# В**ступ**

В теперішній час одною з найважливіших проблем, пов’язаних із захистом конфіденційності інформації, що передається в інформаційно-телекомунікаційних системах, є потреба в генераторах псевдовипадкових послідовностей, що відповідають високим вимогам до рівномірності розподілу ймовірностей чисел, що формуються ними. Генератор ,слабкий з криптографічної точки зору, може значно ослабити захищеність інформаційної системи. Сьогодні відомо декілька сотень шифрів, які можуть ефективно використовуватися для захисту інформації. Незважаючи на це багато спеціалістів і організацій неперервно працюють над вдосконаленням відомих шифрів і пошуком нових. Президент України недавно заявив : Кіберпростір перетворився на ще одне поле протистояння і боротьби за незалежність держави. І це застереження без перебільшень дуже важливе. Слід нагадати, що група CERT-UA при Держспецзв'язку 2014 року зафіксувала 216 кібернападів ззовні (більше половини з них - на державні установи). У 2015 році число атак збільшилося у півтора разу. Найнебезпечніша атака відбулась 23 грудня, коли було завдано масованого вдару по диспетчеру компанії «Укренерго» і 6 енергокомпаніям, що призвело до відключення енергії у 103 населених пунктах на Заході країни. Американські експерти встановили, що ця масована кібердиверсія була здійснена з боку РФ . При проектуванні сучасної мікроелектронної техніки неможливо обходитись без криптографічних засобів захисту інформації. Це стосується мобільних телефонів, мобільних комп’ютерів, смартфонів та багато іншої мікроелектронної техніки, яка може використовуватися для передачі конфіденційної інформації. Так в 2016 році американська фірма WhatsApp вдосконалила і впровадила нову систему шифрування інформації(дзвінки,повідомлення,медіафайли,що передаються), яка за переконаннями авторів має високу стійкість і надійність. Тому очевидно, що дослідження і вдосконалення криптографічних інструментів захисту інформації є актуальне. Найбільш складна проблема для звичайного користувача – зрозуміти комплексну ефективність шифру, яка формується з його стійкості і простоти використання, продуктивності роботи при захисті інформації. Класична структура роботи зломщика шифрів – дослідження повторень в шифрованому тексті і дослідження статистичних характеристик частоти використання символів тексту.

В основу досліджень взяті блокові шифри, які базуються на матричних перетвореннях – множення квадратної матриці ключа на матрицю стовпчик відкритого тексту [1, 2]. Такі шифри довгий час вважались неперспективними, ручними і низькопродуктивними. З широким поширенням комп’ютерної техніки шифри на основі матричних перетворень отримали свою нішу і перспективу використання. Блокові шифри характеризуються (наприклад система шифрування мережа Фейстеля) такими параметрами: наявність властивостей перестановок, властивостей підстановок і статистичними характеристиками, які в деякій мірі наближаються до рівноймовірних.

# Основна частина

#### Хеш-функції

Хешування - перетворення вхідного масиву даних довільної довжини в вихідний бітову рядок фіксованої довжини, за певним алгоритмом. Таке перетворення також називається хеш-функція ( функція згортки), а його результат - хеш, хеш-код, хеш-сума ( зведення повідомлення).

Застосування хешування:

- Побудова асоціативних масивів;

- Пошук дублікатів в серіях наборів даних;

- Побудова досить унікальних ідентифікаторів для наборів даних;

- Контрольне підсумовування даних - для подальшого виявлення в них помилок (випадкових або навмисних), що виникають при зберіганні і (або) передачі сигналу;

- Зберігання паролів в системах захисту (в хешованому вигляді, - для відновлення з якого потрібна обробка функцією, яка є «зворотною» по відношенню до використаної хеш-функції);

- При виробленні електронного підпису (на практиці, - часто, - підписується не саме повідомлення, а його хеш-образ).

У загальному випадку (принцип Діріхле) - немає однозначної відповідності між вихідними даними і хеш-кодом, тому що значення хеш-функцій менш різноманітні, ніж значення вхідного масиву: різний зміст може дати однаковий хеш-код - що може привести до так званої колізії. Імовірність виникнення колізій грає важливу роль в оцінці якості хеш-функцій.

