# Оцінка Ваги Хеммінга Обернених Чисел Відносно Операції Додавання за Модулем Узагальнених Чисел Мерсенна

Дарія Ядуха
кафедра математичних методів захисту інформації
Фізико-технічний інститут
Національний технічний університет України «Київський політехнічний інститут ім. Ігоря Сікорського»
Київ, Україна
dariya.yadukha@gmail.com

# Hamming Weight Bound for Additive Inverse Modulo Generalized Mersenne Number

Dariya Yadukha
Department of Mathematical Methods of Information Security
Institute of Physics and Technology
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"
Kyiv, Ukraine
dariya.yadukha@gmail.com

***Анотація*—У даній роботі наведено співвідношення для швидкого знаходження ваги Хеммінга обернених відносно операції додавання чисел за модулем особливого виду, а саме числа Мерсенна, числа Кренделла та узагальненого числа Мерсенна.**

***Abstract*—This paper shows relations for the fast finding of Hamming weight of additive-inverse numbers modulo a special form, namely a Mersenne number, a Crandall number and generalized Mersenne number.**

***Ключові слова—число Мерсенна; число Кренделла; узагальнене число Мерсенна; вага Хеммінга.***

***Keywords—Mersenne number; Crandall number; generalized Mersenne number; Hamming weight.***

## I. Вступ

Числа спеціальної форми часто дозволяють швидше виконувати арифметичні операції за модулем. Це широко використовується для практичних цілей і пояснює велику кількість досліджень швидкої спеціальної арифметики. У даній роботі розглянуто способи швидкого знаходження ваги Хеммінга обернених чисел за модулем числа різних спеціальних форм.

## II. Співвідношення для ваги Хеммінга обернених чисел за спеціальними видами модулів

### *A. Числа Мерсенна*

Числа Мерсенна є одним з найбільш відомих типів простих чисел, який досліджувався математиками ще з XVII століття, адже дані числа володіють багатьма властивостями. Відомо, що числа Мерсенна $M_n=2^n-1$ дозволяють швидко реалізовувати обчислення редукції за даним модулем, оскільки трудомісткі операції (множення, ділення) за модулем числа Мерсенна можливо замінити на звичайне додавання [2]. Наприклад, нехай $x$ - деяке число, в такому випадку для знаходження $x \bmod M_n$ достатньо представити число $x$ у виді $x=T\cdot 2^n+U$, тоді $x \equiv T+U \pmod{M_n}$. Це випливає з властивостей чисел Мерсенна, а саме з факту, що множення на степінь двійки за модулем числа Мерсенна є циклічним зсувом [1].

Вагою Хеммінга $n$-бітового числа $s$ називається кількість одиниць у двійковому записі числа і позначається $Ham(s)$. Відомо наступне співвідношення для ваги Хеммінга оберненого за модулем числа Мерсенна [1]:

*Лема 1.* Для довільного числа $A \in \{0,1\}^n$, $A \neq 0^n$ та модуля $M_n=2^n-1$ виконується:

$$Ham(-A \pmod{M_n}) = n - Ham(A),$$

де $-A$ - обернене за операцією додавання число до $A$ за модулем числа Мерсенна.

Співвідношення для ваги Хеммінга для обернених чисел за деяким модулем спрощують обчислення, уникаючи необхідність знаходження самого оберненого, отже, пришвидшуючи виконання операцій. Представлена у лемі 1 рівність використовується при розшифруванні повідомлення у криптосистемі AJPS [1], яка є кандидатом у постквантові криптографічні алгоритми з відкритим ключем у рамках конкурсу NIST [6].

Числа Мерсенна мають багато переваг для застосування у криптографії та інших прикладних областях з використанням операцій у кільці лишків. Часто у практичних цілях використовуються прості числа, а серед чисел Мерсенна на даний момент відомо лише *50* простих чисел, з них лише для перших *46*-ти встановлено точні порядкові номера. Саме тому є мотивація досліджувати різні види узагальнень даних чисел та шукати схожі властивості для них.

