Network Temperature: A Novel Statistical Index for Networks Measurement and Management

CHANGDA WANG and XIAOWEI LI, Jiangsu University
ELISA BERTINO, Purdue University

Being able to monitor each packet path is critical for effective measurement and management of networks. However, such detailed monitoring can be very expensive especially for large-scale networks. To address such problem, inspired by thermodynamics, which uses the statistical characteristics of a large number of molecules' motion but not each molecule's trajectory for analysis, we propose the new concept of *network temperature* together with the notions of *network-specific heat* and *network temperature gradient*. Our approach does not only provide a statistical view of the current network state consisting of all the active packet paths at each time instant, but can be used to represent transitions among network states. Our network temperature-based methods have a broad applicability, such as to DDoS detection, dynamic node importance ranking, network stability and robustness evaluation, reliable packets routing, provenance compression assessment, and so on. Numerical and/or the experimental results show that our methods are effective.

CCS Concepts: • **Networks** → **Network performance evaluation**; *Network measurement*; **Network performance evaluation**;

Additional Key Words and Phrases: Network security, network measurement, network management, active packet paths, data provenance, network temperature

ACM Reference format:

Changda Wang, Xiaowei Li, and Elisa Bertino. 2022. Network Temperature: A Novel Statistical Index for Networks Measurement and Management. *ACM Trans. Internet Technol.* 22, 3, Article 66 (January 2022), 20 pages. <https://doi.org/10.1145/3511093>

66

1 INTRODUCTION

Given the topology of a packet-switched network, packet paths provide a means to understand what is or is not working properly through links troubleshooting, protocol debugging, workload characterization, and performance evaluation [1, 7, 18]—all these functions are critical for secure and reliable networks. Network management applications thus often require the ability to efficiently monitoring all the packet paths. However, in large-scale networks, the quadratic growth in the number of network paths with respect to the number of network nodes makes it impractical to record detailed information about each single packet path [12]. To address such an issue,

This work was supported in part by National Natural Science Foundation of China under Grants 62072217 and 61672269. Authors' addresses: C. Wang and X. Li, Jiangsu University, No. 301 Xuefu Road, Zhenjiang, Jiangsu, China, 212013; emails: {changda, xw82}@ujs.edu.cn; E. Bertino, Purdue University, 305 N. University Street, West Lafayette, IN 47907, USA; email: bertino@purdue.edu.

This work is licensed under a Creative Commons Attribution International 4.0 License.

© 2022 Copyright held by the owner/author(s).
1533-5399/2022/01-ART66
<https://doi.org/10.1145/3511093>

compressed packet path recording schemes, also known as provenance schemes, have been proposed [8, 14, 15]. However, lossless compression schemes can only mitigate the packet paths' size expansion but cannot always limit the size to an analyzable range especially when such analyses have to be carried out within certain time constraints [16].

Although the requirement of monitoring each packet path is not new, it is worth pointing out that most packet paths are not being scrutinized or even checked, unless security or performance anomalies are detected. As a result, coarse-grained packet paths schemes are commonly deployed. In the extreme case, when the network operation is defined as a sequence of network state transits, all packet paths active at a given time, which defines one of those network states are mapped onto a single value. Network entropy, which has various definitions such as Shannon entropy, Rényi entropy, Tsallis entropy, and Von Neumann entropy [6], is one such approach. Because those entropy-based approaches preserve the basic packet paths characteristics, many network anomalies can be efficiently detected through network entropy changes [3, 17, 20]. However, network entropy is a numerical value only, which provides little information to understand the patterns of those network state transits.

Inspired by thermodynamics that analyzes large numbers of molecules moving in a system through their statistical characteristics but not each molecule's trajectory, we propose the novel concept of *network temperature* for network measurement and management, in which packet paths are comparable to molecule trajectories. Compared to network entropy, network temperature not only provides a statistical view of a network state as composed of all the active packet paths at a given time, but it also can be used to model the patterns of those network states changes.

Clearly, network temperature is not the thermodynamic temperature of a network. We choose such a name because, mathematically, network temperature and thermodynamic temperature share the same mathematical formulations, though the parameters' physical meanings in these formulas are different. Furthermore, referring to thermodynamic temperature's characteristics, we also introduce the notions of *network-specific heat* and *network temperature gradient*. Network temperature, network-specific heat, and network temperature gradient together are a thermodynamics alike system for measuring the current state of a network as well for modeling the network state transition patterns.

Regarding networks measurement and management, by far, we observed that network temperature is suitable for measuring packets transmission pattern in a network. Higher network temperature generally means that most of nodes are busy generating/transmitting/forwarding/receiving packets; and/or Shannon's entropy of the packets transmission keeps increasing. As a result, lower network temperature is always preferred concerning the reliability of packets transmission. Furthermore, like in thermodynamics, network-specific heat represents the robustness of a network with respect to network temperature, i.e., lower network-specific heat indicates a small change in packets transmission pattern that will lead to an obvious network temperature change. Therefore, higher network-specific heat is always preferred concerning the robustness of packets transmission. Network temperature gradient describes in which direction and at which rate the network temperature changes the most rapidly at a particular location, which provides an effective way to decrease/increase the network temperature in a network. As the meaning of decreasing network temperature is clear, we argue that increasing network temperature may lead to a new approach of **DoS (denial-of-service)** attacks.

The contributions of the article are as follows:

- We propose the concept of network temperature together with the notions of network-specific heat and network temperature gradient.

- We propose and formally prove the properties of network temperature, network-specific heat, and network temperature gradient. We provide numerical results that confirm these properties.
- We discuss several applications of network temperature, network-specific heat, and network temperature gradient to network measurement and management. Many such applications are critical for network security and reliability. We have conducted experiments, on publicly available datasets, to validate such applications and show that our approach is effective for these applications.

The rest of the article is organized as follows: Section 2 provides preliminaries. Section 3 formally defines network temperature, network-specific heat, and network temperature gradient and discuss their properties. Section 4 introduce possible applications of our network temperature-based methods. Section 5 reports experimental results. Section 6 concludes the article and outlines future work.

2 PRELIMINARIES

As preliminary for our work, we give a brief overview of thermodynamic temperature. Furthermore, the relationships between thermodynamic entropy and information entropy are discussed.

2.1 Thermodynamic Temperature

Thermodynamic temperature is an objective comparative measurement of heat, which measures the kinetic energy of atoms and molecules. The higher the temperature, the faster the molecules of matter move. Although thermodynamic temperature is commonly measured by units such as Celsius (denoted °C), Fahrenheit (denoted °F), and Kelvin (denoted K), its definition is independent of those measurement units. The formal definition of thermodynamic temperature, Equation (1), is the derivative of the system's heat quantity with respect to the system's thermodynamic entropy.

$$T = \frac{dQ}{dS}, \quad (1)$$

where dQ is the incremental reversible transfer of heat energy into the system and dS is the corresponding thermodynamic entropy changes.

