Collisions for the LPS expander graph hash function

Jean-Pierre Tillich[1] and Gilles Zémor[2]

[1]INRIA, Équipe SECRET

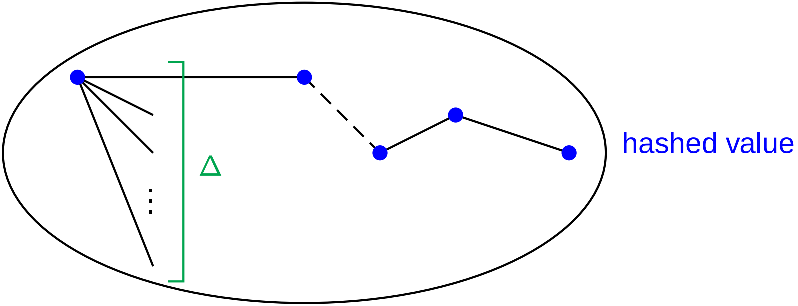
[2]Bordeaux Mathematics Institute

Eurocrypt 2008, Istanbul

Hash functions from graphs

Take a large graph $\mathcal{G}$, (e.g. 2^{1000} vertices), regular of small degree Δ.

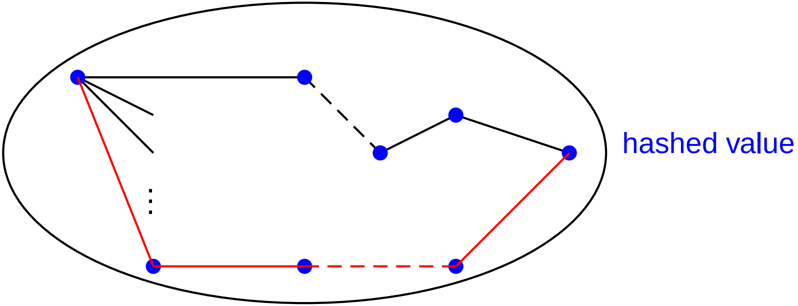
- Input text $\in \{0, 1, \ldots, \Delta - 2\}^* \longrightarrow$ non-bactracking walk from fixed vertex
- hashed value $\longrightarrow$ endpoint.



Hash functions from graphs

Take a large graph $\mathcal{G}$, (e.g. 2^{1000} vertices), regular of small degree Δ.

- Input text $\in \{0, 1, \ldots, \Delta - 2\}^* \longrightarrow$ non-bactracking walk from fixed vertex
- hashed value $\longrightarrow$ endpoint.



- Collisions $\longrightarrow$ cycles.

Hash functions from expander graphs

- Graph should be easy to describe.
- No short cycles.
- Suggestion (Charles, Goren, Lauter 06): use known expander graphs. Advantage: rapidly-mixing property. Distribution of hashed values is almost uniform for short $O(\log \#\{\text{vertices}\})$ uniform inputs.

  In particular: use the Lubotzky, Phillips, Sarnak (LPS) Ramanujan graphs.
- *Strength of the function rests on supposed difficulty of finding explicit short cycles*.

Hash functions from expander graphs

- Graph should be easy to describe.
- No short cycles.
- Suggestion (Charles, Goren, Lauter 06): use known expander graphs. Advantage: rapidly-mixing property. Distribution of hashed values is almost uniform for short $O(\log \#\{\text{vertices}\})$ uniform inputs.

  In particular: use the Lubotzky, Phillips, Sarnak (LPS) Ramanujan graphs.
- *Strength of the function rests on supposed difficulty of finding explicit short cycles*.
- History of the large graph hashing strategy: later on.

The LPS Ramanujan graphs

Graph $\mathcal{G}$ is a *Cayley graph*. Vertices are elements of a group G and $x \longleftrightarrow y$ is an edge iff $y = xs$ for s in a fixed set $\mathcal{S}$ (of generators).

The LPS Ramanujan graphs

Graph $\mathcal{G}$ is a *Cayley graph*. Vertices are elements of a group G and $x \longleftrightarrow y$ is an edge iff $y = xs$ for s in a fixed set $\mathcal{S}$ (of generators).

