

УДК 621.391:004.056

DOI: 10.31673/2412-9070.2023.022530

О. Г. ВАРФОЛОМЕЄВА, канд. техн. наук, доцент;
А. Г. ЗАХАРЖЕВСЬКИЙ, канд. техн. наук, здобувач,
Державний університет телекомунікацій, Київ

МОДЕЛЮВАННЯ ТА АНАЛІЗ ТРАФІКУ В КАНАЛАХ ТЕЛЕКОМУНІКАЦІЙНИХ ПІДСИСТЕМ СИСТЕМ КОМПЛЕКСНОЇ БЕЗПЕКИ

Обґрунтовано потребу в моніторингу телекомунікаційних підсистем систем комплексної безпеки. Запропоновано імітаційну модель, що дає змогу формувати елементи часового ряду, у ролі яких виступають моменти початку передавання повідомлень. Передбачено можливість формування як стаціонарних, так і нестаціонарних потоків подій. Наведено результати досліджень основних методів аналізу часових рядів. Показано, що використовувані підходи не дають змоги з необхідною оперативністю та вірогідністю виявляти можливі аномальні стани трафіку. Запропоновано для оцінювання різних станів трафіку розглядати телекомунікаційну систему як динамічну систему. Наведено аналітичні вирази для формування динамічних змінних, на основі яких сформовано фазові портрети. Здобуті результати експериментів підтвердили інформативність запропонованого варіанта відображення трафіку. Показано, що оцінювання стану телекомунікаційних систем із використанням фазових портретів дає змогу виявити аномальні стани трафіку, а також спрогнозувати поведінку динамічної системи в майбутньому.

Ключові слова: система комплексної безпеки; телекомунікаційна підсистема; імітаційна модель; часовий ряд; методи аналізу часових рядів; динамічна система; динамічні змінні; фазові портрети.

Вступ

Постановка проблеми. Комплексні системи безпеки об'єктів різного призначення і різної значущості — це автоматизовані ієрархічні складні системи. Одним з основних елементів таких систем є телекомунікаційні підсистеми (ТКПС).

Згідно зі встановленими вимогами за автоматичного режиму функціонування ТКПС систем комплексної безпеки час проходження сигналів сповіщення та екстреної інформації на муніципальному та об'єктовому рівнях має становити не більш як 8 с, а коефіцієнт надійності має бути не менш як 0,995 [1]. Порушення встановлених нормативів може призвести до непоправних наслідків. Для вжиття необхідних заходів щодо підвищення якості обслуговування виникає потреба в розв'язанні проблеми оперативного та вірогідного визначення аномальних станів трафіку. Пропонується розглянути низку підходів для вирішення цієї проблеми. Як вихідні дані розглядатимемо потоки подій, у ролі яких виступають моменти початку передавання повідомлень у ТКПС. Сукупність моментів початку передавання повідомлень являє собою часовий ряд, характеристики якого відображають стан трафіку. Пропонується на основі розробленої імітаційної моделі дослідити можливості основних методів аналізу часових рядів для розв'язання задачі виявлення аномалій трафіку.

Формулювання завдання дослідження. Часовий ряд істотно відрізняється від простої вибірки даних, оскільки під час аналізу беруть до уваги взаємозв'язок вимірів із часом, а не тільки статистичну різноманітність і статистичні характеристики вибірки [2]. Часові ряди відображають як внутрішню динаміку об'єкта або процесу, так і їхні взаємні зв'язки та мінливість цих зв'язків у часі. За аномалій трафіку виявляється різка зміна властивостей спостережуваного ряду, що порушує стаціонарність досліджуваного процесу. Виникає потреба в обґрунтуванні вибору методу аналізу часових рядів, що дає змогу оперативно й вірогідно виявляти аномальні стани трафіку. Для аналізу часових рядів, які є випадковими процесами, використовують традиційні методи статистичного аналізу випадкових величин і функцій. Найпоширенішими з них є кореляційний і спектральний аналіз, згладжування та фільтрація даних, моделі авторегресії та прогнозування. Однак ці підходи найчастіше виявляються малоефективними під час аналізу нестаціонарних часових рядів.

