NIEKOWALNE EKSTRAKTORY LOSOWOŚCI

KONRAD DURNOGA

Wydział Matematyki, Informatyki i Mechaniki
Uniwersytetu Warszawskiego

4 czerwca 2013

Rozprawa poświęcona jest analizie ekstraktorów losowości. Jej główny rezultat stanowi bezwarunkowa i efektywna konstrukcja ekstraktora pewnego szczególnego typu, zwanego *ekstraktorem niekowalnym*. Jest to poprawienie wyniku z pracy Dodisa *i in.* [8] opublikowanej niedawno na prestiżowej konferencji FOCS.

EKSTRAKTORY LOSOWOŚCI

Ważne miejsce we współczesnej informatyce zajmują algorytmy randomizowane. W wymiarze praktycznym istotną kwestią staje się zapewnienie tym algorytmom źródła losowych bitów "wysokiej jakości". Jest to kluczowy problem zwłaszcza w przypadku zastosowań kryptograficznych, gdzie na założeniu jednostajnej losowości różnych elementów kryptosystemu może opierać się jego bezpieczeństwo. Fizyczne źródła losowości, powszechnie dostępne w komputerach osobistych, mogą nie zapewniać wystarczających statystycznych własności generowanego strumienia bitów. Tu pojawia się potrzeba konstrukcji *ekstraktorów losowości* – deterministycznych funkcji przekształcających niedoskonałe źródła losowości na takie, które są w statystycznym sensie bliskie rozkładom jednostajnym.

Rozwój dziedziny związanej z poprawianiem statystycznych własności rozkładów dyskretnych dokonał się w zasadzie w całości na przestrzeni ostatniego ćwierćwiecza, choć jej początki sięgają lat 50-tych ubiegłego stulecia i pionierskich prac von Neumanna o symulowaniu rzutów symetryczną monetą przy użyciu monety, na której orzeł wypada ze stałym prawdopodobieństwem $\neq 1/2$. Od tego czasu problematyka ta była przedmiotem badań czołowych naukowców, czego owocem są liczne prace na temat ekstraktorów losowości. Sam termin *ekstraktor* został zaproponowany przez Nisana i Zuckermana [16]. Za fundamentalne uznaje się prace Chora i Goldreicha [4], Cohena i Widgersona [5] oraz Zuckermana [22]. Jedne z najwcześniejszych konstrukcji ekstraktorów pochodzą od Chora i Goldreicha [4] oraz Impagliazzo *i in.* [13]. Przełomowym osiągnięciem ostatniej dekady był wynik Bourgaina [3], laureata medalu Fieldsa, wskazujący istnienie ekstraktorów nawet dla źródeł losowości o niskim współczynniku entropii. Rozprawa w znaczącej części zajmuje się analizą aspektów obliczeniowych klasycznego już dziś przykładu ekstraktora Chora-Goldreicha.

Ekstraktor losowości formalnie definiuje się w terminach entropii oraz statystycznej odległości rozkładów prawdopodobieństwa. Wielkością mierzącą stopień losowości dyskretnej zmiennej losowej X jest, znana z teorii informacji, tzw. *min-entropia* $\mathbf{H}_\infty(X)$, którą określamy jako:

$$\mathbf{H}_\infty(X) := \min_x \log_2 \frac{1}{\Pr(X = x)} ,$$

gdzie x przebiega przez wszystkie elementy zbioru nośnika zmiennej X. Spośród wszystkich rozkładów prawdopodobieństwa na n bitach min-entropia przyjmuje wartość maksymalną dla rozkładu jednostajnego – $\mathbf{H}_\infty(U_{\{0,1\}^n}) = n$. Do określenia odległości dwóch zmiennych losowych X i X' o rozkładzie na zbiorze $\mathcal{X}$ używamy standardowej definicji dystansu statystycznego $\Delta(X, X')$:

$$\Delta(X, X') := \max_{\mathcal{S} \subseteq \mathcal{X}} |\Pr(X \in \mathcal{S}) - \Pr(X' \in \mathcal{S})| = \tfrac{1}{2} \sum_{x \in \mathcal{X}} |\Pr(X = x) - \Pr(X' = x)| .$$

Zapis $X \approx_\epsilon X'$ oznacza $\Delta(X, X') \leq \epsilon$ dla pewnego $\epsilon \geq 0$. Przy tym typowo wymaga się, by ϵ był *zaniedbywalny* jako funkcja pewnego parametru n, tzn. $\epsilon = \epsilon(n)$ powinien dążyć do 0 szybciej niż odwrotność dowolnego wielomianu w punkcie $n \to \infty$.

