Security of Infocommunication Systems and Internet of Things

2023 Vol 1, No 2

https://doi.org/10.31861/sisiot2023.2.02001

Received 07 October 2023; revised 25 December 2023; accepted 27 December 2023; published 30 December 2023

Flows of Information within a Network with Limitations on the Quantity of Flows Allowed at Each Node

Ruslan Politanskyi[*] and Serhii Haliuk

Department of Radioengineering and Information Security, Yuriy Fedkovych Chernivtsi National University, Chernivtsi, Ukraine

*Corresponding author (E-mail: r.politanskyi@cnnu.edu.ua)

ABSTRACT This paper explores a classic problem in transport network research: the analysis of networks with specified total traffic values for each node. We employ linear algebraic methods to derive a comprehensive set of solutions, ensuring statistical reliability and enabling robust analysis of the results. A mathematical model is presented for determining solution sets in fully connected, loop-free networks with three and four nodes. Based on this model, we developed software to calculate the statistical distribution of entropy values within the network. Furthermore, we investigate the statistical properties of information flow entropy for networks with and without constraints that permit uniform flow distribution. This characteristic holds practical significance for analyzing network dynamics and predicting flow redistribution processes from initial unbalanced states, which inherently proceed towards higher entropy. The findings presented in this paper hold additional practical implications for network imbalance detection and adaptability to diverse network topologies. The results can serve as a foundation for algorithms designed to quantify the degree of network imbalance induced by substantial external influences that do not significantly alter the overall network load. This capability proves particularly valuable in identifying covert DDoS (Distributed Denial of Service) attacks that aim to reduce network bandwidth by supplanting legitimate traffic. While the proposed method has been demonstrated on fully connected networks, it demonstrates potential for adaptation to networks with a wide range of topological structures. This includes networks with partial connectivity or loosely connected networks, which constitute a significant proportion of real-world networks. The significance of the method is further amplified by advancements in cloud computing technologies, which offer substantial computational power and enable the accumulation of extensive statistics regarding information flow distributions across networks of diverse purposes. Such advancements create opportunities for the integration of the developed network analysis technique with machine learning and artificial intelligence technologies, fostering enhanced automation, scalability, and adaptability.

KEYWORDS network of information flows, Gauss method, system of solutions of linear equations.

I. INTRODUCTION

Network science unlocks powerful tools to understand real-world systems brimming with interconnected nodes and dynamic flows. The intricate dance between network structure and information exchange determines both resilience and efficiency, shaping the function of systems like the Internet, transportation networks, and even the brain [1-5].

Qquantitative analysis of network states is very often reduced to problems related to two-dimensional matrices, in which there are restrictions on the values of these flows. Over the past few decades, numerous studies have explored network analysis using sophisticated mathematical tools from topology and tensor analysis [6, 7]. These efforts have yielded diverse theoretical frameworks and approximation methods for characterizing network behavior, which are successfully applied to networks of various origins, and which can be combined by specifying the term "transport network".

The purpose of the work is to develop an algorithm for determining network entropy based on a system of integer solutions that satisfy the condition of equality of input and output flows in each node of a given set of integer values. Entropy is an important indicator of the state of equilibrium of natural systems, and it has values for determining the direction of processes in such systems. The concept of entropy and its application to structural analysis was transferred to artificial systems that arise in economic and financial problems [8]. The importance and performance of algorithms using entropy has led to fundamental research related to the general structure of systems and their dynamics, where systems are represented by complex entities consisting of nodes and connections between them [9, 10].

Next, we will consider that each node of the network is an equal participant and can be both a source and a receiver of the information flow. In this case, the matrix of information flows has the form of a square matrix of dimensions $n \times n$.