The definition of thermodynamic entropy S is given by Equation (2), which was established by Ludwig Boltzmann and J. Willard Gibbs in the 1870s, where p_i is the probability of the microstate i taken from an equilibrium ensemble, and k is Boltzmann's constant.

$$S = -k \sum_i p_i \ln p_i \quad (2)$$

There are many different ways to define a microstate in thermodynamics entropy. One of them, easy to understand, is in classical mechanics, where each microstate is defined as a theoretical *absolutely instantaneous photo* of the location and momentum of each molecule. The physical meaning of thermodynamic entropy is the system's chaotic measurement, i.e., the higher the S , the harder to predict those microstates' transit.

2.2 Thermodynamic Entropy and Information Entropy

The definition of information entropy is given by Equation (3) established by Claude E. Shannon in 1948, where p_i is the probability of the message i taken from a message ensemble and b is the base used for the logarithm. The unit of H is *bit* for $b = 2$ and *nat* for $b = e$, where e is Euler's number.

$$H = - \sum_i p_i \log_b p_i \quad (3)$$

Fig. 1. Maxwell's demon, where the arrows indicate the velocity of molecules motion. The longer the arrow, the higher the molecule's speed is.

Compared to the vague physical meaning of thermodynamic entropy, the physical meaning of information entropy is clear, viz. H is the amount of unknown information about a message consisting of a sequence of symbols. For instance, if we know that a symbol occurring with probability p_i appears in the message, then $-\log_2 p_i$ bits will be reduced from H . As a result, if there are no uncertain symbols in the message, then H will be reduced to zero.

Although thermodynamic entropy S and information entropy H are defined independently, formally the only difference between S and H is the constant k in Equation (3) when $b = e$.

The relationship between thermodynamic entropy and information entropy is one of the most fascinating topics in the multi-disciplinary study of physics and information theory. For instance, does an object, e.g., a person, acquiring information have a fundamental thermodynamic cost? An intuitive clue to such a question is a famous joke: Don't think too much, your head will be very hot. As to the serious answer for the question, the first effort can be traced back to the year 1867 when Maxwell's demon was proposed (see Figure 1), in which the demon can create a temperature difference within an insulated container divided into two parts by a piece of insulation board through the demon's recognition of the molecules' speeds. Such a demon violates the second law of thermodynamics, because the temperature difference created by the demon can be utilized by a heat engine to extract work without consuming any known energy at that time.

The paradox of Maxwell's demon was clarified in 1961 by Landauer [9], who argued that the acquisition of information through a measurement requires the dissipation of at least $kT \ln 2$ joule energy for each bit, where T is a temperature scaled in Kelvin and $k \approx 1.38 \times 10^{-23}$ J/K is Boltzmann's constant. In 2012, Brut et al. in the journal *Nature* reported the experimental verification of the Landauer's principle [4]. Brillouin even concluded that any cause for a bit value change, i.e., measurement, the decision about a yes/no question, erasure, display, and so on, will require the same amount of energy [5]. As a result, there is no way for Maxwell's demon to create a temperature difference between the two parts in Figure 1 without consuming any energy. Such a conclusion also connects the number of bits that a brain processes and the energy that the brain consumes.

All those research works confirm that any information change has a fundamental thermodynamic cost, which indicates that thermodynamics and information theory are interrelated. As a bold vision, what will we get if we substitute information entropy for thermodynamic entropy in the temperature's definition? Can a thermodynamics alike system be built to measure and manage network states transits? The article presents our work about that exploration and its application to functions relevant for the security and reliability of networks.

3 NETWORK TEMPERATURE DEFINITION

In this section, we define network temperature, network-specific heat, and network temperature gradient in the context of a packet-switched network (network, for short). We also report numerical results for the three proposed concepts.

3.1 Network Model

We model a network as a graph of nodes, where each node has a unique address or ID.

Definition 1 (Network Model). The network model is a directed graph $G(N, E)$, where $N = \{n_i | 1 \leq i \leq |N|\}$ is the set of nodes and $E = \{e_{ij} | 1 \leq i, j \leq |N|\}$ is the set of directed edges connecting nodes; $|N|$ and $|E|$ denote the cardinality of set N and set E , respectively; a directed edge e_{ij} , $i, j \in \mathbb{N}^+$, specifies that there is a one-hop connection from node n_i to node n_j .

According to Definition 1, the behavior of a network in a time interval of duration Δt can be represented by a matrix $\{a_{ij}\}_{N \times N}$, where a_{ij} represents the number of packets transmitted from n_i to n_j within Δt . $\{a_{ij}\}_{N \times N}$ is updated periodically or on demand to reflect both the network topology and network load changes. If $a_{ij} = \phi$, then the corresponding e_{ij} does not exist. Clearly, in such a network if every node can reach the other nodes through multi-hop packet transmissions, then $|N|$ and $|E|$ satisfy the relationship $|N| - 1 \leq |E| \leq |N| \cdot (|N| - 1)$.

The network operation is defined as a series of network states transits, where a network state at a given time consists of all the active packet paths at this time. Based on the network model of Definition 1, the snapshot of a network state is k packets being transmitted over $|E|$ directed edges, where a packet being transmitted over an edge e_{ij} means that the link from n_i to n_j is occupied by the transmission of the packet.

3.2 Network Entropy

We select Shannon entropy to define network entropy because all of Rényi entropy, Tsallis entropy, and Von Neumann entropy reduce to Shannon entropy under certain conditions [11], i.e., Shannon entropy is the most fundamental among all the known definitions of entropy.

Definition 2 (Network Entropy). The network entropy H for a network $G(N, E)$ is defined as follows:

$$H = - \sum_i p_i \log p_i, \quad (4)$$

where p_i denotes the occurrence probability of network state i taken from a state space $\mathcal{S}$ and the base of $\log p_i$ is 2.

Assume that there are k packets distributed over $|E|$ edges at the moment, where each edge is a section of an active packet path. The cardinality of the state space $\mathcal{S}$ is then formulated as follows:

$$|\mathcal{S}| = \binom{|E|}{k} \cdot k!. \quad (5)$$

If all the network states are equiprobable, then the network entropy defined by Equation (4) will reduce to Hartley's entropy [13]:

$$H = \log |\mathcal{S}|, \quad (6)$$

which is the largest value that the network entropy in Definition 2 can take.

3.3 Network Temperature

Network temperature is likewise defined by the mathematical expression of thermodynamic temperature (see Equation (1)), in which we substitute the network entropy for thermodynamic entropy and take the view that the number of packets distributed over directed edges, viz. the number of packets being transmitted, represents the counterpart for heat energy.

It is easy to understand that the network entropy is comparable to thermodynamic entropy in the definition. As to why take the view that the number of packets distributed over directed edges represents the counterpart for heat energy, the reason is that the functional relationship model for

the number of packets distributed over directed edges and the network entropy in Definition 2 is exactly the same as that of the heat energy and thermodynamic entropy.

Definition 3 (Network Temperature). The network temperature T for a network $G(N, E)$ is defined as follows:

$$T = \frac{\Delta P}{\Delta H}, \quad (7)$$

where ΔP is the incremental quantity of packets in transmission and ΔH is the corresponding network entropy change.

Note that it is meaningless to discuss network temperature with respect to a single packet transmission, because network temperature reflects the statistical characteristics of the paths of all the packets in transmission as an ensemble at a given time.

