Property-Preserving Hash Functions for Hamming Distance from Standard Assumptions

Nils Fleischhacker, Kasper Green Larsen, and Mark Simkin

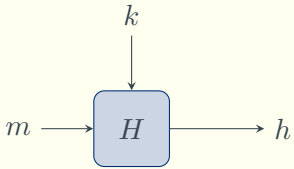
Trondheim, 31. May 2022

RUB

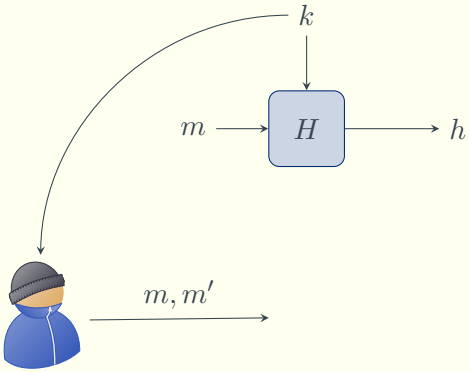
HG HORST GÖRTZ INSTITUT

CASA
CYBER SECURITY IN THE AGE OF LARGE-SCALE ADVERSARIES

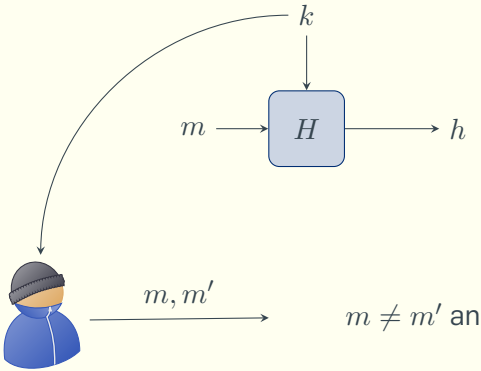
Property Preserving Hash-Functions [BLV19]



Property Preserving Hash-Functions [BLV19]

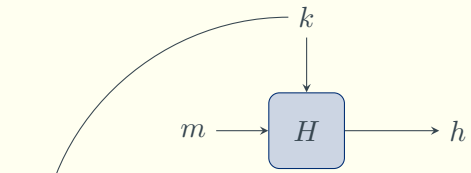


Property Preserving Hash-Functions [BLV19]



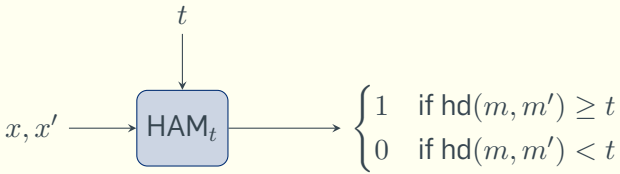
$m \neq m'$ and $H_k(m) = H_k(m')$

Property Preserving Hash-Functions [BLV19]

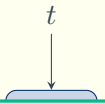


$$P(m, m') \neq \mathsf{Eval}(P, H_k(m), H_k(m'))$$

Hamming Distance Predicate



Hamming Distance Predicate



Observation [FS21]

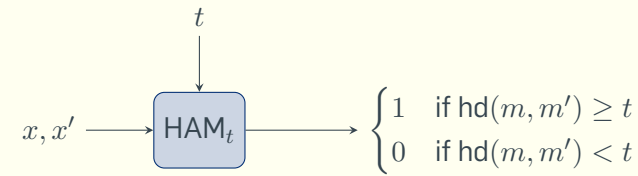
Let $x = x_1 x_2 \dots x_\ell$, $\quad X := \{2i - x_i \mid 1 \le 1 \le \ell\} \subset \{1, \dots, 2\ell\}$

and $x' = x'_1 x'_2 \dots x'_\ell$, $\quad X' := \{2i - x'_i \mid 1 \le 1 \le \ell\} \subset \{1, \dots, 2\ell\}$