II. FORMULATION OF THE PROBLEM

The rows of the matrix assigned to the values of the incoming traffic at each node, and the columns of the values of the outgoing traffic for each node of the system. Another simplification of the problem is the absence of loop flows (if the node simultaneously receives and generates an information flow). This means that the diagonal elements of the matrix of information flows are equal to zero. Then for the matrix of information flows of dimensions $n \times n$:

$$A_n = \begin{bmatrix} 0 & f_{1,2} & & f_{1,n-1} & f_{1,n} \\ f_{2,1} & 0 & \cdots & f_{2,n-1} & f_{2,n} \\ & \vdots & \ddots & \vdots & \\ f_{n-1,1} & f_{n-1,2} & \cdots & 0 & f_{n-1,n} \\ f_{n,1} & f_{n,2} & & f_{n,n-1} & 0 \end{bmatrix}, \quad (1)$$

where $f_{i,j}$ – information flow between nodes i and j.

The system of linear equations connected information flows of sources and destination nodes will have the following form:

$$\begin{cases} 0 + f_{1,2} \cdots f_{1,k} = \sum_{j=1}^{k} f_{1,j} = s_1 \\ f_{2,1} + 0 \cdots f_{2,k} = \sum_{j=1}^{k} f_{2,j} = s_2 \\ \vdots \\ f_{k-1,1} + f_{k-1,2} \cdots f_{k-1,k} = \sum_{j=1}^{k} f_{k-1,j} = s_{k-1} \\ f_{k,1} + f_{k,2} \cdots 0 = \sum_{j=1}^{p} f_{k,j} = s_p \\ 0 + f_{2,1} \cdots f_{k,1} = \sum_{j=1}^{p} f_{j,1} = d_1 \\ f_{1,2} + 0 \cdots f_{k,2} = \sum_{j=1}^{p} f_{j,2} = d_2 \\ \vdots \\ f_{1,k-1} + f_{2,k-1} \cdots f_{k,1} = \sum_{j=1}^{p} f_{j,k-1} = d_{k-1} \\ f_{1,k} + f_{2,k} \cdots 0 = \sum_{j=1}^{p} f_{j,k} = d_k \end{cases} \quad (2)$$

The coefficient in (2) are restricted to the values -1, 0, and 1. Notably, this system has $n^2 - n$ unknown variables, of the $2 \cdot n - 1$ equations of the system, since exhibits linear dependence on the remaining equations, as mathematically expressed below:

$$\begin{cases} s_1 + \cdots + s_k = S \\ d_1 + \cdots + d_k = S \end{cases} \quad (3)$$

Therefore, there are only $m = k^2 - 3 \cdot k + 1$ independent variables among the solutions of (2). The value of S means the total load of the network.

III. MATHEMATICAL MODELS FOR INFORMATION FLOWS AND CALCULATION OF ENTROPY

To lay a foundation for modelling more intricate networks, we commence by examining a network comprising three nodes ($k = 3$). This network possesses six variables ($k(k-1) = 6$), representing the potential flows between each pair of nodes. Thus, the flow matrix of the system with aforementioned constrictions has the following form:

$$A_3 = \begin{bmatrix} 0 & f_{1,2} & f_{1,3} \\ f_{2,1} & 0 & f_{2,3} \\ f_{3,1} & f_{3,2} & 0 \end{bmatrix} \quad (4)$$

The system of linear equation for three nodes is next:

$$\begin{cases} f_{1,2} + f_{1,3} = s_1 \\ f_{2,3} + f_{2,3} = s_2 \\ f_{3,1} + f_{3,2} = s_3 \\ f_{2,1} + f_{3,1} = d_1 \\ f_{1,2} + f_{3,2} = d_2 \\ f_{1,3} + f_{2,3} = d_3 \end{cases} \quad (5)$$

The system (5) was solved by algorithm described in detail in [11]:

$$X = \begin{pmatrix} d_2 \\ s_1 - d_2 \\ d_4 \\ d_1 - s_3 \\ d_2 + d_3 - s_1 \\ 0 \end{pmatrix} + \begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \times \lambda, \quad (6)$$

where the only independent parameter λ must satisfy the constraints:

$$\begin{cases} \lambda_{min} = max\{d_2 - s_1, s_3 - d_1, 0\} \\ \lambda_{max} = min\{d_1, d_2 + d_3 - s_1, s_3\} \end{cases} \quad (7)$$

