Collision Algorithms

CS 303 Alg. Number Theory & Cryptography
Jeremy Johnson

Outline

- Birthday Paradox
- Discrete Log Problem
- Shanks Babystep-Giantstep Algorithm
- Collision Theorem
- Randomized Discrete Log Collision Algorithm
Collision Algorithms

- A simple yet surprisingly powerful search method is based on the observation that it is usually much easier to find matching objects that it is to find a particular object.

- Cryptographic applications of collision algorithms are generally based on the following setup. Bob has a box that contains N numbers. He chooses n distinct numbers from the box and puts them in a list. He then makes a second list by choosing m (not necessarily distinct) numbers from the box. The remarkable fact is that if m and n are slightly larger than $\sqrt{N}$, then it is very likely that the two lists contain a common element.
Birthday Paradox

- Pr(someone has the same birthday as you)
  - $1 - (364/365)^m$

- Pr(two people have the same birthday)
  - $1 - \prod_{i=1..m} (365-(i-1))/365$
  - $1 - \prod_{i=1..m-1} (1-i/365) \approx 1 - \exp(-t*(t-1)/(2*365))$
  - Approximately 23 people required for 50%
  - $1 - (364/365)^{40} \approx 10.4\%$
  - $1 - \prod_{i=1..39} (1-i/365) \approx 89.1\%$
Discrete Logarithms

- **Discrete log problem**
  - Given $Z_p^* = \langle g \rangle$
  - $\log_g(y) = x$, if $y = g^x$.

- **Example**
  - $Z_{17}^* = \langle 3 \rangle$: $3^1=3$, $3^2=9$, $3^3=10$, $3^4=13$, $3^5=5$, $3^6=15$, $3^7=11$, $3^8=16$, $3^9=14$, $3^{10}=8$, $3^{11}=7$, $3^{12}=4$, $3^{13}=12$, $3^{14}=2$, $3^{15}=6$, $3^{16}=1$
  - $\log_3(15) = 6$. 
Shanks Babystep-Giantstep Alg

- Given $Z_p^* = \langle g \rangle$, $N = p-1$
  - Find $\log_g(h) = x$, $h = g^x$.
  - Let $n = 1 + \lfloor \sqrt{N} \rfloor$, $n > \sqrt{N}$ (could use $\lceil \lfloor \sqrt{N} \rfloor \rceil$)
  - Create two lists
    - $1, g, g^2, g^3, \ldots, g^n$
    - $h, hg^{-n}, hg^{-2n}, g^{-3n}, \ldots, hg^{-n^2}$
  - Find match between two lists $g^i = hg^{-jn} \Rightarrow x = i + jn$
  - Complexity $\Theta(\sqrt{N} \log(N))$

- Example $Z_{17}^* = \langle 3 \rangle$, $h = 15$, $n =$
  - $3^1 = 3$, $3^2 = 9$, $3^3 = 10$, $3^4 = 13$
  - $15, 15 \times 3^{-4} = 9$, $15 \times 3^{-2 \times 4} = 2$, $15 \times 3^{-3 \times 4} = 8$
  - $x = 2 + 1 \times 4 = 6$. 
Collision Theorem

- An urn contains N balls, of which n are red and N-n are blue. Bob randomly selects m balls (replacing them in the urn).
- \( \text{Pr(} \text{at least one red)} \)
  - \( 1 - (1-n/N)^m \geq 1 - \exp(-mn/N) \)
- If N is large and m and n are close to \( \sqrt{N} \) then
  - \( - (1-n/N)^m \approx 1 - \exp(-mn/N) \)
Collision Theorem Application

- View list of numbers as an urn containing \( N \) numbered blue balls. Select first list and paint red. Select second list. Match occurs when a red ball is selected.

- Eight cards are dealt and then eight cards are selected from a fresh deck. What is the probability that one of the second eight cards matches the 8 dealt cards?

\[
1 - (1 - \frac{8}{52})^8 \approx 73.7\%
\]
Collision Theorem Application

A set contains $N$ objects. Bob randomly chooses $n$ of them, makes a list, replaces them and then chooses $n$ more. How large should $n$ be to get a 50% chance of match? A 99% chance of a match?

\[
\Pr(\text{match}) \approx 1 - \exp\left(-\frac{n^2}{N}\right) = \frac{1}{2}, \, .99
\]

\[
n = \sqrt{N \cdot \log(2)} \approx 0.83 \sqrt{N}
\]

\[
n = \sqrt{N \cdot \log(100)} \approx 2.15 \sqrt{N}
\]
Randomized Collision Algorithm for Discrete Log

- Given $Z_p^* = \langle g \rangle$, $N = p-1$
  - Find $\log_g(h) = x$, $h = g^x$.  
  - Write $x = y-z$
  - Create two lists
    - $g^{y_1}, g^{y_2}, \ldots, g^{y_n}$
    - $h^{z_1}, h^{z_2}, \ldots, h^{z_n}$
  - Try to find match between two lists $g^{y_i} = h^{z_j} \implies x = y_i - z_j$

- Complexity $\Theta(n \log(N))$
- Choose $n = c\sqrt{N}$ to get high probability of a match