Funkcję $\mathrm{Ext}\colon \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ nazywamy (k, ϵ)*–niekowalnym ekstraktorem*, jeśli dla każdej pary niezależnych zmiennych losowych X i Y nad zbiorami, odpowiednio, $\mathcal{X}$ oraz $\mathcal{Y}$, i dowolnej funkcji $A\colon \mathcal{Y} \to \mathcal{Y}$ spełniającej $A(y) \neq y$ dla wszystkich $y \in \mathcal{Y}$, zachodzi:

$$\big(\mathrm{Ext}(X, Y), \mathrm{Ext}(X, A(Y)), Y\big) \approx_\epsilon \big(U_{\mathcal{Z}}, \mathrm{Ext}(X, A(Y)), Y\big)\ , \tag{1}$$

o ile $\mathbf{H}_\infty(X) \geq k$ oraz Y jest rozkładem jednostajnym nad $\mathcal{Y}$.

Intuicyjnie, powyższa definicja oznacza, że jeśli argument x wybrany został z dostatecznie losowego rozkładu to żaden, w tym nawet nieograniczony obliczeniowo, adwersarz nie potrafi odróżnić $\mathrm{Ext}(x, y)$ od wartości losowej przy znanym losowym ziarnie y oraz $\mathrm{Ext}\big(x, A(y)\big)$, gdzie A jest dowolnie ustaloną (a priori) przez adwersarza funkcją nie posiadającą punktów stałych. Jest to również znaczące wzmocnienie klasycznej definicji *ekstraktora*, w której warunek (1) można zastąpić przez $\mathrm{Ext}(X, Y) \approx_\epsilon U_{\mathcal{Z}}$, oraz *silnego ekstraktora*, gdzie przyjmuje on postać: $\big(\mathrm{Ext}(X, Y), Y\big) \approx_\epsilon \big(U_{\mathcal{Z}}, Y\big)$.

Pojęcie *ekstraktora niekowalnego* wprowadzone zostało przez Dodisa i Wichsa [10], którzy wskazali, że hipotetyczna funkcja o takich własnościach może posłużyć do budowy protokołu tzw. *amplifikacji prywatności*. Jednocześnie przedstawili oni argument probabilistyczny dowodzący istnienia ekstraktorów niekowalnych dla szerokiego spektrum parametrów k oraz ϵ. Jednak problemem otwartym pozostawała kwestia podania jawnego przykładu konstrukcji tego typu.

EKSTRAKTOR CHORA-GOLDREICHA

W literaturze poświęconej teorii ekstraktorów odnaleźć można bogactwo rozmaitych metod – technik analizy fourierowskiej, teorii kodów, kombinatoryki czy teorii liczb. Głębokie rezultaty z tej ostatniej pozwoliły m.in. na uzyskanie przez Bourgaina [3] wspomnianego wyżej ekstraktora dla źródeł o niskiej min-entropii. Teorioliczbowy charakter ma również konstrukcja Chora i Goldreicha [4], oparta na logarytmie dyskretnym, którą krótko przytoczymy poniżej.

Niech p będzie nieparzystą liczbą pierwszą, a g generatorem cyklicznej grupy multiplikatywnej $(\mathbb{Z}/p\mathbb{Z})^*$ ciała $\mathbb{Z}/p\mathbb{Z}$. Ponadto, niech $M > 1$ oznacza dowolnie ustalony dzielnik rzędu tej grupy, czyli $p - 1$. Podstawą ekstraktora Chora-Goldreicha jest następująca funkcja

$$f_g(a) \coloneqq \log_g a \mod M\ , \tag{2}$$

wyznaczona przez logarytm dyskretny przy podstawie g w $(\mathbb{Z}/p\mathbb{Z})^*$, dodatkowo zredukowany modulo M.

Dodis *i in.* [8] wykazali, używając oszacowań Weila dla sum charakterów multiplikatywnych nad ciałem skończonym (zob. np. monografię Schmidta [19]), że f_g zadaje ekstraktor niekowalny.

