Data Acquisition and Utilization of Cognitive Protocol Stack Parameters for Efficient Networking

Ankit Srivastava ¹ and Manoj B. S. ²

¹Indian Institute of Space Science and Technology
²Affiliation not available

December 7, 2023

Abstract

The dataset descriptor paper presents an essential and structured documentation of the Internet Protocol Stack Dataset, offering a comprehensive overview of its contents and potential applications in the field of networking and telecommunications. This descriptor paper details the dataset’s organization into tunable parameters, non-tunable parameters, and performance metrics for VoIP and HTTP sessions, providing insights into its significance for network performance evaluation. Researchers, network engineers, and data scientists can utilize this descriptor as a guide to effectively access and utilize the dataset for various purposes, including network optimization, quality of service assessment, and network modeling, ultimately contributing to advancements in network performance and infrastructure development.
Data Acquisition and Utilization of Cognitive Protocol Stack Parameters for Efficient Networking

Ankit Srivastava, Student Member, IEEE  
Department of Avionics  
Indian Institute of Space Science and Technology  
Thiruvananthapuram, India  
srivastava0616@gmail.com

BS Manoj, Senior Member, IEEE  
Department of Avionics  
Indian Institute of Space Science and Technology  
Thiruvananthapuram, India  
bsmanoj@iist.ac.in

Abstract—The rapid evolution of communication networks and the ever-increasing demand for efficient data transfer have led to the development of cognitive networking, which aims to enhance network performance through intelligent and adaptive protocols. To facilitate research and development in this domain, we present a comprehensive dataset detailing the parameters of a Network Protocol Stack which can be used to develop a Cognitive Network Protocol Stack designed for efficient networking. Through the examination of various scenarios involving both mobile and stationary nodes, data has been gathered using a network simulator to assess protocol stack parameters and their impact on network performance metrics. The dataset can be used to evaluate and compare the extent to which each parameter affects these network performance metrics.

Index Terms—Network, protocol stack, performance metrics, network parameters

I. INTRODUCTION

The ever-evolving landscape of modern communication networks presents a complex challenge in maintaining peak performance, particularly in scenarios involving commonly used VoIP and HTTP applications. With a growing dependence on mobile nodes, the need for flexible solutions that can adapt to a variety of network conditions becomes increasingly critical. There is extremely scarce amount of data available on the protocol stack parameters. The currently available parameter setting is done using the Bayesian methods [1]. This data descriptor paper addresses this challenge by introducing an innovative approach that employs data collection and analysis techniques to know the various network parameters and the extent to which they influence the various network performance metrics.

This descriptor paper outlines our proposed methodology, which entails the collection of data on both adjustable and non-adjustable parameters across the protocol stack, utilizing the OMNet++ Discrete Event Simulator (Network Simulator), and recording the corresponding network performance metrics [2]. Specifically, we conducted experiments using the OMNet++ simulation software, covering four cases each for HTTP and VoIP: two stationary nodes, two mobile nodes, multiple stationary nodes and multiple mobile nodes. We meticulously identified tunable and non-tunable network parameters and documented their associated network performance metrics.

By concentrating on the data collection aspect, we aim to establish a comprehensive foundation for further research aimed at optimizing network performance.

II. PROTOCOL STACK DATA

The network protocol stack is a layered realization conceptual framework that standardizes the functions and interactions of networking protocols and services to enable communication between devices in a computer network. The Internet protocol stack is divided into five distinct layers, each responsible for specific tasks in the data transmission process. Starting from the bottom, the Physical Layer deals with the hardware aspects of data transmission, such as cables and electrical signals. The Data Link Layer, Layer 2, present above the Physical Layer is responsible for framing data into packets and managing access to the physical medium. Above that, the Network Layer, Layer 3, handles routing and addressing, ensuring data reaches its destination across a network. The Transport Layer, Layer 4, manages end-to-end communication and reliability of data transfer. Finally, the Application Layer, at the top, deals with user applications and their interactions, providing the interface for users to access network services and resources [3]. Together, these layers create a comprehensive framework for efficient and organized communication in computer networks.

