ROZSZERZONY ALGORYTM POHLIGA-HELLMANA I JEGO ZASTOSOWANIE DO FAKTORYZACJI

Bartosz Źrałek

Instytut Matematyki Uniwersytetu Warszawskiego,
Banacha 2, 02-097 Warszawa
b.zralek@mimuw.edu.pl

Streszczenie. Wskażemy ścisły związek między problemami logarytmu dyskretnego i faktoryzacji. Opiszemy mianowicie uogólnienie algorytmu Pohliga-Hellmana dla grup niecyklicznych $\mathbb{Z}_n^*$, które można zastosować do derandomizacji algorytmu $p-1$ Pollarda. Algorytm ten bowiem w w wersji potrzebuje źródła losowości. Okazuje się, że obliczenia można przeprowadzić deterministycznie bez znaczącego pogorszenia złożoności.

Słowa kluczowe: faktoryzacja, logarytm dyskretny, derandomizacja

1. Wprowadzenie

W kryptografii klucza publicznego, jawnie zapoczątkowanej przez Diffiego i Hellmana [2], bezpieczeństwo kryptosystemu opiera się zazwyczaj na trudności obliczeniowej pewnych problemów teorii liczb - głównie problemu logarytmu dyskretnego w grupie cyklicznej oraz problemu faktoryzacji liczb naturalnych. Te problemy są uważane za trudne, ponieważ przypuszcza się, że nie istnieją algorytmy rozwiązujące je w czasie wielomianowym (ze względu na rozmiar danych). Najszybsze obecnie algorytmy rozkładu na czynniki liczb naturalnych (np. sito ciała liczbowego [8]) mają pesymistyczną złożoność podwykładniczą. Podobnie jest w przypadku problemu logarytmu dyskretnego, choć jedynie dla grup cyklicznych o„korzystnej obliczeniowo" reprezentacji takich jak grupy multyplikatywne ciał skończonych [4] (w tzw. modelu generycznym [3] złożoność ta jest bowiem wykładnicza). W praktyce kryptograficznej nie można jednak ograniczać badań do ogólnych instancji powyższych problemów. Podstawową rolę odgrywają metody rozwiązywania szczególnych przypadków - z punktu widzenia obliczeniowego łatwych, czyli, z punktu widzenia kryptoanalizy, prowadzących do wygenerowania słabych parametrów kryptosystemu. Takimi łatwymi instancjami są przykładowo:

1. problem logarytmu dyskretnego w grupie cyklicznej o gładkim (podzielnym jedynie przez małe liczby pierwsze) rzędzie - efektywnym atakiem jest algorytm Pohliga-Hellmana [10].

2. problem rozkładu liczby n mającej dzielnik pierwszy p, taki że rząd grupy $\mathbb{Z}_p^*$ (równy $p-1$) jest gładki - efektywnym atakiem jest algorytm $p-1$ Pollarda [11]. (Idea Pollarda była prawdopodobnie inspiracją do opracowania przez Lenstrę [9] algorytmu faktoryzacji, w którym poszukuje się takiej „krzywej eliptycznej nad $\mathbb{Z}_n$", by dla pewnego dzielnika pierwszego p liczby n, grupa jej punktów modulo p miała gładki rząd.)

Wskażemy ścisły związek między problemami logarytmu dyskretnego i faktoryzacji. Opiszemy mianowicie uogólnienie algorytmu Pohliga-Hellmana dla grup niecyklicznych $\mathbb{Z}_n^*$, które można zastosować do derandomizacji algorytmu $p-1$ Pollarda. Algorytm ten bowiem w oryginalnej wersji potrzebuje źródła losowości. Okazuje się, że obliczenia można przeprowadzić deterministycznie bez znaczącego pogorszenia złożoności. Przedstawione tu wyniki zostały opublikowane w [13].

