Adversarial Touch Dynamics

Sean Curran - securran@stanford.edu

Goal

Many people have proposed systems for authenticating users based on how the users interact with their touch screens. The goal of this project is to create an adversary which can generate gestures to pose as any user it chooses.

Data

- **Data set was found online**
  - Contains 20,000 touch gestures from 41 users
  - Gestures are transformed in to 34 features
- **Databases**
  - The adversary gets a portion of the raw sensor data and the authentication server gets the remainder.
- **Database Compromise**
  - We make the assumption that the server's kNN feature database is compromised by the adversary, but the server's raw sensor data is kept safe.

Classifier

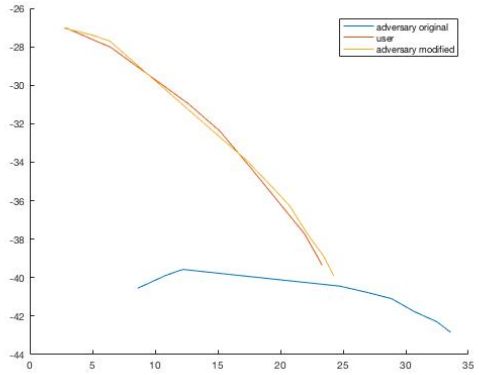
- **kNN** with a reduced, normalized feature set
- **Multiclass SVM** with a Gaussian kernel and a reduced, normalized feature set

Attack

1. **Adversary Collects touch gestures**
   - Example: the adversary releases a mobile application to the public
2. **The server's kNN feature-database is compromised, but no touch gestures are compromised.**
3. **The adversary finds the strokes in its database which have features closest to the user's features taken from the compromised database.**
4. **The adversary uses the information from the compromised database to help its strokes match the user's strokes**

More Specifically:

The adversary uses the start-(x,y) and end-(x,y) from the user's features to shift and rotate adversary's stroke as shown below:



Results

	Classifier	Attack	Relative Performance
kNN - Exclusive	67.48%	45.16%	45.16/67.48 = 66.92%
kNN - Overlap	99.74%	80.81%	80.81/99.74= 81.02%
SVM - Exclusive	51.55%	36.4%	36.4/51.55= 70.61%
SVM - Overlap	94.23%	68.28%	68.28/94.23= 72.46%

The table shows the percentage of gestures correctly matched to 1 of 41 users averaged over all 41 users' data.

Overlap The server has the same compromised database

Exclusive: The server completely renews its compromised database

Conclusion

Forward search was used to select features to improve the performance of the classifiers, but the performance of the classifiers is still not satisfactory. However, the purpose of the project was to attack the classifiers. Looking at the relative performance, the attack is effective against kNN and SVM. Even with the inaccurate classifiers, the adversary's attack was very effective against the old database. One thing is for certain, servers must continually update their machine learning databases.