Дрейс Юрий Александрович, кандидат технических наук, доцент, заведующий кафедрой дистанционного обучения Национального авиационного университета.
E-mail: y.dreis@nau.edu.ua
Дрейс Юрій Олександрович, кандидат технічних наук, доцент, завідувач кафедри дистанційного нав- чання Національного авіаційного університету.
Dreis Yurii, PhD in Eng., Associate Professor, Head of Distance e-Learning Academic Department, National Aviation University (Kyiv, Ukraine).
Гололобов Андрей Юрьевич, аспирант кафедры безопасности информационных технологий Нацио- нального авиационного университета.
E-mail: burn2dust@gmail.com.
Гололобов Андрій Юрійович, аспірант кафедри безпеки інформаційних технологій Національного авіаційного університету.
Gololobov Andrew, postgraduate student of IT-Security Academic Department, National Aviation University.

DOI: 10.18372/2410-7840.18.11085
УДК 004.421.5

ГЕНЕРУВАННЯ ВИПАДКОВИХ ЧИСЕЛ ШТАТНИМИ ЗАСОБАМИ ХОСТІВ МЕРЕЖІ ІНТЕРНЕТ

Володимир Чуприн, Володимир Вишняков, Михайло Пригара

У системах опитування або голосування через мережу Інтернет з метою забезпечення умов для вільного висловлювання думки опитуваних необхідно збереження таємниці голосів. Для цього використовують криптографічний захист інформації, який потребує генерування випадкових чисел. Вкрай бажано, щоб ці числа були дійсно, а не псевдовипадковими. Генератори дійсно випадкових чисел (ГВЧ), що використовують у складі підсистем захисту серверних частин систем голосування, мають відповідати ряду критеріїв, зокрема вимогам НД ТЗІ, оскільки домени безпеки серверного обладнання систем голосування в багатьох випадках мають бути сертифікованими відповідним уповноваженим органом. У той же час вимога щодо обов'язкової сертифікації за критеріями ТЗІ клієнтського обладнання систем голосування, як правило, не висувається, оскільки бажано, щоб виборець мав можливість здійснювати акти волевиявлення за допомогою виключно штатних засобів будь-якого хоста. Саме можливість голосування з будь-якого хоста без будь-яких спеціальних зусиль та засобів є найбільш привабливою властивістю систем дистанційного голосування. У даній роботі пропонується механізм отримання дійсно (не псевдо) випадкових чисел на будь-яких клієнтських вузлах мережі Інтернет без додаткових програмних засобів або фізичних пристроїв. Для цього пропонується скористатися випадковим характером нестабільності частот двох кварцових резонаторів (таймерного та тактового), що входять до складу будь-якого комп'ютера, та випадковим характером потоку запитів, що надходять на порти цього комп'ютера із мережі Інтернет. Проведено експериментальні дослідження запропонованого механізму. Теоретично доведена неможливість передбачати отримані числа в заданих умовах використання, що свідчить про недоцільність будь-яких спроб їх розкриття за допомогою крипто аналізу.

Ключові слова: *Інтернет-голосування, клієнтський хост, технічний захист інформації, генерування випадкових чисел, нестабільність кварцових резонаторів, потік запитів із Інтернет, самоподібний процес.*

ВСТУП. Захист інформації в системах опитування або голосування з використанням мережі Інтернет повинен забезпечувати умови для вільного висловлювання думки опитуваних. Зрозуміло, що тільки у випадку впевненості у неможливості розкриття результату волевиявлення людина може висловлюватись вільно. Загально відомо, що довіру виборців заслуговують лише ті системи, які є прозорими для контролю з боку виборчої спільноти. На прикладі системи опитування «Викладач очима студентів», з якою можна ознайомитись на сайті http://fit.univ.kiev.ua/archives/3246, легко зрозуміти, що у разі розкриття викладачами інформації про те, як проголосували конкретні студенти, з'являється можливість упередженого ставлення до цих студентів на іспитах. Тому через відсутність досконалого захисту інформації про голоси студентів неможливо забезпечити об'єктивність результатів опитування, що дискредитує саму ідею подібного заходу. На рівні державних виборів таємниця голосів є обов'язковою, що вказано, зокрема, у Міжнародному Пакті від 16 грудня 1966 р. «Про громадянські та політичні права». Труд-

нощі щодо захисту інформації в системах електронного голосування описані на ресурсі https://www.frisc.no/wp-content/uploads/2014/05/finse2014-kemmerer-1.pdf у популярній формі. Звідки бачимо, що на кожну чергову ідею розробників щодо захисту системи електронного голосування знаходяться зловмисники, які здатні подолати захист. При цьому система з кожним разом стає все складнішою за рахунок додаткових засобів захисту. Для подолання подібних труднощів Брюс Шнайер у роботі [1] пропонує у наступних розробках стати на шлях спрощення замість ускладнення систем. Саме такий шлях обрано авторами у роботі [2], де описана спроба створення простої і прозорої системи з метою забезпечення можливості контролю усіх програмних засобів будь-яким представником суспільства. Всі прикладні програми цієї системи мають просту структуру і невеличкий обсяг. Крім того, для їх створення обрано найпоширеніші мовні засоби (HTML та JavaScript), що забезпечує легкість перевірок. Для захисту інформації під час передавання обрано шифр Вернама, який забезпечує абсолютний захист за умов наявності на обох кінцях системи зв'язку однакових випадкових бітових послідовностей з довжиною, що є не меншою за довжину повідомлень. Саме те, що неважко реалізувати в системах дистанційного волевиявлення. Як варіант, спочатку на кожному кінці генерують випадкові послідовності по 503 біти, що майже завжди на практиці перевищує сумарну довжину повідомлень, які необхідно захищати, а потім, з використанням алгоритму Диффі-Хеллмана, їх перетворюють у однакові. Для цього виконується операція знаходження степені примітивного елементу поля Галуа $GF(2^{503})$, що потребує витрат часу близько однієї секунди. У порівнянні з тривалістю сеансу голосування такі витрати часу є незначними, а час на дискретне логарифмування з метою розкриття голосів є явно більшим за той, який потрібен зловмисникам. Захист сертифікованого сервера від несанкціонованого доступу у цій системі забезпечується відкритим регламентом роботи, що є обов'язковим для виконання. Адміністратор зобов'язаний надати права доступу для контролю своїх дій і файлів програмного забезпечення усім бажаючим. Для адміністрування сервера створюється користувач, який має право занесення файлів і запуску програм тільки у одній директорії, після чого користувач з повними правами видаляється. У такому стані доступ до сервера, з метою втручання в роботу програми, є неможливим. Управління роботою сервера до повного завершення кампанії волевиявлення надається виключно проконтрольованій прикладній програмі, після чого сервер потребує інсталяції, бо ніякі управляючі дії у цьому стані сервер не сприймає. Гарантований захист інформації в цій системі потребує генерування випадкових бітових послідовностей, які б забезпечували неможливість їх розкриття. Задачі створення механізму генерування таких послідовностей *на клієнтському обладнанні* системи голосування присвячена дана стаття.

Відомо, що для досконалого захисту інформації не рекомендовано обирати стандартні функції генерації псевдовипадкових чисел, що побудовані на засадах математичних перетворень, які у результаті дають числа, що схожі на випадкові, але не є випадковими по-справжньому [3]. Знаючи методи отримання цих чисел, зловмисники можуть подолати систему захисту. Проте реалізація подібних загроз стає неможливою у разі генерування дійсно випадкових чисел, що отримуються шляхом використання властивостей непередбачуваних фізичних процесів. До таких процесів в техніці належить природна нестабільність параметрів елементів електронних схем, наприклад, частоти кварцових резонаторів. На цих засадах побудовано чимало спеціальних пристроїв для генерації по-справжньому випадкових чисел. Існують також генератори, що побудовані виключно штатними технічними засобами комп'ютера, але потребують спеціалізованого програмного забезпечення, наприклад, Intel Digital Random Number Generator [4]. У нашому випадку необхідно, щоб джерело випадкових чисел, що реалізується на хості, для захисту інформації в мережі Інтернет було побудовано з використанням виключно штатних програмно-апаратних засобів хоста і не залежало від типу пристрою, будь-то мобільний телефон чи планшет будь-якого виробника або комп'ютер довільного типу, бо можливість голосування з будь-якого хоста без будь-яких спеціальних зусиль та засобів є найбільш привабливою властивістю системи дистанційного голосування. Єдина спільна вимога до всіх цих пристроїв є забезпечення доступу до мережі Інтернет. Саме такий механізм отримання випадкових чисел пропонується у даній роботі.