Основна частина

Огляд наукових досліджень і дисертаційних праць останнього часу показує, що поряд із класичними підходами стали широко застосовуватися методи фрактального аналізу та теорія динамічних систем.

Пропонується дослідити можливості перспективних методів аналізу для розв'язання задачі виявлення аномалій трафіку. Варіант класифікації методів аналізу часових рядів наведено на рис. 1.

Особливістю теорії фракталів є те, що вона фактично протилежна традиційній алгебраїчній теорії, оскільки розглядає об'єкти з погляду неправильності їхніх форм. Аналіз фрактальних властивостей

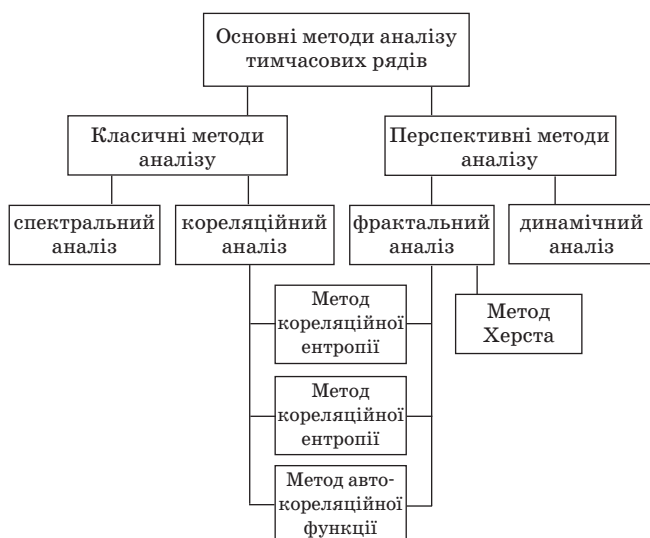


Рис. 1. Класифікація основних методів аналізу часових рядів

ношенням ділення розмаху на стандартне відхилення спостережень. Величина показника Херста характеризує відношення сили тренда (детермінованого чинника) до рівня шуму (випадковий чинник) [3]:

$$\text{звідки } H = \frac{\ln(R/S)}{\ln \alpha N}$$

де R — розмах нагромаджених відхилень ряду спостережень; S — стандартне відхилення ряду спостережень,

$$R/S = (aN)^H,$$

$$S = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - X_{\text{сеп}})^2};$$

α — задана константа, додатне число (автор показника Херст емпірично розрахував цю константу для порівняно короткострокових часових рядів природних явищ як 0,5); N — кількість періодів спостережень; x_i — елемент досліджуваного часового ряду; $X_{\text{сеп}}$ — середнє арифметичне часового ряду спостережень за N періодів,

$$X_{\text{сеп}} = \frac{1}{N} \sum_{i=1}^N x_i.$$

Показник Херста зростає зі зменшенням затримки між двома ідентичними парами значень у часовому ряду. Зазначимо основні властивості показника Херста H :

- якщо $H > 0,5$, то часові ряди вважаються постійними, тобто зберігають поточну тенденцію (якщо в минулому часовий ряд зростав, то й у майбутньому він, найімовірніше, зростатиме і навпаки);
- якщо $H = 0,5$, то часові ряди вважаються випадковими;
- якщо $H < 0,5$, то часові ряди вважаються непостійними, тобто поточна тенденція прагне змінитися на протилежну.

2. Кореляційна розмірність D_c характеризує складність процесу. Математично D_c вказує на мінімальну кількість динамічних змінних, які описують модель даної динамічної системи. Кореляційна розмірність використовується для перевірки на наявність випадкової складової в часовому ряду (мережному трафіку). Крім того, вона може виступати як додатковий параметр під час реалізації інших методів. Стрибокподібні зміни кореляції можуть бути пояснені вибуховим впливом зовнішніх або ж внутрішніх сил на часовий ряд.