In Definition 3, when all packets are being cached in buffers or there is no packet in transmission, i.e., $\Delta P = 0$ and $\Delta H = 0$, the network temperature T is then defined as zero. The reason of such definition is based on thermodynamics, where temperature reflects the average kinetic energy of the molecules in a system. When the molecules lose all the motion freedom, the system's thermodynamic temperature is equal to zero.

Assuming k packets in transmission at a given time, the network temperature T of a network $G(N, E)$ can be approximately calculated by the following equation:

$$\frac{1}{T} = \frac{\Delta H}{\Delta P} \approx \frac{\partial H}{\partial k}. \quad (8)$$

To obtain the analytic expression for H , we apply Stirling's approximation, viz. $x! \approx x^x \cdot e^{-x} \cdot \sqrt{2\pi x}$, to Equation (6). H can then be expressed as:

$$H = \ln \left(\binom{|E|}{k} \cdot k! \right) \cdot \frac{1}{\ln 2} = \left(\ln \binom{|E|}{k} + \ln k! \right) \cdot \frac{1}{\ln 2}.$$

Therefore,

$$T \approx \frac{\partial k}{\partial H} \approx \ln 2 \cdot \left(\ln(|E| - k) + \frac{1}{2k} \right)^{-1}. \quad (9)$$

Note that the domain of Equation (9) has been extended by applying Stirling's approximation. Equation (9) only takes values when $k, |E| \in \mathbb{N}^+$ for network temperature calculations.

THEOREM 1 (THE MONOTONICITY OF NETWORK TEMPERATURE). *Let $G(N, E)$ be the network model of a given network; let T denote the network temperature. Then $T \geq 0$ when $|E| > k$. T is strictly monotonically increasing with the increase of k and is strictly monotonically decreasing with the increase of $|E|$.*

PROOF. Because $k, |E| \in \mathbb{N}^+$, when $|E| > k$, $|E| - k \geq 1$. Therefore, according to Equation (9), $T > 0$. Furthermore, in Definition 3 $T = 0$ when $k = 0$, $T \geq 0$ thus follows. Now, we use $T_d(|E|, k)$ to represent the denominator of Equation (9).

$$T_d(|E|, k + \Delta k) - T_d(|E|, k) = \ln \frac{|E| - (k + \Delta k)}{|E| - k} + \frac{1}{2(k + \Delta k)} - \frac{1}{2k}$$

Note that $k, |E| \in \mathbb{N}^+$,

$$\ln \frac{|E| - (k + \Delta k)}{|E| - k} < \ln 1 = 0, \quad \frac{1}{2(k + \Delta k)} - \frac{1}{2k} < 0.$$

Therefore, when $\Delta k > 0$, $T_d(|E|, k + \Delta k) - T_d(|E|, k) < 0$. Furthermore, when $\Delta|E| > 0$,

$$T_d(|E| + \Delta|E|, k) - T_d(|E|, k) = \ln \frac{|E| + \Delta|E| - k}{|E| - k} > \ln 1 = 0.$$

Thus, T is strictly monotonically increasing with the increase of k and monotonically decreasing with the increase of $|E|$. $\square$

COROLLARY 1 (THE MONOTONICITY OF NETWORK ENTROPY). *For a network $G(N, E)$, the network entropy H is strictly monotonically increasing with the increase of network temperature T .*

PROOF. According to Equation (5), H is strictly monotonically increasing with the increase of k , viz. the number of packets in transmission. Therefore, when H is increasing, k increases too. Furthermore, by Theorem 1, we know that the increase in k leads to the increase in T , H is then strictly monotonically increasing with the increase in T . $\square$

By Equation (9), network temperatures are calculated on the assumption that each one-hop link is used evenly, which results in the largest entropy H for the network, and thus the derived network temperature is the network temperature's upper bound. However, by applying such an equation to calculate the local network temperature, viz. the network temperature around a node within certain packet transmission hops, the deviation of the observed network temperature and the real network temperature will decrease with the decrease in the number of hops. As a result, the entire network temperature can be approximated by the mean of every subarea's local network temperature. Furthermore, Equation (9) is not the only way to obtain network temperature, it also can be calculated through Definition 3, which has numerical solutions only.

3.4 Network Heat Capacity

Like heat capacity in thermodynamics, network heat capacity is defined as a measurable quantity equal to the ratio of the amount of the packets added to (or removed from) a network system to the resulting network temperature change. Because the network heat capacity always increases with the increase of the network, we then define the notion of *specific heat capacity* (specific heat, for short), which is the heat capacity per unit amount of the current network state. In the light of network temperature's definition, which depends on k packets distributed over $|E|$ edges, *packet* and *directed edge between nodes* are selected as the two different units to define the two types of network-specific heat.

Definition 4 (Specific Heat for Packets). The network-specific heat with packets in a network $G(N, E)$ is defined as follows:

$$C_P = \frac{\Delta k}{|E| \cdot \Delta T}, \quad (10)$$

where Δk is the incremental quantity of packets in transmission; and ΔT is the corresponding network temperature change.

According to Equation (10),

$$C_P \approx \frac{1}{|E|} \cdot \frac{\partial k}{\partial T} \approx \frac{(2k \ln(|E| - k) + 1)^2 (|E| - k)}{2|E| \ln 2 \cdot (2k^2 + |E| - k)}. \quad (11)$$

THEOREM 2. C_P is a positive specific heat.

PROOF. According to Equation (9), we know that $|E| > k$ if the network temperature is measurable. Then both the numerator and the denominator of Equation (11) are positive, $C_P > 0$ thus follows, viz. C_P is a positive specific heat. $\square$

Definition 5 (Specific Heat for Edges). The network-specific heat with edges in a network system $G(N, E)$ is defined as follows:

$$C_{|E|} = \frac{\Delta|E|}{k \cdot \Delta T}, \quad (12)$$

where $\Delta|E|$ is the incremental quantity of directed edges; ΔT is the corresponding network temperature change; and k is the number of the packets being transmitted.

Similarly to the derivation of Equation (11),

$$C_{|E|} \approx -\frac{(k \ln(|E| - k) + 0.5)^2(|E| - k)}{k^3 \ln 2}. \quad (13)$$

THEOREM 3. $C_{|E|}$ is a negative specific heat.

PROOF. According to Equation (9), we know that $|E| > k$ when the network temperature is measurable. Then both the numerator and the denominator of Equation (13) are positive, $C_{|E|} < 0$ thus follows, viz. $C_{|E|}$ is a negative specific heat. $\square$

3.5 Network Temperature Gradient

The network temperature gradient is a location-based vector and describes in which direction and at which rate the network temperature changes the most rapidly at a particular location. The unit of the network temperature gradient is degree per unit length, e.g., a hop or a subnetwork. To associate positions with the nodes, without loss of generality, a plane-rectangular coordinate system is used, and then the network temperature at location (x, y) is represented by $T(x, y)$.

Definition 6 (Network Temperature Gradient). Let (x, y) be the position of a given node in a network $G(N, E)$; the network temperature gradient is defined as follows:

$$\nabla T = \left(\frac{\Delta T}{\Delta x}, \frac{\Delta T}{\Delta y} \right), \quad (14)$$

where ΔT is the network temperature change with respect to the location change $(\Delta x, \Delta y)$.