Як відомо, криптографічні хеш-суми відрізняються від звичайних хеш-сум тим, що крім основних властивостей, що вимагаються від будь-якої хеш-функції:

- здатності перетворення вхідного значення (зазвичай тексту) довільної довжини в вихідне значення фіксованої довжини,

- статистичної рівномірності випадаючих вихідних значень,

- хорошим «розкидом» (розбіжністю приблизно в половині біт) вихідних значень навіть для текстів, різниця між якими дуже незначна (можливо тільки в 1 біті);

до криптографічних хеш-алгоритмів пред'являються додаткові вимоги:

- розрядність вихідних значень повинна знаходитися далеко за межами можливостей повного перебору на сучасній техніці як за швидкістю обробки, так і за обсягом зберігання (на практиці це - розрядність в 128, 160, 256 і більше біт);

- не повинно існувати способу (істотно більш ефективного, ніж повний перебір вхідних значень) обчислити будь-яку пару вхідних текстів, що дають на виході однакове хеш-значення (неважливо яке) - успішна атака на цю вимогу називається «колізією» хеш-функції;

- не повинно існувати способу (істотно більш ефективного, ніж повний перебір вхідних значень) за значенням хеш-функції підібрати будь-який вхідний текст, що дає на виході алгоритму це хеш-значення - успішна атака на цю вимогу називається «оборотністю» хеш-функції.

Три перерахованих вище вимоги до крипостійкості хеш-функцій перетворюють їх за фактом в ідентифікатори фіксованої довжини для текстів, файлів і взагалі довільних блоків інформації. При цьому ці ідентифікатори:

- унікальні на всьому просторі сучасного інформаційного суспільства;

- незворотні, тобто не розкривають абсолютно нічого про зміст вихідного документа.

Це дозволяє використовувати хеш-функції в задачах:

- Перевірки цілісності файлів, архівів, збірок тощо (Досить довіреністю передати одержувачу хеш-суму файлу і тоді будь-які несанкціоновані зміни в файлі негайно змінять його контрольну хеш-суму);

- Зберігання паролів у незворотному вигляді (в сховище поміщаються не самі паролі, а хеш-суми від рядка, наприклад, виду «константа (соль)» + «пароль», що дозволяє при розкритті хешу зберегти власне словникову парольну фразу в секреті);

- Електронно-цифрового підпису документів (власне ЕЦП встановлюється завжди не на сам файл, що по-перше, було б дуже повільно, і по-друге, несло б деякі проблеми з безпекою самого ключа підпису, а на його хеш-суму, забезпечуючи такий ж рівень захисту від модифікацій).

На сьогоднішній день переважну частку застосувань хеш-функцій «беруть на себе» алгоритми MD5, SHA-1, SHA-256. Звичайно, існує і безліч інших менш відомих, або поширених тільки у вузьких співтовариствах алгоритмів (наприклад, RIPEMD, TIGER, PANAMA і ін.)

У криптографії сіль – це випадкова послідовність, що використовується в якості додаткових вхідних даних для хешування. Основна функція солі –захист від атак проти списку вихідних паролів за допомогою райдужних таблиць чи одного проходу повного перебору.

Нова сіль генерується випадковим чином для кожного пароля. У загальному випадку, сіль і пароль з’єднуються та обробляються за допомогою криптографічної хеш-функції ,і результуюча вихідна послідовність(але не оригінальний пароль) зберігається разом із сіллю в базі даних. Криптографічні солі широко використовуються в багатьох сучасних комп’ютерних системах від системи облікових даних Unix до інтернет-безпеки.

Райдужні таблиці представляють собою застосування більш раннього і простішого алгоритму Мартіна Хеллмана. Райдужна таблиця – це попередньо обчислена таблиця для реверсу криптографічної хеш-функції ,як правило ,для злому хеш-паролів. Таблиці здебільшого використовують для відновлення оригінального паролю певної довжини, що складається з обмеженого набору символів. Це практичний приклад time-memory tradeoff - розумного компромісу між часом пошуку по таблиці і займаною пам'яттю (використовується менше часу для комп’ютерної обробки інформації і більше пам’яті, ніж при brute-force атаці, але більше часу обробки і менше пам’яті, ніж при використанні таблиць пошуку).Використання ключа ,виведеного функцією з використанням солі робить такий вид атаки практично нездійсненним.

У криптографії, лавинний ефект відноситься до бажаної властивості криптографічних алгоритмів, зазвичай блокових шифрів та криптографічних хеш-функцій. Лавинний ефект очевидний у випадку, коли трохи змінена вхідна послідовність (наприклад, перевернутий один біт) спричиняє суттєві зміни вихідного сигналу (наприклад, половина вихідних біт обернулися). У випадку високоякісних блокових шифрів, така невелика зміна в будь-якому ключі або відкритому тексту повинна викликати різку зміну шифротексту. Сам термін був вперше використаний Хорстом Фейстелем , хоча концепція бере початок принаймні у дифузії Шеннона.

