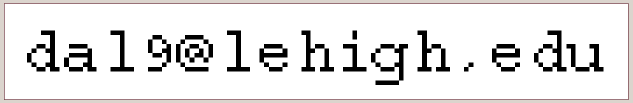
Worldwide Threats: Converging Dangers in a Post 9/11 World

Daniel Lopresti

Department of Computer Science and Engineering



dal9@lehigh.edu *

Moderator: Ron Yoshida, Provost

Panelists: Rick Blum, Electrical & Computer Engineering; Dan Lopresti, Computer Science & Engineering; Janice Bially Mattern, International Relations; Paul McHale, Assistant Secretary of Defense for Homeland Defense; Al Romig '75, VP, Sandia Labs

* *Why do I display my email address in image format? Answer to follow shortly …*

CSE Dept. & Security Research

Faculty in Computer Science & Engineering are engaged in various research projects connected to homeland security and related areas.

- Cybersecurity: data security (biometrics), network security (preventing DoS attacks), e-commerce security (CAPTCHA's).
- Knowledge acquisition from "noisy" or otherwise challenging sources: document analysis, information retrieval, text and data mining, semantic web, heterogeneous database systems.
- Other applications: mobile robotics, distributed sensor networks.

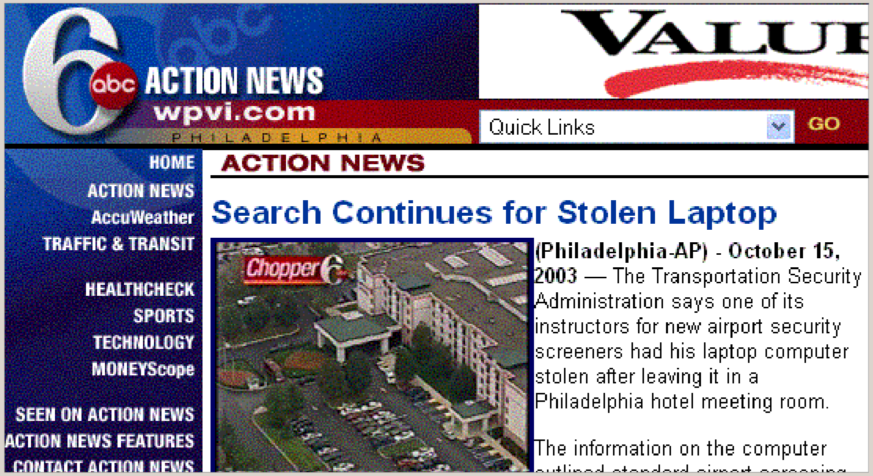
* Topics I'll discuss today.

Protecting Mobile Data

Data is becoming more portable (PDA's, cell phones, laptops, etc.) and theft is a growing concern.

Why aren't passwords enough?

- Very easy to “crack.”
- Thief can just disassemble and reverse-engineer device.



ACTION NEWS

Search Continues for Stolen Laptop

(Philadelphia-AP) - October 15, 2003 — The Transportation Security Administration says one of its instructors for new airport security screeners had his laptop computer stolen after leaving it in a Philadelphia hotel meeting room.

