe

The Device Pairing Problem

- How can a secure channel be initialized between two wireless devices with:
  - No prior association or shared secrets
  - No common trust source
- Example device pairing scenarios:
  - Tablet with a wireless gateway
  - Controller with a game system
  - Cell phone with a Bluetooth headset

Out of Band Device Pairing

- Utilize an out of band (OOB) channel between devices
  - Created using discernible output modalities:
    - Audio
    - Visual
    - Haptic
- OOB channels are physically authenticable
  - Human perceptible

Out of Band Pairing Issues

- Places a usability burden on users
  - Tedious and time consuming
- Users may commit errors
  - Impacts usability
  - Potentially compromises security
- Process is limited by human factors
  - Slow enough for human to follow
  - Maximum security dependent on attention span
- Can users be encouraged to perform better?

The Tom Sawyer Effect

- "The Adventures of Tom Sawyer" by Mark Twain

'AIN'T THAT WORK?'

– Tom is forced to paint a fence
– He pretends to enjoy it rather than resent it
– Friends insist on helping!

COLUMBIA ENGINEERING
The Fu Foundation School of Engineering and Applied Science

The Tom Sawyer Effect

- Apply this principle to security tasks
  - Reframe pairing as a game
  - Introduce gaming elements
- Competitiveness between individuals leads to better performance
  - Example: Random number generation and recognition – Halprin and Naor [SOUPS '09]

Extrinsic Motivation

- Psychological principle
- External incentive provided via a reward system
- Motivators found in games:
  - Competitiveness
  - Scores
  - Timing
  - Emotional rewards
    - Storytelling
    - Character development

Previous Work: Pairing Games

- Designed a game for device pairing to test Tom Sawyer Effect

- Inspired by Hasbro's Simon
  - Well known
  - Easy to understand
  - Suitable for wide range of players
  - Closely related to existing pairing solutions

Adrienne Barr vidgame.net © 2003

Previous Work: Pairing Games

- Implemented on two mobile phones
- Pairing data encoded as sequences of:
  - Brightened colored squares
  - Harmonic tones
- One player mode trains unwitting users
- Two player mode achieves device pairing

COLUMBIA ENGINEERING
The Fu Foundation School of Engineering and Applied Science

Pairing Game Usability

- Performed a within subjects usability study
  - Pairing game
  - Standard numeric transfer

- Positive results:
  - Game more enjoyable
  - Users paired longer patterns successfully

- Negative results:
  - Game took too long!

Efficient Use of Game Elements

- Do individual game elements provide any extrinsic motivation?
- Motivators found in games:
  - Competitiveness
  - **Scores**
  - Timing
  - Emotional rewards
    - Storytelling
    - Character development

