Evaluation of the Effect of Wireshark-based Laboratories on Increasing Student Understanding of Learning Outcomes in a Data Communications Course

Dr. Craig A Chin, Southern Polytechnic State University

Craig A. Chin received his Ph.D. in electrical engineering from Florida International University in 2006. He is currently an Assistant Professor in the electrical and computer engineering technology at Southern Polytechnic State University. His research interests include biomedical signal processing, pattern recognition, and active learning techniques applied to engineering education.

Dr. Leigh Sharma

Dr. Garth V Crosby, Southern Illinois University Carbondale

Dr. Garth V. Crosby is an assistant professor in the Technology Department at Southern Illinois University Carbondale. He obtained his MS and PhD degrees from Florida International University in Computer Engineering and Electrical Engineering, respectively. Dr. Crosby’s primary interests of research are wireless networks, wireless sensor networks, network security and active learning strategies for STEM. He has served as a reviewer for several conferences, magazines and journals publications, including IEEE INFOCOM, and IEEE Transactions on Systems, Man and Cybernetics. He is a senior member of the Institute of Electrical & Electronics Engineers (IEEE), a member of the National Society of Black Engineers (NSBE), and Eta Kappa Nu.
Evaluation of the Effect of Wireshark-based Laboratories on Increasing Student Understanding of Learning Outcomes in a Data Communications Course

Abstract

This paper focuses on a 3000-level course in data communications that is a part of the curriculum in an electrical engineering technology program. Some of the primary learning outcomes of the course section that investigates data packet flow are: to identify the different fields of a data link layer (Ethernet) frame and to describe how the TCP/IP protocol suite utilizes physical, logical and port addresses to send/receive data. In addition, the ability to describe the operation of the Address Resolution Protocol (ARP) in facilitating data link layer communication by using ARP request and ARP reply packets is also a learning outcome. The content relevant to these outcomes is presented in class in a traditional face-to-face lecture format. In these lectures, the data frames are represented in a simplified graphical format.

Recently, two laboratories using the Wireshark packet-sniffer application has been added to deepen student understanding of these learning outcomes. The second laboratory of the sequence involves investigating the function of ARP by examining the contents of an ARP request packet and an ARP reply packet produced when a user uses a web browser to access a file from a website. The laboratory also involves analyzing the different fields of an Ethernet frame. It is expected that the interaction with real-world frames resulting from these laboratories should result in an increased ability to accomplish the above-stated learning outcomes.

The primary objective of the research was to determine the suitability of Wireshark as a learning tool in the undergraduate data communication classroom. The hypothesis is “empirical investigation with the Wireshark packet sniffer application increases student understanding of communication protocols”. This hypothesis is tested through quantitative and qualitative assessment. The quantitative assessment measured the students’ knowledge and understanding based on their performance in pre- and post-tests. Qualitative assessment is done by requiring students to complete a survey that focused on their perception and attitude towards their attainment of the learning objectives. The statistical analysis of pre-test versus post-test scores and the survey responses indicate that student participation in the Wireshark laboratories resulted in an improvement in their ability to interpret and analyze real-world data link frames and in an increase in their understanding of how TCP/IP protocols are used for information exchange between two hosts.

Introduction

ECET 3400 is the first of a two course series in data communications. A bottom-up approach is utilized in this series, with topics relevant to the lower layers of the TCP/IP protocol suite (physical and data link) being examined in ECET 3400 and topics relevant to the upper layers (network, transport and application) being examined in ECET 4820.

As a part of ECET 3400, concepts pertaining to network models, such as frame format, encapsulation and addressing are introduced using a simplified frame format representation. In
this representation, physical, logical, and port addresses are represented by simple numbers or letters (Figure 1). This simplified frame format makes addressing and protocol function more accessible to the student but does not enable him or her to interpret or analyze data link layer frames extracted from real network traffic.

<table>
<thead>
<tr>
<th>27</th>
<th>43</th>
<th>A</th>
<th>D</th>
<th>E</th>
<th>g</th>
<th>Data</th>
<th>T2</th>
</tr>
</thead>
</table>

Physical Addresses  Logical Addresses  Port Addresses  Application Layer Data  Trailer

**Figure 1 – Example of Simplified Data Link Frame Format Representation**

This deficiency motivated the creation of a two laboratory sequence that would support the following general course outcomes:

1. Students should be able to interpret and analyze real-world data link frames.
2. Students should have an understanding of how protocols within TCP/IP operate to accomplish the exchange of information between two hosts.
3. Students should be adequately prepared for ECET 4820.

The primary tool used in these laboratories is a packet sniffer software application called Wireshark\(^1\). A packet sniffer application runs in parallel to an application that sends and receives packets (e.g. a web browser). The packet capture (pcap) component of the packet sniffer creates copies of all frames sent and received at the data link layer of the host entity. The packet analyzer component of the packet sniffer analyzes the contents of each frame by parsing the frame into the header and data portions of the upper layers of the TCP/IP protocol (Figure 2).