Алгоритм розрахунку D_c ґрунтується на обчисленні кореляційного інтеграла, у ролі якого виступає функція

$$C(\delta) = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=i+1}^{n-1} \theta(\delta - |y_i - y_j|),$$

де n — кількість точок, що використовується для визначення розмірності; δ — відстань між точками об'єкта; θ — функція Хевісайда; $|y_i - y_j|$ — відстань між образами i -ї та j -ї точок часового ряду у відновленому просторі. Функція $C(\delta)$ для кожного δ дорівнює нормованій кількості пар точок розглянутого об'єкта, відстань між якими не перевищує δ .

Як ознаку розпізнавання стану трафіку братимемо значення кореляційної розмірності за аналізовані інтервали часу.

часових рядів є перспективним методом аналізу даних. Це зумовлено здатністю фрактального аналізу досліджувати сигнали, які щодо коваріаційної та спектральної теорії є не більш ніж білим шумом. У теорії фрактального аналізу широко використовують метод нормованого розмаху Херста. Застосовують також комбіновані підходи, засновані на методах кореляційного і фрактального аналізу:

- метод кореляційної ентропії;
- метод кореляційної розмірності;
- метод автокореляційної функції.

Розглянемо особливості цих методів і використувані в них характеристики.

1. Показник Херста розраховується з використанням методу нормованого розмаху, або R/S -аналізу, що дає змогу обчислити показник Херста як міру хаотичності ряду, що є безрозмірним відношенням розмаху на стандартне відхилення спостережень.

3. Кореляційна ентропія — інформаційна розмірність, що характеризує міру втрати інформації про систему в часі, кількісну міру хаотичності динамічної системи і того, наскільки швидко ця система перестає бути передбачуваною. Для обчислення кореляційної ентропії розроблено досить ефективний метод, заснований на обчисленні кореляційних сум

$$H_2 = \frac{C(K, m)}{C(K, m+1)},$$

де $C(K, m)$ — кореляційний інтеграл; K — відстань, у межах якої визначається відносна кількість пар точок атрактора; m — розмірність фазового простору.

Цей коефіцієнт показує ступінь хаотичності досліджуваної системи і період часу, на який можна спрогнозувати поведінку системи. Як ознака розпізнавання стану трафіку виступає значення кореляційної ентропії за аналізовані інтервали часу.

4. Автокореляційна функція (АКФ) показує ступінь кореляції між різними інтервалами часового ряду. Тут за ознаку розпізнавання стану мережі братимемо набір значень АКФ за аналізований інтервал часу:

$$r(k) = \frac{\sum_{i=1}^N (X_i - \bar{X})(X_{i+k} - \bar{X})}{(N - \tau)\sigma^2(X)},$$

де $r(k)$ — коефіцієнт автокореляції; X_i — рівні часового ряду; \bar{X} — середнє значення процесу; σ^2 — дисперсія.

У межах теорії динамічних систем припускають, що часовий ряд описує поведінку досліджуваної динамічної системи. Доведено, що наявності такого єдиного часового ряду достатньо для адекватного опису всієї динамічної системи загалом. Це обґрунтовано відомою теоремою Такенса про відновлення атрактора за єдиною динамічною траєкторією [4].

Для проведення досліджень властивостей і характеристик телекомунікаційних систем на основі аналізу трафіку розроблено імітаційну модель, що дає змогу формувати елементи часового ряду, у ролі яких виступають моменти подій. Під час розроблення імітаційної моделі враховували, що моменти початку передавання повідомлень у досліджуваній телекомунікаційній системі є стаціонарним випадковим процесом і підкоряються експоненціальному закону розподілу $f(x) = \lambda e^{-\lambda x}$ з параметром $\lambda > 0$ при $x \geq 0$ [5; 6]. Під час моделювання використовували формулу обчислення випадкових чисел з експоненціальним розподілом виду $x_n = -1/\lambda \ln r_n$, де r_n — число, випадково розподілене від 0 до 1 [5].

У цій моделі передбачено зміну параметра λ за заданим правилом, у результаті чого імітуються не-стаціонарні потоки.

