The Effectiveness of Generative Attacks on a Handwriting Biometric

Daniel P. Lopresti *Jarret D. Raim*

Computer Science & Engineering
Lehigh University
Bethlehem, PA 18015, USA

lopresti@cse.lehigh.edu

Motivation

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

Why aren't passwords enough?

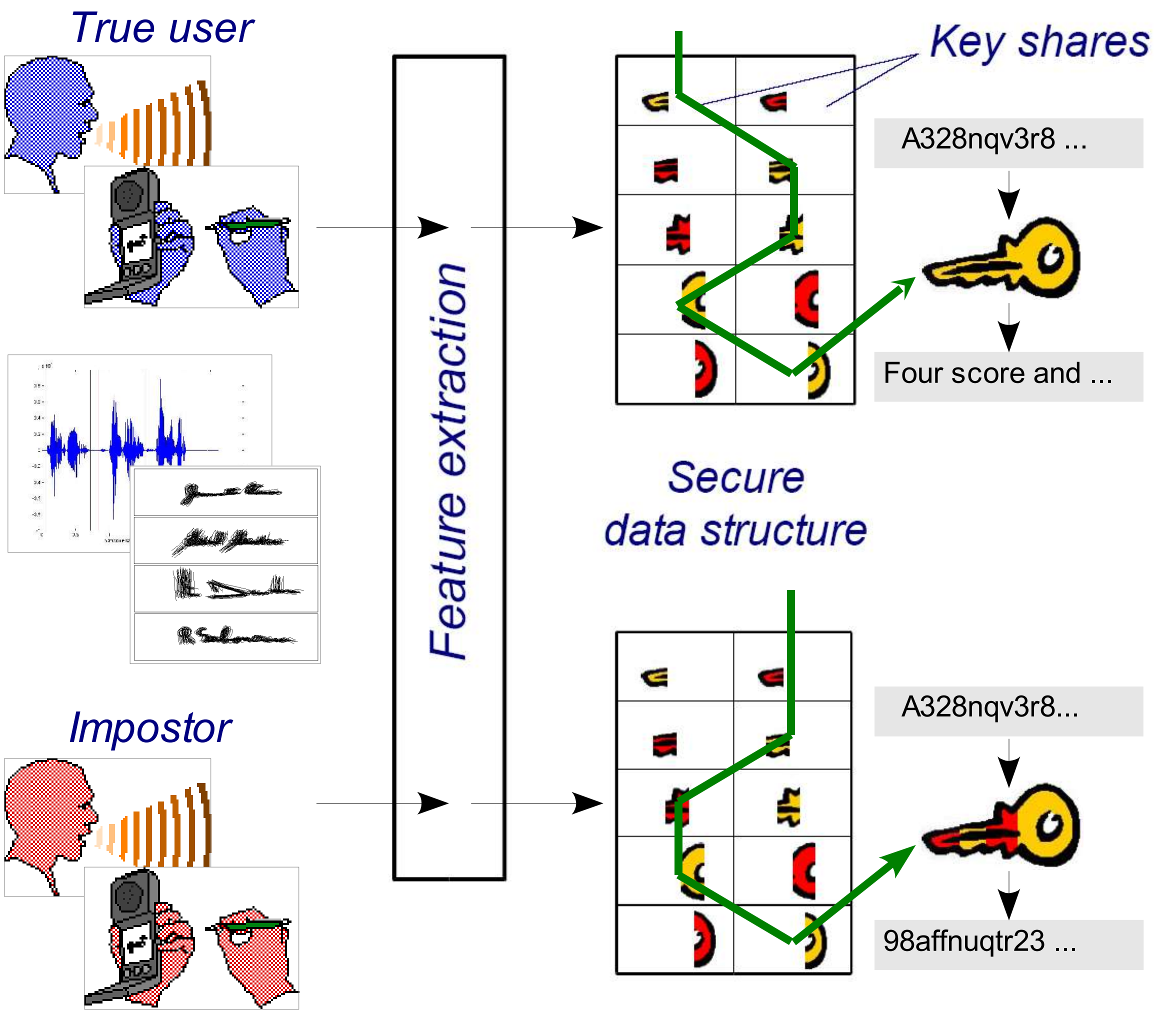
- Very easy to “crack.”
- Thief can disassemble and reverse-engineer device.

Two-pronged solution:

- Biometrics in place of (or in addition to) passwords.
- 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 selects shares to yield key.



"Towards Speech-Generated Cryptographic Keys on Resource-Constrained Devices," F. Monrose, M. Reiter, Q. Li, D. Lopresti, and C. Shih, *Proceedings of the Eleventh USENIX Security Symposium*, August 2002, San Francisco, CA, pp. 283-296.