Therefore, the network temperature gradient at a position $P(x, y)$ can be approximated as:

$$\nabla T \approx \left(\frac{\partial T}{\partial x}, \frac{\partial T}{\partial y} \right). \quad (15)$$

There are two network temperature gradients, viz. the hot gradient and the cooling gradient, where packet transmissions along the hot gradient will increase the network temperature the most rapidly; and the packet transmissions along the cooling gradient achieves the right opposite effect.

3.6 Numerical Examples

To further understand the network temperature and network-specific heat defined above, we provide numerical examples for them in networks composed of 20 nodes using Maple. We intentionally choose relatively small sizes, because network temperature is an average quantity according to Definition 3. The local average network temperature for a node then is more accurate than that of the average network temperature for the entire network.

In the numerical results, we show the changes of network temperature T , network-specific heat C_P and $C_{|E|}$, with respect to the changes of k , viz., the number of packets in transmission, and $|E|$, viz., the number of one-hop data communication links, respectively.

Note that network temperature gradient is not an independent vector, determined by network temperature and a node's location, and thus does not have meaningful mathematical properties

Fig. 2. Numerical examples for network temperature, network-specific heat, and network temperature gradient. (a) The network temperature surface; (b) the network temperature curves with respect to the increase in the numbers of packets, where $|E| = 20, 25, 30$; (c) the network temperature curves with respect to the increase in the number of edges where $k = 9, 13, 17$; (d) the network-specific heat C_P curves with respect to the number of the packets, where $|E| = 20, 25, 30$; (e) the network-specific heat $C_{|E|}$ curves with respect to the number of the edges, where $k = 9, 13, 17$.

such as monotonicity, continuity, and so on. Therefore, we do not provide numerical results for these properties for the network temperature gradient.

3.6.1 Network Temperature. Figure 2(a) shows the network temperature surface in the numerical examples.

Figure 2(b) shows that: (1) T is strictly monotonically increasing with the increase in k ; (2) with the same k , the network with larger $|E|$ has a lower T ; (3) when $k < 0.75|E|$ the T increases slightly with the increases in k ; when $k \geq 0.75|E|$ the T increases drastically with the increases in k .

Figure 2(c) shows that: (1) T is strictly monotonically decreasing with the increases in $|E|$; (2) with the same $|E|$, the network with larger k has a higher T .

Figure 2 endorses Theorem 1, i.e., (1) with the same number of one-hop data communication links in a network, the network temperature increases with the increase in the number of packets being transmitted; (2) with the same number of packets being transmitted, the network temperature decreases with the increase in the number of one-hop data communication links in the network.

3.6.2 Network-specific Heat. Figure 2(d) shows that: (1) $C_P > 0$, which supports the conclusion of Theorem 2; (2) with the same k , C_P increases when $|E|$ increases; (3) for the same $|E|$, C_P increases when k increases for $k < 0.25|E|$ and then decreases when k increases for $k \geq 0.75|E|$ approximately.

Figure 2(e) shows that: (1) $C_{|E|} < 0$, which supports the conclusion of Theorem 3; (2) for the same $|E|$, $C_{|E|}$ increases when k increases; (3) for the same k , $C_{|E|}$ decreases when $|E|$ increases.

4 APPLICATIONS

In this section, we provide insights into network temperature, network-specific heat, and network temperature gradient as well their potential applications.

4.1 DDoS Pre-warning and Detection

Network entropy has been used for **Distributed Denial-of-Service (DDoS)** based on the assumption that the entropy of basic communication parameters change observably during a DDoS attack [10]. However, because such an approach sends a DDoS alarm only when network entropy exceeds the given decision threshold, by this time the attack has already caused damages to network services.

Our network temperature is a more accurate and effective indicator for the DDoS detection. In our approach $H_t = H_0 + at$, where: H_0 is the current network entropy; H_t is the network entropy after time t ; a is the acceleration of network entropy changes. Note that the inverse of network temperature is equal to a , which is the impetus to network entropy changes (see Equation (9)). Therefore, the network temperature change is more sensitive than network entropy under DDoS attacks, and thus it is more accurate.

To show the effectiveness a network temperature-based DDoS detection, we have carried experiments using publicly available data. Results are presented in Section 5.

4.2 Dynamic Node Importance Ranking

When a network is under attack, if we are unable to protect every node with the same strength due to limited resources, ranking and then protecting important nodes is critical. Although many methods have been proposed to evaluate nodes' importance, most of them are network topology based methods only, e.g., **Degree Centrality (DC)**, **Betweenness Centrality (BC)**, **Closeness Centrality (CC)**, **Eigenvalue Centrality (EC)**, and so on. As a result, the important nodes chosen by those methods may be rarely used for packet transmissions, as these nodes are not critical for the network's communication.

Note that the higher the network temperature around a node, the more packets are processed at the node, i.e., the network temperature at a node is a measure of the node's contribution to the network's communication. Therefore, network temperature can be used to rank nodes' importance under such an environment.

ASSERTION 1. *With respect to packet transmissions in a network, a node with a high network temperature is more important than a node with a lower network temperature.*

To show the effectiveness of a network temperature-based approach for detecting importance of nodes in a network, we have carried experiments using publicly available data. Results are presented in Section 5.

4.3 Network Stability and Robustness Evaluation

In physics, when temperature changes, the characteristics of the material, such as volume, extensibility, and conductivity, change accordingly. Likewise, the stability and robustness of a network also change with the changes in the network temperature.

A high network temperature indicates that the network has a heavy load at the moment, and thereby the values of packet loss rate, packet transmission delay, and packet transmission jitters are higher compared to their values when the network temperature is lower.

ASSERTION 2. *With respect to packet transmissions in a network, the network stability will increase when network temperature decreases and vice versa.*

According to Definition 4 and Definition 5, a large absolute value of specific heat C_P or $C_{|E|}$ in a network indicates that network temperature is not sensitive to the fluctuations of the number of packets in transmission and the number of one-hop data communication links, respectively. Furthermore, note that the network stability decreases when the network temperature increases, as stated by the following assertion:

ASSERTION 3. *With respect to packet transmissions in a network, the network robustness will decrease when C_P increases or $C_{|E|}$ decreases and vice versa.*

Because of hardware differences, we cannot simply say that a network with lower network temperature and absolute larger network-specific heat is more stable and robust than another network whose network temperature is high. Nevertheless, for the same network, Assertion 2 and Assertion 3 are rational conclusions.

The biggest difference between C_P and $C_{|E|}$ is that C_P is a positive specific heat (refer to Theorem 2); and $C_{|E|}$ is a negative specific heat (refer to Theorem 3). The physical meaning of the positive specific heat C_P is that a network's temperature will increase if more packets in transmission are added to the network, whereas the negative specific heat $C_{|E|}$ denotes that a network's temperature will decrease if the edges, viz., the one-hop data communication links between nodes, are added into the network.

Positive specific heat is common in physics. However, negative specific heat is a rare phenomenon. An example of negative specific heat is that the temperature of a gravitational system will increase when the gravitational system keeps losing its mass and finally such a gravitational system will become a black hole with a very high temperature and great mass intensity. Likewise, in a network, if the data communication links keep being removed, then the network temperature will increase accordingly and finally no packets can escape such a network, which is also a network black hole for the packets.

