Améliorations au problème du logarithme discret dans $\mathbb{F}_p^*$

Răzvan Bărbulescu

21 avril 2011

L'algorithme de Commeine et Semaev

$(\star)$	PRÉ-CALCULS	
1.1	cribler sur les polynômes de $\{a + bX \mid a, b \in \mathbb{Z}\}$ afin de trouver un nombre suffisant d'entre eux qui soient $L(\frac{1}{3}) - \textit{friable}$ sur $\mathcal{O}_1$ et $\mathcal{O}_2$;	
1.2	calculer les logs virtuels par un algorithme d'algèbre linéaire	
$(\star\star)$	LOG INDIVIDUEL	
2.1	chercher h pour que $Q^h S$ mod p soit $L(\frac{2}{3})$-friable où Q est un premier dont on connaît le log discret; factoriser $Q^h S = q_1 q_2 \dots q_k$;	
2.2	pour tout j, descendre par "special Q" et trouver $log_t q_j$ en fonction des logs virtuels connus.	

L'algorithme de Commeine et Semaev

$(\star)$	PRÉ-CALCULS	
1.1	cribler sur les polynômes de $\{a + bX \mid a, b \in \mathbb{Z}\}$ afin de trouver un nombre suffisant d'entre eux qui soient $L(\frac{1}{3}) - \textit{friable}$ sur $\mathcal{O}_1$ et $\mathcal{O}_2$;	
1.2	calculer les logs virtuels par un algorithme d'algèbre linéaire	
$(\star\star)$	LOG INDIVIDUEL	
2.1	chercher h pour que $Q^h S$ mod p soit $L(\frac{2}{3})$-friable où Q est un premier dont on connaît le log discret; factoriser $Q^h S = q_1 q_2 \dots q_k$;	$L_p(\frac{1}{3}, 1.447)$
2.2	pour tout j, descendre par "special Q" et trouver $log_t q_j$ en fonction des logs virtuels connus.	

L'algorithme de Commeine et Semaev

$(\star)$	PRÉ-CALCULS	
1.1	cribler sur les polynômes de $\{a + bX \mid a, b \in \mathbb{Z}\}$ afin de trouver un nombre suffisant d'entre eux qui soient $L(\frac{1}{3}) - \textit{friable}$ sur $\mathcal{O}_1$ et $\mathcal{O}_2$;	$L_p(\frac{1}{3}, 1.902)$
1.2	calculer les logs virtuels par un algorithme d'algèbre linéaire	$L_p(\frac{1}{3}, 1.902)$
$(\star\star)$	LOG INDIVIDUEL	
2.1	chercher h pour que $Q^h S$ mod p soit $L(\frac{2}{3})$-friable où Q est un premier dont on connaît le log discret; factoriser $Q^h S = q_1 q_2 \dots q_k$;	$L_p(\frac{1}{3}, 1.447)$
2.2	pour tout j, descendre par "special Q" et trouver $log_t q_j$ en fonction des logs virtuels connus.	

L'algorithme de Commeine et Semaev

$(\star)$	PRÉ-CALCULS	
1.1	cribler sur les polynômes de $\{a + bX \mid a, b \in \mathbb{Z}\}$ afin de trouver un nombre suffisant d'entre eux qui soient $L(\frac{1}{3}) - \textit{friable}$ sur $\mathcal{O}_1$ et $\mathcal{O}_2$;	$L_p(\frac{1}{3}, 1.902)$
1.2	calculer les logs virtuels par un algorithme d'algèbre linéaire	$L_p(\frac{1}{3}, 1.902)$
$(\star\star)$	LOG INDIVIDUEL	
2.1	chercher h pour que $Q^h S$ mod p soit $L(\frac{2}{3})$-friable où Q est un premier dont on connaît le log discret; factoriser $Q^h S = q_1 q_2 \dots q_k$;	$L_p(\frac{1}{3}, 1.447)$
2.2	pour tout j, descendre par "special Q" et trouver $log_t q_j$ en fonction des logs virtuels connus.	$L_p(\frac{1}{3}, 1.189)$