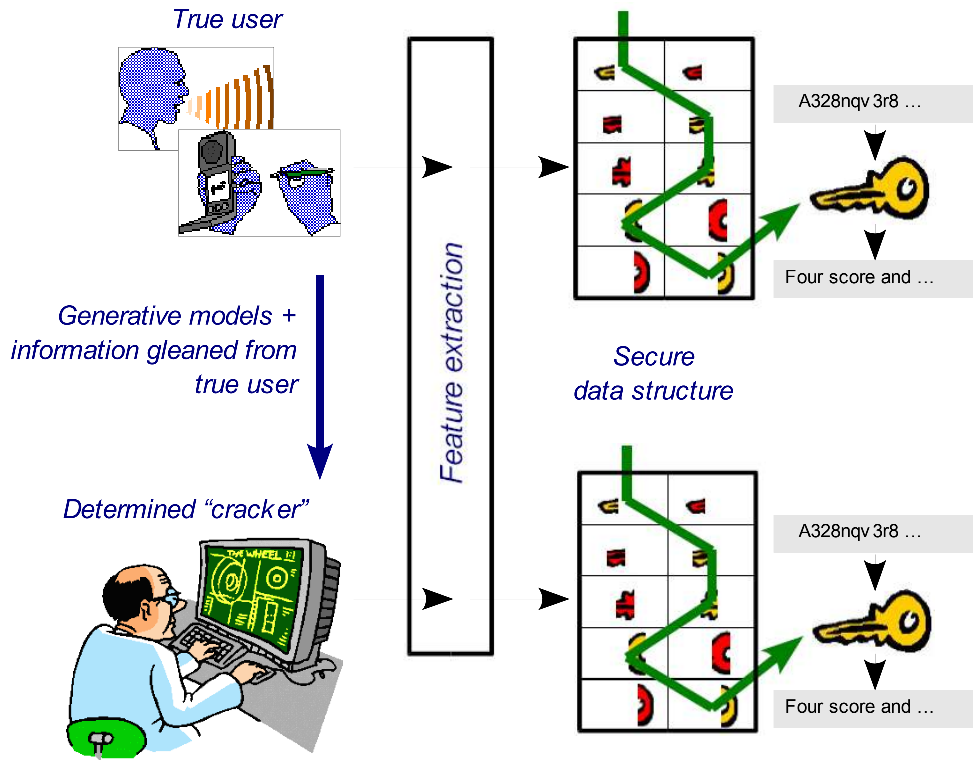
Using Biometrics to Protect Data

Biometrics may be vulnerable:

- Generative models can mimic human behavior.
- If successful, some systems breakable.

Our work:

- Identify potential attacks.
- Analyze risk.



Use our experience to improve biometric security.

Past Work: Speech-Generated Keys

Voice is natural user interface for many devices:

- Keyboard not an option in some cases.
- Unlike static biometrics, passphrases are unlimited.

Main criteria:

- Key (re)generation should be reliable and efficient on resource-constrained devices.
- Key search should be difficult for attacker, even with captured device.

"Towards Speech-Generated Cryptographic Keys on Resource-Constrained Devices," F. Monrose, M. Reiter, Q. Li, D. Lopresti, and C. Shih, *Proceedings of the Eleventh USENIX Security Symposium*, August 2002, San Francisco, CA, pp. 283-296.

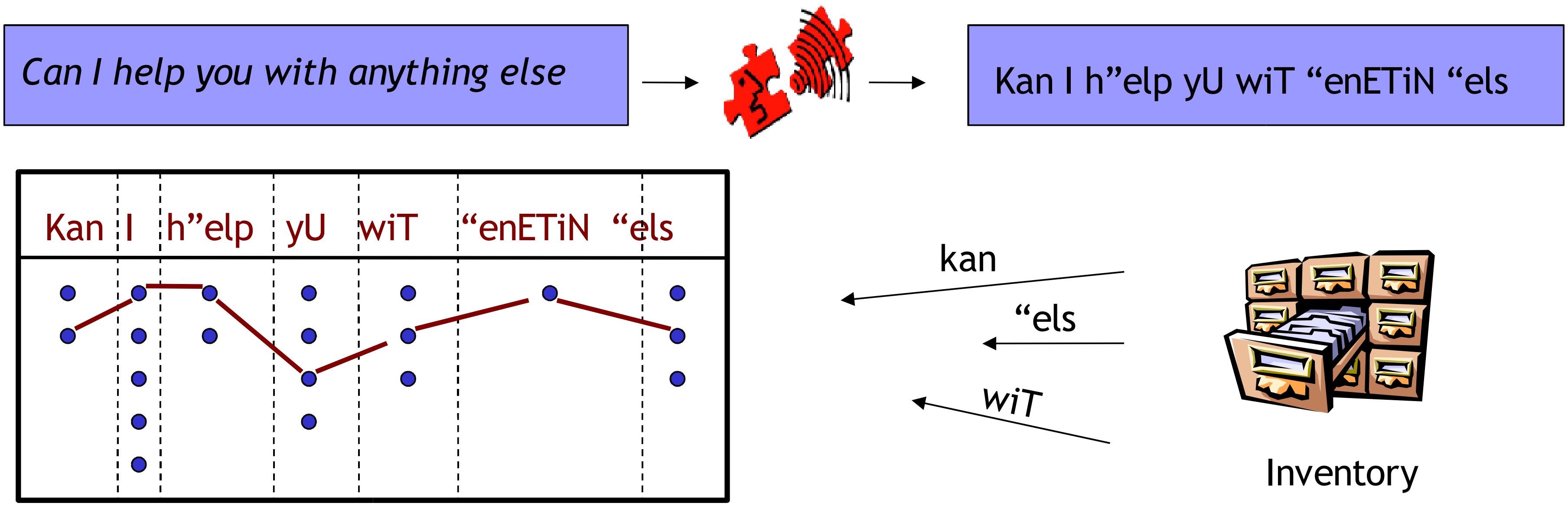
Evaluating Speech-Generated Keys

- Annotated inventory of 1,600 sentences (approx. one hour of speech) recorded by professional voice talent under controlled conditions.
- Five passphrases from same speaker collected one year later (approx. 38 mins of speech).
- Offers opportunity to synthesize candidate passphrases. Our first attempt to answer question:

Is user's key weakened by attacker gaining recordings of user saying phrases other than passphrase?

"Towards Speech-Generated Cryptographic Keys on Resource-Constrained Devices," F. Monrose, M. Reiter, Q. Li, D. Lopresti, and C. Shih, *Proceedings of the Eleventh USENIX Security Symposium*, August 2002, San Francisco, CA, pp. 283-296.

Text-to-Speech Attacks



- Nice, smooth-sounding speech.
- Duration and pitch predicted by TTS backend.
- Poor-quality predictions can impede attack.

"Towards Speech-Generated Cryptographic Keys on Resource-Constrained Devices," F. Monrose, M. Reiter, Q. Li, D. Lopresti, and C. Shih, *Proceedings of the Eleventh USENIX Security Symposium*, August 2002, San Francisco, CA, pp. 283-296.

