IQRA National University Peshawar

Name Noor rahman
Reg Id 14232
MSCS 2017-18
Subject Network Management
Department of Computer Science

The Journal of Supercomputing
Incorporating Evolutionary Computation for Securing Wireless Network against Cyber threats

Review of the paper

The More growth of internet services, the demand for protection and security of the network against sophisticated attacks is continuously increasing. Nowadays, in network security, an Intrusion Detection System (IDS) plays an important role to detect intrusive activity. Evolutionary Algorithms (EAs) based optimization techniques have been broadly applied to tackle network anomaly detection problems. With the purpose of reducing the search dimensionality and enhancing classification performance of IDS model, in the literature several hybrid evolutionary algorithms have been investigated for feature selection but they have few drawbacks such as poor diversity, slow convergence, and stagnation. To resolve these limitations, in this study, we introduce a new hybrid evolutionary algorithm combining the techniques of Grasshopper Optimization Algorithm (GOA) and Simulated Annealing (SA), called GOSA for IDS that extracts the most noteworthy features and eliminates irrelevant ones from the original IDS data sets. In the proposed method, SA is integrated into GOA, while utilizing it to increase the solution quality after each iteration of GOA. Support Vector Machine (SVM) is used as a fitness function in the proposed method to select relevant features which can help to classify attacks accurately. The performance of the proposed (GOSA) method is evaluated on two IDS data sets such as NSL-KDD and UNSW-NB15. From experimental results, we observe that the proposed method outperforms existing state-of-arts methods and attains high detection rate as 99.86 %, an accuracy as 99.89% and low false alarm rate as 0.009 in NSL-KDD and high detection rate as 98.85%, accuracy as 98.96% and low false alarm rate as 0.084 in UNSW-NB15.

In recent years, the cyber security has attained lots of consideration and provided an open area of The cyber security facilitates safety of information systems like hardware, software and re-lated organization, stored data on systems and the services offered by these systems that may be accessed in illegitimate way by intruders, and also may be misused Sometimes, a misuse or harm may be instigated by an operator of the system, intentionally. So, either intentional or accidental harm can be a reason to not conform the security measures. A cyber security guideline is desirable to safeguard computer systems with the intention of enforcing several organizations in addition to companies to secure their systems and data attacks. To deal with this, cyber security requirements are needed into consideration such as confidentiality, integrity, availability, authorization. Intelligent IDS system can play an important role to prevent network from cyber-attacks.

Compared with traditional network protection technology such as firewall, smart and man-centered IDSs that can take advantage to intercept and inform of network intrusion has a prodigious real-world value. It can be deployed either in a central network backbone or inside systems. The major issue is how to increase the efficiency of an intelligent intrusion detection system which becomes a central point of network security

At this time, the use of intelligent IDS is considered an emergent solution for network protection and security in contrast to peripheral attackers. On the contrary, lots of issues have been found in the present IDS, often the low detection rate in contrast to the new attacks, and IDS is heavily overloaded when working with audit data. In the past decades, several machine learning algorithms namely naive Bayes, multi-layer perceptron, support vector machine, and artificial neural network were utilized in IDSs for finding the type of attacks. It can distinguish attacks by examining the factors of the network data. The network data contain noisy attribute, and convinced other attributes which can decrease the detection accuracy, whereas the input of certain attributes increases the accuracy of intrusion detection system. From now, to assist as the input to the learning approaches, the selection of informative attributes is vital for an IDS.

As a wrapper approach, Simulated Annealing (SA) is a popular heuristic approach which is based on the Metropolis Monte Carlo algorithm prolonging local search methods with an explicit scheme to spurt from local optima. The major idea is to consider those movements that yield worse quality solutions than the current solution, so that it can spurt from local optima. In addition, less computational effort, high speed, and a limited number of parameters are some benefits of the SA when compared with other met heuristics such as PSO, GOA, and TS for intrusion detection which work as wrapper methods. In 2017, Sarema hotel. have proposed a capable nature-inspired algorithm, called GOA which is encouraged from the ideal swarm behavior of insects in nature. In addition to, GOA has several merits like easy to implement, fast convergence, less control parameter to adjust, less time of computation, high probability and efficiency in obtaining global optimum.