2. Klasyczny algorytm Pohliga-Hellmana

Przypomnijmy najpierw działanie klasycznego algorytmu Pohliga-Hellmana. Niech G będzie grupą cykliczną rzędu N, generowaną przez g: $G = \langle g \rangle$, $\#G = N$. Niech $h \in G$. Problem logarytmu dyskretnego, w skrócie DLP, w G to znalezienie $l \in \mathbb{Z}$, takiego że $g^l = h$. Przypuśćmy, że znamy nietrywialny rozkład

$$N = N_1 N_2, \ \mathrm{NWD}(N_1, N_2) = 1.$$

Mamy izomorfizm

$$\langle g \rangle \cong \langle g^{N_1} \rangle \times \langle g^{N_2} \rangle.$$

Pomysł polega na sprowadzieniu DLP w $\langle g \rangle$ do DLP w grupie $\langle g^{N_1} \rangle$ i DLP w grupie $\langle g^{N_2} \rangle$, czyli w grupach mniejszych niż G, bo $\#\langle g^{N_1} \rangle = N_2$, $\#\langle g^{N_2} \rangle = N_1$. Jeśli znajdziemy $l_i \in \mathbb{Z}$, takie że $(g^{N_i})^{l_i} = h^{N_i}$, to wystarczy rozwiązać układ

$$\begin{cases} l \equiv l_1 \ (N_2) \\ l \equiv l_2 \ (N_1). \end{cases}$$

Ogólnie, jeśli mamy pełną faktoryzację rzędu N, to możemy sprowadzić DLP w grupie G do DLP w jej podgrupach rzędów p^{v_p}, gdzie p^{v_p} jest maksymalną potęgą liczby pierwszej p dzielącą N.

Podajemy przykład obliczeń algorytmu w podgrupie grupy G rzędu 2^v. Niech $a, b \in G$, przy czym a jest rzędu 2^v, co zapiszemy $\mathrm{rz}(a) = 2^v$,

i niech $b^{2^v} = 1$. Szukamy liczby l, dla której $a^l = b$. Zapisujemy logarytm l binarnie z niewiadomymi cyframi l_i:

$$a^{l_0+l_1\cdot 2+\ldots+l_{v-1}\cdot 2^{v-1}} = b.$$

Po podniesieniu obu stron do potęgi 2^{v-1} otrzymujemy

$$a^{2^{v-1}l_0} = b^{2^{v-1}},$$

czyli

$$(-1)^{l_0} = b^{2^{v-1}}.$$

To nam daje wartość l_0 (0 lub 1). Dalej

$$a^{l_1\cdot 2+\ldots+l_{v-1}\cdot 2^{v-1}} = b \cdot a^{-l_0}$$

i podobnie kolejno obliczamy $l_1, \ldots, l_{v-1}$.

3. Rozszerzony algorytm Pohliga-Hellmana

Rozpatrzmy teraz skończoną grupę przemienną G, ale niekoniecznie cykliczną. Zilustrujemy ideę uogólnienia algorytmu Pohliga-Hellmana na przykładzie najprostszego przypadku: $a, b \in G$, $\mathrm{rz}(a) = 2^v$, $b^{2^v} = 1$. Możemy spróbować zastosować klasyczny algorytm Pohliga-Hellmana, by znaleźć $l \in \mathbb{Z}$, takie że $a^l = b$. Jeśli się nie uda, to znaczy że podgrupa $\langle a, b\rangle$, a więc i G, nie jest cykliczna. Tego rodzaju implikacja znalazłaby zastosowanie w testowaniu pierwszości (dowodzeniu złożoności): dla $G = \mathbb{Z}_n^*$, jeśli G nie jest cykliczna, to n jest liczbą złożoną. Niecykliczność podgrupy $\langle a, b\rangle$ można tu jednak wykorzystać pełniej - rozłożyć liczbę n.