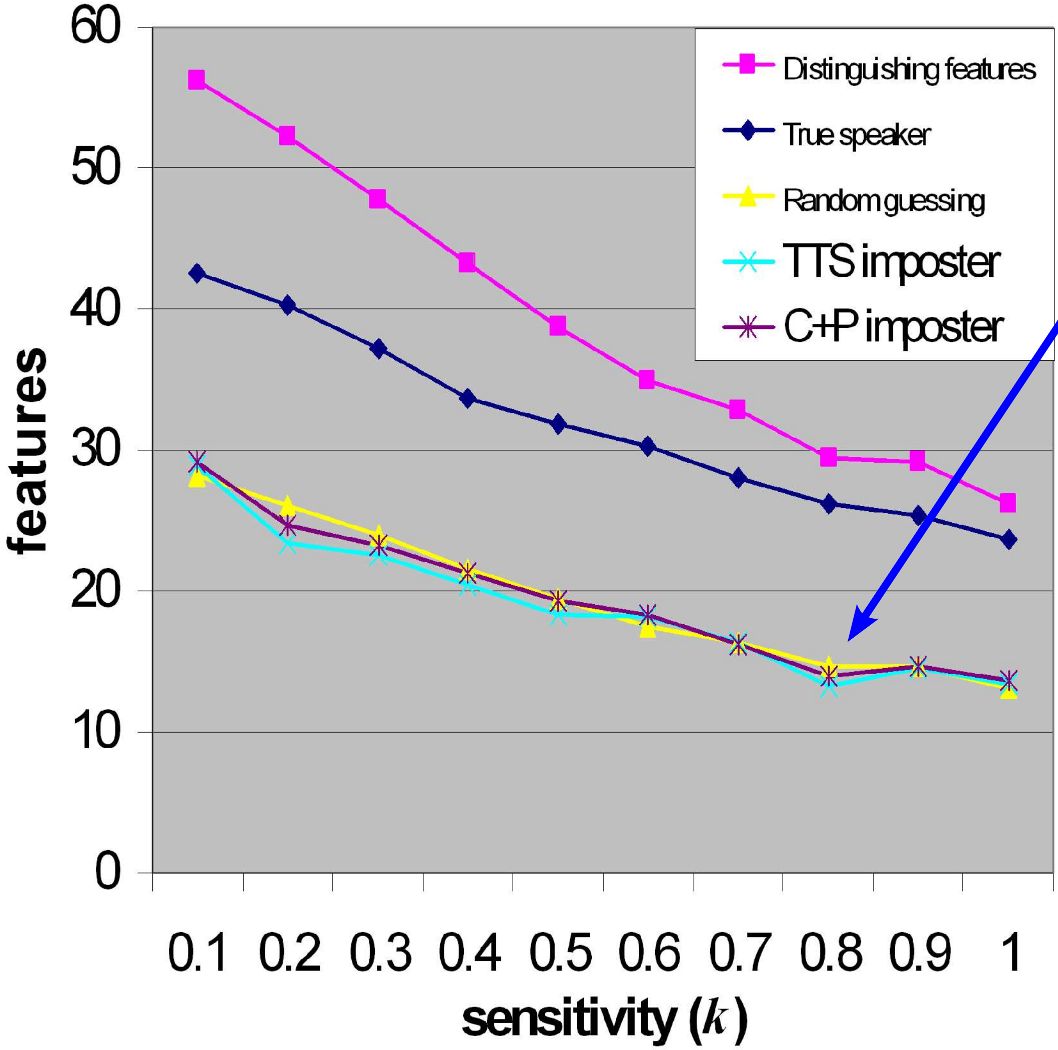
How Much Speech is Needed?

A measure of effort required for generative attack:

Passphrase		Speech Inventory Required	
Number	*Phonemes*	*Sentences*	*Minutes*
0	24	340	13.42
1	52	455	17.85
2	29	1279	51.75
3	27	152	6.12
4	18	415	15.86

"Towards Speech-Generated Cryptographic Keys on Resource-Constrained Devices," F. Monrose, M. Reiter, Q. Li, D. Lopresti, and C. Shih, *Proceedings of the Eleventh USENIX Security Symposium*, August 2002, San Francisco, CA, pp. 283-296.

Results of Text-to-Speech Attacks



TTS is no better than random guessing. Why?

- Speech synthesis too immature at this point.
- We just didn't have enough data.

Either way, we expect attacks to become more worrisome over time.

"Towards Speech-Generated Cryptographic Keys on Resource-Constrained Devices," F. Monrose, M. Reiter, Q. Li, D. Lopresti, and C. Shih, *Proceedings of the Eleventh USENIX Security Symposium*, August 2002, San Francisco, CA, pp. 283-296.

Present Investigations

Is same true for handwriting biometrics (e.g., online signatures), where generative models also exist?

Attack models we studied

Class 1	different user, different passphrase (sometimes called "naive forgery").
Class 2	different user, true passphrase (sometimes called "skilled forgery").
Class 3	true user, different passphrase.
Class 4	concatenation attack (true passphrase constructed from unrelated writing).
Class 5	true user, true passphrase (as baseline).

Biometric Hash from Handwriting

Studied published technique by Vielhauer, et al. for converting handwriting into secure 24-element hash.

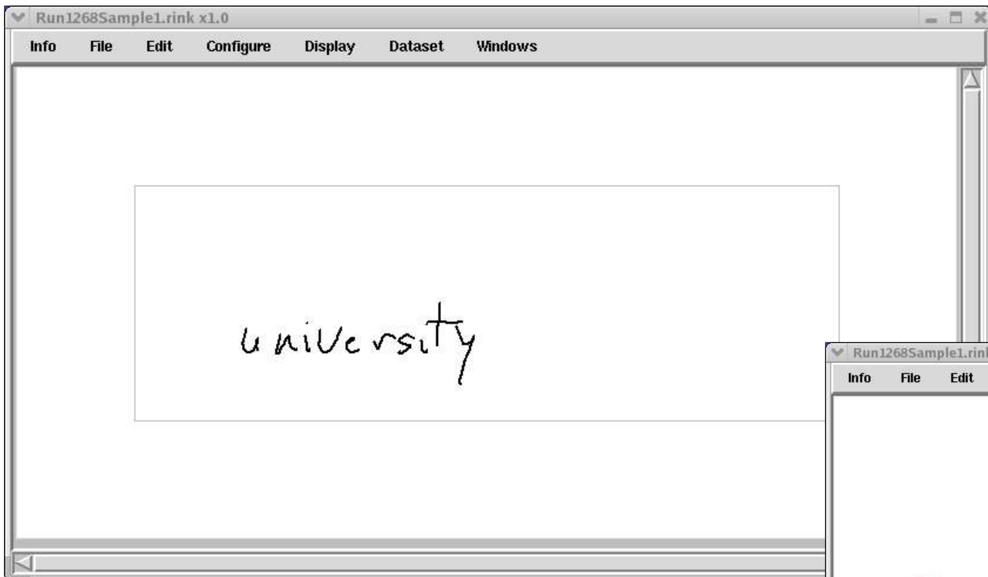
Features extracted from each sample:

