PassGAN

Password cracking for the 21st century

Jarod Gowgiel

Passwords, and why they matter

First (and arguably most important) line of defense for security

Almost 60% of users reuse their passwords

Data breaches can critically impact a number of sites

When databases are breached, stolen passwords are usually hashed

Encrypted so that thieves cannot directly access user information

Password cracking: the process of going from hash back to original password

Current password cracking methods

Hashcat

- Uses GPU processing to iteratively try millions of potential passwords
- Utilize dictionaries and sets of rules (e.g. append digits [0-9] to common passwords)
- Use statistical approaches (Markov attack) to guess passwords based on frequency of certain characters in certain places

John the Ripper

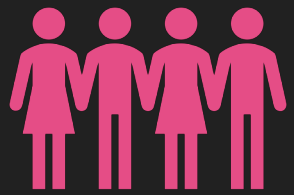
- Older software
- Similar approach to Hashcat, but uses mostly the CPU

Existing applications of machine learning

Medical uses

Analyze images from medical scans to detect tumors and other anomalies

Shown to be potentially more accurate than trained technicians



Marketing and advertising

Deliver targeted advertisements based on social demographics and other information

Machine learning enables this without needing manual oversight

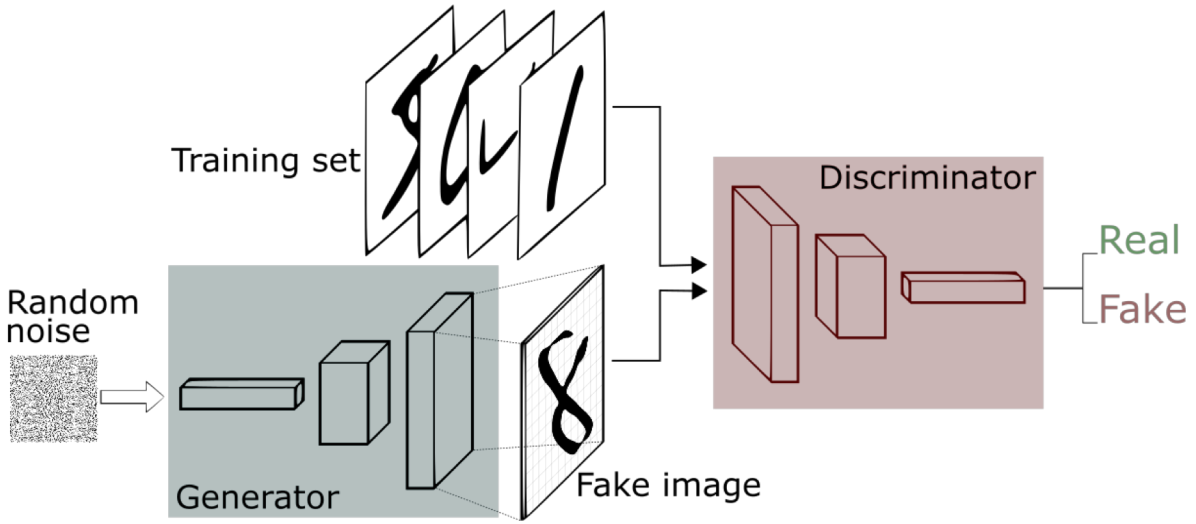
No need for humans crafting profiles for target customers

PassGAN in brief

- Recently developed machine learning approach to password cracking
- Uses neural networks and GAN (generative adversarial learning) to guess passwords based on a given dataset
- Could be used to extend known databases, directly guess user passwords, or applied to evaluate password strength

Generative adversarial networks

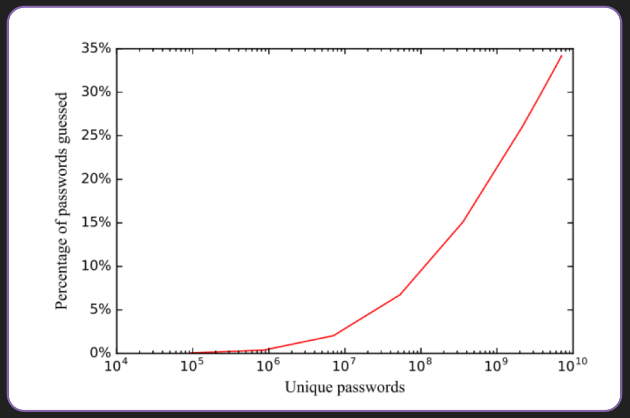
- Two active processes; discriminating and generating
- Discriminator attempts to determine which inputs are from the original data, and which are from the generator
- Generator tries to fool the discriminator by imitating the target data
- This program becomes better at creating very similar outputs to the dataset as iterations progress



https://skymind.ai/wiki/generative-adversarial-network-gan

How useful could PassGAN be?

- Resulting list of password guesses used as input for Hashcat
  - Rules and other mutations applied to this new dictionary
- 50% more matches when compared to Hashcat using just the original password data



Hitaj, Briland, et al. "PassGAN: A Deep Learning Approach for Password Guessing." 2017.

Password strength applications

- Similar methods to PassGAN could be used to evaluate password strength
- Machine learning algorithms evaluating potential user passwords in real time
- Can be packaged and compressed for client side use
  - Negate security risk of sending password candidates to server for more intensive checking
  - Leverage the representative processing power of a trained neural network in a small amount of data
- Provide live feedback for users as they enter a candidate password

Why PassGAN matters

- In the coming years, password cracking through machine learning could be applied on a broader scale
- Important to be aware of emerging technologies
- Test existing user passwords created with previously strong requirements against these new models
- Extend existing dictionaries to imitate privately held stolen databases
  - Producing and applying an extended dataset could reveal potential password weaknesses

Adjustments to account for PassGAN

01

Adjust password requirements to account for advancements in password cracking

02

Use machine learning programs to check password strength

- Leverage advances against attackers

03

Never reuse critical passwords in public instances