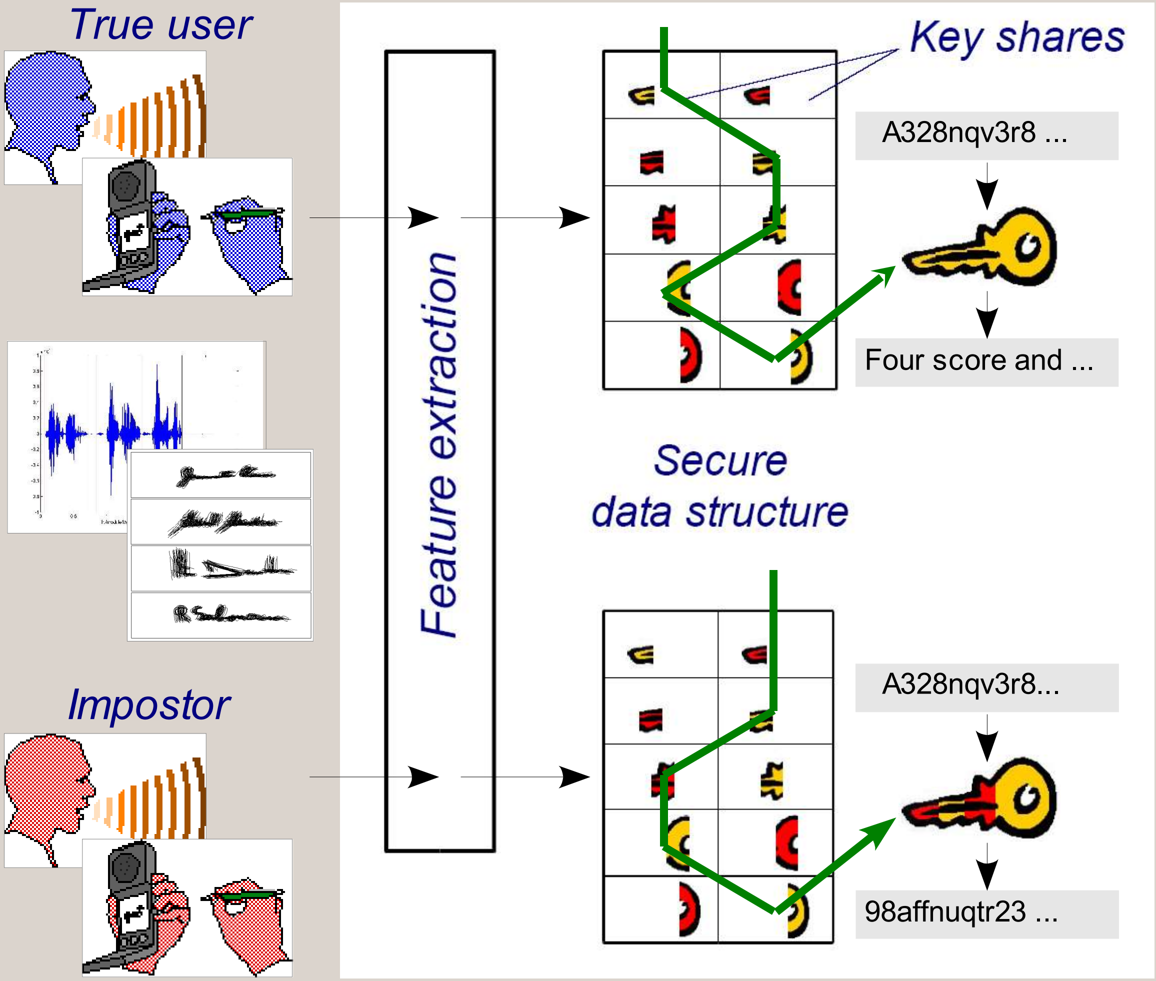
The information on the computer outlined standard airport screening

Two-pronged solution:

- Use biometrics in place of (or in addition to) passwords.
- Use secure data structure to encrypt information.

Using Biometrics to Protect Data

- Cryptographic key broken into shares and mixed with random data.
- Features extracted from user's speech or handwriting.
- Only input from true user will select correct shares to yield proper key.



True user

Key shares

A328nqv3r8 ...

Four score and ...

Feature extraction

Secure data structure

Impostor

A328nqv3r8...

98affnuqtr23 ...

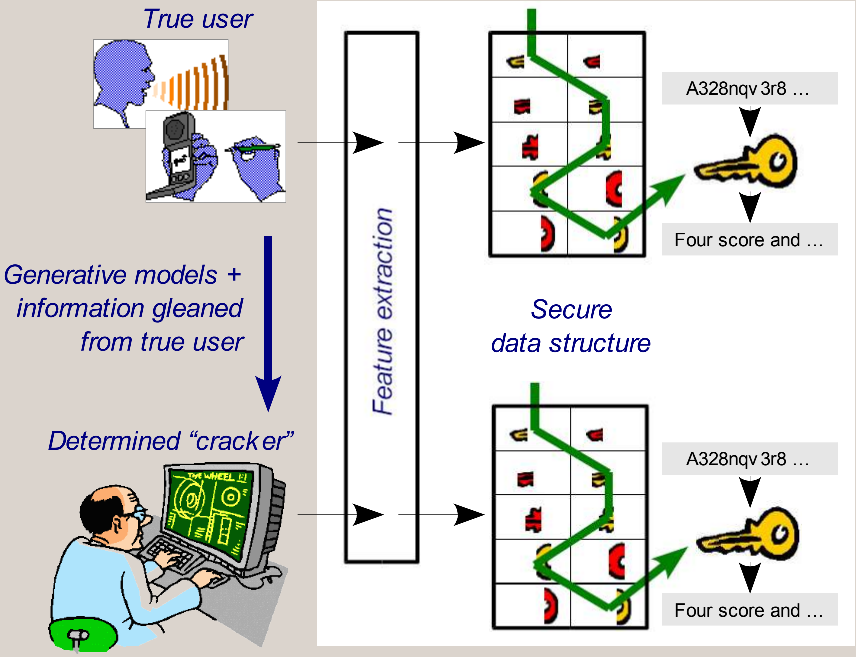
Using Biometrics to Protect Data

Current research:

- Quantify effectiveness.
- Increase number of bits.
- Identify potential attacks.