Twierdzenie 1 (Dodis *i in.* [8], Twierdzenie 4.1)**.** *Niech* $\mathrm{Ext}\colon \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} \to \mathbb{Z}/M\mathbb{Z}$ *będzie określona wzorem*

$$\mathrm{Ext}(x, y) \coloneqq f_g(x + y) = \log_g (x + y) \mod M\ . \tag{3}$$

Wtedy, dla dowolnego k, funkcja Ext *jest (k, ϵ)–niekowalnym ekstraktorem przy $\epsilon = 2Mp^{1/4}2^{-k/2}$.*

Twierdzenie to stanowiło rozszerzenie rezultatu z oryginalnej pracy Chora i Goldreicha [4]. Zawarta tam analiza w istocie implikowała, że (3) jest ekstraktorem losowości, choć w momencie ukazania się tego artykułu pojęcie ekstraktora nie było jeszcze znane. Dodis i Oliveira [9] wskazali, że ta funkcja spełnia warunki definicji silnego ekstraktora.

Rozprawa skoncentrowana jest wokół zagadnienia efektywnego obliczania wartości ekstraktora (3) lub, równoważnie, funkcji f_g danej wzorem (2). Należy tu zwrócić uwagę, że, inaczej niż ma to miejsce w typowych zastosowaniach kryptograficznych opierających swoje bezpieczeństwo na założeniu wysokiej złożoności obliczeniowej problemu logarytmu dyskretnego, w tym przypadku wymagamy, by problem ten był "łatwo" rozwiązywalny. Użycie standardowej metody do wyznaczania logarytmu dyskretnego, czyli algorytmu Pohliga-Hellmana [18], z niewielką modyfikacją polegającą na obliczaniu $\log_g z$ modulo każdy dzielnik pierwszy $q \mid M$, pozwala na znalezienie wartości funkcji $f_g(z)$ w czasie proporcjonalnym do $\mathrm{P}^+(M)$ – największego dzielnika pierwszego M. Jest to czas wielomianowy wyłącznie pod warunkiem, że M jest liczbą *gładką*, tzn. ma jedynie małe dzielniki pierwsze. W tym kontekście kluczowego znaczenia nabiera kwestia wydajnego generowania liczb p oraz $M \mid p - 1$ z dodatkowym wymaganiem gładkości M.

Dodis *i in.* [8] sugerują następującą procedurę znajdowania p oraz M. Najpierw ustalmy liczbę gładką M, np. wybierając odpowiednio dużą potęgę 2, dla której $\log_2 M$ odpowiada w przybliżeniu oczekiwanej liczbie bitów wyjścia ekstraktora. Następnie przeglądamy kolejne wyrazy postępu arytmetycznego $\equiv 1 \pmod{M}$, tzn. $M+1$, $2M+1$, $3M+1$, etc., w poszukiwaniu liczby pierwszej p. Sprawdzenie czy dana liczba jest pierwsza może być zrealizowane za pomocą wielomianowego deterministycznego testu pierwszości [14]. Cała procedura jest również efektywna o ile najmniejsze p w tym postępie nie jest zbyt duże. W przypadku ekstraktorów o tzw. krótkich wyjściach, czyli dla niewielkich, na przykład stałych, wartości M istnienie takiego p wynika z oszacowań dla stałej Linnika. Najlepsze znane obecnie ograniczenie bezwarunkowe otrzymane przez Xylourisa [21] zapewnia znalezienie $p = O(M^5)$. Natomiast w przypadku ogólnym, dla dużych M, konieczne jest odwołanie się do następującej, powszechnie uważanej za prawdziwą (zob. artykuł Granville'a i Pomerance'a [12]), hipotezy.

Hipoteza 2. *Dla dowolnych a oraz M takich, że* $\mathrm{NWD}(a, M) = 1$*, najmniejsza liczba pierwsza $p \equiv a \pmod{M}$ spełnia $p = O(\varphi(M) \ln^2 M)$, gdzie φ oznacza funkcję Eulera.*

Konstrukcja Dodisa *i in.* [8] jest tym samym wynikiem warunkowym. W rozprawie wskazujemy też, że przytoczona wyżej procedura zawiera pewną usterkę. Mianowicie nie gwarantuje ona zachowania odpowiednich proporcji między liczbami M i p. W szczególności dla wartości M bliskich p z Twierdzenia 1 otrzymujemy błąd ϵ, który nie jest zaniedbywalny, przez co samo twierdzenie traci zupełnie siłę wyrazu. Prosta modyfikacja metody generowania M i p pozwala na uniknięcie tego problemu, jednak za cenę wprowadzenia zależności od Hipotezy 2 również w przypadku ekstraktorów o krótkich wyjściach.