Отже, розроблена імітаційна модель дає змогу формувати як стаціонарні, так і нестаціонарні потоки із заданим законом зміни інтенсивності. Результатом моделювання є потік повідомлень абонентів телекомунікаційної системи, що є часовим рядом, зображеним на рис. 2.

Дослідження здійснювалися з використанням таких методів аналізу часових рядів: кореляційної розмірності, кореляційної ентропії, автокореляційної функції та розрахунку показника Херста [6]. Для проведення експериментів було сформовано два варіанти потоків моментів подій у магістральних каналах телекомунікаційних систем. На рис. 3 суцільною лінією подано магістральний канал з квазіпостійною інтенсивністю, а пунктирною лінією — магістральний канал з аномальною зміною інтенсивності на заданій ділянці. Метою експерименту є виявлення можливостей визначення короткочасної зміни інтенсивності у вихідному потоці з використанням основних методів аналізу часових рядів.

Результати експериментів, здобуті за допомогою методу розрахунку показника Херста, унаочнює рис. 4. Розрахунки показують, що показник Херста H магістрального каналу з квазіпостійною інтенсивністю дорівнює 0,4, що відповідає стохастичному ряду, а з аномальною зміною інтенсивності ($H = 1,098$), що відповідає хаотичній системі. Дослідження показали, що момент зміни інтенсивності не фіксується. У ролі ознаки розпізнавання стану мережі виступає значення показника Херста за аналізовані інтервали часу.



Рис. 2. Приклад результату моделювання трафіку

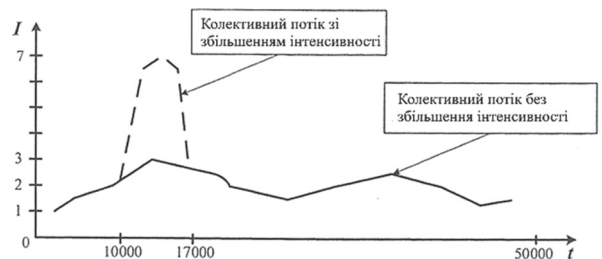


Рис. 3. Графік залежності інтенсивності від часу

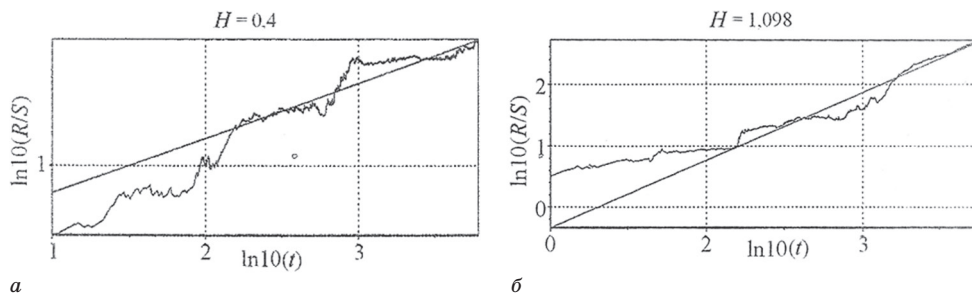


Рис. 4. Результати експериментів здобуто за допомогою методу розрахунку показника Херста для магістрального каналу з квазіпостійною інтенсивністю (а) і аномальною зміною інтенсивності (б)

Результати експериментів, здобуті за допомогою методу розрахунку кореляційної розмірності, зображено на рис. 5. Розрахована кореляційна розмірність для магістрального каналу з квазіпостійною інтенсивністю дорівнює 7,186, а з аномальною зміною інтенсивності — 5,611, що свідчить про те, що ступінь випадковості в першому випадку вищий. Коефіцієнт показує, що в системі відбулися зміни, але без фіксації інтервалу зміни інтенсивності. Як ознака розпізнавання стану мережі виступає значення кореляційної розмірності за аналізовані інтервали часу.