Next, we write down the matrix of information flows for a network without loop flows, which is formed by four nodes:

$$\|A_4\| = \begin{bmatrix} 0 & f_{1.2} & f_{1.3} & f_{1.4} \\ f_{2.1} & 0 & f_{2.3} & f_{2.4} \\ f_{3.1} & f_{3.2} & 0 & f_{3.4} \\ f_{4.1} & f_{4.2} & f_{4.3} & 0 \end{bmatrix} \quad (8)$$

Solutions of this task are as follows:

$$\begin{pmatrix} f_{1.2} \\ f_{1.3} \\ f_{1.4} \\ f_{2.1} \\ f_{2.3} \\ f_{2.4} \\ f_{3.1} \\ f_{3.2} \\ f_{3.4} \\ f_{4.1} \\ f_{4.2} \\ f_{4.3} \end{pmatrix} = \begin{pmatrix} d_2 - (s_3 + s_4) + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 \\ d_1 + d_3 - s_2 + \lambda_1 - \lambda_3 - \lambda_4 - \lambda_5 \\ d_4 - (\lambda_1 + \lambda_2) \\ d_1 - (\lambda_3 + \lambda_4) \\ s_2 - d_1 - \lambda_1 + \lambda_4 + \lambda_5 \\ \lambda_1 \\ \lambda_4 \\ s_3 - (\lambda_2 + \lambda_4) \\ \lambda_2 \\ \lambda_3 \\ s_4 - (\lambda_3 + \lambda_5) \\ \lambda_5 \end{pmatrix}, \quad (9)$$

where $\Lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5)$ is a vector of independent variables with the following constrains (10) and (11):

$$\begin{cases} 0 \le \lambda_5 \le d_3 \\ 0 \le \lambda_3 \le s_4 - \lambda_5 \\ 0 \le \lambda_4 \le d_1 - \lambda_3 \\ 0 \le \lambda_2 \le s_3 - \lambda_4 \\ 0 \le \lambda_1 \le d_4 - \lambda_2 \end{cases}, \quad (10)$$

$$\begin{cases} \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 \ge s_3 + s_4 - d_2 \\ \lambda_1 - \lambda_3 - \lambda_4 - \lambda_5 \ge s_2 - (d_1 + d_3) \\ -\lambda_1 + \lambda_3 + \lambda_4 \ge d_1 - s_2 \end{cases}. \quad (11)$$

Based on the solution of the system (9) with restrictions (10) and (11) the entropy of the network can be calculated. The entropy can be found by using the logarithmic measure of flows normalized by the total load of the network S [12]:

$$H(\|A_4\|) = -\sum_{i=1}^{k} \sum_{j=1}^{k} \frac{f_{i,j}}{S} \cdot log_2\left(\frac{f_{i,j}}{S}\right), \quad (12)$$

where S is a summary network flow.

One of the most important parameters which determines an equilibrium state is the value of the entropy. If any system can freely change its internal state the entropy should come to the maximum value. We use the same analogy for telecommunications networks. At the same time, the network with the maximum entropy is the network with the maximum balanced load, which is distributed as far as possible so that all branches are loaded almost equally.

The problem of determining the entropy and finding its optimal value as well as configuration of the network can also be solved using an analytical method grounded in the complete set of a system's solutions. In doing so, we can exploit the fact that the network has some set of independent variables $\{\lambda_i\}$ that determine all the other flows $\|f_{ij}\|$. We give an analytical expression for the entropy of a three-node network, which is a function of only one independent λ:

$$H_3 = H(\lambda) = -\frac{s_1-d_2+\lambda}{S} log_2\left(\frac{s_1-d_2+\lambda}{S}\right) - \frac{d_1-s_3-\lambda}{S} log_2\left(\frac{d_1-s_3-\lambda}{S}\right) - \frac{d_2-\lambda}{S} log_2\left(\frac{d_2-\lambda}{S}\right) - \frac{d_4+\lambda}{S} log_2\left(\frac{d_4+\lambda}{S}\right) - \frac{d_2+d_3-s_1-\lambda}{S} log_2\left(\frac{d_2+d_3-s_1-\lambda}{S}\right) - \frac{\lambda}{S} log_2\left(\frac{\lambda}{S}\right) \quad (13)$$