Biometrics may be vulnerable:

- Study generative models.
- If successful, many current systems called into doubt.

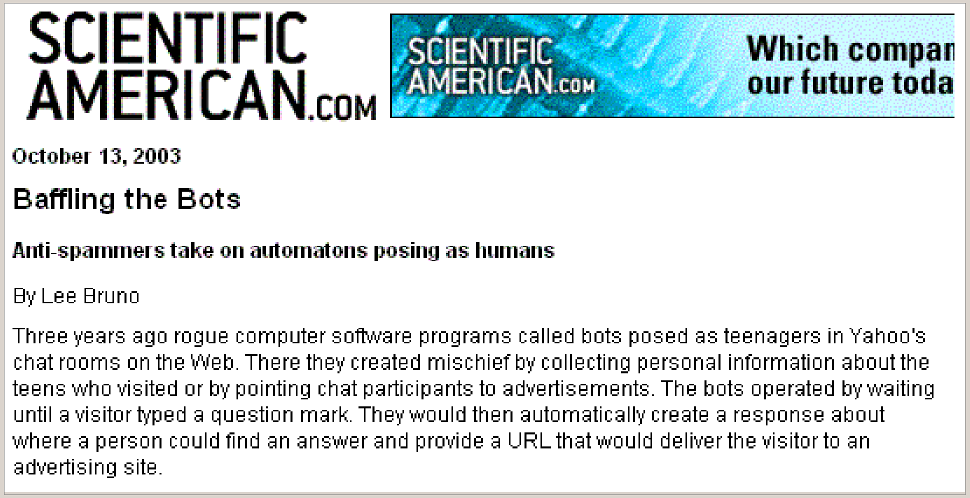


Use our experience to improve biometrics, increase security.

Protecting Online Services

The Internet has become vehicle for distributing valuable content. But malicious programs ("bots") attempt to exploit online services intended for human users.

Idea: create a pattern recognition task that is easy for humans to solve, but hard for machines.



October 13, 2003

Baffling the Bots

Anti-spammers take on automatons posing as humans

By Lee Bruno

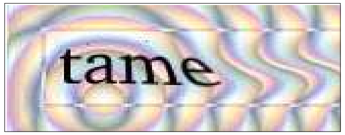
Three years ago rogue computer software programs called bots posed as teenagers in Yahoo's chat rooms on the Web. There they created mischief by collecting personal information about the teens who visited or by pointing chat participants to advertisements. The bots operated by waiting until a visitor typed a question mark. They would then automatically create a response about where a person could find an answer and provide a URL that would deliver the visitor to an advertising site.

CAPTCHA = *Completely Automated Public Turing test to tell Computers and Humans Apart.*

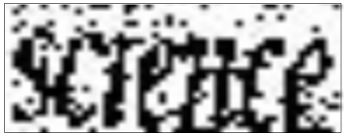
Visual CAPTCHA's

Currently, most CAPTCHA's exploit gap in reading ability between humans and machines when confronted with degraded images of text.

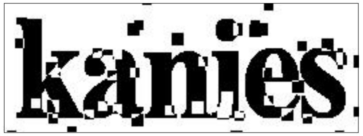
Fortunately, one of our latest hires is Henry Baird, a world expert on optical character recognition and an originator of this line of research.