**Figure 2 – Packet Sniffer Structure**

The instruction style chosen for the laboratories is the traditional expository style\(^2\) as opposed to other laboratory instruction styles found to be better at the development of higher-order cognitive skills. The reason for this choice has primarily to do with the fact that the students are being introduced to the Wireshark tool during this laboratory sequence. This meant that each student has to be taught the rote procedures associated with utilizing this tool, and it is felt that the
student-generated procedures associated with the other laboratory instruction styles would be inefficient or ineffective.

The remainder of this paper will provide the details of the methods used to create the two laboratories and the design of the experiment used to evaluate their effectiveness. The results of the experiment will also be presented, along with a discussion of what can be inferred from the results, in terms of the effectiveness of the laboratories and what improvements can be made for future implementations of these laboratories.

Method

Laboratory Design
The first of the two laboratories is created to familiarize each student to the interfaces displayed by the Wireshark application and to introduce each student to frame analysis using Wireshark. The laboratory introduces various sections of the start-up interface, and also introduces the five major components of the packet capture and analysis interface: the command menus, the packet display filter field, packet-listing window, the packet-header details window, and the packet-contents window display. This laboratory also guided each student through a procedure that enabled him or her to capture a series of frames sent or received by the source host computer during a web-browsing session. It also enabled a student to filter the captured frames in order to select a single frame that encapsulates an http get message for frame analysis. Frame analysis involved: determining the total number of bytes in the frame; locating the headers for the data-link layer, network layer, and transport layer; identifying the sizes of the header fields in bytes; identifying the source/destination IP addresses; and determining the size of the application layer message.

In the second laboratory, each student is required to use Wireshark to investigate the operation of the Address Resolution Protocol (ARP) and to investigate the format of Ethernet frames. The objectives associated with this laboratory are as follows:

a) Describe the operation of ARP in facilitating data link layer communication by using ARP request and ARP reply packets.
b) Identify the different sections of an ARP packet.
c) Identify the different fields of a data link layer (Ethernet) frame.
d) Describe how the TCP/IP protocol suite utilizes physical, logical and port addresses to send/receive data.

The first two sections of the laboratory require each student to examine an ARP request and reply packet respectively. The procedure involves:
i. Starting a web browser and clearing the browser’s cache memory.
ii. Opening Wireshark and starting frame capture.
iii. Using the browser to retrieve any file from a website.
iv. Stopping frame capture, after enough packets have been received, and saving the captured file.
v. Using the filter field of the Wireshark interface to filter out arp packets.
Analysis of an ARP request packet involves identifying the different fields of the packet and stating what the values are for each field. The ARP packet fields include: hardware type, protocol type, hardware length, protocol length, opcode, source/destination hardware addresses, and source/destination protocol addresses. The significance of using a broadcast destination address for the ARP request packet is also examined. A similar procedure is followed for the ARP reply packet.

In the next two sections of the laboratory the contents of an Ethernet frame are examined. The procedure involves a sequence of steps similar to the steps i to iv used for ARP packet analysis. The first section of Ethernet frame analysis involves selecting an Ethernet frame, determining its size, and the size of the payload. It also involves examining a frame that includes zero-padding. The second Ethernet frame analysis section focuses on the Ethernet header, with each student being asked to determine the destination link layer address, source link-layer address, and the upper layer protocol using the Ethernet frame. Each student is also asked to determine whether the frame is a unicast, multicast or broadcast.

**Evaluation Method**

The effectiveness of these laboratories are evaluated via surveys and tests given just prior to the two lab sequence and immediately after the completion of the laboratories. 28 students participate in the experiment, and each student is taught in class on the topics of Ethernet LANs and the purpose and operation of ARP over class sessions that are 1-4 weeks prior to the laboratories.

The survey requires each student rate his or her agreement with five statements by selecting one of five numbers on a Likert-Scale of 1 – 5. The scale values correspond to an individual’s level of agreement or disagreement according to: 1 represents strongly disagree, 2 represents disagree, 3 represents neutral, 4 represents agree, and 5 represents strongly agree.

The five statements presented are:
1. I am able to define the purpose of the Address Resolution Protocol (ARP).
2. I am able to describe the operation of ARP in facilitating data link layer communication by using ARP request and ARP reply packets.
3. I am able to identify the different fields of an ARP packet.
4. I am able to identify the different fields of a data link layer (Ethernet) frame.
5. I am able to describe how the TCP/IP protocol suite utilizes physical, logical and port addresses to send/receive data.

Two sets of survey response distributions are found for the responses given before and after the laboratories. The two response distributions (before and after) corresponding to each statement are compared using a paired t-test.