1. Number of strokes
2. Total writing time (ms)
3. Total number of samples (points)
4. Sum of all local (x,y) minima and maxima
5. Aspect ratio (x/y) * 100
6. Pen-down / total writing time * 100
7. Integrated area covered by x signal
8. Integrated area covered by y signal
9. Average writing velocity in x
10. Average writing velocity in y
11. Average writing acceleration in x
12. Average writing acceleration in y
13. Effective writing velocity in x
14. Effective writing velocity in y
15. Integrated area under x, segment 1
16. Integrated area under x, segment 2
17. Integrated area under x, segment 3
18. Integrated area under x, segment 4
19. Integrated area under x, segment 5
20. Integrated area under y, segment 1
21. Integrated area under y, segment 2
22. Integrated area under y, segment 3
23. Integrated area under y, segment 4
24. Integrated area under y, segment 5

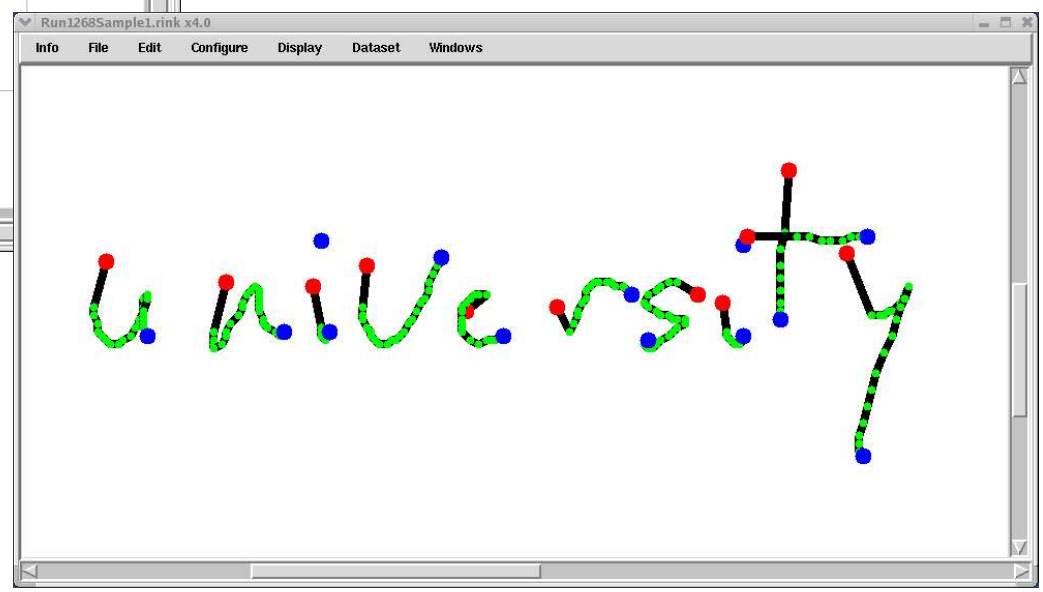
"Biometric Hash based on Statistical Features of Online Signatures," Claus Vielhauer, Ralf Steinmetz, and Astrid Mayerhofer, *Proceedigns of the Sixteenth International Conference on Pattern Recognition*, vol. 1, August 2002, pp. 123-126.

Handwriting Biometric Features #1

Snapshot of our tool for ink capture written in Tcl/Tk:

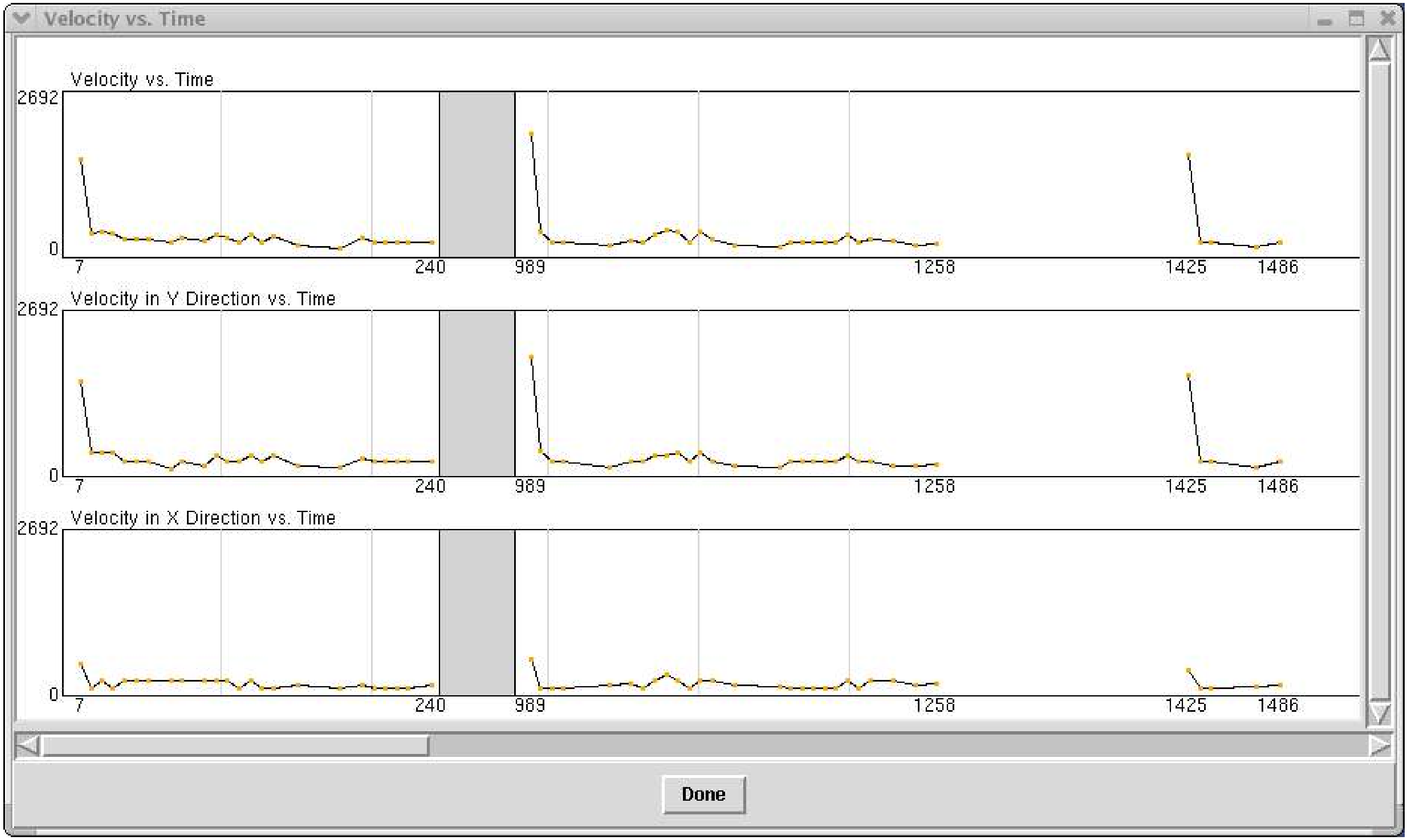


Passphrase



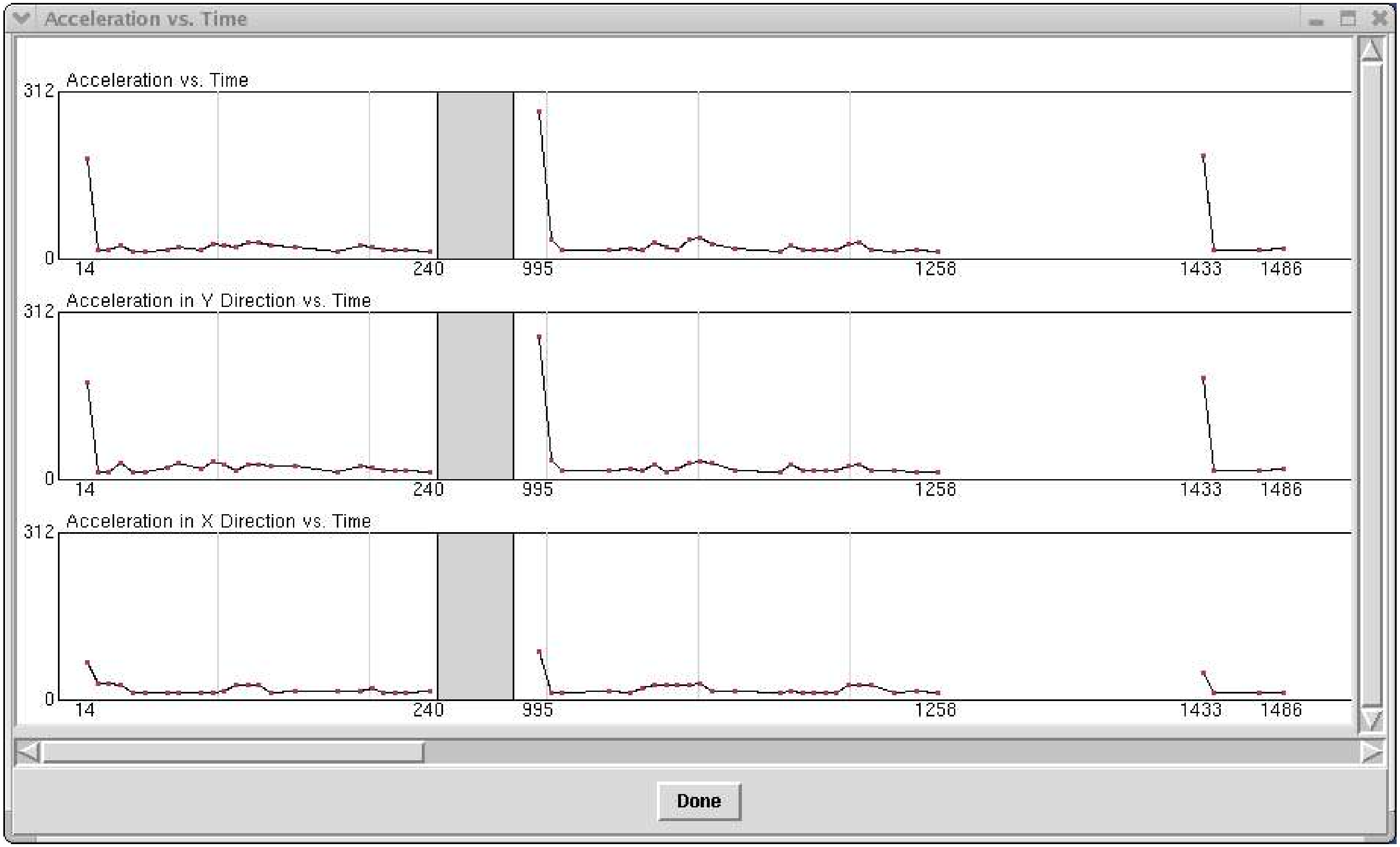
Sampled points

Handwriting Biometric Features #2



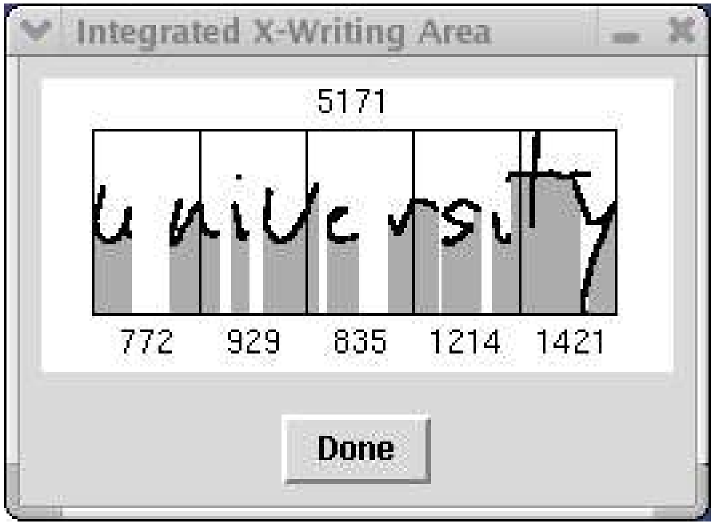
Velocity

Handwriting Biometric Features #3



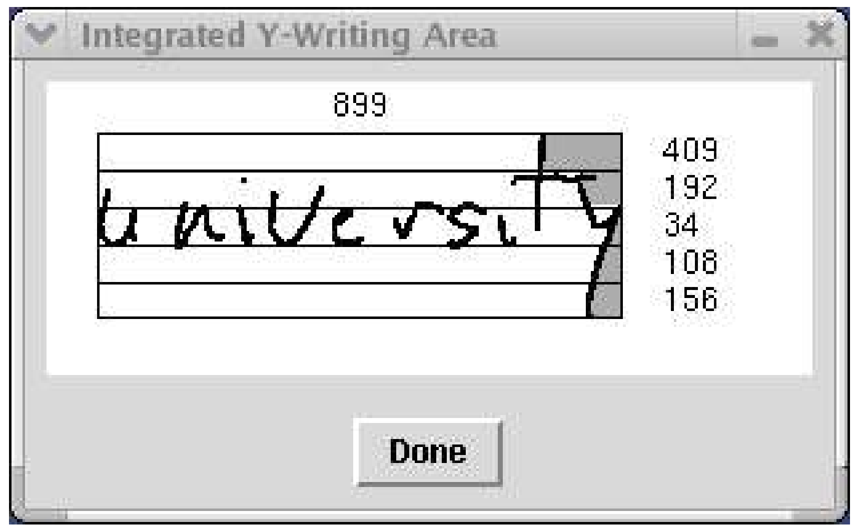
Acceleration

Handwriting Biometric Features #4



Integrating x-writing area (segmented)

Integrating y-writing area (segmented)



Typical Performance Evaluation

Traditional approach: conduct study using human subjects (naive and/or skilled "forgers") and report False Reject Rate (FRR) and False Accept Rate (FAR).

- E.g., Vielhauer, et al. used 10 subjects who provided six samples and also tried to forge writing of other subjects based on static image.
- Average FRR was measured to be 7.0%.
- Average FAR was measured to be 0.0%.

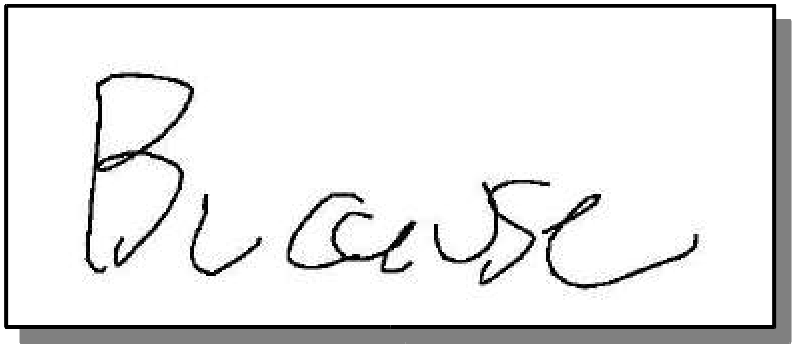
We believe this model misses the more ominous threat.

