INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Hand Me your PIN

Inferring ATM PINs of Users Typing with a Covered Hand

Aman Izardar 2021201028
Diksha Daryani 2021201045

Deep learning Based attack

- The attacker can regenerate ATM pin of the victim even if he/she covers their typing hand with other hand.
- The attacker places a camera that is hidden and can record ATM pin pad videos that captures the hand and finger muscle movements.

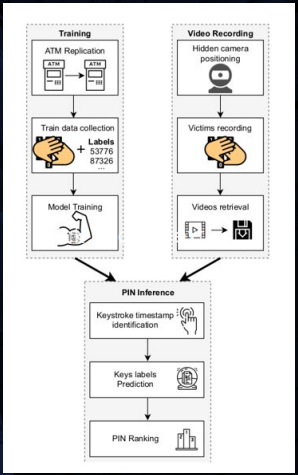
Various Scenarios:

- **Single pin pad scenario :** Where the attacker obtains.the exact copy of the target ATM's pin pad.
- **Pin pad independent scenario :** When it is not possible for the attacker to obtain the exact copy, in that case he/she performs the experiment with a pin pad similar to the target one.
- **Mixed scenario :** This is a combination of the Single Pin pad scenario and Pin pad independent scenario. The experiment will be performed on two different pin pads.

ATTACK PHASES

Attack Phases

1. Training the model
2. Video Recording
3. Guessing the Pin



Training Phase

- The attacker uses a simulation of the ATM machine to train the model.
- Collected two datasets : the videos of the people entering pin. First data collection contains 40 and second collection contains 18 volunteers.
- Total of 5800 random 5 digits pins recorded videos.

Training Phase

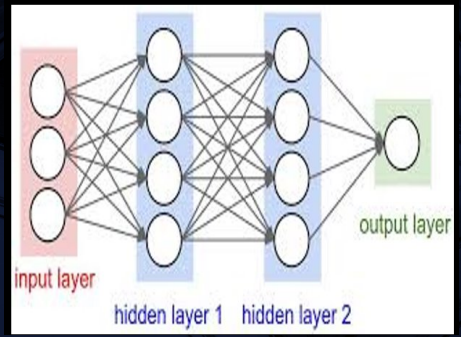
RECORDING AUDIO

- Why?
  - To retrieve the timestamp of the keypress.
- How?
  - Using the feedback sound of the key.

Video Segmentation

- 5 segments for 5-digit pin.
- Length of each segment is 11 frames.
- Middle frame is the keypress frame.

Training Phase



Model Training using deep learning

- Convolutional neural network (CNN).
- Multilayer perceptron.
- Split dataset into train, validation and test sets.

Video Recording Phase

- The attacker places a hidden camera inside the target ATM.
- Records the victim's hand movements.
- Gets the recorded Video.

Pin inference phase

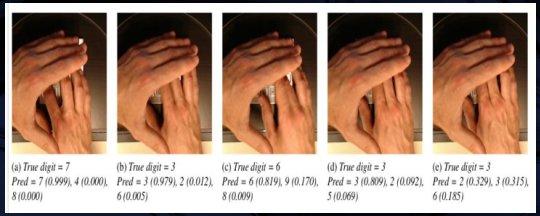
Segment the recorded video.

Get N subsequences of video for a N digit PIN.

Prepare attack set.

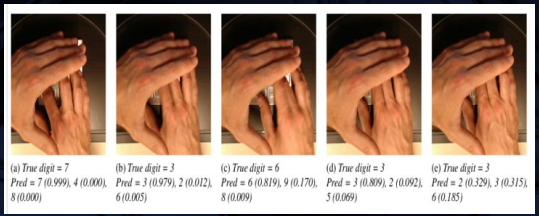
Use the model to generate the pin.

Prediction of the pin



- For each subsequence (containing the key pressed) ,the attacker uses the model to predict the key. Finally, the product of the probabilities for each key would be the probability of the PIN.
- PIN 73633 is the actual pin entered .

