

Re: Hash function, Birthday paradox and probabilities.

Source: <http://sci.tech-archive.net/Archive/sci.math/2006-03/msg02796.html>

- *From:* israel@xxxxxxxxxxx (Robert Israel)
 - *Date:* 16 Mar 2006 02:58:27 GMT
-

In article <1142474371.813464.184320@xx>, <fabrice.gautier@xxxxxxxx> wrote:

I'm wondering about this problem, its coming from a computer science background but I think this is mainly a probability theory problem. The title is probably misleading as I dont really know how to describe the problem in one line, and I'll probably get the vocabulary all wrong, so dont hesitate to fix the way I set the problem :

Lets have:

- a hash function $H()$ that take as input a bit string of length L and return an hash of length N bits (ie an integer in the range $[0, 2^N[)$
- a iterator function F , that takes as input a bit string of length L and return another bit string of length L ,
- $S[0], S[1], \dots S[n]$, bit strings of length L , so that $S[i+1]=F(S[i])$,
- $h[0], \dots h[n]$, so that $h[i]=H(S[i])$
- M a fixed hash

Given all that lets define m , the smallest positive integer so that $h[m]=M$.

I'm wondering how A and H affect the probabilistic properties of m .

You mean F and H ?

....

Now the question is what is the best function $a()$ to minize m (in average), and how to minimize the "spread" of m around its average. (There must be a better word than "spread" but I dont remember)

Re: Hash function, Birthday paradox and probabilities.

If you really want m to be small, make F (which I suppose is the same as "a()") always return a string S such that $H(S) = M$. But I doubt that's what you really want.

I have been trying two things:

- $s(0) = \text{random}()$, $s(i+1) = s(i) + 1$ (modulo 2^{32}),
- $s(i) = \text{random}()$

$s(i) = \text{random}()$ is not compatible with $S[i+1] = F(S[i])$.

Robert Israel israel@xxxxxxxxxxx

Department of Mathematics <http://www.math.ubc.ca/~israel>

University of British Columbia Vancouver, BC, Canada

.