### *B. Узагальнені числа Мерсенна*

Узагальненим числом Мерсенна називається число виду $p=f(2^n),\ n\in N$, де $f$ - нескладний поліном. Наприклад, нехай $f(t)=t^3-t+1$, тоді узагальнене число Мерсенна буде мати вигляд $p=2^{3n}-2^n+1,\ n\in N$, а результатом редукції цілого числа за модулем $p$ будуть $3n$-бітові цілі числа [3]. Оскільки даний тип чисел є узагальненням над числами Мерсенна, то, очевидно, що він є значно більшим та містить більше простих чисел.

Для узагальнених чисел Мерсенна також існує спосіб швидкої редукції за модулем [3]. Даний спосіб має лінійну складність. Крім того, при спеціально підібраних параметрах редукцію за модулем узагальненого числа Мерсенна можна зробити ще більш ефективною, але у такому випадку кількість простих чисел такого виду менша. Також існує метод швидкого множення за модулем даного класу чисел, що працює вдвічі швидше, ніж множення за модулем звичайного числа [5]. Ще однією перевагою узагальнення є те, що операції редукції та множення можна легко розпаралелювати, що робить арифметику за даним модулем простішою для апаратної реалізації.

Співвідношення для ваги Хеммінга оберненого за модулем узагальненого числа Мерсенна наводиться у наступній теоремі.

<u>*Теорема 1.*</u> Нехай $GM_{n,m}$ - узагальнене число Мерсенна виду $2^n-2^m-1$ та $A$- $n$-бітове число, тобто $A=a_{n-1}\,a_{n-2}\ldots a_1\,a_0$, $a_i\in\{0,1\}, i=\overline{0,n-1}$. Тоді виконується:

1) Якщо біт $a_m=0$, то виконується співвідношення:
$$Ham(-A \bmod GM_{n,m}) = n-1-Ham(A).$$
2) У випадку $a_m=1$ потрібно представити число $A$ у вигляді $A=h_1\,||\,h_2\,||\,h_3$, де:
   - $|h_3|=m$, тобто $h_3=a_{m-1}\,a_{m-2}\ldots a_1\,a_0$ - $h_3$ включає молодші $m$ бітів числа $A$;
   - $h_2=a_k a_{k-1}\ldots a_{m-1}a_m$, де $k=\min_i\{a_i=0\mid a_j=1, m\le j<i\}$, тобто $h_2$ включає у себе біти починаючи з $a_m$ та до першого нуля, який зустрінеться після $a_m$, включно;
   - $h_1=a_{n-1}\,a_{n-2}\ldots a_k$, де $k$ - індекс з минулого пункту. Позначимо $|h_1|=l$, тобто $h_1$ містить $l$ бітів.

Тоді співвідношення для оберненого буде наступним:
$$Ham(-A \bmod GM_{n,m}) =$$
$$= l - Ham(h_1) + Ham(h_2) + m - Ham(h_3)$$

*Доведення:* Відзначимо, що модуль $2^n-2^m-1$ має наступний вигляд у двійковому записі: *11…1011…1*, причому саме $m$-й біт дорівнює *0*. Потрібно, маючи число $A$, знайти таке число $B\in\{0,1\}^n$, що $A+B=0 \bmod GM_{n,m}$, тоді $B$ якраз і буде оберненим до $A$, тобто $B=-A \bmod GM_{n,m}$. Слід зауважити, що потрібно шукати таке $B$, що $A+B=GM_{n,m}$, оскільки за умовою $A\in\{0,1\}^n$, тобто випадок $A+B=c\cdot GM_{n,m}$, де $c>1$ - деяка константа, отримати неможливо. Введемо наступні позначення:
$$A = a_{n-1}a_{n-2}\ldots a_m\ldots a_1a_0$$
$$B = b_{n-1}b_{n-2}\ldots b_m\ldots b_1b_0$$
$$GM_{n,m} = g_{n-1}g_{n-2}\cdots g_m\cdots g_1g_0,$$
де $a_i, b_i, g_i\in\{0,1\}, i=\overline{0,n-1}$. Відомо, що $g_m=0$, а інші біти - *1*, тобто $GM_{n,m}=11\ldots101\ldots11$. Потрібно знайти $b_i$ по відомим значенням $a_i, i=\overline{0,n-1}$, тоді можна буде побачити залежність ваги Хеммінга оберненого числа від ваги самого $A$.