Ostatnim parametrem ekstraktora (3), którego wybór wymaga osobnego komentarza jest podstawa logarytmu dyskretnego g. Autorzy oryginalnej pracy [8] nie specyfikują jednak żadnego sposobu konstrukcji takiego g. Istnienie tej luki, zidentyfikowanej przez Durnogę i Źrałka [11], zostało

potwierdzone przez autorów (komunikacja prywatna). W istocie kwestia efektywnego znajdowania generatora grupy multiplikatywnej ciała $\mathbb{Z}/p\mathbb{Z}$ w ogólnym przypadku pozostaje ważnym problemem otwartym. Naturalnie istnieje szybki, niedeterministyczny algorytm wyznaczający pewien generator, ale jedynie przy znanej pełnej faktoryzacji liczby $p-1$ niezbędnej do weryfikacji czy dany losowy element jest pełnego rzędu. W świetle wyniku Ankeny'ego [2] ten proces może być zderandomizowany przy założeniu uogólnionej Hipotezy Riemanna (ERH). Użycie Hipotezy 2 implikuje, że dla znalezionej liczby p rząd grupy $(\mathbb{Z}/p\mathbb{Z})^*$, czyli $p-1$, jest gładki. Deterministyczny algorytm wyznaczania g wymaga jednak dodatkowo zastosowania Hipotezy Riemanna. W rozprawie uzasadniamy, że zakładając jedynie prawdziwość ERH, ale bez odwoływania się do Hipotezy 2, możliwe jest wydajne generowanie odpowiednich wartości p, M oraz g w sposób randomizowany.

W chwili obecnej znanych jest kilka innych przykładów ekstraktora niekowalnego [6, 15]. Są to konstrukcje bezwarunkowe pozwalające na osiągnięcie lepszych parametrów (np. oszacowania na wyraz błędu ϵ czy słabszego warunku na poziom min-entropii źródła k) niż ekstraktor Chora-Goldreicha i używające metod spoza teorii liczb.

WYNIKI

Głównym celem rozprawy jest analiza problemu efektywnego wyboru parametrów p, M i g w ekstraktorze Chora-Goldreicha w formie przedstawionej przez Dodisa *i in.* [8] oraz zaprojektowanie alternatywnej wersji ekstraktora niekowalnego, która nie zakłada prawdziwości ERH ani Hipotezy 2.

Generator Pseudolosowy Online

Pierwszym krokiem do stworzenia bezwarunkowej konstrukcji ekstraktora jest próba budowy algorytmu wyznaczającego wartości funkcji f_g bez danego generatora g. Z pozoru zadanie to wydaje się nie być poprawnie sformułowane ze względu na fakt, że nie jest możliwe obliczanie funkcji, która sama w sobie pozostaje nieznana. Proponowana w rozprawie metoda pozwala na obliczenie wartości pewnej funkcji częściowej $f(a_1), \ldots, f(a_\ell)$ dla wielu argumentów $a_1, \ldots, a_\ell$ gwarantując istnienie takiego, nie danego jawnie, generatora g, że $f(a_i) = f_g(a_i)$ dla wszystkich $i = 1, \ldots, \ell$. Metoda ta jest algorytmem typu *online*, tzn. takim, w którym argumenty przetwarzane są sekwencyjnie i całość wejścia może nie być dostępna na początku działania algorytmu. Algorytmy online mają fundamentalne znaczenie w obliczeniach interaktywnych.