"Biometric Hash based on Statistical Features of Online Signatures," Claus Vielhauer, Ralf Steinmetz, and Astrid Mayerhofer, *Proceedigns of the Sixteenth International Conference on Pattern Recognition*, vol. 1, August 2002, pp. 123-126.

Our Test Data

- Two writers each wrote four different passphrases 20 or more times using Wacom Intuos tablet.
- Additional samples collected independently to support concatenative attacks.
- Dataset is small, but we are not trying to prove biometric is secure: we are studying its weaknesses.

Samples of handwriting we collected:



Concatenative Attack

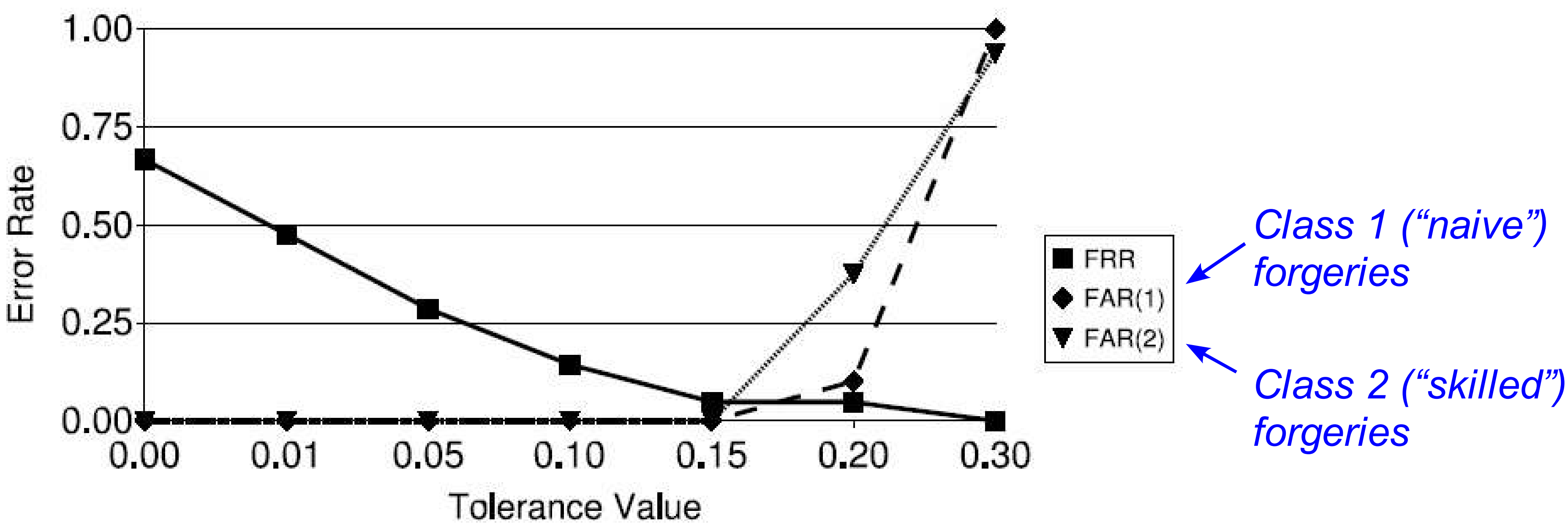
- Separate corpus of writing samples collected and labeled on a per-character basis.
- Provides assortment of n-grams which can be selected to yield targeted passphrase.
- Optimal concatenation can be formulated as dynamic programming problem, much like TTS.