Якщо блокові шифри або криптографічна хеш-функція не проявляє лавинний ефект в значній мірі, то він має погану рандомізацію, і, таким чином, криптоаналітик може робити припущення про вхідний текст, якщо буде мати тільки вихідний. Це може бути достатньо, щоб частково або повністю зламати алгоритм. Таким чином, лавинний ефект є бажаною умовою з точки зору конструктора криптографічного алгоритму або пристрою.

При побудові шифру або хеш-функції значний прояв лавинного ефекту є однією з цілей первинного проектування. Це дозволяє невеликим змінам швидко поширюватися через ітерації алгоритму, таким чином, що кожен біт на виході повинен залежати від кожного біта входу до зупинки алгоритму

Криптографічна хеш-функція - будь-яка хеш-функція, яка є крипостійкою, тобто задовольняє ряд вимог, специфічних для криптографічних додатків.

Більшість криптографічних хеш-функцій приймають на виході рядок будь-якої довжини ,а на виході дають хеш-значення фіксованої довжини. Криптографічна хеш-функція повинна бути здатною витримувати всі відомі типи криптоаналітичних атак. Як мінімум вона мусить мати наступні властивості:

* Стійкість до відновлення прообразу(Для даного хеш-значення h має бути складно знайти повідомлення m,таке ,що h=hash(m).Це поняття пов’язане з поняттям односторонньої функції. Функції, які не володіють цією властивістю можуть вразливі до атак відновлення прообразу. Слід зауважити, що на даний момент немає використовуваних хеш-функцій з доведеною односпрямованістю. )
* Стійкість до відновлення другого прообразу(Для даного вхідного значення m1 має бути неможливо обчислювально знайти відмінне значення m2 таке, щоб hash(m1)==hash(m2).Функції, що не володіють цією рисою є вразливими до атаки відновлення другого прообразу )
* Стійкість до колізій(Повинно бути неможливо обчислювально знайти два відмінних повідомлення m1 і m2 таких, щоб hash(m1)=hash(m2).Така пара називається криптографічною хеш-колізією. Ця властивість іноді називають сильна колізійна стійкість. Вона вимагає, що хеш-значення має бути хоча б у 2 рази довше, ніж вимагає стійкість до повного перебору ; в іншому випадку колізії можуть бути знайдені за допомогою атаки «днів народження». Атака «днів народження» - вид криптографічної атаки, що використовує парадокс днів народження для злому шифрів чи пошуку колізій хеш-функцій. Зазвичай саме знаходження способу побудови колізій криптоаналітиками служить першим сигналом старіння алгоритму і необхідності його швидкої заміни.)

Дані вимоги не є незалежними:

- Оборотна функція нестійка до колізій першого і другого роду.

- Функція, нестійка до колізій першого роду, нестійка до колізій другого роду; зворотні твердження невірні.

Слід зазначити, що досі ще не доведене існування незворотних хеш-функцій, для яких обчислення будь-якого прообразу заданого значення хеш-функції теоретично неможливо. Зазвичай знаходження зворотного значення є лише обчислювально складним завданням.

Наявність цих властивостей означає, що зловмисний противник не зможе замінити чи модифікувати вхідні дані без зміни хеш-значення. Тому, якщо два повідомлення мають однаковий хеш-код , то з високим ступенем вірогідності можна стверджувати, що вони ідентичні.

Функції, що відповідають цим критеріям все ж можуть мати інші небажані властивості. На теперішній час популярні хеш-функції можуть бути вразливими до атак довжини-розширення(коли дано значення hash(m) та довжина len(m),вибираючи підходящі m` зловмисник може обчислити hash(m || m`),де || означає конкатенацію ).Ця слабкість може бути використана для злому наївних схем автентифікації на основі хеш-функцій.

В ідеалі можна побажати ще сильніших умов. Для противника має бути неможливо знайти два повідомлення з істотно схожими значеннями хеш-коду; або вивести яку-небудь корисну інформацію про дані, з огляду лише на їх хеш-значення. Тому криптографічна хеш-функція повинна поводитися ,наскільки це можливо, як рандомна функція, зберігаючи при цьому детермінованість та ефективність обчислення.

Алгоритми контрольних сум, такі як CRC32 та інші циклічні перевірки надмірності призначенні для задоволення набагато слабших вимог і ,як правило, непридатні в якості криптографічних хеш-функцій. Наприклад CRC був використаний для перевірки цілісності повідомлень в стандарті шифрування WEP,але легко був відкритий тип атак, що використовував лінійність контрольних сум проти цього стандарту.