A. Data Collection

In our comprehensive investigation of network performance dynamics within Voice over Internet Protocol (VoIP) and Hyper Text Transfer Protocol (HTTP) scenarios, we conducted an extensive network simulation study using the OMNET++ simulator [4]. Our simulation environment was designed to encompass a wide array of scenarios. These scenarios included configurations with multiple nodes in both mobile and stationary states, as well as two-node setups under similar conditions. For all the mobility cases, we have considered the Gaussian-Markov mobility model with the speed to be 20 mps and for all the multiple node cases, we have considered 9 nodes. We carefully selected tunable parameters from multiple layers of the protocol stack and segregated them from non-tunable ones. For each scenario, we employed commonly encountered real-world values for various parameters within the network protocol stack and noted the values of performance metrics, effectively emulating communication protocols’ behavior.
under diverse real-world conditions [5]. These parameters were systematically varied over their typical value ranges, generating a dataset comprising 1000 data points for each case, which we subsequently stored in a CSV file for our analysis. Throughout these simulations, we tracked performance metrics such as latency, throughput, and packet loss to quantitatively assess the influence of different parameters on network behavior. Additionally, we included non-trainable parameters as benchmarks to establish a baseline for comparison. By methodically varying parameters to mirror their real-world variations, we curated a robust dataset that serves as the basis of our conclusions of the relative influence of each of them on the performance metrics. The data gleaned from these simulations offers valuable insights into the relationships between protocol stack parameters and network performance metrics.

III. DATA ORGANISATION AND DESCRIPTION

The dataset is divided and listed into three broad categories, 1. Tunable parameters, followed by 2. Non-tunable parameters and then 3. Performance metrics. Furthermore, under each of these broad categories, we list the parameters serially in a row based on the layer to which they belong, i.e.- application layer parameters followed by those of transport layer, network layer, data link layer and the physical layer respectively. This structured organisation makes it easy to compare and analyze the impact of tunable parameters on performance metrics for all scenarios, while also considering the influence of non-tunable parameters in the network simulation study. The structured dataset is described below under the following two categories: VoIP session parameters and HTTP session parameters

A. VoIP session parameters

The VoIP session parameters are categorized into the following categories and are discussed here:

- Input performance metrics
- Tunable Parameters
- Non-tunable Parameters

1) Input performance metrics:

- Mean Opinion Score (MOS) (Application Layer): A subjective measure of the overall quality of a VoIP call, typically rated on a scale from 1 to 5, with higher scores indicating better call quality.
- Playout Loss Rate (Application Layer): The percentage of lost or discarded audio packets during transmission, which can degrade call quality.
- Playout Delay (Jitter) (Application Layer): Variation in the time it takes for audio packets to arrive at their destination, which can result in uneven voice quality.
- Packet Loss Rate (Transport Layer): The percentage of VoIP data packets that do not reach their intended destination, often leading to call quality issues.
- End-to-End Delay (Transport Layer): The total time it takes for an audio packet to travel from the sender to the recipient, including transmission and processing delays. Excessive delay can affect real-time conversations.

2) The tunable parameters:

- Message Length (Application layer): The size or length of the data or message being sent or received by the application, often measured in bytes.
- Send Interval (Transport layer): The time interval between sending data segments or packets, which can impact the rate at which data is transmitted and its reliability.
- Packet Queue Capacity (Network layer): The maximum number of packets that a network node’s buffer or queue can hold at a given time, affecting how effectively packets are routed.
- MAC ACK Timeout (Data link layer): The duration a sender waits for an acknowledgment (ACK) from the receiver to confirm successful packet transmission.
- MAC Header Length (Data link layer): The length of the Media Access Control (MAC) layer header added to data frames for addressing and control purposes.
- Radio Bandwidth (Physical layer): The range of frequencies available for data transmission in a wireless communication system.
- Radio Transmitter Power (Physical layer): The strength of the signal emitted by the transmitter, impacting the signal’s reach and quality.
- Radio Transmitter Header Length (Physical layer): The length of control and addressing information added to the transmitted signal.
- Radio Receiver Sensitivity (Physical layer): The minimum signal strength required for a receiver to detect and process a transmission.
- Radio Receiver Energy Detection (Physical layer): The ability of a receiver to detect the presence of a signal in the environment.
- Radio Signal-to-Noise and Interference Ratio (SNIR) Threshold (Physical layer): The threshold level at which a signal can be successfully distinguished from noise and interference.
- Transmitter Communication Range (Physical layer): The maximum distance over which a transmitter can effectively communicate with a receiver.
- Transmitter Interference Range (Physical layer): The distance within which a transmitter may cause interference with other communication systems.
- Transmitter Preamble Duration (Physical layer): The time duration of a signal preamble used for synchronization and signal detection at the receiver end in a wireless communication system.