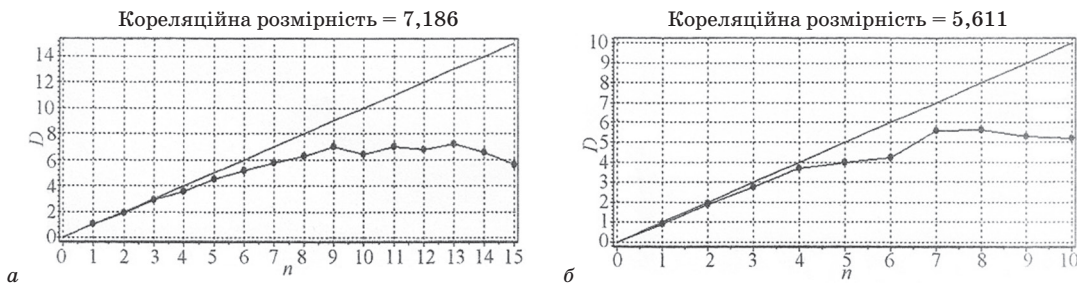


Рис. 5. Результати експериментів, отримані за допомогою методу розрахунку кореляційної розмірності для магістрального каналу з квазіпостійною інтенсивністю (а) і аномальною зміною інтенсивності (б)

Результати експериментів, здобутих за допомогою методу розрахунку кореляційної ентропії, наведено на рис. 6. Розрахунки показали, що кореляційна ентропія для магістрального каналу з квазіпостійною інтенсивністю дорівнює 0,443, а з аномальною зміною інтенсивності — 0,557. Отже, у другому випадку прогноз поведінки можливий на більш тривалий інтервал часу, ніж у першому. У ролі ознаки розпізнавання стану мережі виступає значення кореляційної ентропії за аналізовані інтервали часу.

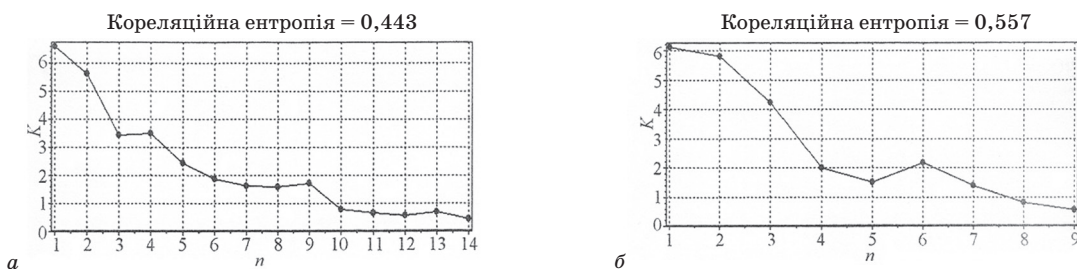


Рис. 6. Результати експериментів, отримані за допомогою методу розрахунку кореляційної ентропії для магістрального каналу з квазіпостійною інтенсивністю (а) і аномальною зміною інтенсивності (б)

Результати експериментів, здобуті за допомогою методу розрахунку автокореляційної функції, подано на рис. 7. Форма АКФ не містить інформації про зміну інтенсивності у вихідному потоці, але дає змогу визначити необхідний параметр затримки Δt (за першим нулем АКФ), який може використовуватися під час реалізації інших методів. У ролі ознаки розпізнавання стану мережі виступає набір значень АКФ за аналізований інтервал часу.

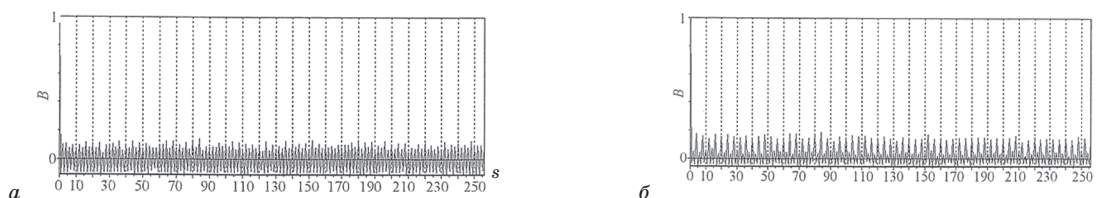


Рис. 7. Результати експериментів, отримані за допомогою методу розрахунку автокореляційної функції для магістрального каналу з квазіпостійною інтенсивністю (а) і аномальною зміною інтенсивності (б)