Перевірка стійкості розробленої криптографічної хеш-функції

1.Стійкість для відновлення образів.

Довжина одного блоку хеш-коду – 32 , символів, доступних для формування послідовності – 42.Тобто,можливих значень – 4232 ≈ 8.8 \*1051.Припустимо,що можна перебирати значення зі швидкістю 1012 штук за секунду. Тоді для повного перебору даної кількості значень буде необхідно 8.8\*1039секунд = 244\*1034 годин = 2.78\*1032 років. Оскільки алгоритм новий, досі ще не існує ефективніших способів, ніж повний перебір для знаходження потрібного значення. Тобто хеш-функція досить добре протистоїть атакам відновлення образів.

2.Стійкість до колізій.

Найпоширеніша атака проти цієї властивості – атака днів народження. Одним із наслідків парадоксу днів народження є те, що для n-бітового блокового шифру повторювані появи блоку шифротексту можуть очікуватися з вірогідністю близько 0,63 при наявності лише 2n/2 випадкових відкритих текстів, зашифрованих на одному ключі (незалежно від розміру ключа).Тобто, щоб з високою ймовірністю можна було знайти колізії, необхідно перебрати 2n/2 варіантів. У даному випадку n =192,тому потрібно перебрати 286 можливих комбінацій ≈7.7 \*1025. Припустимо, що можна перебирати значення зі швидкістю 1012 штук за секунду. Тоді для повного перебору даної кількості значень буде необхідно 7.7\*1013секунд = 213\*108 годин = 24\*105 років. Тобто хеш-функція досить добре протистоїть колізійним атакам.

3.Лавинний ефект.

Результати виконання хешування даних, що мінімально відрізняються одне від одного.

|  |  |  |
| --- | --- | --- |
| Вхідне значення | Хеш-значення | К-ть змінених бітів |
| MAHEROVSKYA | W3.L1HZMST1FX-EE2W;29.,U8VHM,POU | 0 (точка відліку) |
| MAHEROVSKYB | Z8H3DW511N147GJ5Q3;GU-0F;LK,DILU | 87 із 192 |
| MAHEROVSKYC | ,D;RP5HMGH1ZNYO,8G;U9140SBNCUBIU | 68 із 192 |
| MAHEROVSKYD | 'IJ9.KT1VB1O;ATTWT;2U48RH1QS54FU | 93 із 192 |
| MAHEROVSKYE | 2N'X7Z'M451DDSYKE0;G97CC6XT2M;CU | 74 із 192 |
| MAHEROVSKYF | 5SLFJ8B1J'12T4;B,D;UUAG;.NWI;W9U | 87 із 192 |
| MAHEROVSKYG | 8X1;VNNMYZ1X3M22KQ;29DKOQDZYEP6U | 67 із 192 |
| MAHEROVSKYH | B,NL1,Z17T1MJ-7Z2;;GUGO9F3,8VI3U | 82 із 192 |
| MAHEROVSKYI | E133DB5MMN1BZGCQQA;U9JS 4Z'O6B0U | 70 із 192 |
| MAHEROVSKYJ | H6PRPQH1.H109YHH8N;2UMWLZP2-N4;U | 85 із 192 |
| MAHEROVSKYK | KB59.'TMAB1VPAM8W ;G9P 6OF5E-; U | 61 із 192 |
| MAHEROVSKYL | NGRX7E'1P51K'SR'E7;UUS-XD58UFWXU | 86 із 192 |
| MAHEROVSKYM | QL7FJTBM-'19F4WW,K;29V2I2.B4WPUU | 67 із 192 |
| MAHEROVSKYN | TQT;V2N1DZ1-VM.NKX;GUY63XREK7IRU | 85 із 192 |
| MAHEROVSKYO | WV9L1HZMST1T5-0E24;U9.AUMHH OBOU | 37 із 192 |
| MAHEROVSKYP | Z V3DW511N1ILG55QH;2U-EFB7KA'4LU | 80 із 192 |
| MAHEROVSKYQ | ,'BRP5HMGH17.YA,8U;G91I00;NQG;IU | 63 із 192 |
| MAHEROVSKYR | '4X9.KT1VB1,BAFTW1;UU4MRVTQ0XWFU | 91 із 192 |
| MAHEROVSKYS | 29DX7Z'M451RRSKKEE;297QCKJTG8PCU | 66 із 192 |
| MAHEROVSKYT | 5EZFJ8B1J'1G14PB,R;GUAU;99WWPI9U | 88 із 192 |
| MAHEROVSKYU | 8JF;VNNMYZ15HMU2K-;U9DYO-'Z60B6U | 69 із 192 |
| MAHEROVSKYV | BO.L1,Z17T1 X-ZZ2B;2UG,9TV,MH43U | 72 із 192 |
| MAHEROVSKYW | ETH3DB5MMN1P7G-QQO;G9J0 IL',Y;0U | 67 із 192 |
| MAHEROVSKYX | HY;RPQH1.H1ENY3H8.;UUM4L7B2C9W;U | 92 із 192 |
| MAHEROVSKYY | K;J9.'TMAB13;A88W8;29P86,15SQP U | 71 із 192 |
| MAHEROVSKYZ | N2'X7E'1P51YDSD'EL;GUSCXRX821IXU | 89 із 192 |
| MAHEROVSKY | Q7LFJTBM-'1NT4IW,Y;U9VGIGNBIIBUU | 65 із 192 |
| MAHEROVSKY. | TC1;V2N1DZ1C3MNNK5;2UYK35DEYZ4RU | 81 із 192 |
| MAHEROVSKY, | WHNL1HZMST11J-SE2I;G9.OU 3H8A;OU | 36 із 192 |
| MAHEROVSKY; | ZM33DW511N1WZGX5QV;UU-SFPZKORWLU | 86 із 192 |
| MAHEROVSKY- | ,RPRP5HMGH1L9Y,,82;291W0EPN-2PIU | 69 із 192 |
| MAHEROVSKY' | 'W59.KT1VB1APA1TWF;GU4 R3FQEJIFU | 87 із 192 |
| MAHEROVSKY0 | 2.RX7Z'M451''S6KES;U97-CY5TU BCU | 69 із 192 |
| MAHEROVSKY1 | 507FJ8B1J'1UF4BB,';2UA2;N.W4B49U | 89 із 192 |
| MAHEROVSKY2 | 85T;VNNMYZ1JVMG2KC;G9D6OCRZKS;6U | 63 із 192 |
| MAHEROVSKY3 | BA9L1,Z17T185-LZ2P;UUGA91H, 3W3U | 85 із 192 |
| MAHEROVSKY4 | EFV3DB5MMN1;LGQQQ,;29JE W7'AKP0U | 68 із 192 |
| MAHEROVSKY5 | HKBRPQH1.H1S.YVH89;GUMILL;2Q.I;U | 88 із 192 |
| MAHEROVSKY6 | KPX9.'TMAB1HBA 8WM;U9PM6AT50CB U | 71 із 192 |
| MAHEROVSKY7 | NUDX7E'1P516RS''EZ;2USQX'J8GT4XU | 91 із 192 |
| MAHEROVSKY8 | QZZFJTBM-'1.144W,6;G9VUIU9BW4;UU | 72 із 192 |
| MAHEROVSKY9 | T-F;V2N1DZ1QHM9NKJ;UUYY3J'E6LWRU | 87 із 192 |