The phenomenon mentioned above shows that thermodynamic temperature and network temperature have some inherent relationships. Without the reduction to absurdity indirect method, it is an impossible mission to prove a quite intuitive result such as the parallel axiom in geometry. It seems that continuing to remove the one-hop data communication links between nodes will reduce the entire network's ability for packet transmissions is also an axiom that cannot be directly proved within the knowledge system of network communications. However, such a result is soundly proved by the theory of network temperature, which turns an empirical rule to a theorem.

4.4 Reliable Packets Routing

From the previous discussion, we know that network stability with respect to packet transmissions decreases when network temperature increases. Since each node has its own network temperature calculated through the subarea where the node is located, the packets routing mechanism should avoid crossing the already very hot nodes to prevent packets loss during transmission.

The notion of network temperature gradient defined in the article represents in which direction and at which rate the network temperature changes the most rapidly. To schedule the "coldest" route for a packet, each node can simply choose the next packet transmission hop along the cooling network temperature gradient. Clearly, such a method is a greedy algorithm that seeks the coldest packet path, and therefore the generated packet path is not assured to be the globally coldest route in the network. However, if the network temperature gradient function's curve only has one spike,

then the globally coldest packet route can be found, else only a local optimal cooling packets route is obtained.

ASSERTION 4. *Packet transmission along the network temperature gradient toward the cooling direction at each node results in a locally optimal reliable route.*

4.5 Provenance Compression Assessment

Typical provenance information about an individual data packet contains the source that originated the data and the nodes that forwarded and/or aggregated the data. For instance, if a provenance includes any tampered nodes or undependable data aggregations, then the corresponding data item will be discarded or marked with a low trust level to restrict its use [21]. However, the provenance size expands rapidly with the increase in the number of packet transmission hops. In a large-scale network, a data item's provenance size can be several times larger than the data item itself.

To ease the cost of provenance storage as well as of transmission, many compression approaches have been proposed to limit provenance size; we refer the reader to Reference [16] for a survey. However, to date there are no definitive results on the upper bound for provenance lossless compression. In Reference [8], an approach is proposed based on arithmetic coding. Under such an approach the encoded provenance is at most one bit longer than the provenance's entropy. Because Shannon's entropy is the upper bound of lossless compression in information theory, we believe that the approach in Reference [8] attains the upper bound already. Nevertheless, the later approaches [14, 15] achieve even higher lossless provenance compression rate, which is a surprising result.

One possible explanation for the above result is as follows: Under a given packet path's probabilistic model, Shannon's entropy is the upper bound of the lossless compression. The problem is that the packet path's probabilistic model, which is the basis for Shannon entropy's calculation, is not a static one but varies depending on the networks. Therefore, the provenance lossless compression upper bound is not a fixed value but a variable one related to changes in the network's topology and network load.

Provenance compression and network temperature are connected through network entropy H , where H represents the unknown information about k packets distribution among $|E|$ edges in a network (see Definition 2). On average, the number of bits that H can assign to a packet is as follows:

$$\eta = \frac{H}{k} = \frac{H(|E|, k)}{k}. \quad (16)$$

Furthermore, the rate of the increase of H with respect to the increase in k is modeled by the following the analytic expression:

$$\frac{\Delta H}{\Delta k} \approx \frac{\partial H(|E|, k)}{\partial k},$$

where the right part is approximately equal to $1/T$; ΔH and Δk are incremental quantities of H and k at duration Δt , respectively. Therefore,

$$\begin{aligned} \eta \cdot k = H &= \int_{\Delta t} \frac{\partial H(|E|, k)}{\partial k} dk \approx \int_{\Delta t} \frac{1}{T} dk \\ \eta &\approx \frac{1}{k} \int_{\Delta t} \frac{1}{T} dk. \end{aligned} \quad (17)$$

According to Equation (17), η represents the average provenance length in the network, because the physical meaning of H is the minimum number of bits for describing a packet path in the network. Furthermore, Equation (17) provides the mathematical function modeling the dependency of η on T , which also implies the relationship between the provenance lossless compression and the network state defined by network temperature at a given time. In short, we have the following assertions:

ASSERTION 5. *For a given lossless compression provenance encoding scheme, the average provenance size in a network will increase when the network temperature increases.*

An increase in network temperature would then indicate that more bandwidth and storage space are required for provenances' transmission and storage, respectively. From this point of view, a low network temperature is to be preferred while the received data need to be assessed with respect to trustworthiness.

Because network temperature is an ensemble of network topology and network load that determines the probabilistic model for provenances' lossless compression, we have the following assertion:

ASSERTION 6. *The upper bound of lossless provenance compression is the provenance's Shannon entropy in the light of the current network temperature.*

5 EXPERIMENTS

In this section, we show how to measure network temperature on different networks. Furthermore, to show that network temperature has a better performance for detecting DDoS attacks and can be used to rank nodes importance according to network load changes, experimental results on publicly available datasets are presented.

5.1 Network Temperature Measurement

Network temperature measurement is classified as both *direct measurement* and *indirect measurement*, where direct measurement is when all the required parameters are directly provided by a network for the network temperature's calculation; indirect measurement is when some required parameters need to be indirectly measured before the network temperature can be calculated.

5.1.1 Direct Measurement. Assume that all the parameters required to calculate network temperature can be found in a time-dependent packets transmission matrix, Equation (18), in which rows correspond to packet origins and columns to packet destinations in one hop. In addition to $a_{ij} = \phi$ representing no one hop connection from n_i to n_j , the other entries in the matrix are non-negative integers representing the number of packets being transmitted since the time t_0 .

$$\mathbf{M}(t) = \{a_{ij}\}_{n \times m} \quad (18)$$

The initial matrix is $\mathbf{M}(t_0)$, in which $a_{ij} = 0$ if n_i and n_j are connected by a directed edge e_{ij} in the network topology, otherwise $a_{ij} = \phi$.

By using data $\mathbf{M}(t)$ and $\mathbf{M}(t + \Delta t)$, the network temperature can be approximately calculated through Equation (9).

On a network with a few routers, e.g., Internet, $\mathbf{M}(t)$ can be derived from the network's routing matrix. On a network with a central node, e.g., a **wireless sensor network (WSN)** that has a **base station (BS)**, building $\mathbf{M}(t)$ requires all the provenances, which include each packet path with a timestamp at the BS.

5.1.2 Indirect Measurement. If the time-dependent packets transmission matrix $\mathbf{M}(t)$ is not available, then the indirect measurement approach needs to be adopted.

Fig. 3. (a) experiment testbed; (b) network temperature thermometer TM (from left to right: battery, Raspberry Pi); (3b) Zigbee sensor node, Dongle).

On a network that uses SDN technology, the data plane can be programmed to collect the information such as packets routing, data links active frequencies, and network topology changes, which indirectly provide all the parameters required for the network temperature's calculation. In our lab, we have utilized P4 and ONOS together to collect such information. We refer the interested readers to ONOS+P4 Tutorial at wiki.onosproject.org for further details.

On a distributed network without any historical packet routing information, a packet paths sniffing approach is needed. In our lab, we have devised an apparatus, called TM, to sniff the communications among the nodes in a WSN, where no provenance data are available.