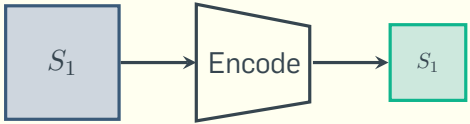
$$\mathsf{HAM}_t(x, x') = 1 \iff |S \triangle S'| \ge 2t$$

where $S \triangle S' = (S \cup S') \setminus (S \cap S')$

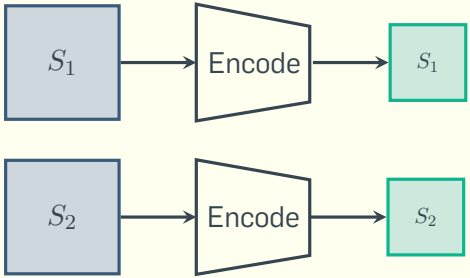
Hamming Distance Predicate



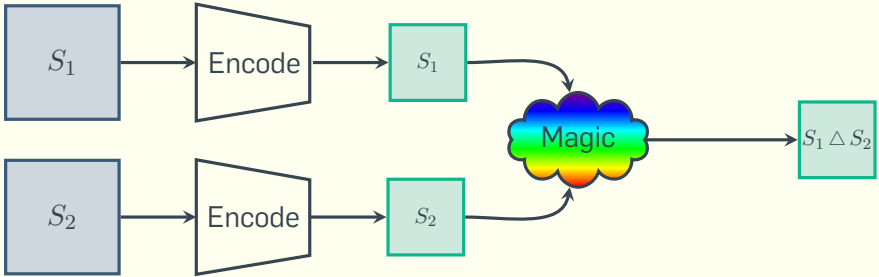
Robust Set Encodings



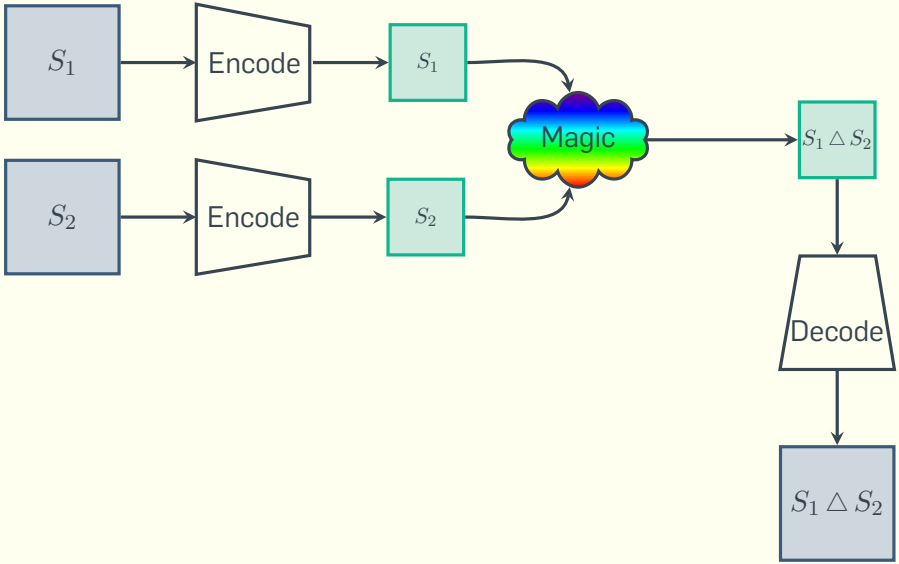
Robust Set Encodings



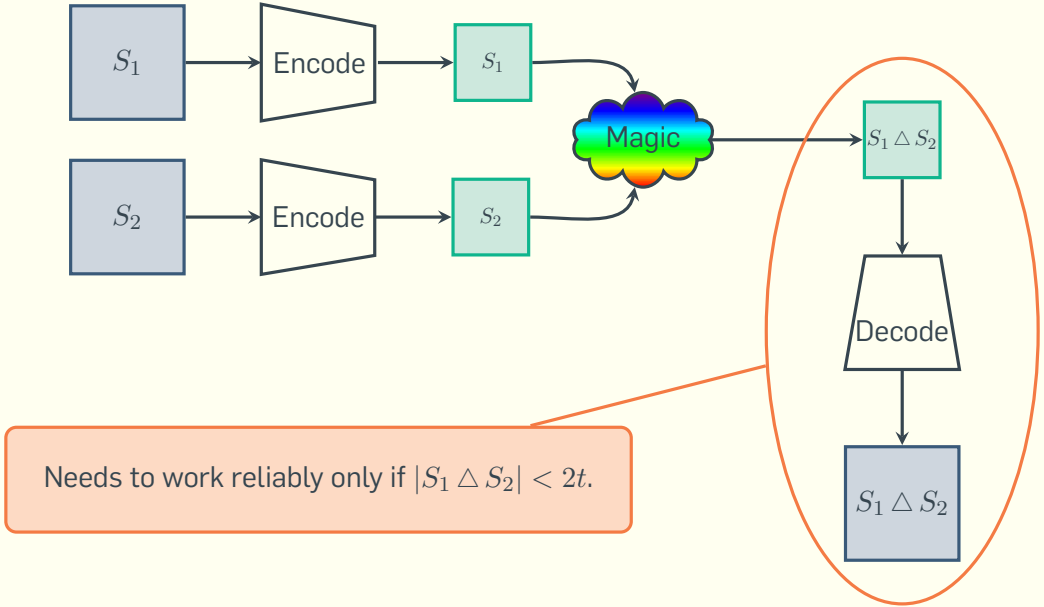
Robust Set Encodings



Robust Set Encodings



Robust Set Encodings



Needs to work reliably only if $|S_1 \triangle S_2| < 2t$.

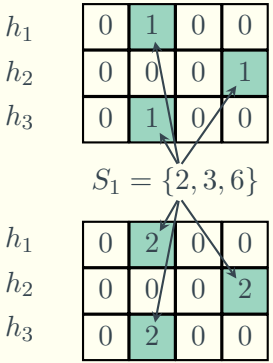
Invertible Bloom Lookup Tables [GM11]

h_1	0	0	0	0
h_2	0	0	0	0
h_3	0	0	0	0

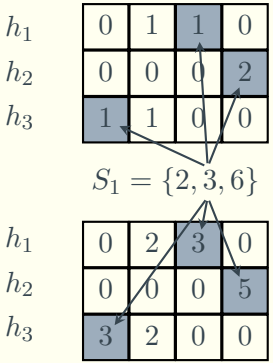
$$S_1 = \{2, 3, 6\}$$

h_1	0	0	0	0
h_2	0	0	0	0
h_3	0	0	0	0

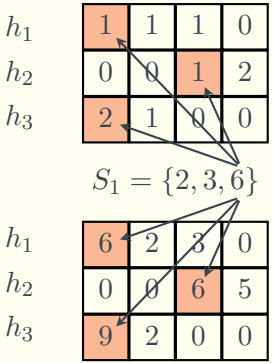