МЕТОЮ даної роботи є розробка механізму та надання рекомендацій щодо отримання випадкових чисел на стороні клієнтського обладнання для захисту інформації у системах опитування або голосування через мережу Інтернет для випадків, коли є небажаним доповнення клієнтських хостів додатковими фізичними пристроями

або встановлення будь-якого додаткового програмного забезпечення, тобто коли існує потреба обмежитись виключно штатними програмно-апаратними засобами, щоб не накладати зайвих обмежень на вибір місця або пристрою для голосування.

ВИХІДНІ ПЕРЕДУМОВИ

1. Для генерування дійсно випадкових чисел, як правило, використовують певні характеристики фізичних процесів, котрим притаманна властивість непередбачуваності. Якщо мова йде про ГВЧ у складі незахищеного хоста, що має бути використаний у системі голосування через Інтернет, то будь-яке втручання в роботу ГВЧ з боку будь-якого зловмисника вважається не суттєво ймовірним (оскільки зловмисник в умовах демократичного суспільства не має змоги здійснювати оперативний контроль пересування виборців, а також забороняти використання тих чи інших засобів доступу до Інтернет).

2. У складі апаратних засобів будь-якого комп'ютеризованого пристрою, що забезпечує можливість доступу до мережі Інтернет, мають місце два незалежних кварцових резонатори, один з яких має частоту 32,768 кГц, а інший – частоту генерування не менше, ніж 14318,18 кГц. Перший резонатор використовується в якості таймера. Другий – для формування тактових сигналів процесора. Обидва резонатори мають нестабільність частоти генерування від 10 до 100 *ppm* (або 10^{-4} – 10^{-5}) і не є синхронізованими між собою. Нестабільність частоти резонаторів може бути використана для отримання випадкових чисел. Проте за певних умов [3] існує можливість визначити функцію розподілу цих випадкових чисел і, отже, використати частотний метод крипто аналізу, що негативно впливає на рівень довіри до стійкості цього методу. Це повинно бути враховано у запропонованому механізмі отримання випадкових чисел.

3. Якщо комп'ютерна система використовується у мережі Інтернет в якості хоста, то, як свідчить практика, на увідні порти цієї системи безперервним потоком надходять численні бажані та небажані запити від різного роду джерел, що можуть не мати відношення до прикладних задач, що вирішуються за допомогою засобів цієї комп'ютерної системи (спами, реклама, об'яви, технологічний супровід встановлених програмних продуктів і т.ін.). Потік цих запитів, згідно результатів багатьох наукових досліджень [5,6], має випадковий характер з невідомою функцією розподілу. Усі запити, що потрапляють на увідні порти комп'ютерної системи, оброблюються засобами встановленої операційної системи. Кількість і тривалість обробки цих запитів, особливо коли їх поява пов'язана з неконтрольованими діями в мережі, передбачити неможливо. От же, фізичний процес обробки запитів в комп'ютерній системі може бути використаний для отримання дійсно випадкових чисел.

4. Кількість комп'ютерних операцій, які синхронізовані тактовими сигналами процесора (наприклад, додавання одиниці), у певних проміжках часу (наприклад, 1 ms), які синхронізовані кварцовим резонатором таймера, являє собою деяке випадкове число, конкретне значення якого залежить від двох вище названих фізичних процесів. Цією особливістю, яка має місце у кожному хості Інтернету, пропонується скористатись з метою отримання дійсно випадкових чисел.

5. До отриманої послідовності випадкових чисел ніщо не заважає додавати послідовності псевдовипадкових чисел від існуючих штатних генераторів.

МЕХАНІЗМ ОТРИМАННЯ ВИПАДКОВИХ ЧИСЕЛ

Пропонується наступний механізм отримання випадкових чисел, що базується на використанні обох кварцових резонаторів, що є складовими елементами будь-якого комп'ютеризованого пристрою, у поєднанні з фізичним процесом обробки зовнішніх запитів засобами операційної системи, що надходять із мережі Інтернет. Перший резонатор (таймер), послідовно у реальному часі формує більш/менш однакові проміжки часу («неоднаковість» пов'язана із нестабільністю частоти таймера), а сигнали другого, більш швидкісного, тактового резонатору використовуються для «заповнення» цих проміжків шляхом формування подій, кількість котрих підраховується у цих проміжках часу, при цьому процес підрахунку може перериватись через тимчасове переключення процесора на обробку більш пріоритетних внутрішніх або зовнішніх запитів. Іншими словами, програмний лічильник підраховує кількість елементарних дій, що синхронізуються тактовим резонатором, протягом проміжків часу, що формує таймер. Внаслідок нестабільності частоти сигналів, що генеруються кожним із резонаторів (які, у свою чергу, не є синхронізованими між собою), а також внаслідок випадковості переривань, які пов'язані із випадковістю запитів, кількість елементарних дій, що попадають у проміжки часу, що генеруються таймером, є випадковою величиною. Дослідження ймо-

вірнісних характеристик цієї випадкової величини, особливо її функції розподілу, являє теоретичний і практичний інтерес, зокрема з точки зору визначення можливостей застосування частотного методу крипто аналізу генератора випадкових чисел, що реалізує запропонований механізм. У даному випадку єдиний шлях дослідження – експеримент.

ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ МЕХАНІЗМУ ОТРИМАННЯ ВИПАДКОВИХ ЧИСЕЛ

Мета експериментів – показати, що послідовність випадкових чисел, що утворюються як результат роботи запропонованого механізму, має непередбачуваний характер.

Для реалізації експериментів створено на мові *JavaScript* та розміщено на сервері у вигляді *HTML* документів в режимі вільного доступу з будь-якого хоста мережі Інтернет ряд програм, що дозволяють перевірити статистичні характеристики запропонованого механізму на кожному з цих хостів. Обрання мови *JavaScript* для реалізації механізму забезпечує його працездатність навіть на тих пристроях, де неможливо встановити ніяке додаткове програмне забезпечення. Для завантаження і виконання усіх цих програм цілком достатньо штатних програмно-апаратних засобів хостів. Дані програми слід розглядати як інструментальні засоби, які дають змогу кожному бажаючому користувачу Інтернету на своєму пристрої (це може бути будь-який комп'ютер з будь-якою операційною системою або мобільний телефон чи навіть телевізор з функцією *SmartTV*) здійснювати перевірки статистичних характеристик отриманих за допомогою даного механізму послідовностей випадкових чисел. Доступ до трьох перших програм може бути реалізовано через одне з наступних посилань:

- http://91.198.50.7:11111/exp.html;
- http://91.198.50.7:11111/EXPPC.html;
- http://91.198.50.7:11111/exptv.html.

В залежності від типу пристрою кожен користувач може обрати для завантаження та виконання той варіант програми, який відповідає швидкодії свого пристрою. Перше посилання призначене для найбільш швидкісних сучасних комп'ютерів. У разі коли комп'ютеру близько 10 років, то слід скористатись другим посиланням. Третє посилання призначено для телевізорів з функцією *SmartTV*. Усі три програми по 999 разів в однакових проміжках часу тривалістю 0,001 с, виконують операцію додавання одиниці. Цей експеримент показує, що кількість нарахованих одиниць на одному й тому ж комп'ютері в однакових проміжках часу є різною. Результат дії програми *exp.html* на комп'ютері типу *Pentium* 2020*M* показано на рисунку 1.