1) Якщо $a_m=0$, то значення $b_i$ будуть наступні:
$$\begin{cases} b_m = 0 \\ b_i = 1-a_i, \quad \forall i\neq m \end{cases}$$

Щоб перевірити правильність даного твердження, помітимо, що для того, щоб $B=-A \bmod GM_{n,m}$ необхідно виконання рівності:

| | $a_{n-1}$ | $a_{n-2}$ | … | $a_{m+1}$ | $a_m$ | $a_{m-1}$ | … | $a_1$ | $a_0$ |
|---|---|---|---|---|---|---|---|---|---|
| + | $b_{n-1}$ | $b_{n-2}$ | … | $b_{m+1}$ | $b_m$ | $b_{m-1}$ | … | $b_1$ | $b_0$ |
| | 1 | 1 | … | 1 | 0 | 1 | … | 1 | 1 |

Підставляючи значення $b_i$ бачимо, що це дійсно виконується:

| | $a_{n-1}$ | $a_{n-2}$ | … | $a_{m+1}$ | 0 | $a_{m-1}$ | … | $a_1$ | $a_0$ |
|---|---|---|---|---|---|---|---|---|---|
| + | $1-a_{n-1}$ | $1-a_{n-2}$ | … | $1-a_{m+1}$ | 0 | $1-a_{m-1}$ | … | $1-a_1$ | $1-a_0$ |
| | 1 | 1 | … | 1 | 0 | 1 | … | 1 | 1 |

Таким чином, у даному випадку для знаходження оберненого потрібно усі біти, окрім $m$-го, замінити на протилежні, а $m$-й біт залишити без змін. Тоді вага Хеммінга оберненого, тобто $Ham(-A \bmod GM_{n,m})$ буде рівна $n-Ham(A)-1$, де $n$ відповідає максимально можливій вазі $n$-бітового вектору, а віднімаючи *1* враховується незмінний $m$-й біт.

2) У випадку $a_m = 1$ різниця з попереднім випадком у тому, що для отримання $g_m = 0$ потрібно щоб $b_m = 1$, а це утворює біт переносу на старші біти. Це і обумовлює розділення двійкового запису числа на три частини, тобто $A = h_1 \,||\, h_2 \,||\, h_3$. Тоді вагу Хеммінга $A$ можна представити як суму $Ham(h_1) + Ham(h_2) + Ham(h_3)$. Позначимо $B = -A \bmod GM_{n,m} = h_1^* \,||\, h_2^* \,||\, h_3^*$. Тоді $h_3^*$ буде знаходитись аналогічно пункту 1), замінюючи всі біти $h_3$ на протилежні. Очевидно, що $Ham(h_3^*) = m - Ham(h_3)$.

Наймолодший біт числа $h_2$, тобто $a_m = 1$, і, як було сказано раніше, $b_m = 1$, а отже біт переносу переходить на старший біт $a_{m+1}$. Можливі два випадки:

- $a_{m+1} = 0$, тоді, оскільки $g_{m+1} = 1$, $b_{m+1} = 0$;
- $a_{m+1} = 1$ - у такому випадку, враховуючи, що $g_{m+1} = 1$, маємо $b_{m+1} = 1$ і знову отримуємо біт переносу, тобто виконуємо аналогічні дії для $a_{m+2}$.

Бачимо, що біт переносу буде з'являтись на кожному кроці, аж поки у послідовності біт не зустрінеться $0$. Саме тому частина $h_2$ включає в себе біти від $a_m$ до першого нуля, який зустрінеться у послідовності. Слід зазначити, що $0$ точно буде, оскільки розглядаються числа менші за модуль $GM_{n,m}$. Бачимо, що при знаходженні $h_2^*$ біти $h_2$ не змінювались (якщо $a_{m+c} = 0$, то $b_{m+c} = 0$, і якщо $a_{m+c} = 1$, то і $b_{m+c} = 1$, де $c \in \mathrm{N}$ - деяка константа), а отже $Ham(h_2^*) = Ham(h_2)$.

Частина $h_1^*$ повинна бути такою, щоб результат суми за модулем $2$ з $h_1$ був $11\ldots1$. Аналогічно до попередніх результатів, $h_1^*$ отримується шляхом заміни бітів $h_1$ на протилежні. У такому випадку, якщо $|h_1| = l$, то $Ham(h_1^*) = l - Ham(h_1)$.

