DTTF/NB479: Dszquphsbqiz Day 28

- Announcements:
  - Choosing presentation dates (at end)

- Questions?

- This week:
  - Hash functions, SHA
  - Birthday attacks
  - Digital signatures (Monday)

Birthday paradox

- What’s the chances that two people in our class have the same birthday?

- Exact solution: use fractions
- Approximate solution:

$$1-e^{\frac{-r^2}{2N}}$$

Where r = 26 people, and N = 365 choices

The birthday paradox doesn't mean that there's a high probability that someone else has my birthday

- What's the chance that one of the other students has **your** birthday?
- Note: the chance of someone matching me is low, but there are lots of ways to get pairs of matches in general.

Likewise, the birthday paradox doesn’t mean that finding a collision with a known digest is easy

- What’s the chance that one of the other students has **your** birthday?
- **Key**: the chance of someone matching me is low, but there are lots of ways to get pairs of matches in general.

Strongly collision-free: Can’t find any pair $m_1 \neq m_2$ such that $h(m_1)=h(m_2)$ easily

(Sometimes we can settle for weakly collision-free: given m, can’t find m’ ≠ m with h(m) = h(m’).

We can calculate how many messages we need to hash to have a good chance of finding a collision

- How many people are needed to get the probability of having 2 with the same birthday to be above 50%?
- Derive for general N (not just days in a year)

Birthday attacks on SHA-1?

- How many digests are possible when h is an n-bit hash? This is N.
- The birthday paradox says I can choose r = sqrt(n) messages and there’s a good possibility that 2 will match.
  - For a 60-bit hash, r = ???
  - For a 160-bit hash, r = ???

Multicollisions are harder to find, but not as hard as expected.

- What if instead of finding a just pair of collisions, we need to find 8 collisions?

Multicollisions

- Recall: given *r* people and *N* (say, 365) birthdays. If $r \approx N^{1/2}$, then there's a good chance that 2 people will have the same birthday

- Generalization: given *r* people and *N* birthdays. If $r \approx N^{\frac{(k-1)}{k}}$ for some **k**, then there's a good chance that **k** people will have the same birthday.
- So for 160-bit hashes, how many messages do we need to generate to get an 8-collision?

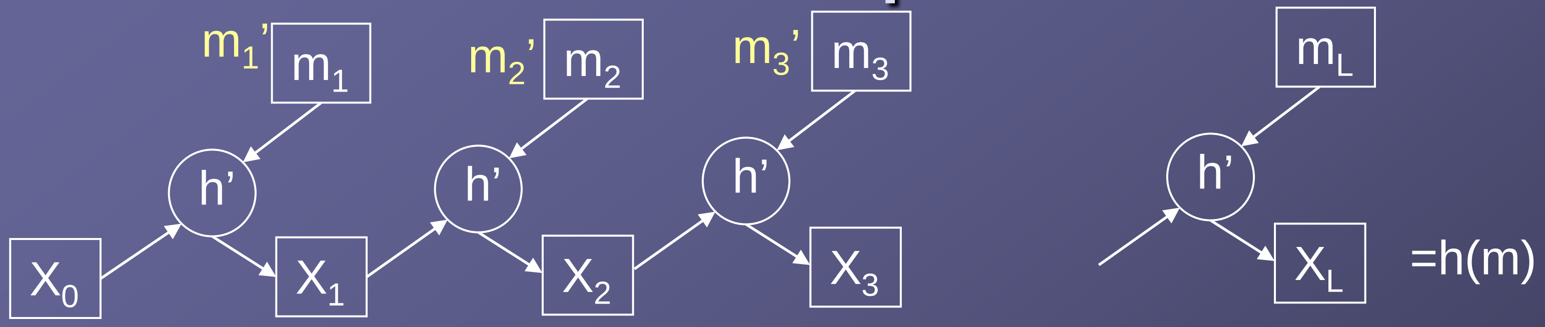
- That's lots more than 2^{80}!
- However, there's a big underlying assumption: the hash function is random!
- Is SHA-1 random?

