Biometric Authentication

Ing. Jan Kozlovský, supervised by Ing. Karel Petránek

FIM UHK

The work describes three methods how to authenticate users by their biometric identifiers. Keystroke dynamics (the typing rhythm of a written password) is used as the biometric identifier. Three machine learning algorithms are implemented: neural network, k-nearest neighbor algorithm and Bayesian classifier. All three algorithms are compared with respect to the ability to learn, prediction accuracy, safety and the suitability for a real-world application.

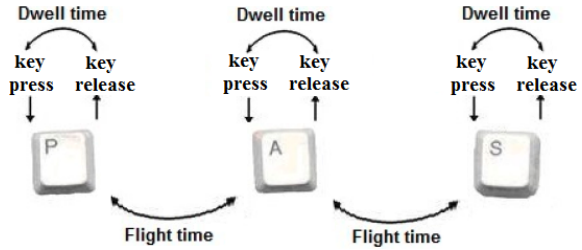
Introduction and implementation

Motivation:

- The usage of computer technology still grows, so does computer threats
- High computer performance can be used for machine learning algorithms to ensure computer security by recognizing biometric identifiers of the user

Idea:

- Use user's keystroke pattern as a biometric identifier
- Two representations for each key: **dwell** time, **flight** time
- Implement and compare three mentioned machine learning algorithms
- Verify the reliability of algorithms by „uknown" users



Preliminaries:

- Approximately 1 200 keystroke pattern samples of 3 users for learning and testing
- Approximately 450 more keystroke pattern samples of 10 „uknown" users for verification

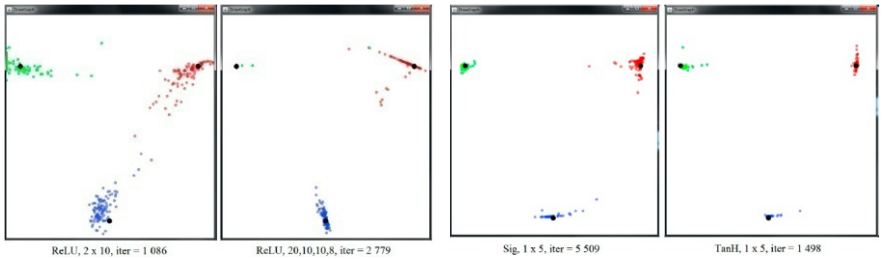
Implementation goals:

- Neural networks: to find optimal network structure
- K-nearest neighbor: to find optimal euclidean threshold distance
- Bayesian classifier: test and compare Gaussian, multinomic and Bernoulli Bayesian classifier

Results

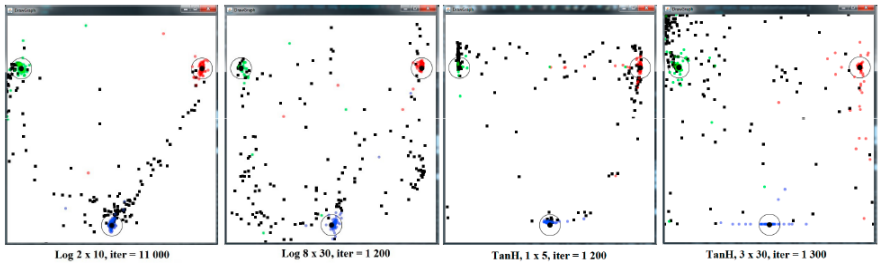
Classification of 3 „known" users:

- Circa 100 neural network structures were tested. Logarithmic and hyperbolic tangent multilayer network was chosen as optimal
- Bayesian classifier not suitable for this solution
- K-nearest neighbor algorithm provided the best results of all algorithms



Veryfication of algorithms by „uknown" users:

- Neural network wrongly authenticated 5 – 20% „uknown users"



- The best result is given by k-nearest neighbor algorithm after finding optimal euclidean threshold distance

Possible development:

- Evaluating keystroke dynamic in realtime on most used patterns, e.g. „the", „pro", „for", „tro", „dear", „hello"

Main sources:

[1] MONROSE, Fabian a Aviel D. RUBIN. Keystroke dynamics as a biometric for authentication. Future Generation Computer Systems 16 (2000)

[2] DOMINGOS, Pedro. University of Washington: Machine Learning. Coursera [online]

[3] HINTON, Geoffrey. University of Toronto: Neural Networks for Machine Learning. Coursera [online]

[4] COVER, T. a P. HART. Nearest neighbor pattern classification. IEEE Transactions on Information Theory [online]