Original passphrase

Synthesized passphrase

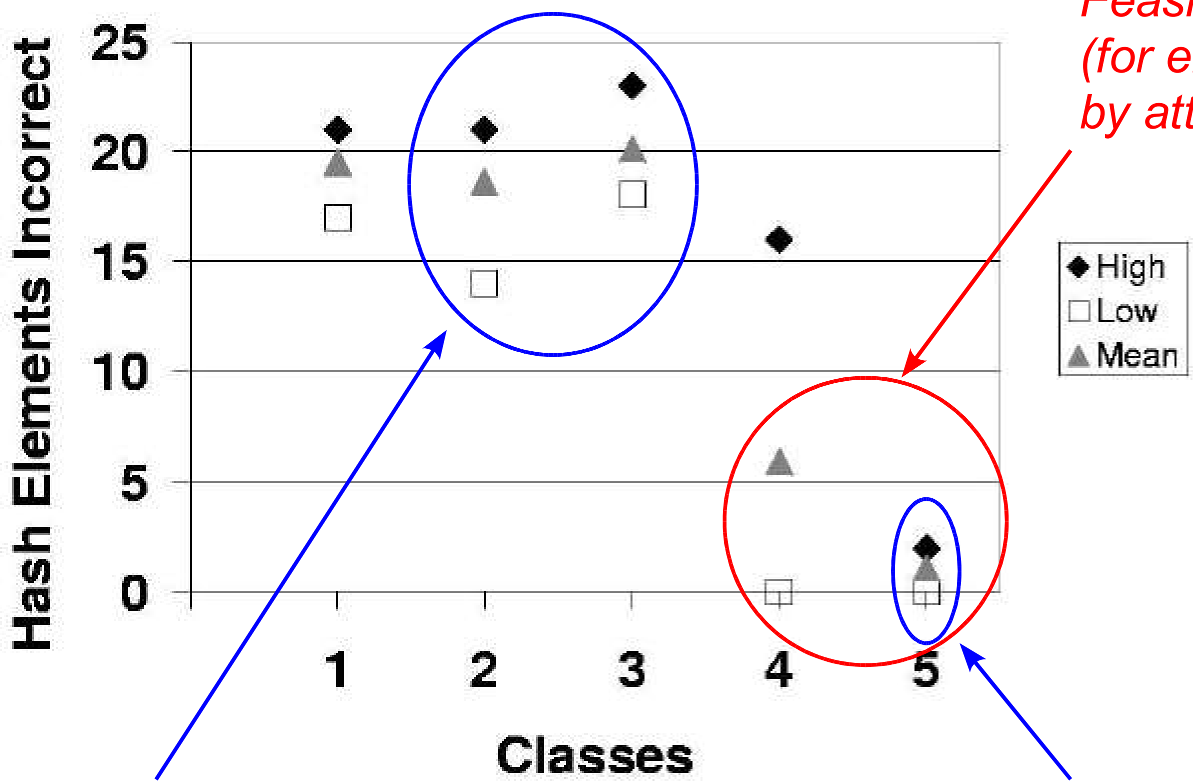
Determining Hash Tolerance

- Training set varied from 15 to 25 samples per class.
- Cross-validation performed using 5 to 10 samples.
- Various tolerances tested, most promising was 0.15.



Count of Incorrect Hash Elements

Feasible search space (for error correction, or by attacker).