Niech więc $G = \mathbb{Z}_n^*$, gdzie n jest, powiedzmy, iloczynem pq dwóch nieparzystych liczb pierwszych. Z chińskiego twierdzenia o resztach,

$$\mathbb{Z}_n^* \simeq \mathbb{Z}_p^* \times \mathbb{Z}_q^*.$$

Przy tym izomorfizmie poczynimy utożsamienia $a \leftrightarrow (a_1, a_2)$, $b \leftrightarrow (b_1, b_2)$. Załóżmy bez utraty ogólności, że $\mathrm{rz}(a_1) = 2^v$ (po ewentualnej zamianie p i q). Istnieje l całkowite, takie że $a_1^l = b_1$, ponieważ grupa $\mathbb{Z}_p^*$ jest cykliczna. Znów rozwijamy l binarnie z niewiadomymi cyframi l_i:

$$a_1^{l_0+l_1\cdot 2+\ldots+l_{v-1}\cdot 2^{v-1}} = b_1,$$

co daje

$$(-1)^{l_0} = b_1^{2^{v-1}}.$$

Stąd

$$p \mid \mathrm{NWD}((-1)^{l_0} - b^{2^{v-1}}, n).$$

Próbując $l_0 = 0, 1$, albo otrzymamy nietrywialny dzielnik liczby n, albo okaże się, że $(-1)^{l_0} = b^{2^{v-1}}$. Jeśli nie otrzymaliśmy dzielnika, powtarzamy to postępowanie dla l_1 itd. Tym sposobem, albo otrzymamy nietrywialny dzielnik liczby n, albo $a^l = b$ i podgrupa $\langle a, b\rangle$ jest cykliczna (generowana przez a). Ogólnie zachodzi twierdzenie

Twierdzenie 1. (Rozszerzony algorytm Pohliga-Hellmana) *Niech $\mathcal{B} \subset \mathbb{Z}_n^*$ i niech będzie dana dla każdego $b \in \mathcal{B}$ pełna faktoryzacja liczby* $\mathrm{rz}(b)$. *Wówczas możemy znaleźć (deterministycznie) generator podgrupy $\langle\mathcal{B}\rangle$ lub (gdy nie jest cykliczna) nietrywialny dzielnik liczby n w czasie $O(\#\mathcal{B} \cdot q(\ln n)^C)$, gdzie $q = p_+(\prod_{b\in\mathcal{B}} \mathrm{rz}(b))$, a C jest stałą.*

W powyższym $p_+(\cdot)$ oznacza największy dzielnik pierwszy. Dla porównania, złożoność klasycznego algorytmu Pohliga-Hellmana w grupie cyklicznej G rzędu N to $O(p_+(N)(\ln N)^C)$, przy wielomianowym czasie operacji elementarnych w G.

4. Algorytm $p - 1$ Pollarda

Przechodzimy do opisu standardowego (probabilistycznego) algorytmu faktoryzacji $p - 1$ Pollarda. Niech n będzie nieparzystą liczbą naturalną, różną od 1 i nie będącą potęgą liczby pierwszej. Niech p będzie dzielnikiem pierwszym liczby n. Niech wreszcie b będzie (losową) liczbą całkowitą względnie pierwszą z n. Dla wielokrotności M liczby $p-1$ mamy $b^M \equiv 1\ (p)$ na mocy małego twierdzenia Fermata. Innymi słowy,

$$p \mid \mathrm{NWD}(b^M - 1, n),$$

przy czym prawą stronę obliczamy za pomocą algorytmu Euklidesa.

Może się zdarzyć, że $\mathrm{NWD}(b^M - 1, n) = n$. Jednak można pokazać, że dla co najmniej połowy elementów $b \in \mathbb{Z}_n^*$ ciąg $\mathrm{NWD}(b^{\frac{M}{2^i}} - 1, n)$ zawiera nietrywialny dzielnik liczby n.

Jeśli liczba $p - 1$ jest B-gładka (inaczej: wszystkie dzielniki pierwsze liczby $p - 1$ są $\leq B$), to można przyjąć

$$M = \prod q^{\left[\frac{\ln n}{\ln q}\right]},$$

gdzie q przebiega liczby pierwsze $\leq B$.