W rozprawie wykazane jest następujące twierdzenie:

Twierdzenie 3. *Istnieje deterministyczny algorytm online, który dla danego ciągu argumentów $a_1, \ldots, a_\ell \in (\mathbb{Z}/p\mathbb{Z})^*$ oblicza $f_g(a_1), \ldots, f_g(a_\ell)$, gdzie f_g jest epimorfizmem danym przez (2) dla pewnego, a priori nieznanego, generatora g grupy $(\mathbb{Z}/p\mathbb{Z})^*$ zależnego od $a_1, \ldots, a_\ell$. Algorytm znajduje $f_g(a_i)$ dla $i = 1, \ldots, \ell$ przed pobraniem kolejnego argumentu a_{i+1} z wejścia. Pojedyncza wartość $f_g(a_i)$ obliczana jest w czasie $O\big(\mathrm{P}^+(M) \cdot s^{-1} \operatorname{poly}(\log p)\big)$ przy użyciu $O\big(s \operatorname{poly}(\log p)\big)$ bitów pamięci, gdzie parametr $0 < s \leq \sqrt{\mathrm{P}^+(M)}$ może być wybrany dowolnie.*

Należy przy tym zaznaczyć, że algorytm z powyższego twierdzenia nie może być bezpośrednio użyty do obliczania wartości ekstraktora niekowalnego (3), który zdefiniowany jest w terminach statystycznej odległości

rozkładów. Dla ciągów wartości pojawiających się na wyjściu ekstraktora możliwe jest jednak, analogicznie jak w przypadku ekstraktorów, udowodnienie pewnej własności nieodróżnialności od ciągu wartości losowych.

Niekowalny Ekstraktor – Konstrukcja Bezwarunkowa

Osiągnięcie celu bezwarunkowej konstrukcji ekstraktora niekowalnego możliwe jest po zaskakująco prostej modyfikacji oryginalnego rozwiązania Chora-Goldreicha. Pomysł polega na zastąpieniu logarytmu dyskretnego występującego w definicji (3), który stanowił źródło opisanych wyżej problemów, przez potęgowanie. Operacja ta przekształca otrzymane argumenty na elementy pewnej grupy G zdefiniowanej jako:

$$G := \{a^{(p-1)/M} \mid a \in (\mathbb{Z}/p\mathbb{Z})^*\} \tag{4}$$

Podstawowy wynik rozprawy sformułowany jest w poniższym twierdzeniu:

Twierdzenie 4. *Niech p, M, k oraz ϵ będą takie jak w Twierdzeniu 1. Wtedy funkcja* $\mathrm{Ext}_G : \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} \to G$ *dana wzorem*

$$\mathrm{Ext}_G(x, y) := (x + y)^{(p-1)/M}$$

jest efektywnie obliczalnym (k, ϵ)–niekowalnym ekstraktorem.

W oczywisty sposób definicja Ext_G nie wymaga znajomości generatora grupy $(\mathbb{Z}/p\mathbb{Z})^*$. Ponadto wartości tej funkcji, w odróżnieniu od ekstraktora (3), można szybko obliczać również w przypadku, gdy M ma duże dzielniki pierwsze. Ta obserwacja pozwala na efektywne wygenerowanie parametrów p oraz M bez dodatkowego warunku na gładkość liczby M. Opierając się na głębokim wyniku Alforda *i in.* [1] można dowieść następującego wniosku:

Twierdzenie 5. *Dla wszystkich dostatecznie dużych z, z' oraz dowolnego $l > 0$ spełniających $(z + l)^4 < \frac{1}{2}z'$, istnieje co najmniej $\frac{1}{3}z'/(\varphi(M) \log z')$ liczb pierwszych w przedziale $(\frac{1}{2}z', z']$, które należą do postępu arytmetycznego $\equiv 1 \pmod{M}$ dla każdego modułu M z przedziału $(z, z + l)$, poza co najwyżej $O\big(l/\log(z + l)\big)$ wyjątkowymi wartościami M.*

To twierdzenie natychmiast prowadzi do randomizowanego algorytmu, o oczekiwanej wielomianowej złożoności, generacji parametrów ekstraktora Ext_G.

Efektywna Bijekcja $G \to \mathbb{Z}/M\mathbb{Z}$

Z praktycznego punktu widzenia pewną wadą ekstraktora Ext_G z Twierdzenia 4 jest fakt, że wartości pojawiające się na jego wyjściu są elementami podgrupy G rzędu M określonej przez (4). Grupa ta, interpretowana jako zbiór ciągów bitów, ma nieregularną strukturę, przez co może nie być odpowiednia w zastosowaniach oczekujących właśnie losowych ciągów bitów. Rozprawa zajmuje się również kwestią przekształcenia elementów G na takie ciągi bitów. To zagadnienie jest ściśle związane ze znanym problemem tzw. ekstrakcji klucza.