Таким чином, результати проведених досліджень класичних підходів до аналізу мережного трафіку із заданою зміною інтенсивності свідчать, що здобуті дані підтверджують лише ступінь складності мережного трафіку, і тільки в окремих випадках вони дають змогу спрогнозувати подальшу поведінку системи. Однак жоден із розглянутих методів не дає змоги виявити інтервали зміни інтенсивності потоку подій, що виникають, у телекомунікаційній системі.

Пропонується підхід, за якого досліджуваний трафік відображає розглянуту телекомунікаційну систему як динамічну систему (X, A) , де $X\{n\}$ — множина моментів запитів у мережному трафіку, які виступатимуть як вихідні дані для формування динамічних змінних, а A — відображення фазових траєкторій у двовимірному фазовому просторі, причому $A: X \rightarrow X$. У ролі множини $X\{n\}$ виступають запити на передавання повідомлень (див. рис. 1) : $X\{n\} = \{x(t_1), x(t_2), \dots, x(t_i), \dots, x(tn)\}$. Для формування динамічних змінних використовують потік запитів, перетворений на інтервальний часовий ряд вигляду $N(\tau_i) = \sum_{t \in \tau_i} x(t_i)$, де τ_i — інтервал, в якому здійснюється підрахунок моментів запитів; N_i — кількість моментів запитів в інтервалі τ_i , $N_i = N(\tau_i)$.

Варто зазначити, що базовою теоремою теорії динамічних систем є теорема Такенса, яка стверджує, що фазові траєкторії динамічної системи можна відновити за скалярним часовим рядом, якщо як динамічні змінні, яких бракує, використовують той самий ряд, узятий із певним запізненням [7]. Це слугує підставою вибору як динамічних змінних диференціальних характеристик потоку запитів.

Отже, пропонується сформувати пари динамічних змінних на основі різницевих рівнянь першого і другого порядків для відображення станів системи на площині [8].

Перший варіант формування динамічних змінних являє собою їх подання у вигляді різницевого рівняння першого порядку $\Delta_1 y(k) = N_{i+1} - N_i$, де $k = 1, 2, \dots, (n - 1)$; N_i — кількість запитів в інтервалі часу τ_i .

Другий варіант передбачає обчислення різниці $\Delta_1 y^*(k)$ кількості моментів запитів між віддаленими інтервалами зі зміщенням l : $\Delta_1 y^*(k) = N_{i+1} - N_i$.

У третьому варіанті обчислюється різниця $\Delta_2 y(k)$ від різниці значень сусідніх інтервалів $\Delta_1 y(k)$: $\Delta_2 y(k) = \Delta_1 y(k+1) - \Delta_1 y(k)$.

Четвертий варіант, що передбачає формування динамічних змінних, аналогічний до третього, за винятком того, що замість різниці $\Delta_1 y(k)$ буде використовуватися різниця $\Delta_1 y^*(k)$: $\Delta_2 y^*(k) = \Delta_1 y^*(k+1) - \Delta_1 y^*(k)$. З усього переліку динамічних змінних запропоновано таку пару: $\Delta_1 y(k)$ і $\Delta_1 y^*(k)$ [9].

Під час застосування запропонованих змінних для оцінювання інтенсивності було здобуто результати, зображені на рис. 8. Результати експерименту засвідчили, що розроблений підхід дає змогу виявляти зміни системи за допомогою введених динамічних змінних.

