

Оцінка характеристик бітового потоку як випадкової величини

О. М. Романов

Науково-дослідний інститут Міністерства оборони України, м. Київ, Україна
Corresponding author. E-mail: rolex@i.ua

Paper received 22.06.20; Accepted for publication 16.07.20.

<https://doi.org/10.31174/SEND-NT2020-233VIII28-15>

Анотація. Априорно невідомий бітовий потік з каналу зв'язку та передачі даних представлено як випадкову величину. За обмеженою кількістю спостережень оцінено його основні характеристики – математичне очікування й дисперсію кількості одиниць у вибірці. Розроблено спеціальне програмне забезпечення, яке дозволяє завантажити довільну вибірку бітового потоку, оцінити її основні статистичні характеристики та відобразити результати. Отримані результати представлені у таблиці й графічно.

Ключові слова: бітовий потік, аналіз, ймовірність, вибірка, математичне очікування, дисперсія, оцінка.

Вступ. При проведенні аналізу априорно невідомого бітового потоку його можна представити як випадкову величину. Виникає необхідність отримання її статистичних характеристик за обмеженою кількістю спостережень.

Огляд публікацій за темою досліджень вказує на розвиненість методичного апарату аналізу мережевих протоколів передачі даних і їх ланцюгів на вищих рівнях моделі OSI [1, 2, 3], а також визначення виду модуляції на першому рівні моделі OSI [4], існування окремих підходів до визначення виду і параметрів завадостійкого кодування [5]. Публікацій, які висвітлюють інші аспекти аналізу бітових потоків на першому рівні ієрархії моделі OSI, немає.

Метою роботи є оцінка математичного очікування і дисперсії кількості одиниць у вибірці з априорно невідомого бітового потоку.

Матеріали і методи. Априорно невідомий бітовий потік, що приймається з каналу зв'язку та передачі даних, представимо генеральною сукупністю $Y(n)=\{y_0, y_1, \dots, y_{z-1}\}$, $y_i \in \{0, 1\}$, де n – номер біту (рис. 1). Статистична ймовірність того, що $y_n=1$, – p_1 , відповідно, ймовірність того, що $y_n=0$, – $p_0=1 - p_1$. Ймовірності p_1 і p_0 можуть змінюватись залежно від інформації, що передається в каналі зв'язку і передачі даних, та методів кодування, що застосовуються.

1011000111010100010001100011.....110001010101000110101

Рис. 1. Приклад бітового потоку, що приймається з каналу зв'язку та передачі даних

Кількість одиниць N_i у довільній вибірці довжиною L , $L < Z$ з генеральної сукупності $Y(n)$ – дискретна випадкова величина. Кількість одиниць N_i , унормована до довжини вибірки L , $F=N_i/L$ – також дискретна випадкова величина, яка може приймати значення від 0 до 1 і не залежить від довжини вибірки.

Визначимо діапазон можливих значень випадкової величини F . Для цього проведемо k спостережень. У кожному досліді братимемо вибірку довжиною L із генеральної сукупності $Y(n)$, підраховуватимемо в ній кількість одиниць N_i і нормуватимемо її до довжини вибірки: $F=N_i/L$. Для забезпечення незалежності вибірок початок вибірки у кожному наступному спостереженні зсуватимемо на L відносно початку попередньої вибірки.

Спроможну незміщену оцінку математичного очікування випадкової величини F обчислимо за форму-

лою $\tilde{m}_F = \tilde{M}[F] = \frac{\sum_{i=1}^k f_i}{k}$ [6], де f_i – значення випадкової величини F , що спостерігалось в i -му досліді.

Спроможну незміщену оцінку дисперсії обчислимо за

формулою $\tilde{D}_F = \tilde{D}[F] = \frac{\sum_{i=1}^k (f_i - \tilde{m}_F)^2}{k - 1}$ [6]. Оцінку середньоквадратичного відхилення, яке більш нагля-

дно характеризує розсіювання випадкової величини, виразимо: $\tilde{\sigma}_F = \sqrt{\tilde{D}_F}$ [6].