This model enables the determination of the maximum entropy value by locating the stationary points of the function $H(\lambda)$. To address the constraint (7), we can leverage the following variable transformation [13]:

$$\lambda = t(\lambda_{max}, \lambda_{min}, \tilde{\lambda}) = \frac{\lambda_{max}+\lambda_{min}}{2} + \frac{\lambda_{max}-\lambda_{min}}{2} \cdot \frac{2\tilde{\lambda}}{1+\tilde{\lambda}^2} \quad (14)$$

Entropy can now be expressed in terms of a new variable using the composition of the function that defines it and the transformation function: $H_3 = \tilde{H}(\tilde{\lambda}) = H\left(t(\lambda_{max}, \lambda_{min}, \tilde{\lambda})\right)$. The advantage of this transformation is that the new parameter is unbounded and can be defined on the entire set of integers.

IV. OBTAINED RESULTS AND THEIR ANALYSIS

An obvious parameter of the problem is the total information flow in the network, which is traditionally measured in erlangs (3). But it is also obvious that the network can be in an unbalanced state in advance, if different requirements for the total incoming and outgoing traffic are set for each of the nodes. Firstly, let's analyze a uniformly distributed load of the network. This is a case, when all loads in every node are equals:

$$\begin{cases} s_i = S/k \\ d_i = S/k \end{cases} \quad (15)$$

Then, all flows are equal also:

$$f_{ij} = \frac{s_i}{k-1} = \frac{S}{k(k-1)} \quad (16)$$

In this case, the entropy has the maximum value:

$$H_{k,max} = log_2(k^2 - k) \quad (17)$$

The entropy of a system with three equivalent nodes and summary traffic $S = 300$ is shown in Fig. 1.

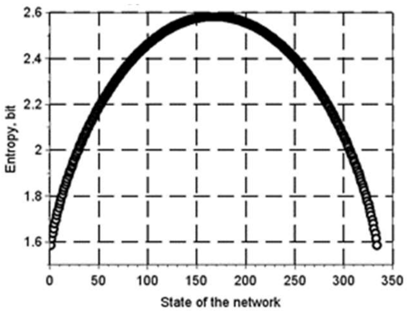

FIG. 1. Entropy of the balanced network with $n = 3$ and $S = 1000$.

In this scenario, entropy exhibits characteristics similar to that of a system comprising two states forming an event group.

If the network becomes less balanced, then it is obvious that its maximum entropy decreases to the smallest possible value of 0 (not shown on Fig. 1).

The dependences of the number of states (different solutions) and entropy for different states of the network are shown in Fig. 2. For the entropy graph if the specified load conditions in the network allow multiple solutions, then the largest entropy value is indicated.

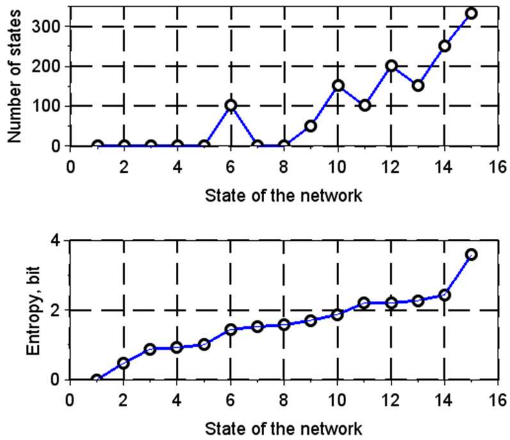

FIG. 2. Entropy and number of states of the unbalanced network with $n = 3$ and $S = 1000$.