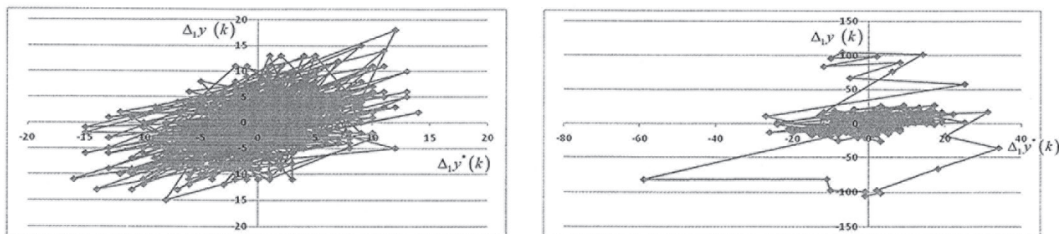


Рис. 8. Результати експериментів, отримані за допомогою методу розрахунку фазових портретів для магістрального каналу з квазіпостійною інтенсивністю (а) і аномальною зміною інтенсивності (б)

Висновки

Таким чином, результати проведених досліджень основних методів аналізу часових рядів показали, що вони не дають змоги з потрібною оперативністю і вірогідністю виявляти можливі аномальні стани трафіку. Показано, що запропонований підхід, де телекомунікаційна система розглядається як динамічна система, дає змогу відстежити ці аномалії завдяки відображенню динамічних змінних у двовимірному фазовому просторі, а отже, підвищити якість обслуговування в ТКПС.

Список використаної літератури

1. Taqqu M., Willinger W., Sherman R. Proof of a fundamental result in self-similar traffic modeling // *Computer Communication Review*. 2007. Vol. 27, no. 2. P. 5–23.
2. On the self-similar nature of Ethernet traffic / W. Leland, M. Taqqu, W. Willinger, D. Wilson // *IEEE/ACM Transactions on Networking*, February, 2013. 2(1). P. 1–15.
3. Paxson V., Floyd S. Wide-area traffic: The failure of Poisson modeling // *IEEE/ACM Transactions on Networking*. 2015. № 3. P. 226–244.
4. Feldmann A., Gilbert A.C., Willinger W. Data Networks as Cascades: Investigating the multifractal nature of Internet WAN traffic: Proc. 1998 ACM SIGCOM. P. 42–55. Che H., Li S.-Q., Fast algorithms for

measurement-based traffic modeling // IEEE Journal on Selected Areas in Communications. June 2018. Vol. 16. P. 612–625.

5. **Mondragon R. J., Pitts J. M.** *Chaotic Maps for Traffic Modelling and Queueing Performance Analysis // Performance Evaluation. 2001. Vol. 43, No. 4. P. 223–240.*

6. **Erramilli A., Sing R. P., Pruthi P.** *Modeling packet traffic with chaotic maps // Royal Institute of Technology, ISRN KTH/IT/R-95/18-SE, Stockholm-Kista Sweden, August, 2014.*

7. **Erramilli A., Singh R., Pruthi P.** *Chaotic maps as models of packet traffic // ITC 14, The Fundamental Role of Teletraffic in the Evolution of Telecommunications Networks (J. Labetoulle and J. Roberts, eds.), June 6-10 2019.*

O. G. Varfolomeeva, A. G. Zakharzhevskiy

TRAFFIC MODELING AND ANALYSIS IN CHANNELS OF TELECOMMUNICATION SUBSYSTEMS OF COMPLEX SECURITY SYSTEMS

In the article, necessity of monitoring telecommunication subsystems of complex security systems is substantiated. A simulation model is proposed, which makes it possible to form elements of a time series, in the role of which the moments of the beginning of the transmission of messages act. The possibility of forming both stationary and non-stationary flows of events is foreseen. The results of research on the main methods of time series analysis are presented. It is shown that the used approaches do not allow detecting possible anomalous traffic states with the necessary efficiency and reliability. It is proposed to consider the telecommunication system as a dynamic system for the evaluation of different traffic states. Analytical expressions for the formation of dynamic variables, on the basis of which phase portraits are formed, are presented. The obtained results of the experiments confirmed the informativeness of the proposed variant of traffic display. It is shown that the assessment of the state of telecommunication systems using phase portraits makes it possible to detect abnormal traffic conditions, as well as to predict the behavior of a dynamic system in the future.

Keywords: complex security system, telecommunication subsystem, simulation model, time series, time series analysis methods, dynamic system, dynamic variables, phase portraits.