За нерівністю Чебишева ймовірність відхилення величини F від свого математичного очікування не менше, ніж на додатне число a , обмежена зверху величиною D_F/a^2 : $P(|F - m_F| \geq a) \leq \frac{D_F}{a^2}$ [6].

Покладаючи $a=3\sigma_F$, отримуємо ділянку практично можливих значень випадкової величини $m_F \pm 3\sigma_F$, ймовірність попадання в яку величини F не менше 0,88.

Результати і їх обговорення. Для аналізу бітових потоків у середовищі програмування Visual Studio 2019 мовою програмування C# розроблено СПЗ, яке дозволяє завантажити довільну вибірку бітового потоку, здійснити обчислення за наведеними виразами та наглядно відобразити отримані результати.

Для проведення експериментів було обрано бітовий потік з каналу зв'язку, організованого за стандартом [7], з корисним навантаженням.

Отримані оцінки математичного очікування і середньоквадратичного відхилення нормованої кількості одиниць у вибірці для різних довжин вибірок і кількості спостережень представлені у табл. 1.

При кількості спостережень $k=450$, довжині вибірки у кожному спостереженні $L=450$ отримано такі оцінки нормованої кількості одиниць: $\tilde{m}_F = 0,505$, $\tilde{\sigma}_F = 0,027$. Практично можливі значення – від 0,424 до 0,586. На рис. 2 наведені результати обчислення

нормованої кількості одиниць у вибірці за результатами спостережень, а на рис 3 – гістограма її розподілу при $k=450$ і $L=450$.

Таблиця 1. Оцінки математичного очікування і середньоквадратичного відхилення нормованої кількості одиниць у вибірці

		Довжина вибірки L				
		100	200	300	400	500
Кількість спостережень k	100	$\tilde{m}_F = 0,519$ $\tilde{\sigma}_F = 0,064$	$\tilde{m}_F = 0,518$ $\tilde{\sigma}_F = 0,050$	$\tilde{m}_F = 0,514$ $\tilde{\sigma}_F = 0,043$	$\tilde{m}_F = 0,511$ $\tilde{\sigma}_F = 0,036$	$\tilde{m}_F = 0,513$ $\tilde{\sigma}_F = 0,034$
	200	$\tilde{m}_F = 0,518$ $\tilde{\sigma}_F = 0,065$	$\tilde{m}_F = 0,511$ $\tilde{\sigma}_F = 0,043$	$\tilde{m}_F = 0,511$ $\tilde{\sigma}_F = 0,037$	$\tilde{m}_F = 0,508$ $\tilde{\sigma}_F = 0,032$	$\tilde{m}_F = 0,507$ $\tilde{\sigma}_F = 0,029$
	300	$\tilde{m}_F = 0,514$ $\tilde{\sigma}_F = 0,061$	$\tilde{m}_F = 0,511$ $\tilde{\sigma}_F = 0,041$	$\tilde{m}_F = 0,508$ $\tilde{\sigma}_F = 0,035$	$\tilde{m}_F = 0,507$ $\tilde{\sigma}_F = 0,031$	$\tilde{m}_F = 0,506$ $\tilde{\sigma}_F = 0,027$
	400	$\tilde{m}_F = 0,511$ $\tilde{\sigma}_F = 0,059$	$\tilde{m}_F = 0,508$ $\tilde{\sigma}_F = 0,041$	$\tilde{m}_F = 0,507$ $\tilde{\sigma}_F = 0,034$	$\tilde{m}_F = 0,506$ $\tilde{\sigma}_F = 0,029$	$\tilde{m}_F = 0,505$ $\tilde{\sigma}_F = 0,027$
	500	$\tilde{m}_F = 0,513$ $\tilde{\sigma}_F = 0,059$	$\tilde{m}_F = 0,507$ $\tilde{\sigma}_F = 0,040$	$\tilde{m}_F = 0,506$ $\tilde{\sigma}_F = 0,033$	$\tilde{m}_F = 0,505$ $\tilde{\sigma}_F = 0,029$	$\tilde{m}_F = 0,505$ $\tilde{\sigma}_F = 0,026$