W rozprawie przedstawiony jest algorytm do wyznaczania wartości pewnej bijekcji $G \to \mathbb{Z}/M\mathbb{Z}$. Szczególne znaczenie miałoby skonstruowanie podobnej bijekcji dla grup stosowanych w protokole Diffie-Hellmana, w których problem obliczania logarytmu dyskretnego jest uznawany za trudny.

Proponowany w rozprawie algorytm działa w czasie proporcjonalnym do największego dzielnika pierwszego $\mathrm{P}^+(M)$ liczby M, a więc jest efektywny jedynie dla tych M, które są gładkie. Tym samym zakładamy, że problem znajdowania logarytmów dyskretnych w G jest "łatwy". Mimo, że w takim przypadku istnieje naturalny i prosty do obliczenia izomorfizm między G i $\mathbb{Z}/M\mathbb{Z}$, zadany właśnie przez logarytm dyskretny, to jednak jego konstrukcja wymaga znajomości generatora grupy G. Rozważana metoda obliczania bijekcji nie zakłada, że taki generator jest dany. Ten przypadek nie był, wg wiedzy autora, wcześniej badany w literaturze i może być interesujący również poza kontekstem ekstraktorów.

W rozprawie wykazane jest następujące twierdzenie:

Twierdzenie 6. *Istnieje bijekcja $\sigma\colon G \to \mathbb{Z}/M\mathbb{Z}$ taka, że dla danego $a \in G$ odpowiadającą wartość $\sigma(a)$ można obliczyć w sposób deterministyczny w czasie $O\big(\mathrm{P}^+(M)\,\mathrm{poly}(\log p)\big)$.*

Jednocześnie konstrukcja może być zmodyfikowana tak, by uzyskać algorym obliczający bijekcję dla wielu argumentów:

Twierdzenie 7. *Istnieje deterministyczny algorytm online obliczający wartości pewnej bijekcji $G \to \mathbb{Z}/M\mathbb{Z}$. Algorytm wykorzystuje $O(s\,\mathrm{poly}(\log p))$ bitów pamięci, a jego łączny czas działania dla ℓ argumentów to*

$$O\big(\mathrm{P}^+(M)(\ell s^{-1}+1)\,\mathrm{poly}(\log p)\big)\ ,$$

gdzie $0 < s \leq \sqrt{\mathrm{P}^+(M)}$ może być ustalone dowolnie.

Obliczanie Logarytmów Dyskretnych na Klasycznej Maszynie Turinga

Algorytm do rozwiązywania wielu instancji problemu logarytmu dyskretnego w pewnym ograniczonym modelu obliczeniowym stanowi ostatni wynik przedstawiony w rozprawie. Jest to w pewnym sensie produkt uboczny rozważań na temat konstrukcji efektywnej bijekcji $G \to \mathbb{Z}/M\mathbb{Z}$ opisanej powyżej. Może być jednak uważany za interesujący rezultat ze względu na fakt, że dotyczy on fundamentalnego w teorii liczb problemu i pytania o jego złożoność w najbardziej klasycznym modelu, tzn. na jednotaśmowej maszynie Turinga.

Punktem wyjścia dla dalszej analizy jest generyczny, tzn. działający dla dowolnej grupy cyklicznej rzędu pierwszego q, algorytm znajdowania logarytmów dyskretnych w takiej grupie, czyli metoda małych-wielkich kroków Shanksa [20]. Jego charakterystyczną cechą jest odowiedniość: "czas $\frac{q}{s}$ przy pamięci s", która oznacza, że za cenę poświęcenia ok. s bitów pamięci czas działania algorytmu może zostać ograniczony do ok. $\frac{q}{s}$ kroków, przy dowolnym $s \leq \sqrt{q}$. Dzięki temu dla $s \approx \sqrt{q}$ otrzymujemy istotne przyspieszenie w porównaniu z naiwną metodą znajdowania logarytmu dyskretnego. W typowych zastosowaniach q oraz s są jednak dużymi liczbami, co wymaga przechowywania znacznych ilości danych, które w praktyce nie będą mogły zmieścić się jednocześnie w pamięci operacyjnej. To uzasadnia analizę algorytmów z tzw. pamięcią zewnętrzną, której nośniki są wolniejsze niż te używane do budowy pamięci operacyjnych oraz nie zapewniają dostępu do dowolnej komórki pamięci w stałym czasie. Zbliżona argumentacja była motywacją prowadzącą do zaprojektowania algorytmów sortowania zewnętrznego takich, jak algorytm sortowania przez scalanie. Od strony teoretycznej abstrakcją urządzeń z pamięcią tego typu są klasyczne maszyny Turinga.