Узагальнюючи отримані результати, маємо:

$$Ham(-A \bmod GM_{n,m}) =$$
$$= Ham(B) = Ham(h_1^*) + Ham(h_2^*) + Ham(h_3^*) =$$
$$= l - Ham(h_1) + Ham(h_2) + m - Ham(h_3),$$

що і треба було довести.

*C. Числа Кренделла*

Ще одним узагальненням чисел Мерсенна є числа Кренделла виду $CR_n = 2^n - c$, де $c \in \mathrm{N}$. Очевидно, що при $c = 1$ отримуємо класичні числа Мерсенна [4]. Редукція за модулем цього класу чисел є також набагато швидшою, і обчислюється наступним чином [5]:

$$x \equiv (x \bmod 2^n) - c\left\lfloor \frac{x}{2^n} \right\rfloor (\bmod CR_n)$$

Наступна теорема демонструє спосіб знаходження ваги Хеммінга оберненого за модулем числа Кренделла.

<u>*Теорема 2.*</u> Нехай $CR_n$ - число Кренделла, тобто число виду $2^n - c$, де $c \in \mathrm{N}$ - деяка константа, та $A$ - $n$-бітове число, тобто $A = a_{n-1}a_{n-2}\ldots a_1a_0$, $a_i \in \{0,1\}, i = \overline{0,n-1}$. Позначимо модуль як $CR_n = r_{n-1}r_{n-2}\ldots r_1r_0$ та $B = -A \bmod CR_n = b_{n-1}b_{n-2}\ldots b_1b_0$ - обернене до $A$ число, де $r_i, b_i \in \{0,1\}, i = \overline{0,n-1}$. Тоді вага Хеммінга оберненого до $A$ за модулем $CR_n$ (тобто вага $B$) буде обчислюватись наступним чином:

1) Якщо число $a_{\lceil \log_2 c \rceil - 1}\ldots a_1a_0$ менше за число $r_{\lceil \log_2 c \rceil - 1}\ldots r_1r_0$, то:

$$Ham(-A \bmod CR_n) = n - \lceil \log_2 c \rceil - Ham(h_1) + Ham(h_2^*),$$

де $A = h_1 \,||\, h_2$, причому $h_2$ - молодші $\lceil \log_2 c \rceil$ бітів числа $A$; $B = -A \bmod CR_n = h_1^* \,||\, h_2^*$, аналогічно $h_2^*$ - це молодші $\lceil \log_2 c \rceil$ бітів числа $B$.

2) Якщо ж число $a_{\lceil \log_2 c \rceil - 1}\ldots a_1a_0$ є більшим за число $r_{\lceil \log_2 c \rceil - 1}\ldots r_1r_0$, то:

$$Ham(-A \bmod CR_n) =$$
$$= n - \lceil \log_2 c \rceil - |h_2| - Ham(h_1) + Ham(h_2) + Ham(h_3^*),$$

де $A = h_1 \,||\, h_2 \,||\, h_3$, причому $h_3$ - молодші $\lceil \log_2 c \rceil$ бітів числа $A$; $h_2$ включає у себе біти числа $A$ починаючи з $\lceil \log_2 c \rceil$ біту та до першого нуля, який зустрінеться після $a_{\lceil \log_2 n \rceil - 1}$; $h_1$ - старші біти, що залишились, тобто $h_1 = a_{n-1}a_{n-2}\ldots a_w$, де $a_{w-1}$ - старший біт $h_2$; $|h_2|$ - кількість бітів у числі $h_2$.

*Доведення:* Слід зауважити, що модуль $CR_n$ має вигляд $111\ldots11{*}\ldots{*}$, де під символом * розуміється або $0$, або $1$ в залежності від значення константи $c$. Кількість таких неоднозначно визначених бітів можна обмежити значенням $\lceil \log_2 c \rceil$. Для того, щоб $B$ був оберненим до $A$ за модулем $CR_n$ необхідно щоб виконувалось наступне. Оскільки значення $r_{\lceil \log_2 c \rceil - 1}\ldots r_1r_0$ не фіксовані, причому від них прямо залежать значення $b_{\lceil \log_2 c \rceil - 1}, b_{\lceil \log_2 c \rceil - 2}, \ldots, b_1, b_0$, то вага Хеммінга числа $B$ буде залежати від $Ham(h_2^*)$, для обчислення якого потрібно вирахувати значення $b_{\lceil \log_2 c \rceil - 1}, b_{\lceil \log_2 c \rceil - 2}, \ldots, b_1, b_0$ та знайти кількість одиниць серед них. Значення $b_{\lceil \log_2 c \rceil - 1}, b_{\lceil \log_2 c \rceil - 2}, \ldots, b_1, b_0$ обчислюються наступним чином:

$$\begin{cases} b_0 = (r_0 + a_0) \bmod 2 \\ b_i = (r_i + a_i + x_i) \bmod 2, \quad x_i = \begin{cases} 1, & a_{i-1} + b_{i-1} + x_{i-1} > 1 \\ 0, & \text{інакше.} \end{cases} \end{cases},$$
$$i = \overline{1, \lceil \log_2 c \rceil - 1}, \quad x_0 = 0.$$

Тоді можна знайти кількість одиниць серед $\lceil \log_2 c \rceil$ молодших бітів числа *B*, позначимо це значення як *z*.

*1)* Очевидно, що якщо $B = -A \bmod CR_n = h_1^* \| h_2^*$, *то* $Ham(B)=Ham(h_1^*)+Ham(h_2^*)$. У даному випадку $Ham(h_2^*)=z$. Важливою умовою є те, що число $a_{\lceil log_2 c \rceil-1} \dots a_1 a_0$ менше за число $r_{\lceil log_2 c \rceil-1} \dots r_1 r_0$ - у такому випадку при додаванні старших бітів $h_2^*$ та $h_2$ не виникає біту переносу, який би мав враховуватись при додаванні молодших бітів $h_1^*$ та $h_1$, а отже значення $h_1^*$ мало би змінитись для виконання умови $A+B=0 \bmod CR_n$. Оскільки біту переносу немає, то маємо $h_1+h_1^*=11\dots1$, тоді $Ham(h_1^*)=|h_1|-Ham(h_1)$. Оскільки $A$ - $n$-бітове число, а $|h_2| = \lceil log_2 c \rceil$, то $|h_1| = n - \lceil log_2 c \rceil$. Тоді маємо:

$$Ham(-A \bmod CR_n) = Ham(B) = Ham(h_1^*) + Ham(h_2^*) =$$
$$= n - \lceil log_2 c \rceil - Ham(h_1) + Ham(h_2^*)$$

2) Аналогічно, якщо $B = -A \bmod CR_n = h_1^* \| h_2^* \| h_3^*$, то

$$Ham(B) = Ham(h_1^*) + Ham(h_2^*) + Ham(h_3^*).$$

У даному випадку $Ham(h_3^*) = z$. Оскільки $a_{\lceil log_2 c \rceil-1} \dots a_1 a_0$ є більшим за $r_{\lceil log_2 c \rceil-1} \dots r_1 r_0$, то при додаванні старших бітів $h_3$ та $h_3^*$ з'являється біт переносу. Враховуючи те, що $r_{n-1} r_n \dots r_{\lceil log_2 c \rceil} = 11\dots1$, біт переносу буде з'являтись на кожному кроці до того моменту, як у $A$ зустрінеться $0$ - таким чином утворюється частина $h_2$. Зрозуміло, що на позиціях $i$, де $a_i=1$, буде $b_i=1$, а при першій зустрічі нуля, позначимо його $a_t=0$, буде і $b_t=0$. Маємо рівність $h_2=h_2^*$. Отже, вага Хеммінга не змінюється: $Ham(h_2^*)=Ham(h_2)$. $Ham(h_1^*)$ отримується аналогічно як у доведенні пункту 1), тобто: $Ham(h_1) = |h_1| - Ham(h_1) = n - t - 1 - Ham(h_1)$.

Підсумовуючи отримане, маємо:

$$Ham(-A \bmod CR_n) = Ham(B) =$$
$$= Ham(h_1^*) + Ham(h_2^*) + Ham(h_3^*) =$$
$$= |h_1| - Ham(h_1) + Ham(h_2) + Ham(h_3^*) =$$
$$= n - \lceil log_2 c \rceil - |h_2| - Ham(h_1) + Ham(h_2) + Ham(h_3^*),$$

що і потрібно було довести.