Among all the states of the network, the state with the maximum entropy value, when the flows are evenly distributed among all nodes, is the most. In this particular case, the entropy according to formula (17) is equal to $log_2(6) \approx 2.6$.

Based on the obtained results, the entire set of possible states can be obtained (microstructural analysis of the entire network). This makes it possible to analyse the network from the point of view of uniform loading of all its branches, to exclude unjustified underloading of equipment in one part of it, and, conversely, excessive load on other nodes and devices of the system. Another possible advantage is that the determination of the microstate of the network and its deviation from the equilibrium state, which is characterized by the maximum entropy, provides the possibility of dynamic analysis, which shows the possible direction of changes in the flows in the network to a state of more uniform load. Determining microstates of the network, which is in a state close to equilibrium, provides an opportunity to identify problem areas in the network, where there is an unpredictable decrease or, conversely, an increase in the value of the information flow. Such problem areas in the network can be related to the points of occurrence of cyber threats in real telecommunication systems.

V. CONCLUSION

The work developed an algorithm for determining the set of microstates of a network that has a given load of incoming and outgoing traffic in each of its nodes. Unlike previous works, the algorithm determines microstates of networks where there are no loop flows, which brings it closer to the possibility of real application. Based on the developed algorithm, software for its implementation was created, and results were obtained for networks with three and four nodes. Another important result is obtaining analytical expressions for the entropy of the entire network, which determines its equilibrium state. It is shown that for a network of three nodes, the entropy resembles the classical entropy model obtained for Markov sources with two states, for which the dynamics of the system is determined by the probability of transition from one state to another. Based on this, the possibility of implementing algorithms based on entropy for the analysis, synthesis and control of telecommunication systems was analysed.

AUTHOR CONTRIBUTIONS

R.P. – conceptualization, methodology, writing-original draft preparation writing-review and editing, methodology; S.H. – design of results, work with reviewers.

COMPETING INTERESTS

The authors declare no conflict of interest.

REFERENCES

[1] R. Ahlswede, N. Cai, S.Y. Li, R.W. Yeung,. "Network information flow. IEEE Transactions on information theory", vol. 46(4), pp. 1204-1216, 2000.

[2] A. S. Avestimehr, S. N. Diggavi, N.C. David, "Wireless network information flow: A deterministic approach", IEEE Transactions on Information theory, vol. 57(4), pp. 1872-1905, 2011.

[3] L. R. Ford Jr, D.R. Fulkerson. Flows in networks, vol. 54, Princeton university press, 2015.

[4] G. Bertagnolli, R. Gallotti, M. De Domenico, "Quantifying efficient information exchange in real network flows", Commun Phys 4, 125, 2021.

[5] K. Ponanan, T. Srichanthamit, W. Watanabe, S. Watanabe, H. Suto, Hidetsugu, "A Framework of Supporting System for Optimizing Information Flow in International Trade Transaction", Transactions of Japan Society of Kansei Engineering, vol. 18, 2018.

[6] O. Lemeshko, M. Yevdokymenko, O. Yeremenko. "Optimization routing model of delay-sensitive traffic in infocommunication networks", Control, Navigation and Communication Systems. Academic Journal, vol. 2, № 72, pp. 152-159, 2020.

[7] O.V. Lemeshko, T.V. Vavenko "Usovershenstvovanije potokovoj modeli mnogoputevoj marshrutizatsii na osnove balansirivki nagruzki", Problemi telekomunikatsij. – № 1(6), pp. 12-29, 2012.

[8] Drzazga-Szczęśniak, E.A.; Szczepanik, P.; Kaczmarek, A.Z.; Szczęśniak, D. "Entropy of Financial Time Series Due to the Shock of War", Entropy, vol. 25(5), pp. 823-835, 2023.

[9] Xi, Y.; Cui, X. "Identifying Influential Nodes in Complex Networks Based on Information Entropy and Relationship Strength", Entropy, vol. 25(5), pp. 754-771, 2023.