3) The non-tunable parameters:

- Incoming Data Rate (Application Layer): The rate at which data is received by an application, often measured in bits per second (bps).
- Packet Sent (Transport layer): The number of data packets sent by the transport layer.
• Packet Received (Transport layer): The number of data packets successfully received by the transport layer.
• Queue Bit Length (Network layer): The total length, in bits, of data stored in a network node’s queue.
• Queue Length (Network layer): The number of packets or data units waiting in a network node’s queue.
• Queuing Time (Network layer): The time that packets spend in a queue before being forwarded.
• Tail Drop Loss (Network layer): Packets that are dropped from the end of the queue when it is full.
• Dropped Packet Loss (Network layer): The number of packets that are discarded due to various reasons, such as congestion or errors.
• Time to Live (Network layer): A value in IP header that limits the time a packet can exist in the network, preventing infinite loops.
• Incoming Packet Length (Data link layer): The length of incoming data packets at the data link layer, often measured in bits or bytes.
• Residual Energy Capacity (Data link layer): The remaining energy available in a device’s battery or power source, which can be crucial for wireless communication systems.
• Center Frequency (Physical layer): The specific frequency at which a wireless communication channel or radio signal is centered, defining its position within the frequency spectrum.

B. HTTP session parameters

The HTTP session parameters are categorized into the following categories and are discussed here:

• Input performance metrics
• Tunable Parameters
• Non-tunable Parameters

1) Input performance metrics:

• Throughput (Application layer): The rate at which data is transferred between a web server and a client, typically measured in bits per second (bps) or bytes per second (Bps). Higher throughput indicates faster data transfer.
• Latency (Application layer): The time it takes for a request from a client to reach a web server and receive a response. Lower latency is desirable for faster loading times of web pages and applications.

2) The tunable parameters:

• Start Time (Application layer): The time at which an application session or process begins its operation.
• Number of Requests per Session (Application layer): The count of requests or interactions made by an application during a single session.
• Think Time (Application layer): The duration of inactivity or delay between two consecutive user interactions or requests within an application session.
• Request Length (Transport layer): The size or length of data in a request sent by the transport layer to the destination.

• Reply Length (Transport layer): The size or length of data in a reply received by the transport layer from the destination.
• Idle Interval (Transport layer): The period of inactivity or lack of data transmission within the transport layer.
• Queue Packet Capacity (Network layer): The maximum number of packets that a network node’s buffer or queue can hold at a given time, affecting how effectively packets are routed.
• Reconnect Interval (Network layer): The time interval at which a network device attempts to establish a connection or reestablish a session after a disconnect.
• MAC Ack Timeout (Data link layer): The duration a sender waits for an acknowledgment (ACK) from the receiver to confirm successful packet transmission.
• MAC Header Length (Data link layer): The length of the Media Access Control (MAC) layer header added to data frames for addressing and control purposes.
• Transmitter Communication Range (Physical layer): The maximum distance over which a transmitter can effectively communicate with a receiver.
• Transmitter Interference Range (Physical layer): The distance within which a transmitter may cause interference with other communication systems.

3) The non-tunable parameters:

• Incoming Data Rate (Application layer): The rate at which data is received by an application, typically measured in bits per second (bps).
• Packets Sent (Transport layer): The number of data packets sent by the transport layer.
• Packets Received (Transport layer): The number of data packets successfully received by the transport layer.
• Window Size (Transport layer): The number of unacknowledged packets that a sender can transmit before waiting for acknowledgments.
• Queue Bit Length (Network layer): The total length, in bits, of data stored in a network node’s queue.
• Queue Length (Network layer): The number of packets or data units waiting in a network node’s queue.
• Queuing Time (Network layer): The time that packets spend in a queue before being forwarded.
• Tail Drop Loss (Network layer): Packets that are dropped from the end of the queue when it is full.
• Dropped Packet Loss (Network layer): The number of packets that are discarded due to various reasons, such as congestion or errors.
• Time to Live (Network layer): A value in packet headers that limits the time a packet can exist in the network, preventing infinite loops.
• Incoming Packet Length (Data link layer): The length of incoming data packets at the data link layer, often measured in bits or bytes.
• Residual Energy Capacity (Data link layer): The remaining energy available in a device’s battery or power source, which can be crucial for wireless communication systems.
• Center Frequency (Physical layer): The specific frequency at which a wireless communication channel or radio signal is centered, defining its position within the frequency spectrum.