5.1.3 Experimental Results on Temperature Measurement. To carry experiments about the indirect measurement of the network temperature, we have used a WSN testbed from our previous work [14] is reused in here. In this testbed each Zigbee node has a TI CC2530 microcontroller, 2.4 GHz radio, 8 KB RAM, and 256 KB flash for data logging, and one of them was used as the BS and connected to a laptop computer through a USB port. We placed the testbed in an indoor environment (area of size $15 \times 10 \text{ m}^2$) and set the transmission power to the minimum level to ensure multihop communications (see Figure 3(a)). Our network temperature-independent measurement apparatus, TM, consists of a Dongle, a Zigbee sensor node, and a Raspberry Pi (3b), as shown in Figure 3(b). We devise that TM services as a rudimentary signal analyzer that eavesdrops the nearby wireless signals and then sends the captured data to the Raspberry Pi node for further analysis to derive the required parameters for network temperature's calculation.

To perform both direct measurement and indirect measurement on the same testbed, we used a simple provenance scheme [16] to record packet trajectories and let each node keep the latest 200 provenances for the packets originated at or passed through the node.

Figure 4(a) reports the network temperature from testbed experiments through direct measurement. Compared with the numerical results in Figure 2(b), the curves in Figures 4(a) and 2(b) have almost the same trend. Figure 4(b) reports the specific heat C_P at testbed experiments through direct measurement. Compared with the numerical results in Figure 2(d), the curves in Figures 4(b) and 2(d) share almost the same trend.

Figure 4(c) reports the local network temperature at a node with respect to packet transmission with 1, 2, and 3 hops through indirect measurement by TM. The curve T, marked by a square symbol, represents the network temperature measured by direct measurement; the curve T(2m), marked by a circle symbol, represents the network temperature measured by indirect measurement, where TM is placed about 2 meters away from the node being measured; the curve T(5m), marked by a triangle symbol, represents the network temperature measured by indirect

Fig. 4. (a) the network temperature measured by the direct measurement; (b) the network-specific heat with packets measured by the direct measurement; (c) the local network temperature scaled by the indirect measurement.

measurement, where TM is placed about 5 meters away from the node being measured. The experimental data show that: (1) the curve T and the curve $T(2m)$ have high overlap; (2) the curve $T(5m)$ falls below the curve $T(2m)$ and shares almost the same trend with the curve $T(2m)$.

Because TM eavesdrops the wireless signals in the testbed and then analyzes the eavesdropped data at the Raspberry Pi node, while TM moves away from a node being monitored, the probability that packet transmissions around such a node are detected by TM decreases accordingly. This is the reason why the curve $T(5m)$ falls below the curve $T(2m)$. Note that while TM is close enough to a node, the direct measurement and the indirect measurement return almost the same values, as shown by the curves of T and $T(2m)$ in Figure 4(c).

5.2 DDoS Detection

As shown in Equation (8), network temperature is decided by the network entropy's changing rate. When we treat the network entropy changes as a dynamic system, the changing rate is the acceleration of network entropy. Although network entropy is a well-known index for DDoS attacks, some of them cannot be directly detected, because the network entropy changes are too small within an observation time window. Hence, we are inspired to use network temperature that is more sensitive to network entropy changes to detect those DDoS attacks.

To show that network temperature has a better performance in detecting some of DDoS attacks, we experimentally compared our proposed **network temperature (NT)** with approaches based, respectively, on: **information entropy (IE)**, **generalized information entropy (GE)** [19], **generalized information divergence (GID)** [2], and **cross entropy (CRE)** [19]. Our experiments were carried out on a dataset of publicly available network intrusion detection data, downloaded from www.unb.ca/cic/datasets. The topology of the test network can be found at www.unb.ca/cic/datasets/ids-2018.htm; and the detail information of those attacks can be found at iscxdownloads.cs.unb.ca/iscxdownloads/ISCX-SlowDoS-2016/attacks.txt. All the network temperatures in the experiments are scaled through the direct measurement.

Let $P = \{p_1, p_2, \dots, p_n\}$ and $Q = \{q_1, q_2, \dots, q_n\}$ denote IP address probability distributions in the test data, where $\sum_{i=1}^n p_i = \sum_{i=1}^n q_i = 1, n \in \mathbb{N}^+$. IE, GE, GID, and CRE are generated as follows:

- **IE**. It is calculated by Equation (3) in Section 2.2, where $b = 2$.
- **GE**. It is generated by Equation (19).

$$H_\alpha(x) = \frac{1}{1-\alpha} \log_2 \sum_{i=1}^n p_i^\alpha, \alpha \geq 0 \quad (19)$$

Note that in Equation (19) $H_1(x) = -\sum_i p_i \log_2 p_i$, viz., GE degrades to IE when $\alpha \rightarrow 1$.

- **GID**. It is generated by Equation (20).

$$D_{\alpha}(P||Q) = \frac{1}{1-\alpha} \log_2 \sum_{i=1}^n p_i^{\alpha} q_i^{1-\alpha}, \alpha \geq 0 \quad (20)$$

Note that in Equation (20) $D_1(P||Q) = \sum_{i=1}^n p_i \log_2 \frac{p_i}{q_i}$, viz., GID degrades to relative entropy when $\alpha \rightarrow 1$. In this article, $\alpha = 15$, which is an empirical data recommended by Reference [2].

- **CRE**. It is generated by Equation (21).

$$H(P, Q) = - \sum_{i=1}^n p_i \log_2 q_i \quad (21)$$

In the test data, there are two different kinds of attacks: (1) High-volume HTTP attacks, which include DoS improved GET (goldeneye), DDoS GET (ddossim), and DoS GET (hulk); (2) Low-volume HTTP attacks, which include slow-send body (slowbody2), slow send body (rudy), slow send headers (slowloris), slow-send headers (slowheaders), and slow-read (slowread).

The key metrics for comparing these techniques is how early they can correctly identify the attack.

All test data are in the pcap format. We first convert those pcap data into the csv format by Wireshark and then use Python to analyze them. IE, GE, GID, CRE, and NT are first encoded as Python libraries. Due to page limits, we only choose the goldeneye of High-volume HTTP attacks and the slowloris of Low-volume HTTP attacks to show that network temperature has a better performance in detecting them. Goldeneye sends large numbers of HTTP Get requests to exhaust the targeted web servers. Slowloris opens multiple connections to a targeted web server and keeps them open as long as possible by continuously sending partial but never actually completes HTTP requests, which depletes the available connections on the web server.

Figure 5 shows the curves of IE, GE, GID, CRE, and NT for the glodeneye attacks. As shown in Figure 5(e), there are two clear spikes in the curve of NT; the first one indicates that the glodeneye attack happens at about 160th second after the starting point in the test data, and the second one indicates that a flash event [19] happens at about 490th after the starting point in the test data. Although in Figure 5(d) CRE also shows some clear spikes, no spike can be observed during the interval when goldeneye happens, which means that CRE has a lower accuracy in detecting the goldeneye attack.

Table 1 lists a series of key data segmentations shown in Figure 5, viz., the two peak values of NT and their four neighbor values, where the starting point of the time is the 10,145th second in the above test data, and all the indicators were renewed every 5 seconds; the subscript cr of an indicator denotes the indicator's changing rate in the current time window. According to the goldeneye marks in the test data, the empirical valve for goldeneye attacks detection is $NT_{cr} \geq 20$.