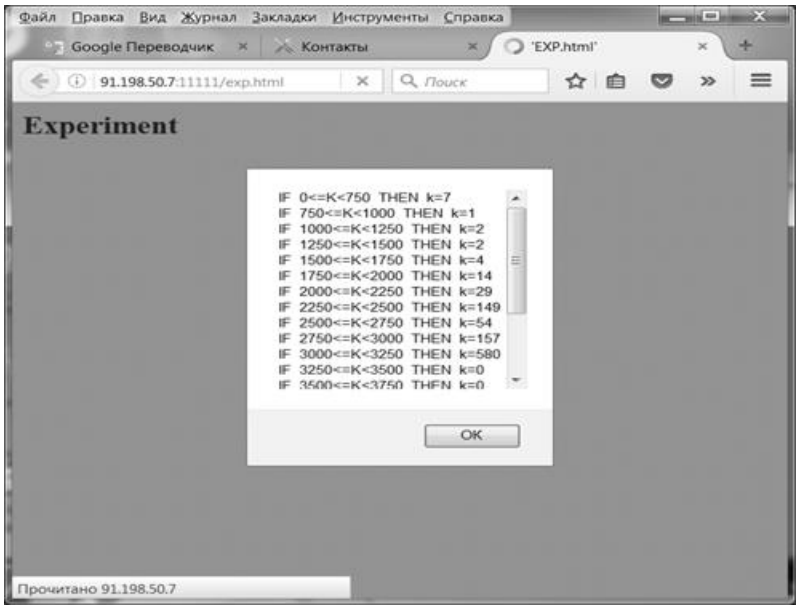

Рис. 1. Вигляд результату виконання програми через перше посилання на комп'ютері з тактовою частотою 2,1 ГГц

Ця програма висвітлює на екрані у першому рядку кількість випадків (із 999 можливих), коли результат підрахунку був в межах від 0 до 750 одиниць, у другому рядку кількість випадків з результатом в межах від 750 до 1000 одиниць і в кожному

наступному рядку ліва і права межі послідовно збільшуються на 250 одиниць, аж поки результат підрахунку дійде до значення 4750. В останньому рядку висвітлюється кількість випадків нарахування більше ніж 4750 одиниць.

Програма *EXPPC.html*, яку можна завантажити через друге посилання буде висвітлювати в першому рядку кількість випадків із 999 можливих, коли результат підрахунку був в межах від 0 до 75 одиниць, далі в межах від 75 до 100 одиниць і в кожному наступному рядку ліва і права межі збільшуються на 25 одиниць до значення 800. В останньому рядку висвітлюється кількість випадків нарахування більше ніж 800 одиниць.

Програма *exptv.html*, яку можна завантажити через третє посилання на телевізорах з функцією SmartTV, відрізняється від *EXPPC.html* меншою кількістю рядків, що висвітлюються на екрані, бо у SmartTV не буває випадків нарахування більше ніж 475 одиниць.

Чисельні експерименти на різних типах комп'ютерів показують, що розбіжність між кількістю нарахованих одиниць на одному й тому ж комп'ютері в однакових проміжках часу, що виробляє таймер, є досить значною. Однією з причин даної розбіжності є нестабільність частоти кварцових резонаторів, що є невід'ємними частинами будь-якого комп'ютера. Як вже вказувалось, у склад кожного сучасного комп'ютера входять два незалежних кварцових резонатори, один з яких має частоту 32,768 кГц і використовується в якості таймера. В нашому випадку від цього резонатора формуються інтервали тривалістю 0,001 с. Другий резонатор з частотою не менше, ніж 14318,18 кГц, використовується для формування тактових сигналів процесора. Обидва ці генератори мають нестабільність від 10 до 100 *ppm* (або $10^{-4} – 10^{-5}$). Якби ці резонатори були синхронізовані між собою, то за умов відсутності переривань даної прикладної програми в будь-якому інтервалі, який визначається частотою першого резонатора, фіксувалась би завжди однакова кількість дій, що однозначно залежить від частоти другого резонатора. Однак резонатори, що функціонують у складі кожного комп'ютера, не синхронізовані між собою, а їх нестабільність є одним з факторів отриманої розбіжності. У разі, якщо б цей фактор був єдиним, випадкова кількість нарахованих одиниць у рамках даного експерименту мала б характер близький до гаусової випадкової величини. З теоретичної точки зору у цьому випадку існувала б можливість оцінити на кількісному рівні основні статистичні характеристики функції розподілу нарахованих одиниць. Зокрема, можливо було б зробити оцінку значущості коефіцієнту кореляції за відомими критеріями (оцінити вибірковий коефіцієнт кореляції, перевірити гіпотези щодо значущості цього коефіцієнту, побудувати довірчі інтервали з використанням перетворення Фішера та розподілу Стьюдента, визначити рівень значущості коефіцієнту кореляції). Як результат, це надало б змогу спробувати використати частотний метод крипто аналізу. Але графіки, які побудовані для різних типів комп'ютерів на основі експериментальних даних і показані на рисунках 2-4, свідчать, що характер випадкової величини не є гаусовим (більше нагадує закон Релея). Надамо пояснення цьому факту.

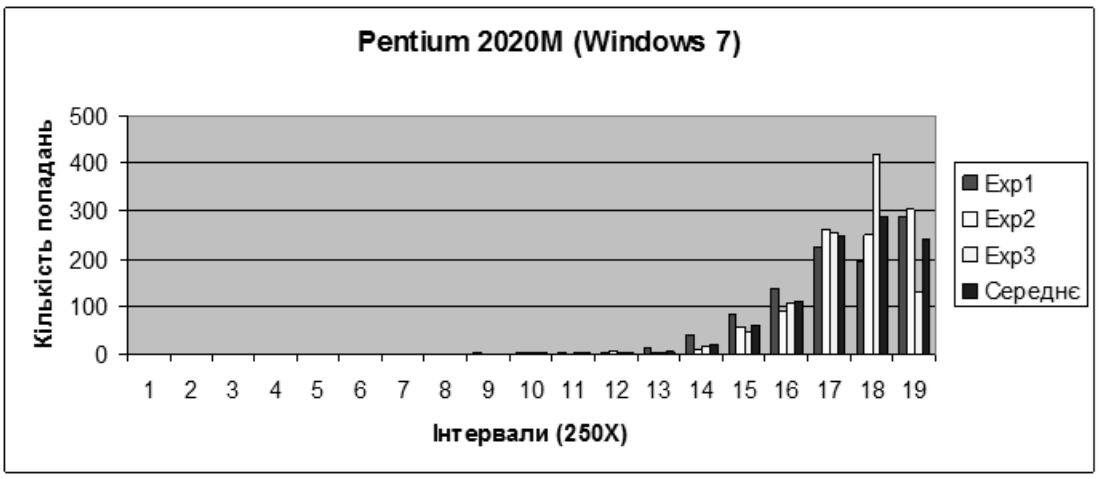

Рис. 2. Залежність кількості попадань числа нарахованих одиниць за проміжки часу 0,001 с в різні інтервали значень на комп'ютері з тактовою частотою 2,1 ГГц

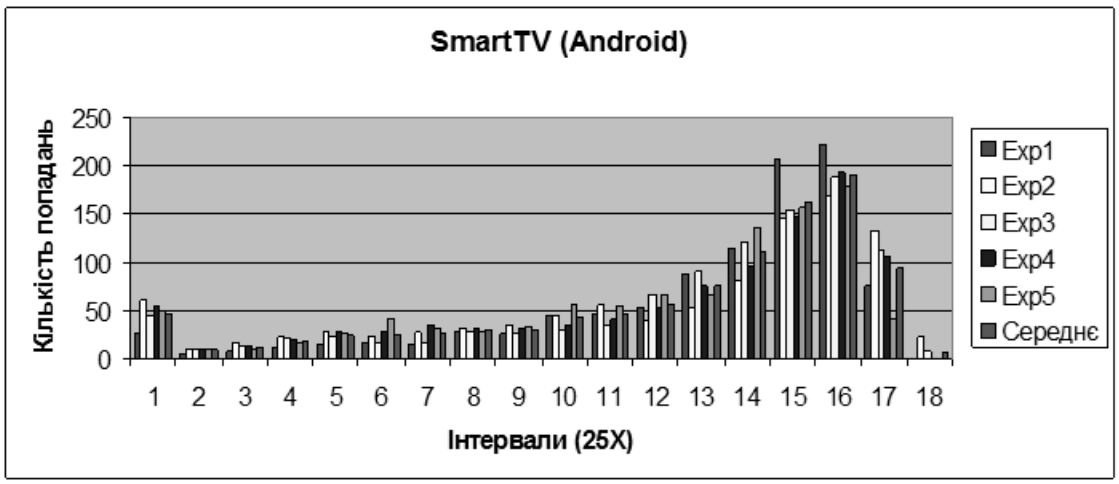

Рис. 3. Залежність кількості попадань числа нарахованих одиниць за проміжки часу 0,001 с в різні інтервали значень на телевізорі з функцією SmartTV