Complexité de l'étape de friabilisation

Soit C la borne de friabilité de l'étape 2.1. Soit $\theta > 0$ un paramètre tel que $C = L_p(\theta, a)$ pour un $a > 0$. Un test *ECM* coûte $L_C(\frac{1}{2}, \sqrt{2}) = L_p(\frac{\theta}{2}, *)$. Ainsi le temps moyen de l'étape 2.1 est:

$$P_{smooth}(L_p(1), L_p(\theta))^{-1} \cdot L_p(\frac{\theta}{2}) = L_p(max\{1-\theta, \frac{\theta}{2}\}, *). \quad (1)$$

Cela est minimal pour $\theta = \frac{2}{3}$. Il nous reste à choisir a et de poser $C = L_p(\frac{2}{3}, a)$.

L'idée de l'admissibilitè

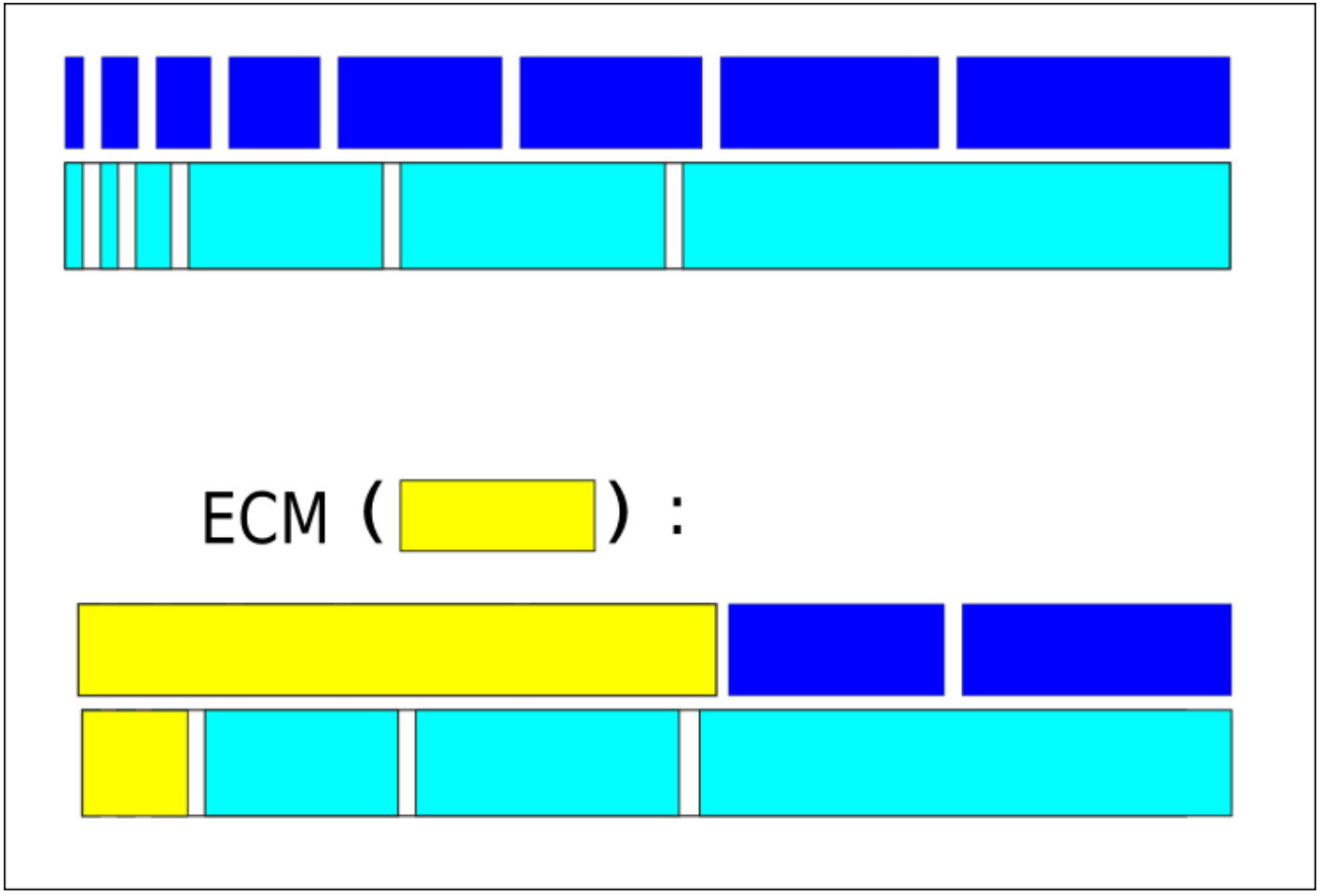
But : Trouver un nombre $L_p(\frac{2}{3}, a)$-friable pour une valeur de $a > 0$ non imposée.