The test consisted of five questions. Each question was designed to evaluate a student’s mastery of the four objectives a) – d) associated with the second laboratory. The actual test questions can be found in the appendix. The total pre-test and post-test scores for the students are collected, and the pre-test and post-test scores are compared using a paired t-test.
Results

Survey Results
The distributions of survey responses, both before and after the laboratory sequence, are illustrated via clustered histograms, one for each statement.

Figure 3 – Clustered Histograms of Responses to Statement 1 of Survey

Figure 4 – Clustered Histograms of Responses to Statement 2 of Survey
Figure 5 – Clustered Histograms of Responses to Statement 3 of Survey

Figure 6 – Clustered Histograms of Responses to Statement 4 of Survey
These data indicate an increase in students’ perceived mastery over the course material following the laboratory sequence. Additionally, there was less variability in students’ responses, suggesting this increase in perceived mastery was characteristic of the class as a whole and not only a few students.

A paired t-test to examine change in survey responses from pre- to post-testing was significant ($p < .0001$, 2-tailed), indicating students’ perceived mastery improved after completing the laboratory sequence.

**Test Score Results**
The box plots of pre-test and post-test scores are presented in Figure 8.
The maximum test score possible was 25. A paired t-test to examine change in scores from pre- to post-testing was significant ($p < .0001$, 2-tailed), indicating students’ performance improved after completing the laboratory sequence.

**Discussion**

The results show that student participation in these laboratories has resulted in improvements in their ability to interpret and analyze real-world data link frames and in their understanding of how TCP/IP protocols are used for information exchange between two hosts. This assertion is supported by student perception, and pre-test and post-test scores.

The improvement in student perception of their ability to accomplish the tasks presented via survey statements is illustrated by the transformation in histogram shape from approximately uniform distributions prior to the laboratories to distributions that are centered on the “agree” and “strongly agree” portion of the scale. This improvement in perception is further validated by the paired t-test results, which indicate a statistically significant difference between survey responses taken before and after the laboratory sequence.

Furthermore, the statistically significant difference between pre-test and post-test scores, as shown by the paired t-test, implies that there has been an improvement in student understanding and proficiency after completing the laboratory sequence.
Given that the mean value of the post-test scores was 18.53 (74%), it is clear that there is room for improvement. The average post-test scores were evaluated on a per-question basis, (Table 3) and this revealed poor performance on question 2 (54%) in particular. This question required each student to fill in ARP request and reply packets based on information given in the question. The main causes of student errors based teacher observation are:

i. Students failed to translate the IP addresses presented in decimal format in the question into hexadecimal format as required for the answer.

ii. Students failed to rearrange the source and destination fields from request packet to reply packet.

Table 3 – Average Post-Test Scores per Question

<table>
<thead>
<tr>
<th>Question Number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Score (%)</td>
<td>74.6</td>
<td>54.3</td>
<td>82.7</td>
<td>82.1</td>
<td>78.6</td>
</tr>
</tbody>
</table>

These errors can be remedied by emphasizing the conversion process of IP addresses from digital to hexadecimal values in the second laboratory and placing guidance in the laboratory manual with respect to the rearrangement of source and destination fields in ARP packets.

References

Appendix – Copy of Pre-test/Post-Test Questions

Test (25 points)

Instructions: The duration of this assignment is 20 minutes. Credit will be given for your work in addition to the correct answer. Please show your work in the space provided. The back of each page may be used to continue your solution.

Question 1 (2 points)
What is the purpose of the Address Resolution Protocol (ARP)?

Question 2 (6 points)
Consider a host A with IP address 10.220.7.73 and MAC address 00:26:82:02:8F:0D that wants to send data to another host B with IP address 10.220.7.1 and MAC address 00:0F:B7:21:4F:60 (Note: MAC address of B is not known by A initially). With reference to Figure 9.8 below, fill in the ARP packet fields with hexadecimal values based on what you would expect for an ARP request and an ARP reply.
Question 3 (4 points)
Name the 7 sections for the Ethernet MAC frame shown below.

Question 4 (7 points)
Consider the Ethernet MAC frame below with the preamble, SFD and CRC fields removed.

00 26 82 02 8f 0d 00 0f b7 21 4f 60 08 00 45 30 00 34 3f b4 00 00 31 06 6a ab 4a 7d 82 93 0a dc 07 49 00 50 c4 56 d5 b7 34 f7 68 84 4c 23 80 12 37 dc d4 1f 00 00 02 04 05 88 01 01 04 02 01 03 03 06

Answer the following questions
a) What is the destination MAC address for this frame? (2 points)
b) What is the source MAC address for this frame? (2 points)
c) What is the type for this frame? (1 point)
d) What is the length of this frame in bytes? (1 point)
e) What is the length of the data and padding field in bytes? (1 point)
Question 5 (6 points)
For the internetwork in Figure 1, process b on computer A is sending a message to the process g running on computer T. Show the contents of the frames at the data link layer for each hop along the route required to carry the message from computer A to T.

Figure 1