Prediction of the pin



(a) *True digit = 7*
Pred = 7 (0.999), 4 (0.000), 8 (0.000)

(b) *True digit = 3*
Pred = 3 (0.979), 2 (0.012), 6 (0.005)

(c) *True digit = 6*
Pred = 6 (0.819), 9 (0.170), 8 (0.009)

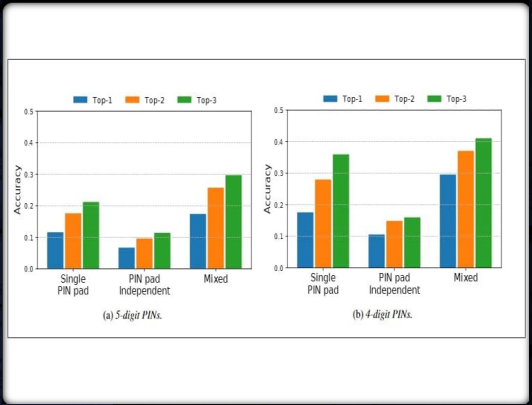
(d) *True digit = 3*
Pred = 3 (0.809), 2 (0.092), 5 (0.069)

(e) *True digit = 3*
Pred = 2 (0.329), 3 (0.315), 6 (0.185)

- The proposed algorithm suggests 73632 as the most probable PIN (probability = 21.32%), 73633 as the second most probable PIN (probability = 20.43%), and 73636 as the third most probable PIN (probability = 11.96%).
- The algorithm predicts the correct PIN in the second attempt.

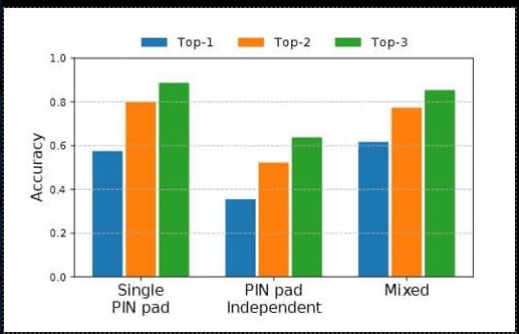
Results

- For a single digit guess, model's accuracy reaches 63.8% for Top-3 guesses.
- For a 5-digit pin Pin-pad independent case model's accuracy is 11.4%.
- For a 5-digit pin mixed scenario model's accuracy is 30% for top-3 guesses.
- For a 4-digit pin model's accuracy is above 41%.

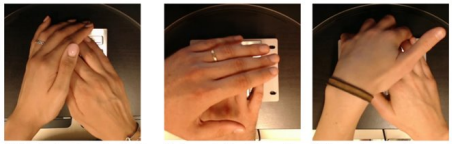


Conclusion

- Covering the pin pad with the other hand is not a sufficient defense.
- Keypad difference aspect is quite significant.
- Similar type of keypad increases model's performance.



Defense or counter-measures



(a) *Side: hand resting on the side of the palm.*

(b) *Over: raised hand not touching the surface.*

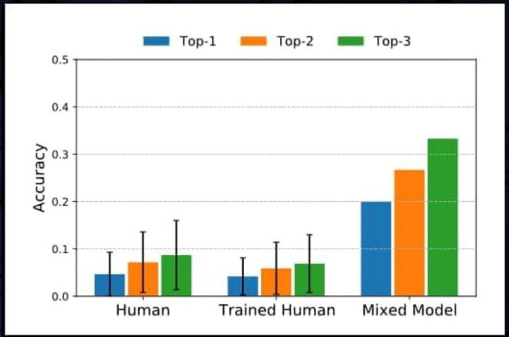
(c) *Top: hand resting on fingers and vertically covering the PIN pad.*

Different covering strategies for the non-typing hand.

- Different hand covering strategies (over hand strategy significantly reduces the pin accuracy).
- Use Longer pins.
- Virtual/Randomized keypad.
- Screen protectors.

Human vs Deep learning

- Designed a questionnaire.
- A total of 78 participants participated.
- They were asked to guess the pin.



Results

- Participants could infer only 4.49% of the pins in the first attempt and 7.92% within three attempts.
- Provided the same set of videos to trained model.
- Human accuracy on a single key classification : 0.351
- Model's Accuracy : 0.687

Major critiques :

Some assumptions which may not be always true are:

- Wearing hand glove images should also be included in the dataset.
- The videos are recorded via the hidden cameras . Strong surveillance and security systems would make it difficult .
- If there is no audio, either we record the screen internally or place a camera outside/on the body of the ATM to record the screen.
- The number of participants are only 58.
- No clear explanation why they selected 11 frames for each subsequence.
- Including more models of ATM keypads would make a more generalized Training set.

Improvements/Extensions:

- Including the left-handed people.
- Increasing the number of pin-pad/keypads in the experiment.
- Increasing the number of participants in the experiment(currently Only 58 are there).
- Including a thermal camera in the experiment.
- The Questionnaire, training of the participants should be long enough to train them.
- Including people from other races.
- Training the model with different hand positions and conditions. Ex Using gloves.
- Increasing the Frames per subsequence.(currently 11 FPSS).

THANKYOU!