Outils : Un générateur aléatoire de nombres de taille p et un algorithme qui teste la β-friabilité en temps $L_\beta(\frac{1}{2}, \sqrt{2})$.

Idée : Soumettre les candidats à un test rapide et ne garder que les admissibles pour le test final. Connue depuis les années 70.

Intuition

Que se passe-t-il si on fait un test ECM avec une borne plus petite que celle qui nous interesse? Dans le graphique le nombre bleu clair est quelconque alors que celui bleu fonçé est friable.



Notations

Soient $0 < \theta, c < 1$ paramètres à choisir. Pour tout $m \leq n$ on note m_1 le plus grand diviseur de m qui est $L_n(\frac{2}{3}, \theta a)$-friable. On appelle M_{ble} l'ensemble des candidats admissibles:

$$M_{ble} = \{m \leq n | \ \frac{m}{m_1} \leq n^{(1-c)}\}. \tag{2}$$

Ensuite on définit l'ensemble des candidats admis:

$$M_{ed} = \{m \in M_{ble} | m \text{ is } L(\frac{2}{3}, a)\text{-friable}\}. \tag{3}$$

Deux théorèmes de friabilité

On note $\psi(x, y)$ le nombre d'entiers inférieurs à x qui sont y-friables. On pose $\psi(x, y, z)$ le nombre d'entiers inférieurs à x et dont tous les facteurs premiers sont dans $[z, y]$.

Théorème (Canfield Erdös Pomerance)

Soit ϵ une constante strictement plus grande que 0. Alors on a:

$$\psi(x, B)/x = u^{-u(1+o(1))}$$

uniformément dans la région du plan où $x \geq 10$ et $B \geq (\log x)^{1+\epsilon}$, sachant que $u = \frac{\log x}{\log B}$ et que la limite dans $o(1)$ se réfère à $u \to \infty$.

Théorème (Pomerance)

Soit $\epsilon > 0$. Si $u = \frac{\log x}{\log y}$ et $z < y^{1-\frac{1}{\log u}}$, alors on a:

$$\psi(x, y, z) = x \cdot u^{-u+o(u)} \tag{4}$$

uniformément pour $(\log x)^{1+\epsilon} < y < exp(\log x)^{1-\epsilon}$, $x \geq 10$.

Le résultat principal pour l'admissibilité

Le test d'admissibilité est le suivant "$ECM(B^{\theta})$ mange une fraction c du nombre, avec $0 < \theta \leq 1$".

Théorème

Soient $0 < c, \theta < 1$, $n \in \mathbb{N}$ et définissons M_{ble} et M_{ed} comme précédemment. Alors:

(i) $\#M_{ble} \leq n \cdot L_n(\frac{1}{3}, -\frac{c}{3\theta a})^{1+o(1)}$;

(ii) $\#M_{ed} \geq n \cdot L_n(\frac{1}{3}, -\frac{c}{3\theta a} - \frac{1-c}{3a})^{1+o(1)}$.

Analyse de complexité 1

Le nombre moyen de candidats à tester pour trouver un admis M_{ed} est $\frac{n}{\#M_{ed}}$. Tous les candidats prennent le test d'admissibilité qui dure $t_{ECM}(L_n(\frac{2}{3}, a\theta))$ pour chaque. Le nombre moyen d'admissibles à examiner pour trouver un admis est $\frac{\#M_{ble}}{\#M_{ed}}$. Le test final dure $t_{ECM}(L_n(\frac{2}{3}), a)$ pour chaque admissible. On obtient la formule:

$$\frac{n}{\#M_{ed}} \cdot t_{ECM}(L_n(\frac{2}{3}, a\theta)) + \frac{\#M_{ble}}{\#M_{ed}} \cdot t_{ECM}(L_n(\frac{2}{3}, a)) \qquad (5)$$