Metoda małych-wielkich kroków sprowadza się do znalezienia kolizji pewnych elementów rozważanej grupy. W modelu RAM, ze swobodnym

dostępem do pamięci, wyszukiwanie tej kolizji zamplementować można przy użyciu tablicy haszującej. Diem [7] przenosi ten algorytm na *wielotaśmową* maszynę Turinga, w której haszowanie zastąpione jest sortowaniem dużych tablic. Pomysł proponowany w rozprawie nie wymaga sortowania, dzięki czemu ta nowa metoda ma przewagę nad algorytmem generycznym w przypadku, gdy taka operacja nie jest wydajna. Na mocy rezultatu Petersena [17] nie istnieje żaden podkwadratowy algorytm sortowania w modelu z *jednotaśmową* maszyną Turinga.

Problem obliczania logarytmu dyskretnego w $(\mathbb{Z}/p\mathbb{Z})^*$ ograniczyć można do przypadku podgrup rzędu będącego liczbą pierwszą, powiedzmy q. Dla $q \mid p-1$ niech $q-1 = s \cdot t$ będzie znanym rozkładem na czynniki liczby $q-1$. Ustalmy zespolony pierwiastek stopnia q z jedynki $\zeta_q := \mathrm{e}^{2\pi i/q}$. Rozważmy grupę $(\mathbb{Z}/q\mathbb{Z})^* = \langle h \rangle$ z pewnym generatorem h oraz jej podgrupę $H := \langle h^s \rangle$. Proponowana w rozprawie metoda wykorzystuje własności niezmiennicze *okresów gaussowskich* – obiektów szeroko znanych w teorii liczb, w tym w jej obliczeniowej części. Definiujemy okres gaussowski η_j dla $j = 0, \ldots, s-1$ jako

$$\eta_j := \sum_{\alpha \in h^j H} \zeta_q^{\alpha} ,$$

gdzie suma przebiega po wszystkich elementach warstwy $h^j H$ in $(\mathbb{Z}/q\mathbb{Z})^*$. Niech $F \in \mathbb{Q}[X]$ będzie wielomianem minimalnym dla η_0, tzn.

$$F(X) := \prod_{j=0}^{s-1} (X - \eta_j) .$$

W rozprawie wykazane jest następujące twierdzenie:

Twierdzenie 8. *Załóżmy, że p nie dzieli wyróżnika wielomianu F. Wtedy istnieje deterministyczna maszyna Turinga (z jedną taśmą roboczą oraz jednokierunkowymi taśmami: wejściową i wyjściową), która przy danych ℓ parach (a_i, b_i) nad $(\mathbb{Z}/p\mathbb{Z})^*$ dla $i = 1, \ldots, \ell$ spełniających $\operatorname{ord}(a_i) = q$ i $b_i \in \langle a_i \rangle$, znajduje kolejno ℓ logarytmów dyskretnych $\log_{a_i} b_i$. Łączny czas działania algorytmu to $O\big((q + \ell \max(s,t)) \operatorname{poly}(\log p)\big)$, a maszyna używa dodatkowo $O(s \log p)$ bitów pamięci (tzn. pamięci na taśmie roboczej bez uwzględnienia danych zapisanych na taśmach wejściowej i wyjściowej).*

W przypadku zbalansowanych podziałów $s \approx t \approx \sqrt{q}$ i wielu instancji $\ell = \Omega(q^{1/2})$ zamortyzowany koszt czasowy obliczania pojedynczego logarytmu dyskretnego na maszynie jednotaśmowej jest tym samym porównywalny z kosztem działania metody Shanksa dla maszyny wielotaśmowej, który ograniczyć można przez $O\big(q^{1/2} (\log q) \operatorname{poly}(\log p)\big)$.

[1] William R. Alford, Andrew Granville, and Carl Pomerance. There are infinitely many Carmichael numbers. *Annals of Mathematics*, 139:703–722, 1994.

[2] Nesmith C. Ankeny. The least quadratic non residue. *Annals of Mathematics*, 55:65–72, 1952.