Specifically: p large prime, ℓ small prime $\equiv 1 \mod 4$, G a group of 2×2 matrices, elements in $\mathbb{F}_p$, generator set $\mathcal{S}$ made up of the matrices

$$S = \begin{pmatrix} a + \iota b & c + \iota d \\ -c + \iota d & a - \iota b \end{pmatrix}$$

where $\iota^2 = -1$ in $\mathbb{F}_p$ and a, b, c, d integers such that

$$\begin{cases} \det S = a^2 + b^2 + c^2 + d^2 & = \ell \\ a > 0, \ \ a & \equiv 1 \pmod 2 \\ b \equiv c \equiv d & \equiv 0 \pmod 2 \end{cases}$$

The LPS Ramanujan graphs (2)

Identify matrices obtained from each other through multiplication by $\lambda \in \mathbb{F}_p$. $\mathcal{S}$ generates a subgroup G of $\mathrm{PGL}_2(\mathbb{F}_p)$, (isomorphic to $\mathrm{PSL}_2(\mathbb{F}_p)$), and $\mathcal{S} = \mathcal{S}^{-1}$. $|\mathcal{S}| = \ell + 1$.

This is the graph $X_{\ell,p}$.

- $\#$Vertices $= p(p^2 - 1)/2$,
- degree $\Delta = \ell + 1$.

Facts:

- no small cycles: smallest has length $2 \log_{\Delta-1} p$
- good expansion properties.

The LPS Ramanujan graphs (3)

Example, $\ell = 5$:

$$S_1 = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \quad S_2 = \begin{pmatrix} 1+2\iota & 0 \\ 0 & 1-2\iota \end{pmatrix} \quad S_3 = \begin{pmatrix} 1 & 2\iota \\ 2\iota & 1 \end{pmatrix}$$
$$S_4 = \begin{pmatrix} 1 & -2\iota \\ -2\iota & 1 \end{pmatrix} \quad S_5 = \begin{pmatrix} 1-2\iota & 0 \\ 0 & 1+2\iota \end{pmatrix} \quad S_6 = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$$

We have: $\mathcal{S} = \mathcal{S}^{-1}$.

$$S_1 S_6 = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} = 5 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{in } G$$

The LPS Ramanujan graphs (3)

Example, $\ell = 5$:

$$S_1 = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \qquad S_2 = \begin{pmatrix} 1+2\iota & 0 \\ 0 & 1-2\iota \end{pmatrix} \quad S_3 = \begin{pmatrix} 1 & 2\iota \\ 2\iota & 1 \end{pmatrix}$$
$$S_4 = \begin{pmatrix} 1 & -2\iota \\ -2\iota & 1 \end{pmatrix} \quad S_5 = \begin{pmatrix} 1-2\iota & 0 \\ 0 & 1+2\iota \end{pmatrix} \quad S_6 = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$$

We have: $\mathcal{S} = \mathcal{S}^{-1}$.

$$S_1 S_6 = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} = 5 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{in } G$$

Input text of length t is put into $1-1$ correspondence with product

$$G_1 G_2 \ldots G_t$$

such that $G_i \in \mathcal{S}$, $G_i G_{i+1} \neq 1$.

Looking for collisions

A collision is equivalent to a short cycle in the graph $X_{\ell,p}$, i.e. a string $G_1 G_2 \ldots G_t$ of elements of $\mathcal{S}$ such that $G_i G_{i+1} \neq 1$ and

$$\prod_{i=1}^{t} G_i = 1 \quad \text{in } G.$$

Looking for collisions

A collision is equivalent to a short cycle in the graph $X_{\ell,p}$, i.e. a string $G_1 G_2 \dots G_t$ of elements of $\mathcal{S}$ such that $G_i G_{i+1} \neq 1$ and

$$\prod_{i=1}^{t} G_i = 1 \quad \text{in } G.$$

The idea.

Looking for collisions

A collision is equivalent to a short cycle in the graph $X_{\ell,p}$, i.e. a string $G_1 G_2 \dots G_t$ of elements of $\mathcal{S}$ such that $G_i G_{i+1} \neq 1$ and

$$\prod_{i=1}^{t} G_i = 1 \quad \text{in } G.$$

The idea.

Lift the graph $X_{\ell,p}$ to the Cayley graph generated by the matrices

$$M(a,b,c,d) = \begin{pmatrix} a+ib & c+id \\ -c+id & a-ib \end{pmatrix}$$

where $i \in \mathbb{C}$ and (as before)

$$\begin{cases} \det S = a^2+b^2+c^2+d^2 & = & \ell \\ a > 0,\ \ a & \equiv & 1 \pmod 2 \\ b \equiv c \equiv d & \equiv & 0 \pmod 2 \end{cases}$$

The universal cover of $X_{\ell,p}$

The set of products of $M(a,b,c,d)$'s (lifted generators of $\mathcal{S}$) is

$$\Omega = \left\{ \begin{pmatrix} a+ib & c+id \\ -c+id & a-ib \end{pmatrix} \middle| (a,b,c,d) \in E_w \ \text{ for some } w > 0 \right\}$$

where E_w is the set of 4-tuples $(a,b,c,d) \in \mathbb{Z}^4$ such that

$$\begin{cases} a^2+b^2+c^2+d^2 & = & \ell^w \\ a>0, \ \ a & \equiv & 1 \pmod 2 \\ b \equiv c \equiv d & \equiv & 0 \pmod 2. \end{cases}$$

Factoring in Ω is easy. If $M = G_1 G_2 \ldots G_t$, find G_t by finding the unique (lifted) generator $S \in \mathcal{S}$ such that MS has entries in $\mathbb{Z}[i]$ divisible by ℓ ! Then $G_t = S^{-1}$.