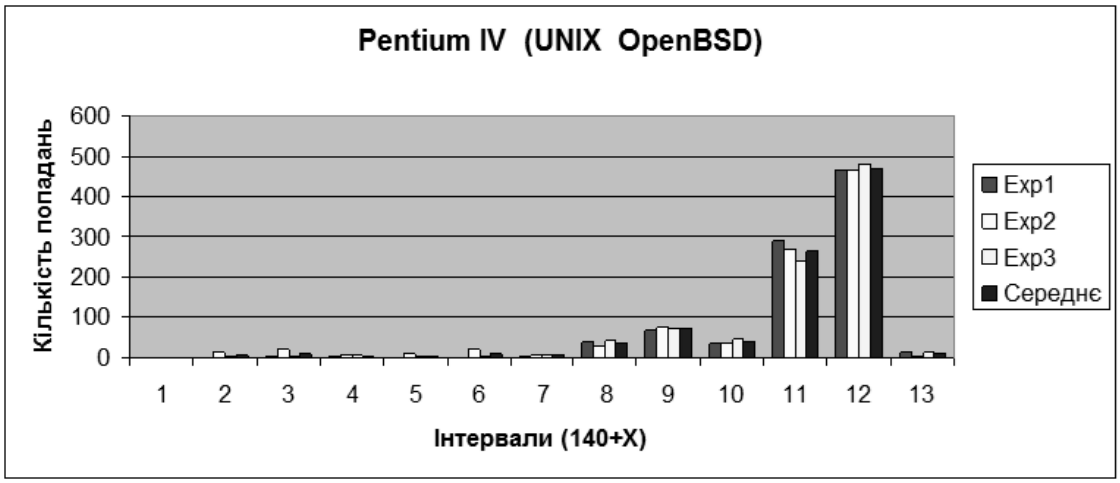

Рис. 4. Залежність кількості попадань числа нарахованих одиниць за проміжки часу 0,001 с в різні інтервали значень на комп'ютері з тактовою частотою 100 МГц

Усі отримані характеристики для різних типів комп'ютерів мають досить чітке обмеження з боку більших значень кількості нарахованих одиниць. Це пояснюється тим, що для кожного комп'ютера з конкретним значенням тактової частоти процесора існує максимальне значення кількості нарахованих одиниць $K_{\max}$, яке може бути налічено в проміжку часу 0,001 с. Це значення можна знайти з виразу

$$K_{\max} = \frac{F_{\max}}{1000k}, \qquad (1)$$

де $F_{\max}$ – максимальне значення тактової частоти процесора, Гц; k – кількість тактів процесора, що відповідають додаванню одної одиниці.

Розтягнення характеристик в бік менших значень нарахованих одиниць пояснюється тим, що в сучасних операційних системах процесорний час розподіляється між різними задачами. Зрозуміло, що через переривання дії нашої програми для виконання більш пріоритетних задач в деяких проміжках часу кількість нарахованих одиниць буде зменшуватись. Таке зменшення може сягати аж до нуля. Обрана тривалість 0,001 с є мінімально можливою, яку забезпечує програмний таймер. При цьому забезпечується максимум швидкості отримання випадкових чисел.

Як бачимо з представлених на рисунках 2-4 гістограм, близько половини результатів розміщено у правій частині, яка прилягає до значення $K_{\max}$. Для більш детального аналізу цієї частини результатів на будь-якому клієнтському пристрої можна скористатись посиланням:

– http://91.198.50.7:11111/expmax.html

Результат дії цього посилання представлено на рисунку 5.

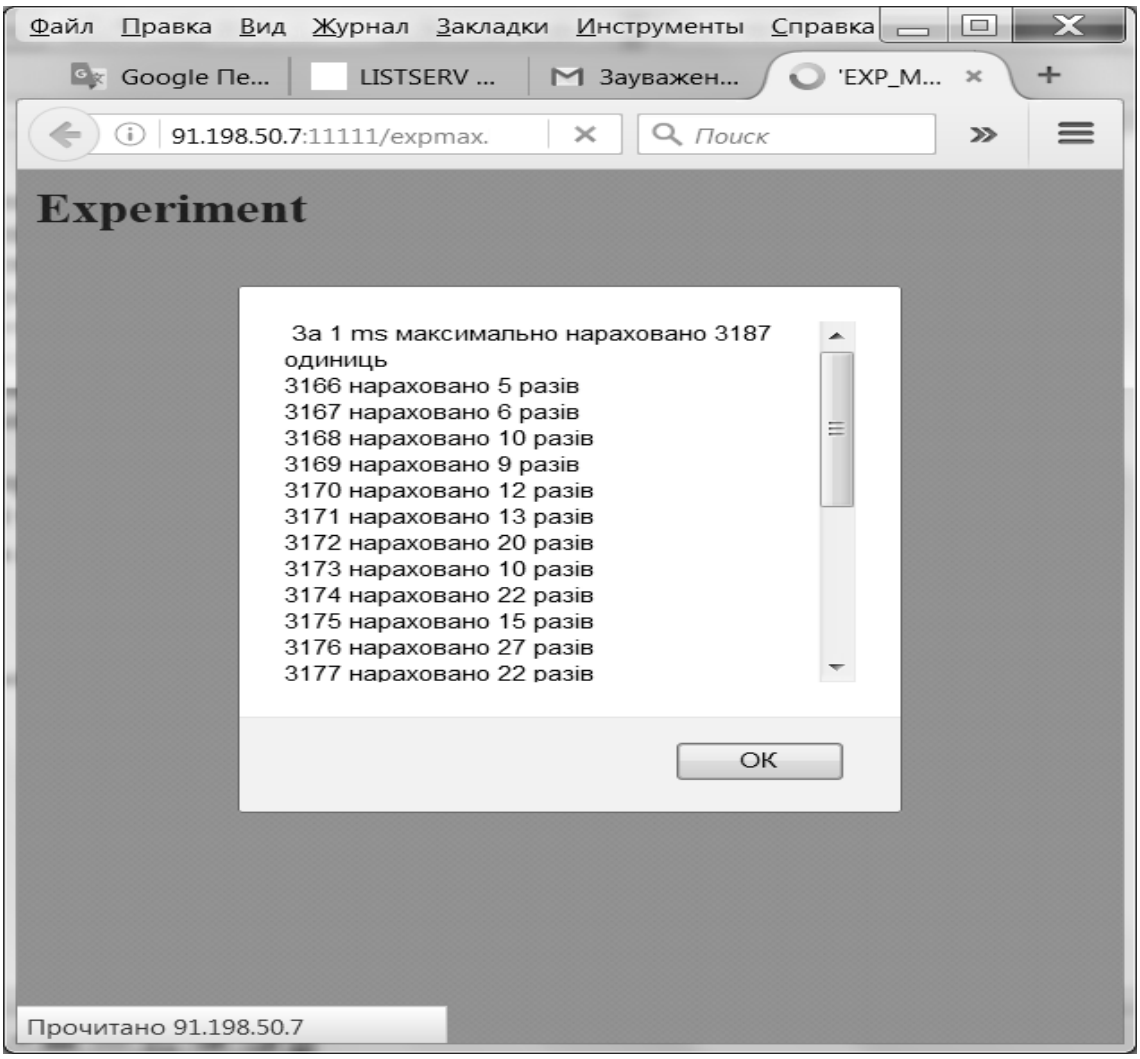

Рис. 5. Результат перевірки нарахувань одиниць, що прилягають до значення $K_{\max}$ на комп'ютері з тактовою частотою 2,1 ГГц

В цьому експерименті протягом однієї секунди виконується 999 вимірів кількості нарахувань одиниць, після чого значення результатів, які прилягають до $K_{\max}$ (це близько половини), висвітлюються на екрані. При повтореннях експерименту спостерігаються суттєві зміни результатів нарахувань, але кожного разу близько десяти чисел будуть мати значення наближене до $F/1000k$, де F – тактова частота процесора. Через нестабільність частоти F у різні секунди маємо зміщення максимальних значень нарахованих одиниць, а також зміну ширини інтервалу зі значеннями, що наближені до максимального. Ці числа будуть мати близькі одне до одного значення, а їх велика кількість у послідовності буде сприяти підвищенню коефіцієнта кореляції.