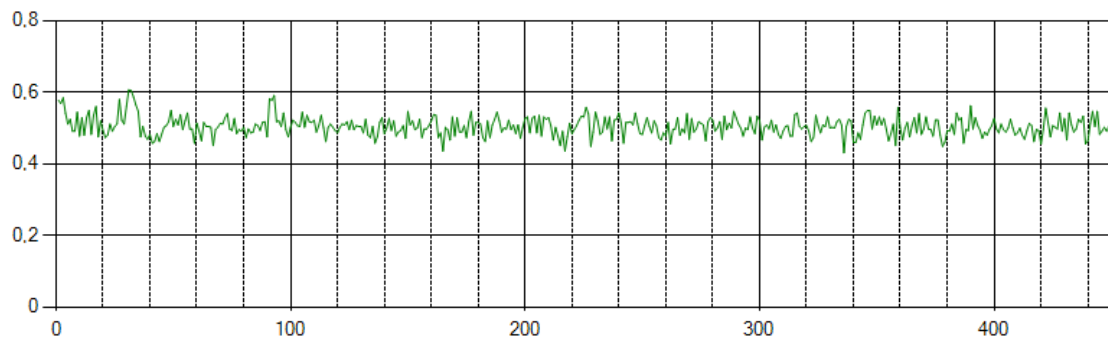


Рис. 2. Нормована кількість одиниць у вибірці за результатами спостережень

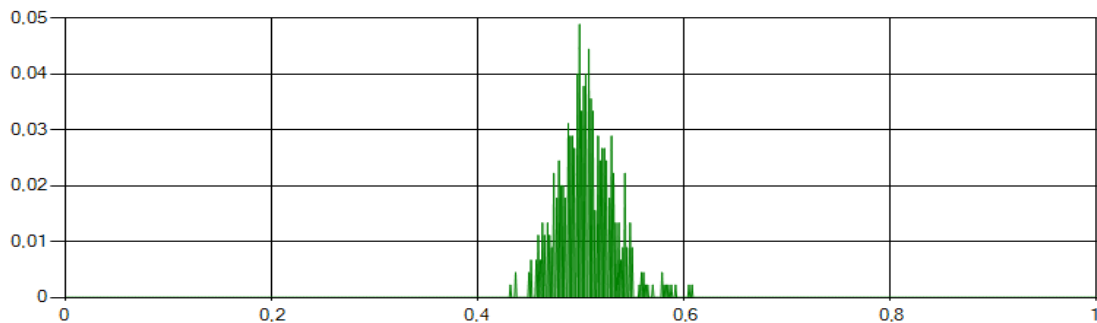


Рис. 3. Гістограма розподілу нормованої кількості одиниць у вибірці

Висновки. Теоретичні залежності і отримані експериментально дані вказують на можливість представлення апріорно невідомого бітового потоку як випадкової величини, яка характеризується математичним

очікуванням і дисперсією. По мірі збільшення довжини вибірки і кількості спостережень оцінка математичного очікування наближається до її істинного значення, а дисперсія зменшується.

