Remarks on Gödel's code as a hash function

M. Mikuš

Department of Applied Informatics and IT
Slovak University of Technology

CECC'09

Inspiration

Scarlett Thomas: PopCo (bestseller of 2004)

- function in the book used Gödel's numbering function and Fletcher Pratt's frequency table
- result computed on pocket calculator
  - 10-digit display, 20-digit precision

But where Mrs.Thomas took the inspiration?

Definition

Let $\Sigma_1 = \{a, \ldots, z\}$ and $\Sigma_2 = \{0, \ldots, 9\}$.
Then Gödel's hash function is a function $f : {\Sigma_1}^* \rightarrow {\Sigma_2}^n$

- we will use simple example, where $n = 10$.

Computing hash code:

- map message letters according to their position in the frequency table
- use Gödel's numbering function on this sequence of numbers
- cut n most significant digits

Definition

Let $\Sigma_1 = \{a, \ldots, z\}$ and $\Sigma_2 = \{0, \ldots, 9\}$.
Then Gödel's hash function is a function $f : {\Sigma_1}^* \rightarrow {\Sigma_2}^n$

- we will use simple example, where $n = 10$.

Computing hash code:

- map message letters according to their position in the frequency table
- use Gödel's numbering function on this sequence of numbers
- cut n most significant digits

Definition

Let $\Sigma_1 = \{a, \ldots, z\}$ and $\Sigma_2 = \{0, \ldots, 9\}$.
Then Gödel's hash function is a function $f : {\Sigma_1}^* \rightarrow {\Sigma_2}^n$

- we will use simple example, where $n = 10$.

Computing hash code:

- map message letters according to their position in the frequency table
- use Gödel's numbering function on this sequence of numbers
- cut n most significant digits

Definition

Let $\Sigma_1 = \{a, \ldots, z\}$ and $\Sigma_2 = \{0, \ldots, 9\}$.
Then Gödel's hash function is a function $f : {\Sigma_1}^* \rightarrow {\Sigma_2}^n$

- we will use simple example, where $n = 10$.

Computing hash code:

- map message letters according to their position in the frequency table
- use Gödel's numbering function on this sequence of numbers
- cut n most significant digits

Example

message: $M =$'articles'
frequency table:

E T A O I N S R H D L U C
M F Y W G P B V K X Q J Z [1]

compute:
$2^A \times 3^R \times 5^T \times 7^I \times 11^C \times 13^L \times 17^E \times 19^S =$
$2^3 \times 3^8 \times 5^2 \times 7^5 \times 11^{13} \times 13^{11} \times 17^1 \times 19^7 =$
$= 2073462302902019636598357946398633451432775509400$
(code)
$2073462302 = f(M)$ (hash code)

[1] from http://www.math.cornell.edu/ mec/2003-2004/cryptography/subs/frequencies.html

Hash function properties

- preimage resistance
- second preimage resistance
- collision resistance

Preimage resistance

- rational reconstruction method
  - doesn't work here, assumptions are not fulfilled

Collision resistance

First observation:

- special primes: 2 and 5
  $2^7 \times 3^a \times 5^6 \times X = Y$
  $2^8 \times 3^a \times 5^7 \times X = 10 \cdot Y$
  e.g.
  arising $\Leftrightarrow$ erasing
  barter $\Leftrightarrow$ matter $\Leftrightarrow$ waiter
- solution: start the numbering function with 3

Finding another collisions

- main principle: find a message with "almost the same" code
- "almost the same" means $(1 + \varepsilon)$-times larger
- if $\varepsilon < 10^{-10}$ then the hash codes will be the same

How to make a small change

- possible way: shift two consecutive letters: first by -s, other by +s
- the change of the code will be $(p_{i+1}/p_i)^s$
- distance could be more than 1 and also the shift s

How to make a small change

E T A O I N S R H D L U C
M F Y W G P B V K X Q J Z

1-shift: b**ag**s b**ow**s
2-shift: **ar**ch **in**ch
2-shift: **ba**gs **ke**gs
3-shift: ext**in**ct ext**ra**ct

With simple dictionary search we were able to find many meaningful changes.

Algorithm

1. for i = 1 to L: store ratios $r_i = p_{i+1}/p_i$
2. sort all the ratios and group those r_j that have the same prefix
3. compute ratios in these groups and repeat step 2

The higher L, the less steps are required. With $L = 500$, after three steps we got 10 collisions:
$r = 1.00000000001781 =$
$463/461 * 2777/2789 * (2999/3001)^2 * 3041/3037$

Collision example

An influx of party drugs has killed two teenagers but she has a suspect handsome rich newcomer a published author seeking smalltown atmosphere To build her case she moves closer to Mac She risks everything for her investigation even when it means letting her guard down and falling for her suspect Mac Callan lives and breathes for undercover work She still hears his words said in a warm tone innocent His last mission ended in near disaster and he has chance to prove his value She decided and started to pack her things Her movements were soft and quick
hash code: 4187651997

Collision example

An influx of party drugs has killed two teenagers but she has a suspect handsome rich newcomer a published author seeking smalltown atmosphere To build her case she moves closer to Mac She risks everything for her investigation even when it means letting her guard down and falling for her suspect Mac Callan lives and breathes for undercover work She still hears his words said in a **fi**rm tone innocent His last mission ended in near disaster and he has chance to prove his value She decided and started to pack her things Her movements were **de**ft and quick
hash code: 4187651997

Another example

ten people working at cge went down in the plane in the atlantic nine of them were accompanied by a child or father the bitter outcome was a consequence of one of the normally happy rituals of corporate life they were on the flight as a reward for winning a competition for top sales performance a linguet had been at cge for ten years linguets career started in banks moving between strasbourg and the paris region before joining the company our family suffered a fatal blow said c linguet mother of two very well known bankers father of those young men ...
hash code: 5256844069

Another example

ten people working at cge went down in the plane in the atlantic nine of them were accompanied by a child or **mo**ther the bitter outcome was a consequence of one of the normally happy rituals of corporate life they were on the flight as a reward for winning a competition for top sales performance a linguet had been at cge for ten years linguets career started in banks moving between strasbourg and the paris region before joining the company our family suffered a fatal blow said c linguet **fa**ther of two very well known ba**rb**ers **mo**ther of those young men ...
hash code: 5256844069

Summary

Finding meaningfull collisions in Gödel's hash function is

- feasible with birthday attack (with reasonable n)
- tricky but easy – with 2 and 5 in numbering function
- more difficult without 2, but still possible

$\Rightarrow$ *it is very good "classroom example" of a hash function*

Summary

Finding meaningfull collisions in Gödel's hash function is

- feasible with birthday attack (with reasonable n)
- tricky but easy – with 2 and 5 in numbering function
- more difficult without 2, but still possible

$\Rightarrow$ *it is very good "classroom example" of a hash function*

Summary

Finding meaningfull collisions in Gödel's hash function is

- feasible with birthday attack (with reasonable n)
- tricky but easy – with 2 and 5 in numbering function
- more difficult without 2, but still possible

$\Rightarrow$ *it is very good "classroom example" of a hash function*

Summary

Finding meaningfull collisions in Gödel's hash function is

- feasible with birthday attack (with reasonable n)
- tricky but easy – with 2 and 5 in numbering function
- more difficult without 2, but still possible

$\Rightarrow$ *it is very good "classroom example" of a hash function*

Thank you for your attention.