New Pairing Methods

- Designed two new pairing methods
  - Plain Comparison
  - Scored Comparison
- Both used:
  - Four color quadrants
 - Five digit numbers in each
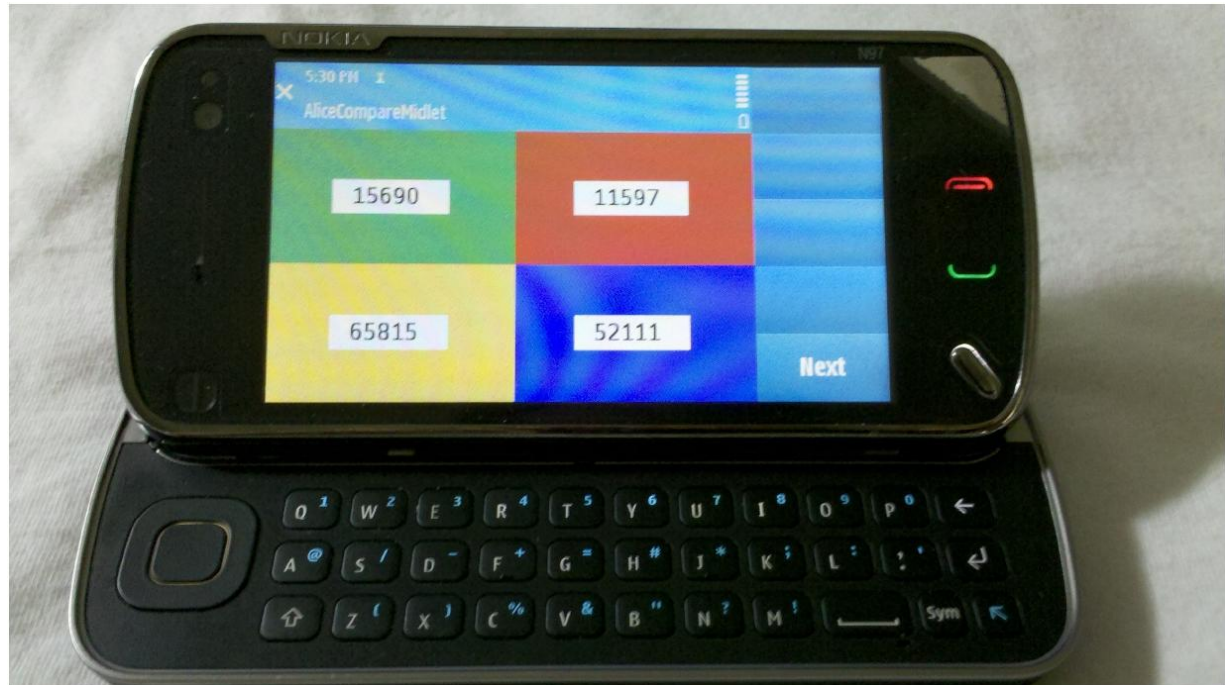
Screenshot of Scored Comparison on a Nokia N97 Phone
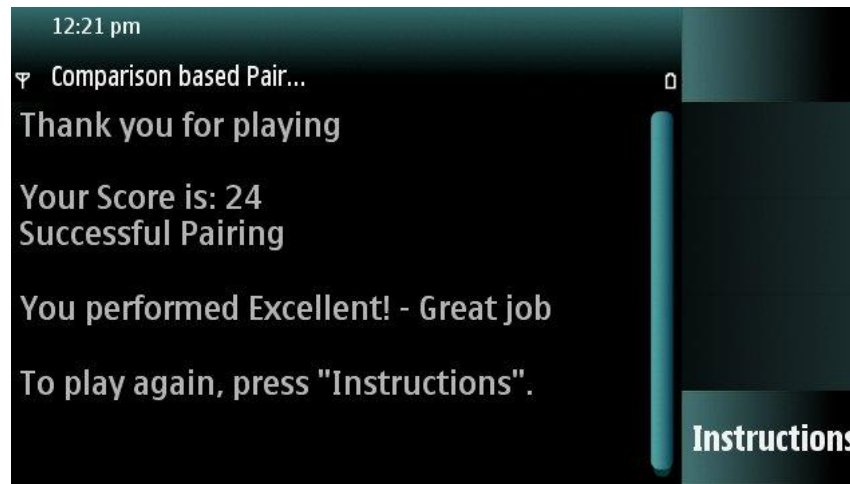
Scoring Pairing Performance

- Want to quantify how well users did
  - But devices unable to verify pairing values
- Solution: Insert scoring values
  - Indistinguishable from pairing values
  - Approximate users' ability to match pairing values

	Pairing Values	Scoring Values
Impact on pairing success/failure?	**Yes**	No
Impact on score?	No	**Yes**