[3] Jean Bourgain. More on the sum-product phenomenon in prime fields and its applications. *International Journal of Number Theory*, 1(1):1–32, 2005.

[4] Benny Chor and Oded Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity. *SIAM J. Comput.*, 17(2):230–261, April 1988.

[5] Aviad Cohen and Avi Wigderson. Dispersers, deterministic amplification, and weak random sources (extended abstract). In *FOCS*, pages 14–19. IEEE Computer Society, 1989.

[6] Gil Cohen, Ran Raz, and Gil Segev. Non-malleable extractors with short seeds and applications to privacy amplification. In *IEEE Conference on Computational Complexity*, pages 298–308. IEEE, 2012.

[7] Claus Diem. On the complexity of some computational problems in the turing model, 2012. preprint.

[8] Yevgeniy Dodis, Xin Li, Trevor D. Wooley, and David Zuckerman. Privacy amplification and non-malleable extractors via character sums. In *Proceedings of the 2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, FOCS '11, pages 668–677, Washington, DC, USA, 2011. IEEE Computer Society.

[9] Yevgeniy Dodis and Roberto Oliveira. On extracting private randomness over a public channel. In Sanjeev Arora, Klaus Jansen, José D. P. Rolim, and Amit Sahai, editors, *RANDOM-APPROX*, volume 2764 of *Lecture Notes in Computer Science*, pages 252–263. Springer, 2003.

[10] Yevgeniy Dodis and Daniel Wichs. Non-malleable extractors and symmetric key cryptography from weak secrets. In *Proceedings of the 41st annual ACM symposium on Theory of computing*, STOC '09, pages 601–610, New York, NY, USA, 2009. ACM.

[11] Konrad Durnoga and Źrałek Bartosz. On randomness extractors and computing discrete logarithms in bulk, 2013. preprint.

[12] Andrew Granville and Carl Pomerance. On the least prime in certain arithmetic progressions. *Journal of the London Mathematical Society*, 2(2):193–200, 1990.

[13] Russell Impagliazzo, Leonid A. Levin, and Michael Luby. Pseudo-random generation from one-way functions. In *Proceedings of the twenty-first annual ACM symposium on Theory of computing*, STOC '89, pages 12–24, New York, NY, USA, 1989. ACM.

[14] Hendrik W. Lenstra. Primality testing with Gaussian periods. In Manindra Agrawal and Anil Seth, editors, *FST TCS 2002: Foundations of Software Technology and Theoretical Computer Science, 22nd Conference Kanpur, India, December 12-14, 2002, Proceedings*, volume 2556 of *Lecture Notes in Computer Science*, page 1. Springer, 2002.

[15] Xin Li. Non-malleable extractors, two-source extractors and privacy amplification. In *FOCS*, volume 0, pages 688–697, Los Alamitos, CA, USA, 2012. IEEE Computer Society.

[16] Noam Nisan and David Zuckerman. More deterministic simulation in logspace. In *Proceedings of the twenty-fifth annual ACM symposium on Theory of computing*, STOC '93, pages 235–244, New York, NY, USA, 1993. ACM.

[17] Holger Petersen. Sorting and element distinctness on one-way turing machines. In Carlos Martín-Vide, Friedrich Otto, and Henning Fernau, editors, *LATA*, volume 5196 of *Lecture Notes in Computer Science*, pages 433–439. Springer, 2008.

[18] Stephen Pohlig and Martin Hellman. An improved algorithm for computing logarithms over $GF(p)$ and its cryptographic significance. *IEEE Transactions on Information Theory*, 24(1):106–110, 1978.

[19] Wolfgang M. Schmidt. *Equations over finite fields: an elementary approach*. Lecture Notes in Mathematics. Springer-Verlag, 1976.

[20] Daniel Shanks. Class number, a theory of factorization, and genera. In *1969 Number Theory Institute (Proc. Sympos. Pure Math., Vol. XX, State Univ. New York, Stony Brook, N.Y., 1969)*, pages 415–440. Amer. Math. Soc., Providence, R.I., 1971.

[21] Triantafyllos Xylouris. *Über die Nullstellen der Dirichletschen L-Funktionen und die kleinste Primzahl in einer arithmetischen Progression*. PhD thesis, Mathematisch-Naturwissenschaftliche Fakultät der Universität Bonn, 2011.

[22] David Zuckerman. General weak random sources. In *FOCS '90*, pages 534–543, 1990.