Figure 6 shows the curves of IE, GE, GID, CRE, and NT for the slowloris attack. Since every HTTP request has a certain duration, to keep occupying the available connection on the targeted web server, the slowloris attack will periodically renew the uncomplete connections as shown in Figure 6(e), where the curve of NT accurately reflects such a period characteristic of the slowloris attacks. Although the curve of GID also shows such a period, the calculation of NT using Equation (7) has a much higher efficiency than that of GID using Equation (20).

In addition to the time starting point, all the data shown in Figure 6 were calculated as the ones in Table 1. To use the slowloris marks, the starting point is selected as the 55,205th second in the

Fig. 5. Information entropy (IE), generalized information entropy (GE), generalized information divergence (GID), cross entropy (CE), and network temperature (NT) for goldeneye attacks. (a) curve of IE; (b) curve of CE; (c) curve of GID; (d) curve of CRE; (e) curve of NT, where the first spike indicates the goldeneye attack and the second spike indicates a flash event.

Table 1. The Values of IE, GE, GID, CRE, NT and their Changing Rates

Time (s)	IE	GE	GID	CRE	NT	IE _{cr}	GE _{cr}	GID _{cr}	CRE _{cr}	NT _{cr}
155	1.740	0.878	-0.258	0.064	124.268	0.614	0.439	0.160	2.291	1.062
160	2.400	1.302	-2.624	0.054	3,234.000	1.379	1.483	10.171	0.840	26.024
165	2.717	1.923	4.453	0.004	1,865.465	1.132	1.477	1.697	0.073	0.577
490	1.428	0.978	-1.960	0.066	390.608	1.102	1.026	1.343	2.477	1.575
495	3.144	1.869	-0.870	0.014	3,186.000	2.201	1.910	0.444	0.209	8.157
500	3.236	1.595	0.992	0.005	-592.000	1.029	0.853	1.140	0.333	0.186

test data. Note that visualization is a better way to illustrate data's period than data themselves; the data lists are then omitted in here.

As a general conclusion, if a network event causes very small network entropy changes rate and/or no clear period can be observed from such network entropy change rates, then such a network event cannot be detected by the network temperature-based methods.

5.3 Node Importance Ranking

To show that network temperature can be used to rank the nodes importance with respect to network load changes, we use the Abilene network's historical data downloaded from maths.adelaide.edu.au/matthew.roughan/data/ as an example, where the same dataset was used in Reference [22]. Here, a node's importance is defined as its influence on the other nodes under the current network load [23]. The topology of Abilene network is shown in Figure 7.

We calculated the network temperature for each node through the first day's data in the dataset. The duration of the calculation is 5 minutes, which is the minimum update duration provided in the dataset. For page limit, only the network temperatures in three time slots are provided in Table 2:

Fig. 6. Information entropy (IE), generalized information entropy (GE), generalized information divergence (GID), cross entropy (CE), and network temperature (NT) for slowloris attacks. (a) curve of IE; (b) curve of GE; (c) curve of GID; (d) curve of CRE; (e) curve of NT, which shows the clear period of slowloris attacks.

Fig. 7. The topology of Abilene network, which provides 24 weeks of 5-minute averages, 12 routers (12×12 matrices).

(1) middle night, from 00:00 to 00:55; (2) morning, from 09:00 to 09:55; (3) afternoon, from 15:00 to 15:55, where in each row the nodes are listed in descending order according to their network temperatures. For example, the first row of Table 2 indicates that at the time 00:05 AM, node n_{12} has the highest network temperature followed by the nodes $n_{11}, n_6, \dots, n_1$.

When we use network temperature to measure a node's importance in the Abilene network, Table 2 shows that a node's importance varies with the network load changes. For comparison, we used two different approaches for determining the centrality of the nodes: an approach based on **CC (Closeness Centrality)** measurement and an approach based the **EC (Eigenvector Centrality)** measurement. The first identified node n_2 is treated as the most important one, whereas the second identified node n_6 is treated as the most important one. We note that those two approaches only use the network's topology and thus are unable to take network load changes into account.

To judge Assertion 1 in Section 4.2, viz., with respect to packet transmissions in a network, a node with a high network temperature is more important than a node with a lower network temperature, the **Susceptible-infected (SI)** model, which was used as a reference model for the

Table 2. Node Importance Ranking According to the Node's Network Temperature

Time	Nodes' ranking order											
	1	2	3	4	5	6	7	8	9	10	11	12
00:05	n_{12}	n_{11}	n_6	n_8	n_7	n_9	n_2	n_4	n_5	n_3	n_{10}	n_1
00:15	n_9	n_7	n_2	n_{12}	n_6	n_4	n_3	n_5	n_{10}	n_{11}	n_1	n_8
00:25	n_8	n_{12}	n_3	n_6	n_7	n_9	n_5	n_2	n_{11}	n_{10}	n_4	n_1
00:35	n_6	n_2	n_8	n_{12}	n_4	n_7	n_{11}	n_5	n_9	n_3	n_{10}	n_1
00:45	n_6	n_{12}	n_4	n_5	n_7	n_9	n_2	n_8	n_3	n_{10}	n_{11}	n_1
00:55	n_2	n_6	n_4	n_{12}	n_9	n_8	n_3	n_7	n_{11}	n_{10}	n_5	n_1
09:05	n_2	n_6	n_{12}	n_9	n_3	n_8	n_5	n_{11}	n_4	n_7	n_{10}	n_1
09:15	n_2	n_4	n_6	n_8	n_9	n_{12}	n_7	n_5	n_3	n_{11}	n_{10}	n_1
09:25	n_9	n_2	n_4	n_8	n_{12}	n_7	n_5	n_6	n_3	n_{10}	n_{11}	n_1
09:35	n_3	n_{12}	n_8	n_2	n_9	n_5	n_6	n_{10}	n_4	n_{11}	n_7	n_1
09:45	n_9	n_{12}	n_6	n_5	n_{10}	n_2	n_8	n_7	n_4	n_{11}	n_1	n_3
09:55	n_9	n_{12}	n_7	n_8	n_4	n_{10}	n_3	n_5	n_2	n_{11}	n_1	n_6
15:05	n_{12}	n_6	n_2	n_8	n_9	n_4	n_5	n_{11}	n_3	n_7	n_1	n_{10}
15:15	n_6	n_7	n_{12}	n_9	n_{11}	n_8	n_4	n_3	n_5	n_{10}	n_2	n_1
15:25	n_6	n_{12}	n_7	n_3	n_2	n_{11}	n_8	n_4	n_5	n_9	n_{10}	n_1
15:35	n_{12}	n_4	n_2	n_6	n_3	n_8	n_{10}	n_9	n_{11}	n_7	n_5	n_1
15:45	n_4	n_8	n_{12}	n_6	n_5	n_9	n_3	n_2	n_7	n_{11}	n_1	n_{10}
15:55	n_5	n_{12}	n_6	n_9	n_7	n_3	n_4	n_8	n_{11}	n_{10}	n_1	n_2

effectiveness of nodes ranking methods [23], is applied. Our SI model test results on the same dataset validate the insight that a node with higher network temperature is more important than another node with low network temperature. Furthermore, a group of nodes with higher average network temperature is more important than another group of different nodes with low average network temperature.