New Pairing Methods
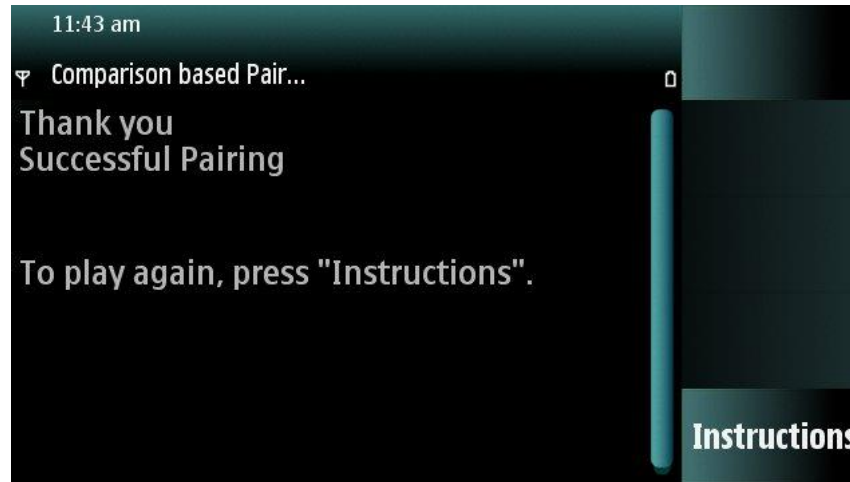
Plain Comparison Result Screen

Scored Comparison Result Screen

COLUMBIA ENGINEERING
The Fu Foundation School of Engineering and Applied Science

Scored Pairing User Study

- Performed a between subjects study
  - Two groups of 21 users
- Each participant used one of the pairing methods five times
- Asked to complete post conditional questionnaire
  - Demographic information
  - System Usability Scale (SUS) questions

User Demographics

Demographic Information	Plain Comparison	Scored Comparison
Age		
17 – 25	52.4%	52.4%
26 – 29	33.3%	38.1%
30 – 40	14.3%	9.5%
Gender		
Male	42.9%	61.9%
Female	57.1%	38.1%
Education		
High School	14.2%	19.0%
Bachelor's	28.6%	38.1%
Master's	52.4%	33.3%
Doctorate	4.8%	9.5%

Study Participant Demographic Data

COLUMBIA ENGINEERING
The Fu Foundation School of Engineering and Applied Science

System Usability Scale

- Ten questions on a five point Likert scale
    1. I think that I would like to use this method frequently.
    2. I found the method unnecessarily complex.
    3. I thought the method was easy to use.
    4. I think that I would need the support of a technical person to be able to use this method.
    5. I found the various functions in this method were well integrated.
    6. I thought there was too much inconsistency in this method.
    7. I would imagine that most people would learn to use this method very quickly.
    8. I found the method very cumbersome to use.
    9. I felt very confident using the method.
    10. I needed to learn a lot of things before I could get going with this method.

Study Results: Efficiency

	Plain Comparison		Scored Comparison	
	Execution Time (seconds)	Standard Deviation	Execution Time (seconds)	Standard Deviation
Phone A	16.4	13.3	22.0	11.2
Phone B	16.7	13.6	22.1	11.3
Average	**16.6**	13.5	**22.1**	11.3

- Scored Comparison was slower overall
- Twice as many comparison tasks in Scored
  - Individual values compared more rapidly

Study Results: Error Rates

	Plain Comparison		Scored Comparison	
	Safe Error Rate	Fatal Error Rate	Safe Error Rate	Fatal Error Rate
Phone A	6.0%	6.0%	1.5%	1.8%
Phone B	7.0%	3.7%	2.2%	3.8%
Average	**6.5%**	**4.8%**	**1.8%**	**2.8%**

- Presence of a score reduced error rates substantially

Study Results: User Feedback

	Plain Comparison		Scored Comparison	
	SUS Score	Standard Deviation	SUS Score	Standard Deviation
Average	69.2	**15.2**	74.3	**11.7**

- Both methods assessed positively
- More agreement among users of Scored Comparison

Conclusions

- Providing a score as incentive increased users' awareness of pairing decisions

- Individual game elements may help users complete security tasks
  - Without too much efficiency impact

- Future work: Apply concept to other usable security problems

COLUMBIA ENGINEERING
The Fu Foundation School of Engineering and Applied Science

Thank you!