Табл.1.Результати виконання хешування

Як видно з наведеної вище таблиці, внаслідок зміни одного біта у вхідній послідовності були спричинені значні зміни у вихідній. Тобто дана хеш-функція явно демонструє лавинний ефект.

4.Рівномірність розподілу частот.

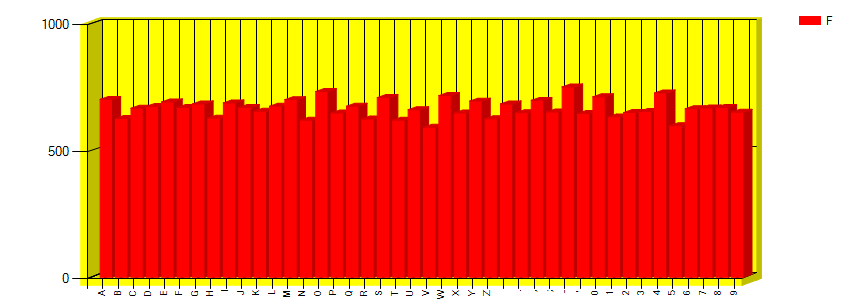


Рис.1.Частотний розподіл символів у тексті, шифрованому хеш-функцією.

Визначимо середнє інтегральне відхилення за допомогою формули 3.1.

Середнє інтегральне відхилення дорівнює 11.23 % (для порівняння – стандартний метод Хілла дає відхилення близько 30%).По рис.1. можна побачити, що частотний розподіл символів близький до рівномірного, водночас процес хешування не вимагає значних ресурсів для виконання.

# Список Літератури

# А. П. Алферов, Основы криптографии. Брюс Шнайер, Прикладная криптография.