Invertible Bloom Lookup Tables [GM11]



Invertible Bloom Lookup Tables [GM11]



Invertible Bloom Lookup Tables [GM11]



Invertible Bloom Lookup Tables [GM11]

h_1	1	1	1	0
h_2	0	0	1	2
h_3	2	1	0	0

$$S_1 = \{2, 3, 6\}$$

h_1	6	2	3	0
h_2	0	0	6	5
h_3	9	2	0	0

Invertible Bloom Lookup Tables [GM11]

h_1	1	1	1	0
h_2	0	0	1	2
h_3	2	1	0	0

	1	2	0	0
	0	1	1	1
	1	1	1	0

$S_1 = \{2, 3, 6\}$ $S_2 = \{2, 4, 6\}$

h_1	6	2	3	0
h_2	0	0	6	5
h_3	9	2	0	0

	6	6	0	0
	0	4	6	2
	6	2	4	0

Invertible Bloom Lookup Tables [GM11]

h_1	1	1	1	0		1	2	0	0		0	−1	1	0
h_2	0	0	1	2	−	0	1	1	1	=	0	−1	0	1
h_3	2	1	0	0		1	1	1	0		1	0	−1	0

$S_1 = \{2, 3, 6\}$ $S_2 = \{2, 4, 6\}$ $\approx S_1 \triangle S_2 = \{3, 4\}$

h_1	6	2	3	0		6	6	0	0		0	−4	3	0
h_2	0	0	6	5	−	0	4	6	2	=	0	−4	0	3
h_3	9	2	0	0		6	2	4	0		3	0	−4	0