Аналіз отриманих вибірок свідчить про нестаціонарний характер послідовностей випадкових чисел, що формуються запропонованим механізмом, через що математичне обґрунтування їхньої непередбачуваності не уявляється можливим. Тим не менш, у роботі [9] надаються рекомендації щодо оцінки рівню значущості коефіцієнту кореляції експериментально отриманих вибірок випадкової величини. Уданому випадку врахуємо той факт, що значення коефіцієнту кореляції залежить від кількості врахованих молодших бітів в числах отриманої послідовності.

Відомо, що ідеальна випадкова послідовність чисел повинна мати нульовий коефіцієнт кореляції в умовах нескінченої тривалості експерименту. Для обчислення значень коефіцієнтів кореляції для різних значень врахованої кількості молодших бітів в отриманих числах на будь-якому клієнтському пристрої можна скористатись посиланням:

– http://91.198.50.7:11111/expco.html

Результат дії цього посилання представлено на рисунку 6.

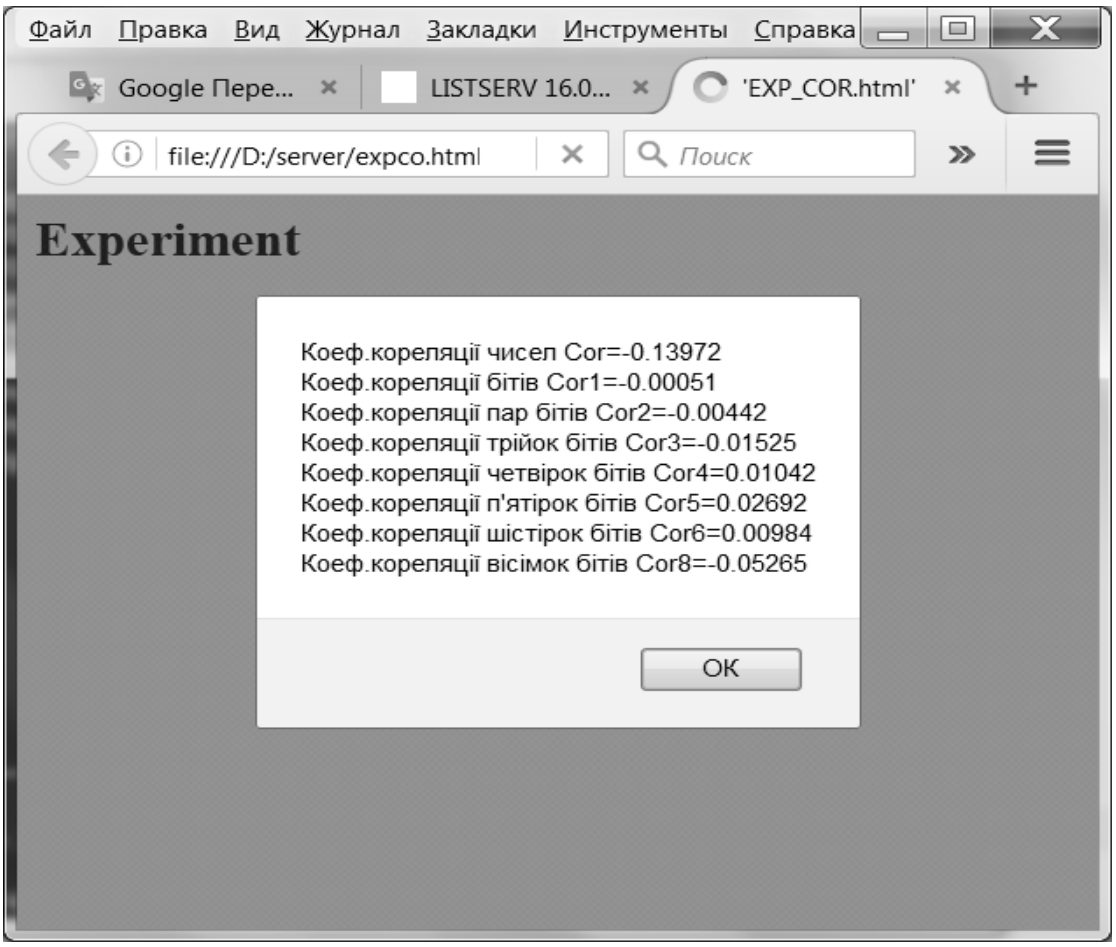

Рис. 6. Результат обчислення значень коефіцієнтів кореляції на комп'ютері з тактовою частотою 2,1 ГГц

Для обчислення коефіцієнтів кореляції у програмі, що реалізує даний експеримент, обрано довжину послідовності чисел у 10 тисяч, а також широко відому розрахункову формулу:

$$r = \frac{\sum_{i=2}^{n}(x_i - \bar{x}_2)(x_{i-1} - \bar{x}_1)}{\sqrt{\sum_{i=2}^{n}(x_i - \bar{x}_2)^2 \cdot \sum_{i=2}^{n}(x_{i-1} - \bar{x}_1)^2}}, \qquad (2)$$

де r – коефіцієнт кореляції,

$$\bar{x}_1 = \frac{\sum_{i=2}^{n} x_{i-1}}{n-1}, \quad \bar{x}_2 = \frac{\sum_{i=2}^{n} x_i}{n-1},$$

де x_i – елемент послідовності чисел $x_1, x_2, ..., x_i, ..., x_n$.

Оскільки для захисту інформації необхідні послідовності випадкових бітів, то з кожного числа нарахованих одиниць для розрахунку коефіцієнтів обрано також послідовності молодших бітів (від одного до восьми). Більше ніж 8 бітів обирати недоцільно, бо вони відповідають числам, що перевищують значення F_{max} для деяких хостів. Для прикладної задачі, що описана у роботі [2], потрібна послідовність з 503 випадкових бітів. У разі, коли з кожного числа нарахованих одиниць використати тільки по одному молодшому біту, то процес отримання необхідної послідовності займатиме часу близько 0,5 с. У разі використання m молодших бітів, часу буде потрібно в m разів менше. Але при цьому може збільшуватись коефіцієнт кореляції. В залежності від тактової частоти процесора клієнтського пристрою значення K_{max} може бути від декількох сотень (для застарілих планшетів) до декількох тисяч (для сучасних пристроїв), але в усіх випадках для кількості молодших бітів від одного до чотирьох значення коефіцієнта кореляції не перевищує 0,2. Також в усіх випадках коефіцієнт кореляції для послідовностей молодших бітів буде меншим ніж для послідовності чисел в цілому. Для комп'ютерів, що мають менші значення F_{max}, коефіцієнти кореляції мають більші значення. Виходячи з рекомендацій що наведені у роботі [7], а саме за умов $|r|>0{,}5$ можна вважати, що існує помітна кореляція, для $|r|>0{,}7$ можна вважати, що кореляція є суттєвою, а для випадків $|r|<0{,}3$ не має сенсу вести розмови про наявність

кореляційної залежності. Це означає, що отримані результати свідчать про можливість використання від одного до трьох молодших бітів послідовності випадкових чисел, які отримані за допомогою механізму, що запропонований в даній роботі, для формування випадкових послідовностей бітів на усіх клієнтських хостах, які мають доступ до Інтернету. При цьому підключення до Інтернету створює додаткові умови для довіри випадковому характеру отриманих чисел, бо якщо комп'ютер підключений до Інтернету, то на його увідні порти безперервним потоком надходять численні непередбачувані запити від різного роду джерел. Внаслідок чого процес додавання одиниць нашою програмою буде призупинятись в непередбачувані моменти часу. У цьому випадку логічно припустити, що кількість і тривалість обробки цих запитів передбачити неможливо.