Te spostrzeżenia są podstawą następującego twierdzenia:

Twierdzenie 2. (Algorytm $p-1$ Pollarda) *Niech p będzie takim dzielnikiem pierwszym danej liczby naturalnej n, że liczba $p-1$ jest B-gładka. Wówczas liczbę p można znaleźć probabilistycznie w czasie $O(B(\ln n)^C)$, gdzie C jest stałą.*

Zachodzi również twierdzenie mocniejsze:

Twierdzenie 3. (Deterministyczny algorytm $p-1$ Pollarda) *W twierdzeniu 2 słowo „probabilistycznie" można zastąpić słowem „deterministycznie".*

5. Dowód poprawności derandomizacji algorytmu $p-1$ Pollarda

Naszkicujemy dowód twierdzenia 3. Niech n będzie złożone, bez dzielników pierwszych $\leq (\ln n)^2$, i niech

$$\mathcal{B} = \{2, 3, \ldots, [(\ln n)^2]\}.$$

Jednym z głównych składników rozumowania jest dolne oszacowanie funkcji ψ zliczającej liczby gładkie, $\psi(x,y) = \#\{n \leq x : \ p_+(n) \leq y\}$. Mamy bowiem

$$\#\langle\mathcal{B}\rangle \geq \psi(n, (\ln n)^2) > \sqrt{n},$$

a więc zbiór $\mathcal{B}$ generuje „dużą" podgrupę grupy $\mathbb{Z}_n^*$. Trzeba rozpatrzyć dwa przypadki:

1. Grupa $\langle\mathcal{B}\rangle$ jest cykliczna. Wtedy nietrywialny dzielnik liczby n znajdujemy za pomocą znanej (z prac [6, 5, 7] na temat testowania pierwszości) metody.
2. Grupa $\langle\mathcal{B}\rangle$ nie jest cykliczna. Wtedy nietrywialny dzielnik liczby n znajdujemy za pomocą rozszerzonego algorytmu Pohliga-Hellmana.

Załóżmy najpierw, że grupa $\langle\mathcal{B}\rangle$ jest cykliczna. Wówczas zbiór $\mathcal{B}$ zawiera taki element b, że

$$\mathrm{NWD}(b^{\frac{\mathrm{rz}(b)}{s}} - 1, n) > 1$$

dla pewnej liczby pierwszej $s \mid \mathrm{rz}(b)$. Takie b będziemy nazywać świadkiem Fermata-Euklidesa (złożoności liczby n). Istotnie, przypuśćmy, że w $\mathcal{B}$ nie

ma świadka Fermata-Euklidesa. Niech q będzie najmniejszym dzielnikiem pierwszym liczby n. Wtedy

$$\#\langle\mathcal{B}\rangle = \mathrm{NWW}_{b\in\mathcal{B}}\ \mathrm{rz}(b) = \mathrm{NWW}_{b\in\mathcal{B}}\ \mathrm{rz}_q(b) \mid q-1,$$

gdzie $\mathrm{rz}_q(\cdot)$ oznacza rząd w grupie $\mathbb{Z}_q^*$. W szczególności $\#\langle\mathcal{B}\rangle < q \leq \sqrt{n}$, co przeczy otrzymanej wcześniej nierówności przeciwnej.

W pozostałym przypadku niecyklicznej grupy $\langle\mathcal{B}\rangle$ wystarczy zastosować twierdzenie 1.