Lifting the identity

Finding a collision is now reduced to lifting the identity element in G to a matrix of Ω with reasonable length w. Means find

$$\begin{pmatrix} a+ib & c+id \\ -c+id & a-ib \end{pmatrix}$$

such that the integers a, b, c, d satisfy

$$\begin{cases} a^2+b^2+c^2+d^2 & = & \ell^w \\ a>0,\ \ a & \equiv & 1 \pmod 2 \\ b \equiv c \equiv d & \equiv & 0 \pmod 2 \end{cases}$$

and b, c, d, multiples of p.

Lifting the identity (2)

set $b = 2px$, $c = 2py$, $d = 2pz$. The search for solutions of $a^2 + b^2 + c^2 + d^2 = \ell^w$ becomes

$$a^2 + 4p^2(x^2 + y^2 + z^2) = \ell^{2k}$$

and

$$(\ell^k - a)(\ell^k + a) = 4p^2(x^2 + y^2 + z^2).$$

Set $a = \ell^k - 2mp^2$, arbitrary m (in practice $m = 1, 2$). We get

$$x^2 + y^2 + z^2 = m(\ell^k - mp^2).$$

Solve through taking random z, check whether right hand side $-z^2$ is sum of two squares.

fast computation of collisions

Limiting factor is number of random choices of z to get a sum of two squares ($\log p$). Then decompose into sum of two squares ($\log p$).

In practice: overall complexity small power of $\log p$. No problem for p 1000-bit prime.

History

A similar scheme (Z. 91) with $G = \mathrm{SL}_2(\mathbb{F}_p)$ and set of generators $\mathcal{S}$ consisting of

$$S_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad S_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

(Graph $\mathcal{G}$ is *directed*).

History

A similar scheme (Z. 91) with $G = \mathrm{SL}_2(\mathbb{F}_p)$ and set of generators $\mathcal{S}$ consisting of

$$S_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad S_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

(Graph $\mathcal{G}$ is *directed*).

(Tillich-Z. 93) collisions through lifting the identity to a product of S_1's and S_2's in $\mathrm{SL}_2(\mathbb{Z})$. Then use Euclidean algorithm to finish factorisation. Problem lies in the (too large) density of the set of products of S_1's and S_2's in $\mathrm{SL}_2(\mathbb{Z})$.

(Bold) comparison with factoring

How does one factor an integer n ?

(Bold) comparison with factoring

How does one factor an integer *n* ?

Take a set $\mathcal{S} = \{2^2, 3^2, 5^2, \ldots, \ell^2\}$ (set of squares of small primes). Generator set of Cayley graph $\mathcal{G}$ over (multiplicative) subgroup of $\mathbb{Z}/n\mathbb{Z}$ (the invertible squares).

Lift random square to a product of elements of $\mathcal{S}$ in $\mathbb{Z}$. Finish with Euclidean algorithm.

Conclusion: Future for Cayley-graph based hashing ?

Goal: defeat density or lifting attacks.

Suggestion for LPS-based hashing: throw away some generators. For $S \in \mathcal{S}$ keep either S or S^{-1} but not both. Keeps part of the expansion properties. Speed of convergence to uniform less easy to estimate but small diameter easy to prove.

Other possibilities: look for other interesting sets of generators of $\mathrm{SL}_2()$ groups with a view to defeating lifting attacks. How does one find short factorisations of 1 in the group ?

(Tillich-Z. 94) $G = \mathrm{SL}_2(\mathbb{F}_{2^m})$ and set of generators $\mathcal{S}$ consisting of:

$$S_1 = \begin{pmatrix} X & 1 \\ 1 & 0 \end{pmatrix} \qquad S_2 = \begin{pmatrix} X & X+1 \\ 1 & 1 \end{pmatrix}$$

For given trusted defining polynomials of $\mathbb{F}_{2^m}$, no known method for producing short factorisations, i.e. reasonable-length collisions.