ТЕОРЕТИЧНЕ ДОСЛІДЖЕННЯ НЕПЕРЕДБАЧУВАНОСТІ ПОТОКУ ЗАПИТІВ НА УВІДНОМУ ПОРТІ ХОСТА

Покажемо, що визначити функцію розподілу (щільність ймовірності) потоку запитів на увідному порту комп'ютерної системи, що підключена до Інтернету, не уявляється можливим. Чисельні спостереження за «поведінкою» потоків запитів на портах вузлового обладнання свідчать про нестаціонарний та пульсуючий характер змін інтенсивності цих потоків у реальному часі, що обумовлено механізмом статистичного мультиплексування індивідуальних потоків клієнтів пакетної мережі [6]. Тому ці потоки мають не прогнозований характер. Проте внаслідок усереднення, як показано у [7], тренди цих потоків отримують ознаки фрактального (самоподібного) процесу. Для фрактального процесу, ступень самоподоби котрого визначається параметром Херста, існує можливість визначити лише два перших статистичних моменти, а функція розподілу не є відомою[8]. Тим не менш, чим вище ступінь самоподоби (у певних нижче вказаних межах), тим вище рівень прогнозованості процесу. От же, інтерес являють теоретичні дослідження непередбачуваності потоку запитів на увідному порті хоста.

Реальний потік запитів не є безперервним процесом, стандартне визначення масштабної інваріантності котрого стосовно безперервного процесу $X(t)$ пов'язане з виконанням наступної рівності [7] :

$$X(t) \overset{d}{=} a^{-H} Z(at), \quad t \in T, \ a > 0, \ 0 < H < 1, \tag{3}$$

де символ $\overset{d}{=}$ розуміється у сенсі рівності щодо статистичного розподілу процесів $X(t)$ та $a^{-H} Z(at)$, а параметр H називають параметром Херста. У більшості випадків реальний потік запитів у пакетних мережах може адекватно представлятися дискретним самоподібним випадковим процесом. Покажемо можливість моделювання потоку запитів у пакетних мережах у вигляді дискретного асимптотично самоподібного (фрактального) процесу. Рівень самоподібності проаналізуємо не через схожість процесів у сенсі їхніх статистичних розподілів, а шляхом дослідження властивостей статистичних абсолютних моментів, зокрема функції автокореляції та індексу дисперсії.

Математичне представлення дискретного фрактального процесу здійснимо наступним чином. В якості моделі потоку запитів розглянемо напівнескінчений відрізок стаціонарного випадкового процесу X дискретного аргументу (часу) $t=0,1,\ldots,k, \ldots$, тобто часовий ряд $\{X_k;\ k = 0;1;2;...\}$, де k - поточний номер часового інтервалу усереднення процесу X. Тоді точкове значення κ-го відліку часового ряду $\{X_k^{(\tau)}; k= 0;1;2; \ldots\}$ при моделюванні потоку запитів має інтерпретуватися як кількість запитів x_k^{τ}, що надійшли у вузол обробки даних протягом k-го інтервалу часу тривалістю τ. Тобто, у даному випадку τ - це інтервал усереднення запитів у потоці. Якщо ряд $\{X_k^{(\tau)}; k= 0;1;2; \ldots\}$ унормувати відносно τ, то отримаємо ряд $\{I_k^{(\tau)}; k= 0;1;2; \ldots\}$, в якому k -й компонент визначає поточну інтенсивність запитів на k -ому кроці його усереднення. Кількість запитів, що надійшли у вузол обробки даних протягом k-го інтервалу часу тривалістю τ, дорівнює максимально можливому значенню індексу i_{max} , що задовольняє нерівності

$$\tau \geq \sum_{i=1}^{i\Delta\tau_{k,i}=\tau} \Delta\tau_{k,i} = \Delta\tau_{k,1} + \Delta\tau_{k,2} + ... + \Delta\tau_{k,i_{\max}}, \tag{4}$$

де $\Delta\tau_{k,i}$ - проміжок часу між сусідніми запитами у потоці, $i = 0, 1, 2, \ldots$ - поточний номер цього проміжку, а k - поточний номер часового інтервалу усереднення процесу $\{X_k^{(\tau)}; k= 0; 1; 2; \ldots\}$.

Отже, k-й компонент ряду $\{I_k^{(\tau)}; k= 0; 1; 2; \ldots\}$, що визначає поточну інтенсивність запитів на k-ому кроці його усереднення, визначено як

$$I_k^{(\tau)} = \frac{X_k^{(\tau)}}{\tau} \tag{5}$$

Якщо береться значення індексу $i_{max} \geq 2$, то маємо справу із усередненим процесом. Якщо ж розглядається послідовність моментів проходження одиничних запитів, то вважаємо, що $i_{max} = 1$.

$$R^{(\tau)}(n) = M\{X_k^{(\tau)} X_{k+n}^{(\tau)}\} = M\{|\frac{1}{\tau}\sum_{i=1}^{i\Delta\tau_i=\tau}\Delta\tau_{k,i} - M\{X_k^{(\tau)}\}|\cdot|\frac{1}{\tau}\sum_{i=1}^{i\Delta\tau_i=\tau}\Delta\tau_{k+n,i} - M\{X_k^{(\tau)}\}|\}, \quad (6)$$

де математичне очікування береться по усім K вибірки $\{X_k;\ k = 0;1;2;...\}$, а n - інтервал кореляції, тобто кількість членів ряду, що розміщені між членами щодо яких визначається кореляційний зв'язок.

Якщо дослідити кореляційну структуру ряду $\{I_k^{(\tau)};\ k = 0;1;2;\ \ldots\}$ згідно з виразом (6), то можливо упевнитись в тім, що потокам запитів у пакетних мережах притаманна властивість довгострокової залежності. Завдяки властивості довгострокової залежності (а також наявності у розподілі фрактального процесу «вагомих хвостів») існує можливість за певних умов прогнозування самоподібного трафіка. Але це не дозволяє виявити вигляд його функції розподілу.

Упевнимось, що досліджуваному потоку запитів притаманні властивості фрактального процесу. Таку упевненість отримуємо шляхом оцінювання значень параметра Херста. Якщо процесу X притаманна властивість самоподоби, то значення абсолютних моментів $\mu^{(\tau)}(q)$ визначається із формули [7]

$$\mu^{(\tau)}(q) = E\left\{\left|x^{(\tau)}\right|^q\right\} = E\left\{\left|\frac{1}{\tau}\sum_{i=k\tau-\tau+1}^{k\tau} x_i\right|^q\right\}, \quad (7)$$

де q – значення моменту статистичного розподілу процесу X (першого, другого і т.д.).

Але у загальному випадку для усіх самоподібних процесів справедливо наступне наближення [7]:

$$r(\tau) \sim \tau^{-\beta}, \qquad \tau \to \infty, 0 < \beta < 1. \quad (8)$$

Отже, абсолютні моменти будуть пропорційними значенню $\tau^{\beta(q)}$. Тому для фіксованих значень q буде дійсним наступне співвідношення:

$$\log \mu^{(\tau)}(q) = \beta(q)\log\tau + C(q). \quad (9)$$

Тобто, для асимптотично самоподібної послідовності X маємо [7] :

$$X \overset{d}{=} m^{1-H} X^{(\tau)}, \quad (10)$$

де символ $\overset{d}{=}$ розуміється у сенсі рівності щодо автокореляційної функції.

Враховуючи (10), вираз для показника ступеню $\beta(q)$ можливо записати у явному вигляді як

Найбільш зручним об'єктом аналізу під час моделювання потоку запитів є автокореляційна функція $R(n)$ процесу X або його коефіцієнт автокореляції $r(n)$. Тому автокореляційна функція має обчислюватися наступним чином:

$$\beta(q) = q(H-1). \quad (11)$$

Співвідношення (10) можливо використати для визначення самоподоби, що зводиться до перевірки лінійності залежності змін $\log \mu^{(\tau)}(q)$ від змін $\log\tau$.

Оцінювання параметра Херста у даній роботі виконується за індексом дисперсії (IDC), що здійснюється наступним чином. IDC визначається як відношення дисперсії кількості оброблених запитів на заданому часовому інтервалі T до математичного очікування цієї величини [7]:

$$F(T) = \frac{Var[N(T)]}{E[N(T)]}, \quad (12)$$

де $N(T)$ – кількість пакетів досліджуваного потоку, що були оброблені на інтервалі T.

Для самоподібних процесів натуральний логарифм $F(T)$-1 як функція від натурального логарифму інтервала T лінійно зростає, оскільки [7]:

$$\ln[F(T)-1] = (2H-1)\ln T + y, \quad (13)$$

де $y = \ln\left[\frac{2K}{\alpha(1-\alpha)} M_\Gamma(\alpha) B^{-\alpha/2}\right]$, $M_\Gamma(x) = \frac{\Gamma(1+x/2)\Gamma(1-x)}{\Gamma(1-x/2)}$.

ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ ПОТОКУ ЗАПИТІВ НА УВІДНОМУ ПОРТІ ХОСТА

Ціль даного експерименту – оцінити ступінь фрактальності (самоподібності) пакетного трафіку, що просувається через типовий прикінцевий вузол мережі Інтернет. Для досліджень обрано прикінцевий вузол типового мережного фрагменту Інтернет низового рівня з приблизно однаковою часткою корпоративних та домашніх клієнтів. Цій умові відповідає вузол Інтернет Державного НДІ автоматизованих систем у будівництві (ДНДІАСБ) (м. Київ). У період з 05.05.16 по 20.05.16 на базі цього вузлу експериментальним шляхом було отримано представницькі вибірки даних. Було прийнято одногодинний формат запису зареєстрованих показань. Тобто, кожен текстовий файл представляв вибірку даних розміром

3600 показань. Такий обсяг вибірки за будь-якими статистичними критеріями може вважатися представницьким.

На рис. 7 представлена добова гістограма однієї із характерних реалізацій пакетного трафіку. Проміжок усереднення – 1 хвилина. Інші реалізації трафіку мають подібні характеристики.

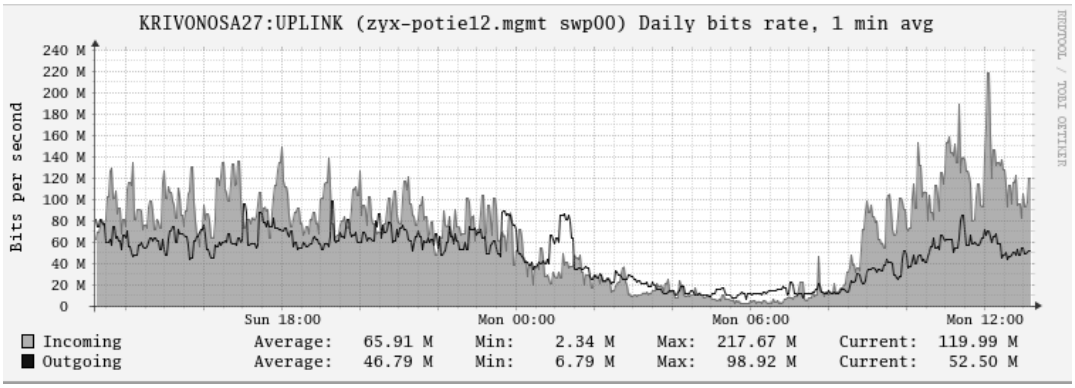

Рис. 7. Зразок експериментально отриманої вибірки даних

Оцінку параметру Херста отримано по ІДС, тобто по індексу дисперсії $F(\tau)$, що обчислено згідно виразу (13) як функція від проміжку кореляції τ. Отриманий у даному експерименті вигляд залежності $ln(F(T)$ -1) від $ln(T)$ показано на рис. 8

Відповідно до виразу (13), точки на рис.8 повинні утворювати пряму лінію (якщо досліджуваний вибірці трафіка притаманні властивості самоподібного процесу), що дозволяє по кутовому коефіцієнту нахилу прямої лінії однозначно визначити параметр Херста. Як бачимо, пряма, що отримана шляхом лінійної регресії, має кутовий коефіцієнт 0.46. Це означає, що параметр Херста $H=0,77$. Таким чином, констатуємо: для розглянутої реалізації досліджуваного трафіку значення параметра Херста, що отримані шляхом аналізу IDC вказують на істотно виражені фрактальні властивості досліджуваного трафіку. Проте не усі досліджені реалізації трафіку виявились самоподібними. А для значної частини реалізацій характерна невисока ступінь самоподібності. Отримані експериментальні дані свідчать: приблизно для 15% вибірок із загального числа досліджених реалізацій пакетного трафіка обчислені значення параметра H виявились меншими за 0,5, а ще для 80% вибірок діапазон значень параметра H знаходився у межах 0,5 - 0,7, і тільки 5% реалізацій мали яскраво виражений фрактальний характер. Звідсіля витікає, що для значної долі експериментально отриманих вибірок навіть такі «слабкі» закономірності як властивість самоподібності, «вагомі хвости» і т.п. не є характерними. Так що потік запитів має непередбачуваний характер.

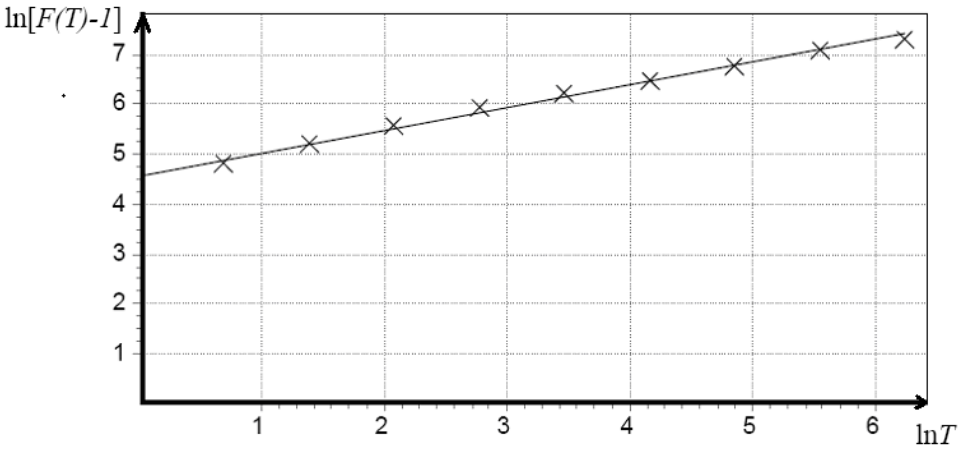

Рис. 8. Оцінювання параметра Херста H за значеннями IDC

ВИСНОВКИ

1. У системах дистанційного голосування через Інтернет вкрай важливо забезпечити можливість волевиявлення з будь-якого не сертифікованого за критеріями ТЗІ хоста без будь-яких спеціальних зусиль та засобів. За цих умов для збереження таємниці голосування у складі хоста має функціонувати генератор випадкових послідовностей, що побудований з використанням виключно

штатних програмно-апаратних засобів хоста. У будь-якій комп'ютерній системі, що приєднана до Інтернету, існує можливість генерувати дійсно (не псевдо) випадкові числа. Для цього можна скористатись випадковим характером нестабільності частот двох кварцових резонаторів (таймерного та тактового), що входять до складу будь-якого комп'ютера, та випадковим характером потоку запитів, що надходять на порти цього комп'ютера із мережі Інтернет.

2. Запропоновано механізм отримання випадкових послідовностей чисел на клієнтському обладнанні систем голосування через Інтернет, що заснований на використанні вищеназваних випадкових фізичних процесів, що супроводжують роботу будь-якого хоста. Особливість механізму: поєднання характеристик випадковості двох різних фізичних процесів в одному механізмі утворення випадкових послідовностей. Таке поєднання збільшує рівень довіри до непередбачуваності отриманих послідовностей чисел.

3. Для генерування випадкових чисел на будь-якому комп'ютері, що підключений до Інтернету пропонується скористатися випадковим характером нестабільності частот двох кварцових резонаторів (таймерного та тактового), що входять до складу будь-якого комп'ютера, та випадковим характером потоку запитів, що надходять на порти цього комп'ютера із мережі Інтернет. Таке генерування може здійснюватися за допомогою програми, яка у реальному часі підраховує кількість елементарних дій, що синхронізуються тактовим резонатором, протягом проміжків часу, що формує таймер. Внаслідок нестабільності частоти обох резонаторів та випадковості актів переривань процесора утворюється часовий ряд випадкових чисел.

4. Отримана послідовність випадкових чисел буде мати непередбачуваний характер внаслідок двох обставин. По-перше, внаслідок не стаціонарності характеристик нестабільності кварцових генераторів хоста. По-друге, використовується факт, що у непередбачувані моменти часу з Інтернету будуть надходити запити від різного роду джерел, кількість і тривалість обробки яких передбачити неможливо. Через що застосування частотного методу крипто аналізу втрачає сенс.