EZ-Gimpy



PessimalPrint



BaffleText

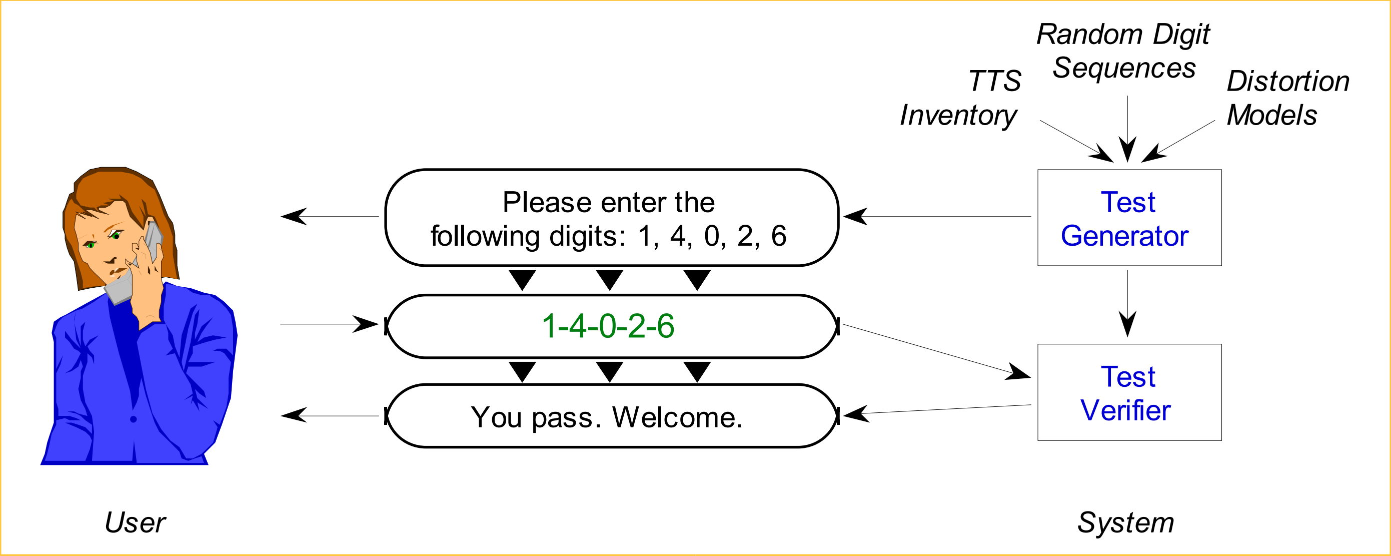
Topic has also attracted strong commercial interest: Yahoo!, PayPal, Microsoft, Google, Avaya, etc.

Spoken Language CAPTCHA's

Speech interfaces are becoming popular, raising similar issues:

- Phone access to financial services, reservations, etc.
- Hands- and/or eyes-free applications (e.g., driving car).
- Internet telephony will soon see SPAM (SPIT) and virus attacks.

A spoken language CAPTCHA:



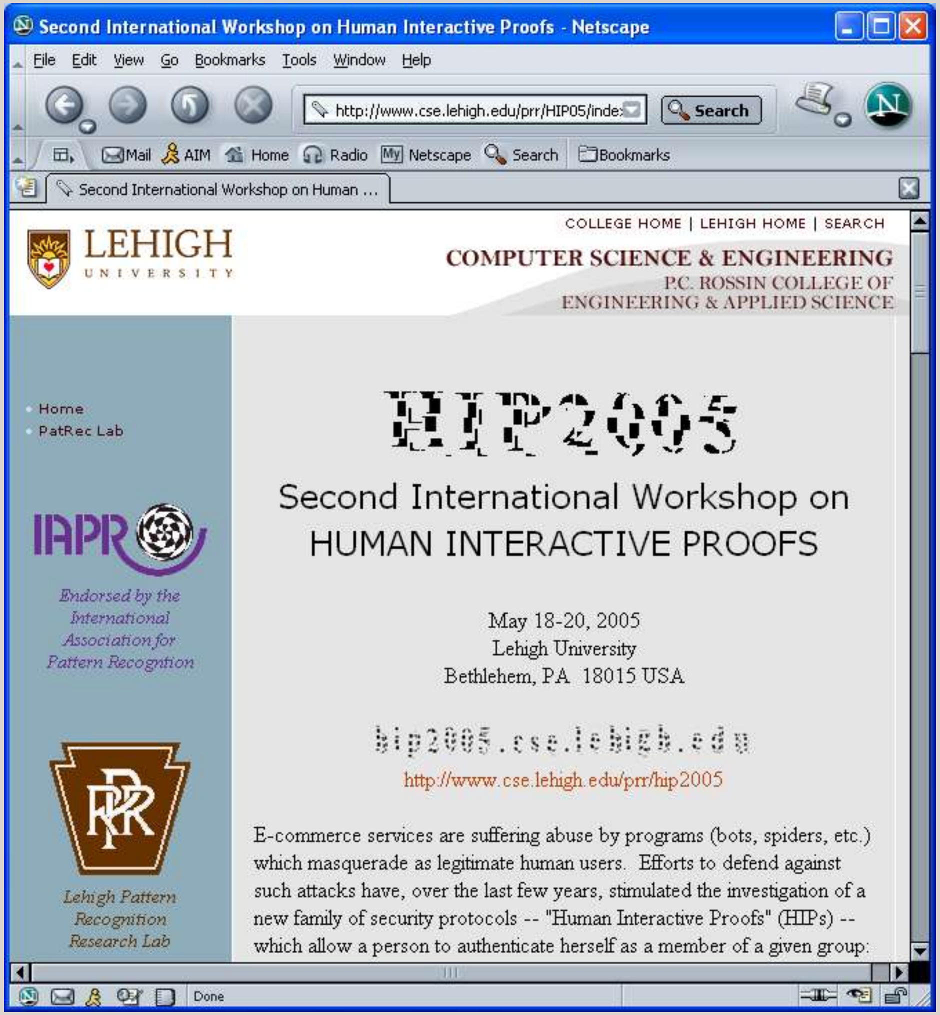
HIP 2005 Workshop at Lehigh

Second International Workshop on Human Interactive Proofs will take place at Lehigh next Spring (co-chaired by Henry Baird and Dan Lopresti).

If interested in more details, please see me or send email:

dal9@lehigh.edu *

* This is why I display my email address in image format – to foil bots!


Second International Workshop on Human Interactive Proofs - Netscape
File Edit View Go Bookmarks Tools Window Help
http://www.cse.lehigh.edu/prr/HIP05/index
Search
Mail AIM Home Radio Netscape Search Bookmarks
Second International Workshop on Human ...
COLLEGE HOME | LEHIGH HOME | SEARCH
LEHIGH UNIVERSITY
COMPUTER SCIENCE & ENGINEERING
P.C. ROSSIN COLLEGE OF ENGINEERING & APPLIED SCIENCE
Home
PatRec Lab
IAPR
Endorsed by the International Association for Pattern Recognition
Lehigh Pattern Recognition Research Lab
Done

HIP2005

Second International Workshop on HUMAN INTERACTIVE PROOFS

May 18-20, 2005
Lehigh University
Bethlehem, PA 18015 USA

hip2005.cse.lehigh.edu
http://www.cse.lehigh.edu/prr/hip2005

E-commerce services are suffering abuse by programs (bots, spiders, etc.) which masquerade as legitimate human users. Efforts to defend against such attacks have, over the last few years, stimulated the investigation of a new family of security protocols -- "Human Interactive Proofs" (HIPs) -- which allow a person to authenticate herself as a member of a given group:

http://www.cse.lehigh.edu/prr/HIP05/index.html

Real-World Knowledge Acquisition

Important intelligence sources include: HUMINT (captured documents), SIGINT (intercepted faxes), OSINT (websites).

Original text

Kingdom To Sign Nuclear Non-proliferation Treaty

Saudi Arabia on Tuesday decided to sign the nuclear weapons non-proliferation treaty, a strong indication it will not seek nuclear warheads for intermediate-range missiles it recently acquired from China.

The official Saudi Press Agency reported that King Fahd made the decision during a Cabinet me

Errors that arise during optical character recognition (OCR) have a devastating impact on information extraction.

OCR of light photocopy

Kin.adom To Si@n Nuclear Non-pi-oliferation Treat),

Saudi Arabia on Tuesday decided to si.-n the nuclear weapons non-proliferation treatv, a strone indication it will not seek nuclear A,arheads for inte,-mediate-ran@e missiles it rccent]3, acquired from China.

The official Saudi Press Aacnc), reported that Kinc., Fahd made, the decision durin@ a Cabinet mectin- in Riyadh, the Saudi capital-

How Big is the Problem?

Some quotes from those in-the-know (government end-users):

"One billion pages – ink on paper – to be manually reviewed. Can't find/retain enough people to do it. Need far greater automation."

"We have an urgent need to automate processing of printed and handwritten documents in a wide variety of languages and, especially, with low image quality."

"U.S. Government is collecting vast databases of heterogeneous multimedia data, including scanned images of printed and handwritten paper documents: content must be extracted from them far more accurately to allow automated downstream processing."

Research Issues Under Study in CSE

Some of the challenges faced by law enforcement, military, and intelligence-gathering agencies we are attempting to address:

- "Brittleness" of current OCR systems: needed are robust techniques that handle noisy inputs, degrade gracefully.
- Implications for later-stage text analysis / information extraction.
- Attempts to hide identity through use of aliases (or perhaps natural spelling variations) by criminals and terrorists.
- Mining intelligence from large collections of documents.
- Integrating knowledge from disparate sources (databases assembled by different agencies).