ЛІТЕРАТУРА

1. Романов О. М. Застосування аналізаторів протоколів при технічному аналізі сигналів систем зв'язку / О. М. Романов // Проблеми кібербезпеки інформаційно-телекомунікаційних систем: II науково-практична конференція, Київ, 23–24 березня 2017 р.: матеріали доповідей. – К.: КНУ ім. Тараса Шевченка, 2017. – С. 177–179.
2. Маркин Ю. В. Методы и средства углубленного анализа сетевого трафика: автореф. дис. ... к-та техн. наук: 05.13.11 “Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей” [Электрон. ресурс] / Ю. В. Маркин. М.: ИСП РАН, 2017. – 30 с. – Режим доступа: <https://www.ispras.ru/dcouncil/docs/diss/2017/markin/autoref-markin-publ.pdf>.
3. Гетьман А. И. Обзор задач и методов их решения в области классификации сетевого трафика / А. И. Гетьман, Ю. В. Маркин, Е. Ф. Евстропов, Д. О. Обыденков // Труды ИСП РАН. – 2017. – Том 29. Вып. 3. – С. 117–150.
4. Воробьева Е. И. Распознавание вида модуляции сигналов в системах радиомониторинга / Е. И. Воробьева, Р. А. Немцов, П. П. Чураков // Вестник Воронежского государственного технического университета. – 2015. – № 11 (4). – С. 72–75.
5. Ревуцкий В. А. Устойчивые к мешающим факторам алгоритмы распознавания вида помехоустойчивых кодов в

- радиотехнических системах: автореф. дис. ... к-та техн. наук: 05.12.04 "Радиотехника, в том числе системы и устройства телевидения" / В. А. Ревуцкий. – Рязань: ГРТУ, 2013. – 19 с.
6. Вентцель Е. С. Теория вероятностей. 4е изд. / Е. С. Вентцель. – М.: Наука, 1969. – 576 с.
7. Digital Video Broadcasting (DVB). Framing structure, channel coding and modulation for 11/12 GHz satellite services: EN 300 421 V1.1.2 (1997-08). – Sophia Antipolis: European Telecommunications Standards Institute, 1997. – 24 p.

REFERENCES

1. Romanov O. M. Use of protocol analyzers in technical analysis of communication system signals / O. M. Romanov // Problems of cybersecurity of information and telecommunication systems: II scientific and practical conference, Kiev, 23–24 march 2017 y.: materials of reports. – K.: Taras Shevchenko KNU, 2017. – P. 177–179.
2. Markin U. V. Methods and tools for in-depth analysis of network traffic: abstract of the diss. ... cand. of techn. sciences: 05.13.11 "Mathematical and software of computers, complexes and computer networks" [Electronic resource] / U. V. Markin. M.: ISP RAS, 2017. – 30 p. – Access mode: <https://www.ispras.ru/dcouncil/docs/diss/2017/markin/autoref-markin-publ.pdf>.
3. Hetman A. I. Overview of tasks and methods for solving them in the field of network traffic classification / A. I. Hetman, U. V. Markin, E. F. Evstropov, D. O. Obydenkov // Works ISP RAS. – 2017. – Vol 29. Iss. 3. – P. 117–150.
4. Vorobeva E. I. Recognition of the type of signal modulation in radio monitoring systems / E. I. Vorobeva, R. A. Nemtsov, P. P. Churakov // Bulletin of Voronezh State Technical University. – 2015. – № 11 (4). – P. 72–75.
5. Revutsky V. A. Interference-resistant algorithms for recognizing the appearance of error-correcting codes in radio engineering systems: abstract of the diss. ... cand. of techn. sciences: 05.12.04 "Radio engineering, including television systems and devices" / V. A. Revutsky. – Ryazan: RSRTU, 2013. – 19 p.
6. Wentzel E. S. Probability theory. 4th ed. / E. S. Wentzel. – M.: Nauka, 1969. – 576 p.

Estimation of bitstream characteristics as a random variable

O. M. Romanov

Abstract. The a priori unknown bitstream from the communication and data channel is represented as a random variable. By a limited number of observations evaluated its main characteristics – mathematical expectation and variance of the number of ones in the sample. Special software has been developed that allows you to upload a random sample of bitstream, evaluate its main statistical characteristics and display the results. The obtained results are presented in the table and graphically.

Keywords: *bitstream, analysis, probability, sample, mathematical expectation, variance, estimation.*