Analyse de complexité 2

D'après le théorème précédent, le temps moyen est inférieur à:

$$L_n(\frac{1}{3}, \frac{1-c}{3a} + \frac{c}{3\theta a} + \sqrt{\frac{4}{3}\theta a}) + L_n(\frac{1}{3}, \frac{1-c}{3a} + \sqrt{\frac{4}{3}a}). \quad (6)$$

On trouve une formule $\theta = \theta(a, c)$ pour minimiser à a et c fixé. Ensuite on trouve une formule $c = c(a)$. On reste avec une fonction en a dont le minimum est pour $a = 0.811$. Cela rend la complexité $L_n(\frac{1}{3}, 1, 296)$.

La stratégie de selection par tests d'admissibilité

Un seul test d'admissibilité rend la complexité $L_n(\frac{1}{3}, 1.296)^{1+o(1)}$.
Une succession de 8 tests la descendent à $L_n(\frac{1}{3}, 1.232)$.
On note m_i le plus grand diviseur de m dans $\psi(n, L_n(2/3, \theta_i a), L_n(2/3, \theta_{i-1} a))$. On pose M^0_{ble} et on définit: $M^i_{ble} = \{m \in M^{i-1}_{ble} |\ m_i \geq n^{c_i}\}$.

Théorème

Avec les notations antérieures on a:

(i) $\#M^i_{ble} \leq n \cdot L_n(\frac{1}{3}, -\frac{c_1}{3\theta_1 a} - \ldots - \frac{c_i}{3\theta_i a})^{1+o(1)}$;

(ii) $\#M_{ed} \geq n \cdot L_n(\frac{1}{3}, -\frac{c_1}{3\theta_1 a} - \ldots - \frac{c_k}{3\theta_k a} - \frac{1-c_1-\ldots-c_k}{3a})^{1+o(1)}$.

Discrete Logarithm Factory

Algorithme

PRÉREQUIS: un entier p_0 et $m_0 = L_{p_0}(\frac{2}{3}, \frac{1}{\delta})$; un fichier permanent contenant des couples (a, b) tels que $a + m_0 b$ est B-friable.

0. *trouver un polynôme irréductible $f \in \mathbb{Z}$ tel que $p \mid f(m_0)$ et $|f|_\infty \leq m_0$; posons $K_1 = \mathbb{Q}$, $K_2 = \mathbb{Q}[X]/\langle f \rangle$.*

($\star$) *PRÉCALCULS*

1.1 *CRIBLE;*

1.2 *ALGÈBRE LINÉAIRE;*

($\star\star$) *LOG INDIVIDUEL*

2.1 *SMOOTHNING: prendre un nombre premier Q dans la base de facteurs du côté rationnel; tirer au hasard h jusqu'à ce que $Q^h S$ est C-friable; factoriser $Q^h S = q_1 q_2 \dots q_k$;*

2.2 *DESCENTE: pour tout $j \leq k$ calculer les log discrets par descente.*

Compromis entre temps et espace

$$K = RR$$

$$b, d, e = var('beta\ delta\ epsilon')$$

$$P = 2 * e - 2/(3 * d * b) - d * e/(3 * b) - b$$

$$Q = 2 * b - b - 1/(3 * b * d) - e * d/(3 * b)$$

$$R = b + 1/(3 * d * b) + e * d/(3 * b)$$

$$solve([P == 0, Q == 0, R == 1.639], b, d, e)$$

Pour $R \leq 1.635$ on n'a plus de solutions réelles. Donc l'echange temps contre espace coûte cher en temps.

Conclusion

Ce stage a abouti sur l'amélioration de l'algorithme de Commeine et Semaev de deux manières différentes.

1. On a accélérer les précalculs de $L_p(\frac{1}{3}, 1.923)$ à $L_p(\frac{1}{3}, 1.639)$ grâce à l'idée de l'algorithme Factorization Factory.
2. On a adapter l'idée de l'admissibilité de Pomerance 1982 pour rendre le logarithme individuel plus rapide: de $L_p(\frac{1}{3}, 1.447)$ à $L_p(\frac{1}{3}, 1.232)$.