The key distinction between parameters for multiple nodes and two nodes is the absence of transmitter interference range in the two nodes case. This is because, for effective communication, the two nodes must always remain within each other’s communication range, which is inherently smaller than the interference range. As a result, the metric related to interference range becomes redundant in two nodes scenario, resulting in one less tunable parameter.

C. Dataset Availability

Dataset is available in the IEEE Dataport in the name - Internet Protocol Stack Dataset.

IV. POSSIBLE APPLICATIONS

The structured Network Protocol Stack dataset presented here contains various tunable parameters, non-tunable parameters, and performance metrics for two different types of network connections: VoIP (Voice over Internet Protocol) and HTTP (Hypertext Transfer Protocol). These parameters and metrics can be used in a variety of ways in the field of networking and telecommunications. Here are some possible uses for this dataset:

• Cognitive Protocol Stack Development: The dataset can be used to train an AI/ML model for creating a cognitive protocol stack wherein the parameters are fine tuned based on the requirements of the user. Cognitive Protocol Stack will help adapt the parameters of the protocol stack such that the changing user conditions are catered and suitable alterations are done to ensure good and uninterrupted connection.

• Network Performance Optimization: Engineers and network administrators can use the tunable parameters and performance metrics to optimize network performance for VoIP and HTTP connections. They can adjust parameters such as message length, send interval, or start time to improve throughput, reduce latency, or enhance call quality.

• Quality of Service (QoS) Assessment: The Mean Opinion Score (MOS) and Playout Loss Rate can help assess the quality of VoIP calls. Network operators can use these metrics to monitor and maintain a high level of service quality for voice communication.

• Troubleshooting and Diagnostics: Network administrators can use the packet loss rate, end-to-end delay, and other performance metrics to diagnose issues in VoIP and HTTP connections. They can pinpoint where problems occur and take corrective actions.

• Protocol and Application Development: Developers working on VoIP or HTTP applications can test and fine-tune their software using this dataset to ensure optimal performance under different network conditions.

• Network Security: Understanding network performance and the impact of various parameters can help in identifying abnormal behavior and security threats in the network. Anomalies in metrics can trigger security alerts.

V. DESCRIPTION OF FILES IN DATASET

The dataset is divided into 8 files (1000 data-points in each). The name of the file indicates the specific case to which the data inside corresponds. The name of each file is as follows:

- Multiple Mobile Node VoIP
- Multiple stationary Node VoIP
- Two Mobile Node VoIP
- Two Stationary Node VoIP
- Multiple Mobile Node HTTP
- Multiple stationary Node HTTP
- Two Mobile Node HTTP
- Two Stationary Node HTTP

The data is available in two different formats for each of the 8 cases - CSV and XLSX. The first row of each file represents whether the parameter is tunable, non-tunable or a performance metric. The second row has the name of the parameter/metric. Third row onward we have 1000 data-points (row 3 to 1003).

VI. CONCLUSIONS

The structured dataset presented in this paper containing tunable parameters, non-tunable parameters and performance metrics for VoIP and HTTP connections offers a valuable resource for a multitude of applications within the field of networking and telecommunications. By organizing the data into these distinct categories, it becomes useful dataset for network optimization, quality of service assessment, troubleshooting, and protocol development. Furthermore, the dataset’s applicability extends to research and development, serving as a foundation for network simulations, modeling, and capacity planning. It also facilitates benchmarking, enabling organizations to make informed choices about networking technology. Machine learning and data analysis can be applied to gain insights into network behavior and predict performance based on parameter settings.

ACKNOWLEDGMENT

We would like to acknowledge the support and encouragement received from the Indian Institute of Space Science and Technology (IIST) Thiruvananthapuram, where this research was conducted.

REFERENCES