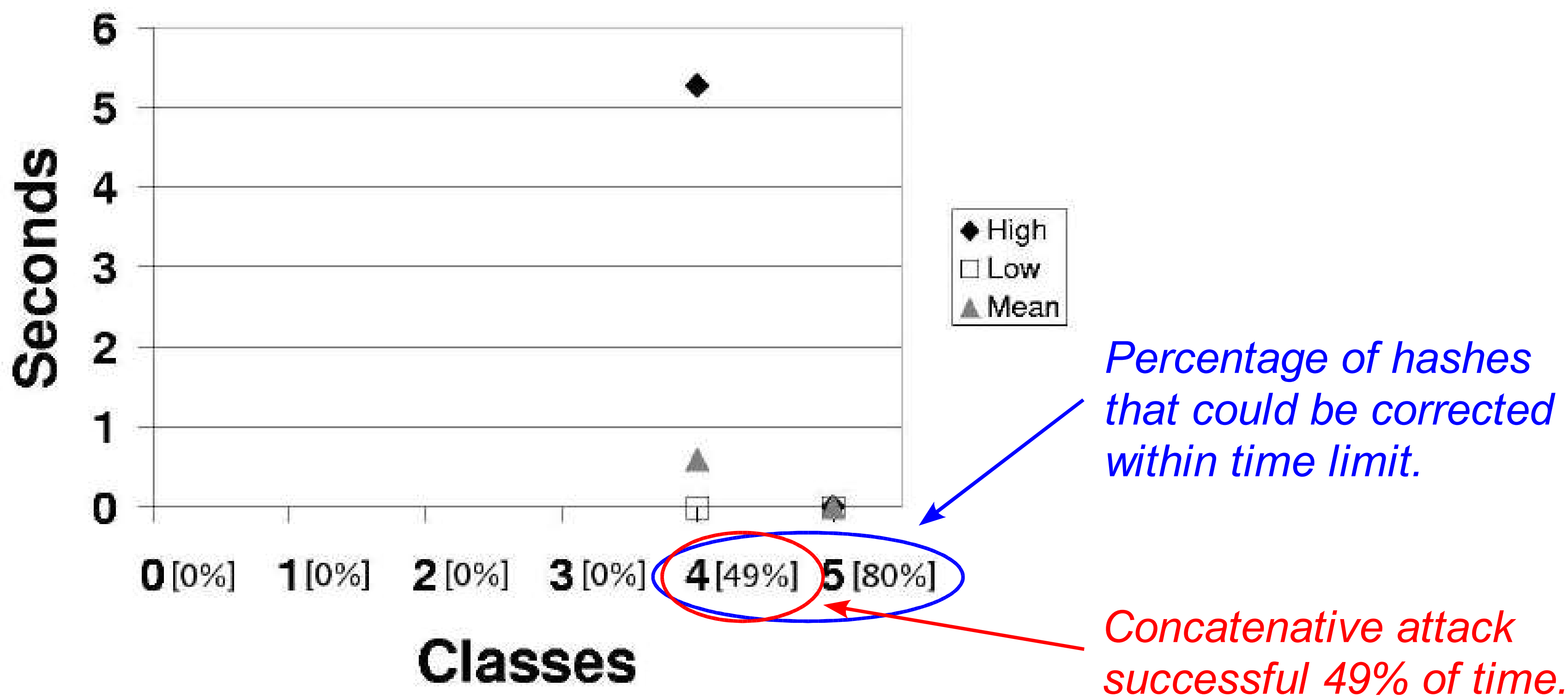


Roughly same number of features sensitive to passphrase (Class 2) versus user (Class 3).

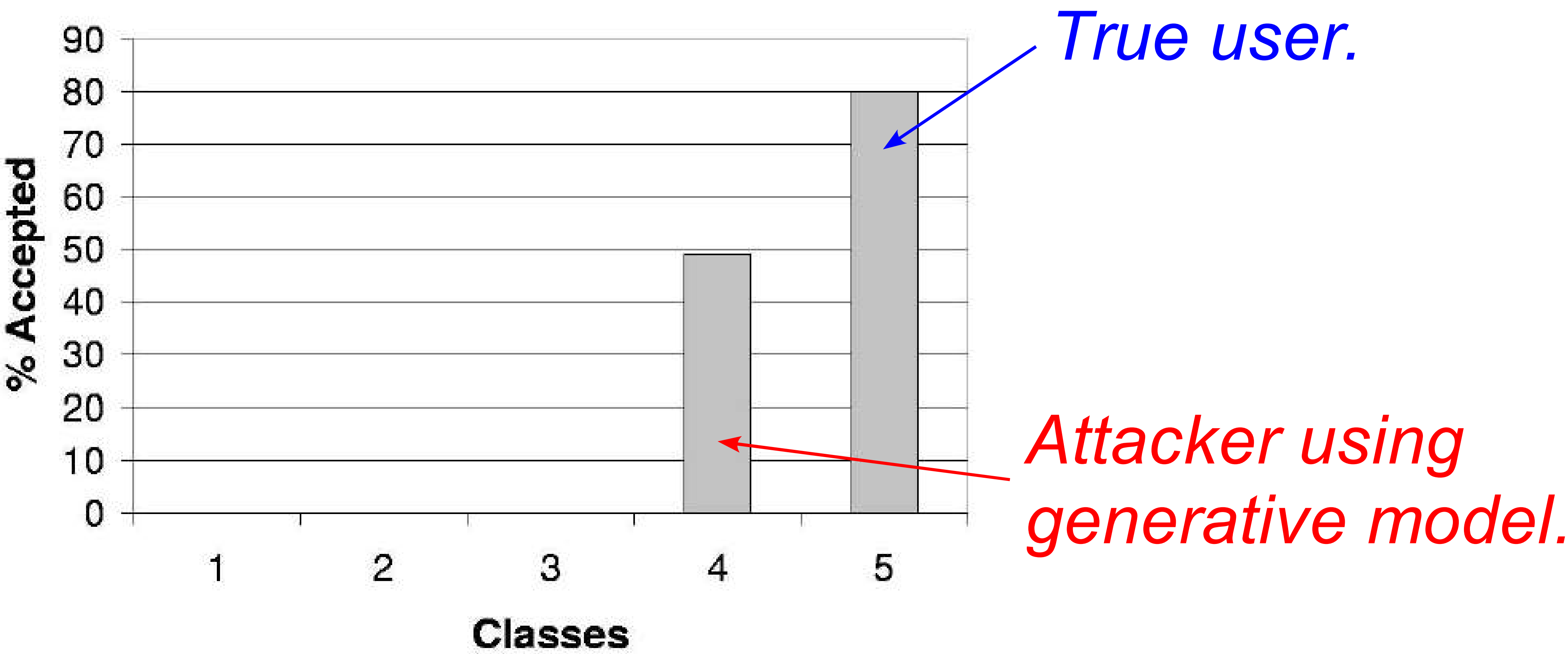
Even true user (Class 5) requires some post-error-correction.

Time to Correct Hashes

- Perform exhaustive search around hash vector.
- Timeout (failure) after 60 second time limit.
- Tests run on Pentium 4 PC, 3.2 Ghz, 1 GB RAM.



Hashes Corrected After Search



Conclusions

- Generative models for human behavior (speech, handwriting) present a threat to security of biometric systems based on such inputs.
- The traditional approach to performance evaluation, i.e., human studies involving “naive” and “skilled” forgers, is inadequate for assessing this threat.
- A published biometric for online handwriting is easily defeated using such an attack.
- Full extent of this threat not yet characterized – much more work needs to be done.

Credits

This work is supported in part by:

- National Science Foundation CNS CYBER TRUST 0430178, “ Using generative models to evaluate and strengthen biometrically enhanced systems” (in collaboration with Fabian Monrose and Mike Reiter).
- The Keystone Alliance for Homeland Security.

Additional results to be presented at *International Conference on Audio- and Video-based Biometric Person Authentication* (AVBPA) in July 2005.