6. Uwagi końcowe

Jak pisaliśmy we wprowadzeniu, jednym z podstawowych zagadnień kryptografii asymetrycznej (jak również oczywiście samej obliczeniowej teorii liczb) jest badanie klasy liczb „efektywnie" rozkładalnych. Klasa ta zawiera liczby naturalne n mające dzielnik pierwszy p, taki że wartość ustalonego wielomianu cyklotomicznego Φ_k w p jest gładka. Algorytmy faktoryzacji wykorzystujące tę własność liczby rozkładanej n nazywane są cyklotomicznymi [1]. Do nich zaliczamy algorytm $p-1$ Pollarda - przy $k=1$. Również dla $k \geq 2$ algorytmy cyklotomiczne (oryginalnie probabilistyczne) można, przynajmniej częściowo, zderandomizować, co wykazano w [13] (warto podkreślić pewną uniwersalność stosowanej tu metodologii, która pozwala także faktoryzować wielomiany nad ciałem skończonym [14]). Sam zaś związek problemu faktoryzacji liczby n z problemem logarytmu dyskretnego w grupie $\mathbb{Z}_n^*$ jest jeszcze głębszy niż przedstawiono to tutaj na tle liczb n o szczególnej własności. W pracy [12] przedstawiono deterministyczną, podwykładniczą (szybszą niż sito ciała liczbowego) redukcję faktoryzacji (dowolnej) liczby n do obliczania logarytmu dyskretnego w grupie $\mathbb{Z}_n^*$.

Literatura

[1] E. Bach, J. O. Shallit, *Factoring with cyclotomic polynomials*, Mathematics of Computation, **52** (1989), 201-219.

[2] W. Diffie, M. Hellman, *New directions in cryptography*, IEEE Transactions on Information Theory, **22** (1976), 644-654.

[3] J. von zur Gathen, *Arithmetic circuits for discrete logarithms*, Lecture Notes in Computer Science, **2976** (2004), 557-566.

[4] D. Gordon, *Discrete logarithms in GF(p) using the number field sieve*, SIAM Journal on Discrete Mathematics, **6** (1993), 124-138.

[5] M. R. Fellows, N. Koblitz, *Self-witnessing polynomial-time complexity and prime factorization*, Designs, Codes and Cryptography, **2** (1992), 231-235.

[6] M. Fürer, *Deterministic and Las Vegas primality testing algorithms*, Lecture Notes in Computer Science, **194** (1985), 199-209.

[7] S. Konyagin, C. Pomerance, *On primes recognizable in deterministic polynomial time*, The Mathematics of Paul Erdös, R. L. Graham, J. Nesetril, eds., Springer-Verlag, 1997, 176-198.

[8] A. Lenstra, H. Lenstra Jr. (Eds.), *The development of the number field sieve*, Lecture Notes in Mathematics, **1554** (1993).

[9] H. Lenstra Jr., *Factoring integers with elliptic curves*, Annals of Mathematics, **126** (1987), 649-673.

[10] S. Pohlig, M. Hellman, *An improved algorithm for computing logarithms over $GF(p)$ and its cryptographic significance*, IEEE Transactions on Information Theory, **24** (1978), 106-110.

[11] J. M. Pollard, *Theorems on factorization and primality testing*, Proceedings of the Cambridge Philosophical Society, **76** (1974), 521-528.

[12] J. Pomykała, B. Źrałek, *On reducing factorization to the discrete logarithm problem modulo a composite*, computational complexity, **21** (2012), 421-429.

[13] B. Źrałek, *A deterministic version of Pollard's $p-1$ algorithm*, Mathematics of Computation, **79** (2010), 513-533.

[14] B. Źrałek, *Using partial smoothness of $p-1$ for factoring polynomials modulo p*, Mathematics of Computation, **79** (2010), 2353–2359.

A GENERALIZATION OF THE POHLIG-HELLMAN ALGORITHM AND ITS APPLICATION TO FACTORING

Abstract. We will show that the discrete logarithm problem and the problem of factoring are closely related. Namely, we will describe a generalization of the Pohlig-Hellman algorithm to noncyclic $\mathbb{Z}_n^*$ groups which can be used to derandomize Pollard's $p-1$ algorithm. The original version of this factoring algorithm needs indeed a source of randomness. It turns out however that the computations can be done deterministically with only slightly worse complexity.

Keywords: factoring, discrete logarithm, derandomization.