Invertible Bloom Lookup Tables [GM11]

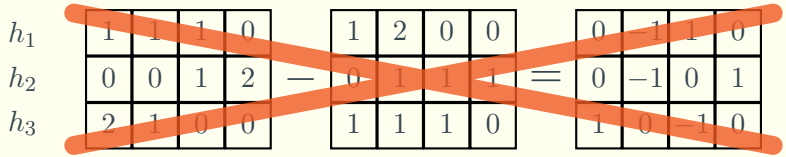
h_1	1	1	1	0		1	2	0	0		0	−1	1	0
h_2	0	0	1	2	−	0	1	1	1	=	0	−1	0	1
h_3	2	1	0	0		1	1	1	0		1	0	−1	0

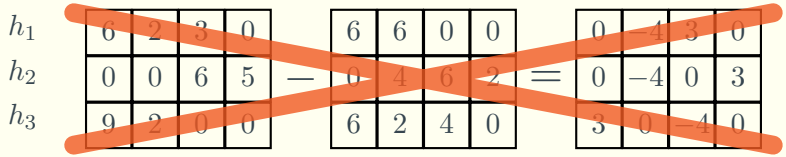
$S_1 = \{2, 3, 6\}$ $S_2 = \{2, 4, 6\}$ $\approx S_1 \triangle S_2 = \{3, 4\}$

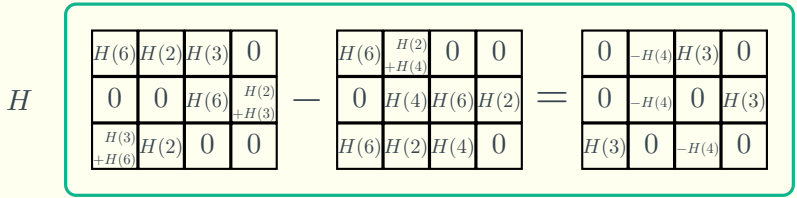
h_1	6	2	3	0		6	6	0	0		0	−4	3	0
h_2	0	0	6	5	−	0	4	6	2	=	0	−4	0	3
h_3	9	2	0	0		6	2	4	0		3	0	−4	0

H	$H(6)$	$H(2)$	$H(3)$	0		$H(6)$	$H(2)+H(4)$	0	0		0	$-H(4)$	$H(3)$	0
	0	0	$H(6)$	$H(2)+H(3)$	−	0	$H(4)$	$H(6)$	$H(2)$	=	0	$-H(4)$	0	$H(3)$
	$H(3)+H(6)$	$H(2)$	0	0		$H(6)$	$H(2)$	$H(4)$	0		$H(3)$	0	$-H(4)$	0

Invertible Bloom Lookup Tables [GM11]







Instantiation

We instantiate the hash function H with Ajtai's hash function [Ajt96] bases on SIS.

Theorem

If the $(n = n(\lambda), 2\ell, q, \sqrt{2\ell + 3}, 2)$-SIS problem is hard then there exists a robust set encoding for universe $[2\ell]$ with output length $\tilde{\mathcal{O}}(t\lambda^2 \log \ell)$.

Corollary

If the $(n = n(\lambda), 2\ell, q, \sqrt{2\ell + 3}, 2)$-SIS problem is hard, there exists a direct access robust PPH for exact Hamming distance for input length ℓ with output length $\tilde{\mathcal{O}}(t\lambda^2 \log \ell)$.

Comparison

	Predicate	Output Length	Assumption
[BLV19]	Gap Hamming	$c \cdot \ell, c < 1$	SSV
[FS21]	Exact Hamming	$\tilde{\mathcal{O}}(t\lambda)$	q-SBDL
This work	Exact Hamming	$\tilde{\mathcal{O}}(t\lambda^2 \log \ell)$	SIS
[Min22]	Exact Hamming	$\tilde{\mathcal{O}}(t^2\lambda^2 \log \ell)$	CRHF

t: threshold, ℓ: input length, λ: security parameter