6 CONCLUSIONS

We propose the novel concept network temperature together with the notions of network-specific heat and network temperature gradient. The properties and applicability of network temperature, network-specific heat and network temperature gradient are systematically analyzed in the article, with conclusions supported by the numerical and/or the experimental results. As future work, we plan to explore other applications of our novel network measurement and management framework.

ACKNOWLEDGMENTS

Our graduate students Mr. Yao Wei and Ms. Qian Chen helped with some of the experiments in the article.

REFERENCES

- [1] Amer Aljaedi and C. Edward Chow. 2016. Pathseer: A centralized tracer of packet trajectories in software-defined datacenter networks. In *Principles, Systems and Applications of IP Telecommunications (IPTComm)*. IEEE, 1–9. DOI:<https://doi.org/10.1109/IPTComm39427.2016.7780246>
- [2] Sunny Behal and Krishan Kumar. 2017. Detection of DDoS attacks and flash events using novel information theory metrics. *Comput. Netw.* 116 (2017), 96–110. DOI:<https://doi.org/10.1016/j.comnet.2017.02.015>

- [3] Przemysław Bereziniński, Bartosz Jasiul, and Marcin Szpyrka. 2015. An entropy-based network anomaly detection method. *Entropy* 17, 4 (2015), 2367–2408. DOI:<https://doi.org/10.3390/e17042367>
- [4] Antoine Bérut, Artak Arakelyan, Artyom Petrosyan, Sergio Ciliberto, Raoul Dillenschneider, and Eric Lutz. 2012. Experimental verification of Landauer’s principle linking information and thermodynamics. *Nature* 483, 7338 (Mar. 2012), 187–189. DOI:<https://doi.org/10.1038/nature10872>
- [5] Leon Brillouin. 1956. *Science and Information Theory*. Academic Press, New York, NY.
- [6] Michael Dairyko, Leslie Hogben, Jephian C. H. Lin, Joshua Lockhart, David E. Roberson, Simone Severini, and Michael Young. 2017. Note on Von Neumann and Rényi entropies of a graph. *Linear Algeb. Appl.* 521 (2017), 240–253.
- [7] Y. Gao, Y. Jing, and W. Dong. 2018. UniROPE: Universal and robust packet trajectory tracing for software-defined networks. *IEEE/ACM Trans. Netw.* 26, 6 (Dec. 2018), 2515–2527. DOI:<https://doi.org/10.1109/TNET.2018.2871213>
- [8] S. R. Hussain, C. Wang, S. Sultana, and E. Bertino. 2014. Secure data provenance compression using arithmetic coding in wireless sensor networks. In *IEEE 33rd International Performance Computing and Communications Conference (IPCCC)*. IEEE, 1–10. DOI:<https://doi.org/10.1109/PCCC.2014.7017068>
- [9] R. Landauer. 1961. Irreversibility and heat generation in the computing process. *IBM J. Res. Devel.* 5, 3 (July 1961), 183–191. <https://doi.org/10.1147/rd.53.0183>
- [10] Ângelo Cardoso Lapolli, Jonatas Adilson Marques, and Luciano Paschoal Gaspary. 2019. Offloading real-time DDoS attack detection to programmable data planes. In *IFIP/IEEE Symposium on Integrated Network and Service Management (IM’19)*. 19–27.
- [11] Dénes Petz. 2001. Entropy, Von Neumann and the Von Neumann entropy. In *John von Neumann and the Foundations of Quantum Physics*. Springer Netherlands, Dordrecht, 83–96. DOI:https://doi.org/10.1007/978-94-017-2012-0_7
- [12] H. H. Song, L. Qiu, and Y. Zhang. 2009. NetQuest: A flexible framework for large-scale network measurement. *IEEE/ACM Trans. Netw.* 17, 1 (2009), 106–119.
- [13] E. Szmidi. 2007. Uncertainty and information: Foundations of generalized information theory (Klir, G.J.; 2006). *IEEE Trans. Neural Networks* 18, 5 (Sept. 2007), 1551–1551. DOI:<https://doi.org/10.1109/TNN.2007.906888>
- [14] Changda Wang and Elisa Bertino. 2017. Sensor network provenance compression using dynamic Bayesian networks. *ACM Trans. Sen. Netw.* 13, 1 (Jan. 2017). DOI:<https://doi.org/10.1145/2997653>
- [15] Changda Wang, Syed Rafiul Hussain, and Elisa Bertino. 2016. Dictionary based secure provenance compression for wireless sensor networks. *IEEE Trans. Parallel Distrib. Syst.* 27, 2 (Feb. 2016), 405–418. DOI:<https://doi.org/10.1109/TPDS.2015.2402156>
- [16] Changda Wang, Wenyi Zheng, and Elisa Bertino. 2016. Provenance for wireless sensor networks: A survey. *Data Sci. Eng.* 1, 3 (01 Sep. 2016), 189–200. DOI:<https://doi.org/10.1007/s41019-016-0017-x>
- [17] Rui Wang, Zhiping Jia, and Lei Ju. 2015. An entropy-based distributed DDoS detection mechanism in software-defined networking. In *IEEE Trustcom/BigDataSE/ISPA*, Vol. 1. IEEE, 310–317. DOI:<https://doi.org/10.1109/Trustcom.2015.389>
- [18] C. Williamson. 2001. Internet traffic measurement. *IEEE Internet Comput.* 5, 6 (Nov. 2001), 70–74. DOI:<https://doi.org/10.1109/4236.968834>
- [19] Ruoyu Yan and Qinghua Zheng. 2009. Using Renyi cross entropy to analyze traffic matrix and detect DDoS attacks. *Inf. Technol. J.* 8 (2009), 1180–1188.
- [20] Chen Yang. 2019. Anomaly network traffic detection algorithm based on information entropy measurement under the cloud computing environment. *Cluster Comput.* 22, S4 (July 2019), 8309–8317. DOI:<https://doi.org/10.1007/s10586-018-1755-5>
- [21] Faheem Zafar, Abid Khan, Saba Suhail, Idrees Ahmed, Khizar Hameed, Hayat Mohammad Khan, Farhana Jabeen, and Adeel Anjum. 2017. Trustworthy data: A survey, taxonomy and future trends of secure provenance schemes. *J. Netw. Comput. Applic.* 94 (2017), 50–68. DOI:<https://doi.org/10.1016/j.jnca.2017.06.003>
- [22] Yin Zhang, Zihui Ge, Albert Greenberg, and Matthew Roughan. 2005. Network anomography. In *5th ACM SIGCOMM Conference on Internet Measurement*. ACM Press, New York, New York, 317–330. DOI:<https://doi.org/10.1145/1330107.1330146>
- [23] Jie Zhao, Yutong Song, and Yong Deng. 2020. A novel model to identify the influential nodes: Evidence theory centrality. *IEEE Access* 8 (2020), 46773–46780. DOI:<https://doi.org/10.1109/ACCESS.2020.2978142>

Received March 2021; revised August 2021; accepted September 2021