[10] Bayrakdar N, Gemmetto V, Garlaschelli D. "Local Phase Transitions in a Model of Multiplex Networks with Heterogeneous Degrees and Inter-Layer Coupling", Entropy, vol. 25(5), pp. 828-863, 2023.

[11] R. L. Politanskyi, O.L. Zarytska, M.V. Vistak, V.V. Vlasenko, "Research of distribution of information flows in a network", Mathematical Modeling and Computing, vol. 8(4), pp. 821-829, 2021.

[12] R. Politanskyi, A. Samila, L. Politanskyi, V. Vlasenko, V Popa, Y. Bobalo, I. Tchaikovsky, "Investigation of High-Speed Methods for Determining the Equilibrium State of a Network Based on the Principle of Maximum Entropy". Lecture Notes in Electrical Engineering, vol. 96, pp. 602-614, 2023.

[13] Mykel J. Kochenderfer, Tim A. Wheeler. Algorithms for Optimization. The MIT Press. Cambridge, Massachusetts. London, England, 2019.

Ruslan Politanskyi

Received a Ph.D. in solid state physics from Yuriy Fedkovych Chernivtsi National University. Received a Dr. of Eng. Science in Telecommunication from National University «Lviv Polytechnic». He is currently a professor at Yuriy Fedkovych Chernivtsi National University. His research interests include signal processing, coding theory, pseudorandom sequence systems with chaotic dynamics, physical principles of processes for information coding, distribution and transferring.

ORCID ID: 0000-0003-0015-7123

Serhii Haliuk

Assistant professor at Radio Engineering and Information Security Department of Yuriy Fedkovych Chernivtsi National University. His research interest covers the development of the different components of hidden communication systems. Author of more than 20 publications.

ORCID ID: 0000-0003-3836-2675

Потоки інформації в мережі з обмеженнями кількості потоків, дозволених на кожному вузлі

Руслан Політанський*, Сергій Галюк

Кафедра радіотехніки та інформаційної безпеки, Чернівецький національний університет імені Юрія Федьковича, Чернівці, Україна

*Автор-кореспондент (Електронна адреса: r.politanskyi@cnnu.edu.ua)

АНОТАЦІЯ У роботі розглядається класична проблема аналізу транспортних мереж, зокрема, вивчається структура мереж з урахуванням передбачених обсягів трафіку для кожного вузла мережі. У роботі пропонується математична модель для визначення рішень у повністю зв'язаних мережах без петель, які складаються з трьох та чотирьох вузлів.

На основі цієї моделі розроблено програмне забезпечення для оцінки статистичного розподілу ентропії у мережі. Додатково проведено дослідження статистичних характеристик ентропії потоку інформації в мережах з обмеженнями та без них, що дозволяє рівномірно розподіляти потік інформації. Ця аналіз має практичне значення для вивчення динаміки мережі та передбачення процесів перерозподілу потоку з початкових незбалансованих станів, які спрямовані на збільшення ентропії. У статті приведені висновки, які можуть мати практичне застосування для виявлення дисбалансу в мережі та адаптації до різних топологій. Отримані результати можуть бути використані для розробки алгоритмів, які визначають ступінь дисбалансу мережі, що виникає внаслідок значних зовнішніх впливів, які незначно змінюють загальне навантаження на мережу. Це особливо актуально для виявлення прихованих атак DDoS, спрямованих на зниження пропускної здатності мережі через заміну легального трафіку. Запропонований метод було випробувано на повністю зв'язаних мережах, але його потенціал адаптації до різноманітних топологічних структур, таких як мережі з частковим підключенням або слабко підключені мережі, також допускається. Враховуючи прогрес у технологіях хмарних обчислень, які забезпечують значну обчислювальну потужність, і здатність накопичувати обширні статистичні дані, метод може бути інтегрований із технологіями машинного навчання та штучного інтелекту, що сприятиме автоматизації, масштабованості та адаптивності в аналізі мережі.

КЛЮЧОВІ СЛОВА інформаційні потоки в мережі, метод Гауса, система розв’язків лінійних рівнянь.