(answer on next slide)

No

(It’s iterative…)

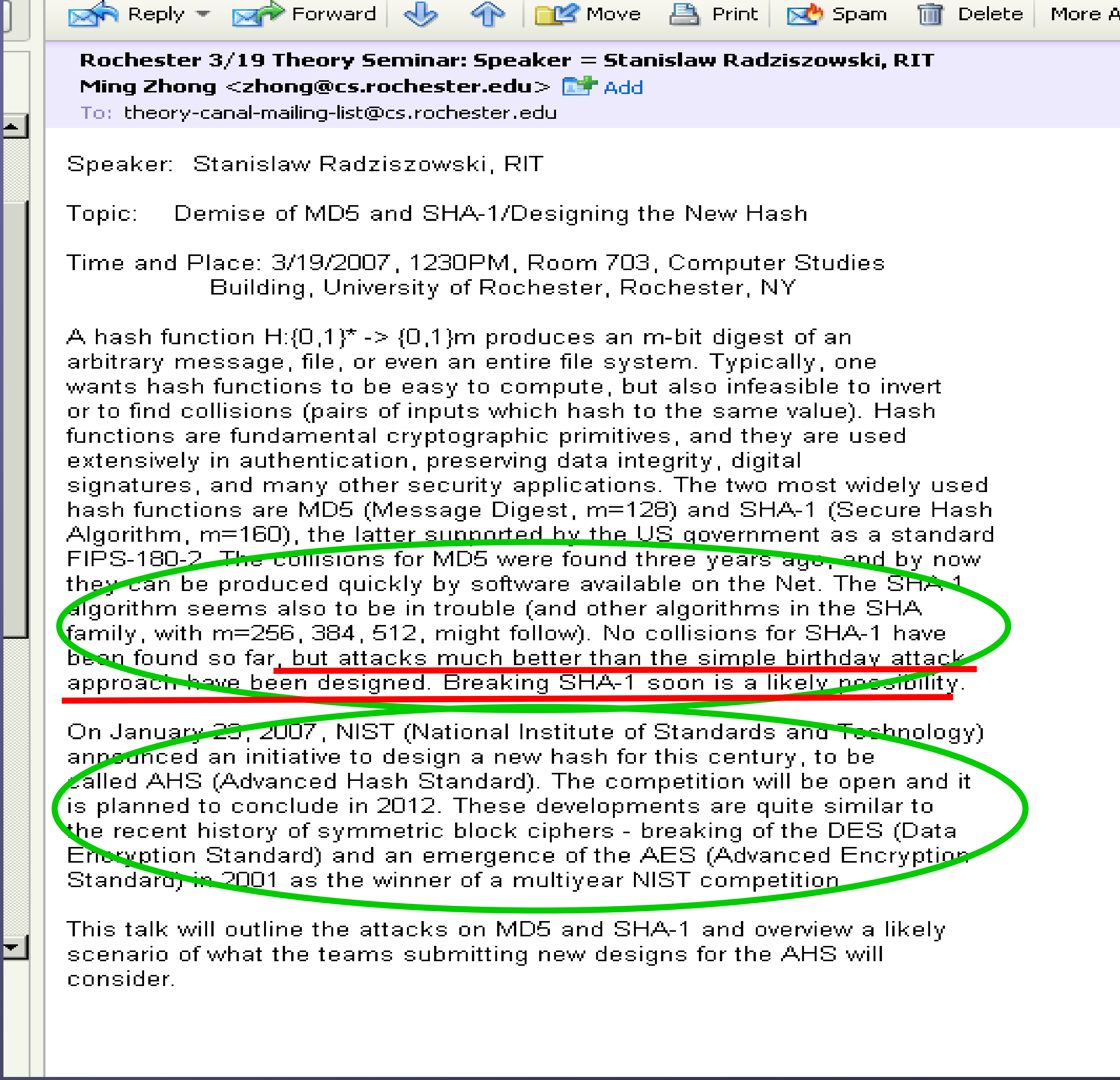
Recall this picture



Consider the following attack:

1. Birthday attack the first block: x1 = h'(x0, m1)
   1. Need to generate $2^{n/2}$ messages
   2. Result: found (m1, m1') such that x1 = h'(x0, m1) = h'(x0, m1')
2. Repeat for x2 and x3, finding pairs (m2, m2') based on x1 and (m3, m3') based on x2.
   1. Need to generate total of 3 * $2^{n/2}$ messages
   2. Result: found 8 combinations (m1, m1') x (m2, m2') x (m3, m3') with same x3.
3. 3 x 2^{80} is lots smaller than 2^{140}.

The Future of SHA-1?



Rochester 3/19 Theory Seminar: Speaker = Stanislaw Radziszowski, RIT
Ming Zhong <zhong@cs.rochester.edu>
To: theory-canal-mailing-list@cs.rochester.edu

Speaker: Stanislaw Radziszowski, RIT

Topic: Demise of MD5 and SHA-1/Designing the New Hash

Time and Place: 3/19/2007, 1230PM, Room 703, Computer Studies
Building, University of Rochester, Rochester, NY

A hash function H:{0,1}* -> {0,1}m produces an m-bit digest of an arbitrary message, file, or even an entire file system. Typically, one wants hash functions to be easy to compute, but also infeasible to invert or to find collisions (pairs of inputs which hash to the same value). Hash functions are fundamental cryptographic primitives, and they are used extensively in authentication, preserving data integrity, digital signatures, and many other security applications. The two most widely used hash functions are MD5 (Message Digest, m=128) and SHA-1 (Secure Hash Algorithm, m=160), the latter supported by the US government as a standard FIPS-180-2. The collisions for MD5 were found three years ago, and by now they can be produced quickly by software available on the Net. The SHA-1 algorithm seems also to be in trouble (and other algorithms in the SHA family, with m=256, 384, 512, might follow). No collisions for SHA-1 have been found so far, but attacks much better than the simple birthday attack approach have been designed. Breaking SHA-1 soon is a likely possibility.

On January 23, 2007, NIST (National Institute of Standards and Technology) announced an initiative to design a new hash for this century, to be called AHS (Advanced Hash Standard). The competition will be open and it is planned to conclude in 2012. These developments are quite similar to the recent history of symmetric block ciphers - breaking of the DES (Data Encryption Standard) and an emergence of the AES (Advanced Encryption Standard) in 2001 as the winner of a multiyear NIST competition.

This talk will outline the attacks on MD5 and SHA-1 and overview a likely scenario of what the teams submitting new designs for the AHS will consider.

The best attack so far…

- On 17 August 2005, an improvement on the SHA-1 attack was announced on behalf of Xiaoyun Wang, Andrew Yao and Frances Yao at the CRYPTO 2005 rump session, lowering the complexity required for finding a collision in SHA-1 to 2^{63}.

SHA-3 is not yet standardized

- 2007: SHA-3 competition announced
- 2009: 51 submissions cut down to 5
- 2011: 5 finalists under evaluation
  - Michael Pridal-LoPiccolo ('11) studied Keccak for senior thesis
- 2013: Keccak chosen!
- Latest on SHA-3: http://www.nist.gov/itl/csd/sha-100212.cfm

For your pleasure…

- What’s the chance that 2 people in a family of 4 have a birthday in the same *month*?

- How big does our class need to be to have:
  - a 99% chance that 2 have the same birthday?
  - a 100% probability (guaranteed) that 2 have the same birthday?

- **Trivia**: If a professor posts grades for his class by using the last 4 digits of each student’s SSN, what’s the probability that at least 2 students have same last 4 digits?
- …for a class at UIUC? (200 students)
- …for a class at Rose? (30 students)