Бажано, щоб для знаходження ваги Хеммінга оберненого не потрібно було знаходити саме його значення, це значно пришвидшить обчислення на практиці, що можна помітити на прикладі реалізації криптосистеми AJPS, де використовується рівність для ваги Хеммінга оберненого числа за модулем числа Мерсенна. Оскільки в теоремі 2 для обчислення ваги Хеммінга $Ham(-A \bmod CR_n)$ потрібно знайти значення молодших $\lceil log_2 n \rceil$ бітів числа $-A \bmod CR_n$, то спробуємо оцінити вагу Хеммінга так, щоб уникнути обчислення оберненого.

*Теорема 3.* Нехай $CR_n$ - число Кренделла, $A$ - $n$-бітове число, $B=-A \bmod CR_n$ - обернене до $A$ число. Тоді для ваги Хеммінга оберненого числа до $A$ за модулем $CR_n$ виконується наступне:

1) Якщо число $a_{\lceil log_2 c \rceil-1} \dots a_1 a_0$ менше за $r_{\lceil log_2 c \rceil-1} \dots r_1 r_0$, то: $n - \lceil log_2 c \rceil - Ham(h_1) \le Ham(-A \bmod CR_n) \le n - Ham(h_1)$, де $A=h_1 \| h_2$, причому $h_2$ - молодші $\lceil log_2 c \rceil$ бітів числа $A$.

2) Якщо ж $a_{\lceil log_2 c \rceil-1} \dots a_1 a_0$ більше числа $r_{\lceil log_2 c \rceil-1} \dots r_1 r_0$, то:

$$|h_1| - Ham(h_1) + Ham(h_2) \le Ham(-A \bmod CR_n) \le$$
$$\le n - |h_2| - Ham(h_1) + Ham(h_2),$$

де $A=h_1 \| h_2 \| h_3$, причому $h_3$ - молодші $\lceil log_2 c \rceil$ бітів числа $A$; $h_2$ включає у себе біти числа $A$ починаючи з $\lceil log_2 c \rceil$ біту та до першого нуля, який зустрінеться після $a_{\lceil log_2 c \rceil-1}$; $h_1$ - старші біти, що залишились, тобто $h_1=a_{n-1}a_{n-2}\dots a_w$, де $a_{w-1}$ - старший біт $h_2$.

Нерівності отримуються очевидно, враховуючи, що $h_2^*$ (аналогічно $h_3^*$ у 2) ) може мати мінімально можливу вагу Хеммінга рівну $0$, коли $h_2^*=00\dots0$, та максимальну можливу у випадку $h_2^*=11\dots1$, а саме $\lceil log_2 c \rceil$.

## Висновки

У даній роботі було описано співвідношення для знаходження ваги Хеммінга обернених чисел відносно операції додавання за модулем чисел, які є узагальненнями чисел Мерсенна, а саме модулем числа Кренделла та узагальненого числа Мерсенна. Отримані результати можна використовувати для спрощення обчислень за модулями цих класів чисел у криптографії або інших практичних реалізаціях, зокрема для узагальнення побудови криптосистеми AJPS з метою розширення класу модулів, що використовуються.

## Література References


[1] Divesh Aggarwal, Antoine Joux, Anupam Prakash, Miklos Santha, "A New Public-Key Cryptosystem via Mersenne Numbers" [Online]. Available: https://eprint.iacr.org/2017/481.

[2] Joppe W. Bos, Thorsten Kleinjung, Arjen K. Lenstra, "Efficient SIMD arithmetic modulo a Mersenne number" [Online]. Available: https://eprint.iacr.org/2010/338.

[3] Jerome A. Solinas, "Generalized Mersenne Numbers", Technical Report CORR Centre for Applied Cryptographic Research, University of Waterloo, 1999.

[4] Richard E. Crandall, "Method and apparatus for public key exchange in a cryptographic system" (oct. 27, 1992). U.S. Paent 5,159,632.

[5] Greg Zaverucha, "Generalized Mersenne Numbers in Pairing-Based Cryptography" , 2006, unpublished.

[6] Post-Quantum cryptography standardization NIST [Online]. Available: https://csrc.nist.gov/projects/post-quantum-cryptography/post-quantum-cryptography-standardization.