5. Виконано експериментальні дослідження обох вищеназваних фізичних процесів з метою обґрунтування їхнього випадкового характеру. Статистична обробка результатів експериментів показала дійсно випадковий характер характеристик цих процесів, що використовуються у запропонованому механізмі.

6. Запропонований механізм отримання випадкових чисел рекомендовано використовувати у системах таємного голосування з повністю відкритим програмним забезпеченням, у яких непотрібна висока швидкість генерації цих чисел та виключена мотивація потенційних зловмисників щодо зовнішнього втручання в роботу хоста.

ЛІТЕРАТУРА

[1]. Schneier B. What's Wrong With Electronic Voting Machines? https://www.schneier.com/essays/ archives/ 2004/11/whats_wrong_with_ele.html

[2]. Вишняков В.М., Пригара М.П., Воронін О.В. Відкрита система таємного голосування // Управління розвитком складних систем. Збірник наукових праць. – 2014. – Вип. 20. – С. 110 -115. http://urss.knuba.edu.ua/files/zbirnyk-20/ 22.pdf

[3]. Сушко С.О., Кузнєцов Г.В., Корабльов А.В. Математичні основи крипто аналізу. Дніпропетровськ: Національний гірничий університет, 2010. - 465 с.

[4]. Intel Digital Random Number Generator (DRNG): Software Implementation Guide, Revision 1.1. Intel Corporation. http://www. webcitation.org/ 6GhG1P7iR

[5]. Ghaderi M. On the Relevance of Self-Similarity in Network Traffic Prediction, 2003. http://www.cs. uwaterloo.ca / cs-archive / CS-2003 / 28 / TR-CS-2003-28.pdf

[6]. Муранов О.С. Дослідження можливостей прогнозування самоподібного трафіка у пакетних мережах // Моделювання та інформаційні технології: Збірник наукових праць Інституту проблем моделювання в енергетиці ім. Г.Є.Пухова. – Київ: ІПМЕ ім. Г.Є.Пухова НАН України, 2007. - № 44. – С. 98-106.

[7]. Alomar Mhamad. Influence of traffic prognostic mechanism on quality of adaptive control of switchboard / Alomar Mhamad, Saleh Alomar, Atef Obeidat // International Journal of Engineering Science (IJES). Bethesda (USA), 2014. – Vol. 7, no. 33. – pp. 1763-1776. (http://www.scirp.org/journal/eng) http://dx.doi.org/10.12988/ces.2014.4797.

[8]. Городецкий А. Я. Информатика. Фрактальные процессы в компьютерных сетях / А. Городецкий, В. Забровский. – Санкт-Петербург : Изд-во СПб ГТУ, 2000. – 96 с.

[9]. Мячин М.Л. Миф о значимости коэффициента корреляции https://sites.google.com/site/ltwood/ projects/stataddons/corrmyth

REFERENCES

[1]. Schneier B. What's Wrong With Electronic Voting Machines? https://www.schneier.com/essays/archives/ 2004/11/whats_wrong_with_ele.html

[2]. Vyshniakov V.M., Prygara M.P., Voronin O.V. 'Open secret ballot system', Managing the development of complex systems, vol. 20, pp. 110-115. http:// urss.knuba.edu.ua/files/zbirnyk-20/22.pdf

[3]. Sushko S.O., Kuznetsov G.V. & Korablov A.V. 2010, Mathematical fundamentals of cryptanalysis, National Mining University, Dnipropetrovsk.
[4]. Intel Digital Random Number Generator (DRNG): Software Implementation Guide, Revision 1.1. Intel Corporation. http://www.webcitation.org/6GhG1P7iR
[5]. Ghaderi M. On the Relevance of Self-Similarity in Network Traffic Prediction, 2003. http://www.cs.uwaterloo.ca / cs-archive /CS-2003 / 28 / TR-CS-2003-28.pdf
[6]. Muranov O.S. 2007, 'Research predictability self-similar traffic in packet networks', Modelling and Information Technology, vol. 44, pp. 98-106.
[7]. Alomar Mhamad. Influence of traffic prognostic mechanism on quality of adaptive control of switchboard / Alomar Mhamad, Saleh Alomar, Atef Obeidat // International Journal of Engineering Science (IJES). Bethesda (USA), 2014. – Vol. 7, no. 33. – pp. 1763-1776. (http://www. scirp.org / journal / eng) http://dx.doi.org/10.12988/ces.2014.4797.
[8]. Gorodetski A. & Zabrovski V. 2000, Computer science. Fractal processes in computer networks, SPb GTU, Sankt-Peterburg.
[9]. Mjachin M.L. The myth of the significance of the correlation coefficient https://sites.google.com/site/ltwood/projects/stataddons/corrmyth

ГЕНЕРАЦИЯ СЛУЧАЙНЫХ ЧИСЕЛ С ИСПОЛЬЗОВАНИЕМ ШТАТНЫХ СРЕДСТВ ХОСТОВ СЕТИ ИНТЕРНЕТ

В системах опроса или голосования в сети Интернет, с целью обеспечения условий для свободного выражения мнения опрошенных, необходимо сохранение тайны голосов. Для этого используют криптографическую защиту информации, требующую генерирования случайных чисел. Известно, что стандартные компьютерные функции для генерирования случайных чисел, которые построены на математических преобразованиях, дают псевдослучайные числа, похожие на случайные, но которые на деле не являются действительно случайными. Зная методы получения таких чисел, злоумышленники могут взломать систему защиты. В данной работе даются рекомендации, касающиеся принципов построения компьютерных программ генерации действительно (не псевдо) случайных чисел в оконечных узлах сети Интернет при условии использования исключительно стандартных программно-аппаратных средств (без каких-либо дополнительных физических устройств). Приведены результаты экспериментальных исследований предложенных принципов генерации случайных чисел. Теоретически доказана невозможность предвидения получаемых чисел, что свидетельствует о бесполезности попыток использования средств криптоанализа для их раскрытия.

Ключевые слова: криптографическая защита, совершенная защита, генерация случайных чисел, защита информации в сети Интернет, самоподобный процесс.

METHOD OF GENERATION OF CASUAL NUMBERS ON THE BASIS OF THE USE OF APPARATUS OF THE COMPUTER PLUGGED IN THE INTERNET

In systems or voting poll on the Internet, with a view to ensuring the conditions for the free expression of respondents thought necessary secrecy of the vote. For this purpose cryptographic protection of information that needs to generate random numbers. It is known that the standard computer functions to generate random numbers, which are based on mathematical transformations, give a pseudorandom generator that resemble random, but random really. Knowing the methods of obtaining such numbers, hackers can crack the security system. In this paper makes recommendations on the principles of construction of computer programs generating real (not pseudo) casual numbers in the terminal nodes of the Internet, provided the exclusive use of standard software and hardware (without any additional physical devices). The experimental results of the proposed principles of random number generation. In theory, proved the impossibility of foreseeing the numbers received, indicating that the futility of trying to use cryptanalysis funds to disclose them.

Keywords: cryptographic priv, perfect priv, generation of random numbers, priv in a network the Internet, self-similar process.

Чуприн Володимир Михайлович, кандидат технічних наук, професор кафедри телекомунікаційних систем Національного авіаційного університету.
E-mail: vladimir@ndiasb.kiev.ua
Чуприн Владимир Михайлович, кандидат технических наук, профессор кафедры телекоммуникационных систем Национального авиационного университета.
Chupryn Volodymyr, PhD in engineering, professor, Department of Telecommunication Systems, National Aviation University.

Вишняков Володимир Михайлович, кандидат технічних наук, доцент, доцент кафедри телекомунікаційних систем Національного авіаційного університету.
E-mail: volodymyr.vyshniakov@gmail.com
Вышняков Владимир Михайлович, кандидат технических наук, доцент, доцент кафедры телекоммуникационных систем Национального авиационного университета.
Vyshniakov Volodymyr, PhD in engineering, associate professor, Department of Telecommunication Systems, National Aviation University.

Пригара Михайло Петрович, аспірант кафедри інформаційних технологій Київського національного університету будівництва і архітектури.
E-mail: misha_prigara@ukr.net
Пригара Михаил Петрович, аспирант кафедры информационных технологий Киевского национального университета строительства и архитектуры.
Prygara Mykhailo, graduate student of Department of Information Technologies, Kyiv National University of Construction and Architecture.