Related to other EAs, GOA is also prone to get stuck in local optima when applied with various domains of problems. Contrary to GOA, Simulated Annealing (SA) contains a stochastic heuristic mechanism to jump out of local optima investigated by Kirkpatrick et al Indeed, SA will absolutely acquire the global opti-mum point from the physical progression of the solids annealing process, if very low deviations of controlling parameter and steadiness in each iteration of cooling are sustained. This technique iteratively attempts to enhance the quality of solution regarding the criteria of the objective function. In this, the improved solution will be acknowledged and on the contrary, solution having worse impact is also acknowledged with some probability based on the Boltzmann probability concept. Amongst all met heuristics methods, GOA and SA have augmented because of their simpler concepts, accelerated good quality solution and e artlessness of employment.

In general, algorithms which are inspired by nature, plausibly divide the research process into two phases: exploration and exploitation. In the exploration phase, agents are stimulated to move involuntarily, while they incline to run locally in the space of exploitation. Widespread exploration and prompt convergence inspire the employment of GOA. The mathematical model used to simulate the swarm behavior of grasshoppers.

The main reason behind GOA and SA hybridization is to develop the social conduct of an agent (swarm), by updating in the best position of agent. In the proposed model, GOA and SA algorithms have integrated to get benefitted from their merits. When there is no progress perceived in an iteration of GOA, the old best position is changed by an updated best position calculated using SA in the same iteration. This updated position truly provides a signal to the leader of the agents to which the previous best solution value was related to, for updating its direction through communication or interaction with each other. The proposed GOSA method responds between GOA and SA and incorporates the merits of exploration proficiency of GOA and local search ability of SA. In this way, the proposed method also increases the classification performance obtaining the tuned parameter of SVM, avoiding stagnation and improving convergence rate and also decreases complexity while searching to generate the attribute sets to enrich the performance of the anomaly-based system. The abstract model of the proposed GOSA is illustrated in Figure 3. The algorithm begins with GOA generating a population named as a candidate solution and introducing an initial temperature. Temperature drops constantly after each iteration in GOA conferring to the described cooling plan. If there is no progress during an iteration in the current best solution, that current solution will be presented to the SA algorithm for enumeration of an improved position of agents. SA solution process continues until a rejection occurs. Consequently, the best solution through SA is chosen then return as the new best solution to GOA. Similarly, the search process is repetitive with GOA and SA utilizing the current best solution until convergence criteria are satisfied. Afterward selected reduced attributes subset with optimal tuning parameter and attributes are used to classify through multi-class SVM classifier whether there is an attack or not which is described in the above subsection. The steps for performing the proposed hybrid GOSA method is illustrated in Algorithm 4.

With the intention of evaluating the efficacy and performance of the anomaly based IDS using proposed GOSA technique, there has been executed the simulation through NSL-KDD and UNSW-NB15 data sets which are available openly for intrusion detection estimation. The reasons behind the selection of these kinds f data sets in the work are that these have utilized primarily in intrusion detection proposals to assess the DS performance and to identify the type of attacks for experimental analysis. Each record in these data ets is recognized as either an attack or normal. Additionally, these data sets have dissimilar data values and everal attribute numbers that meet exhaustive tests to validate attribute selection methods. We have also elected these two sets of data for evaluating the performance of the differential system. From [52], Tavallaee t al. have suggested a premeditated form of the KDD Cup 99. The UNSW-NB15 data set has produced in he Lab of Cyber Range of Australian Centre for Cyber Security (ACCS) from existent normal events and mitated modern-day attacks. This data set comprises 49 features and nine sorts of new attacks. Recently 20], Haji Salem and Babied have carried out experiments on NSL-KDD and UNSW-NB15 data sets to assess he performance of IDS in the identification of attack. Altogether, NSL-KDD [53] and UNSW-NB15[54] data ets are two recognized data sets that are utilized to estimate detection through intrusion techniques. Thus, we have designated data sets to authenticate the proposed method. In the next subsection, we illustrate the performance measures for evaluating the performance of proposed IDS system.

The UNSW-NB15 data set is the most predominant and comprehensive data sets for identification of intrusion and is extensively pragmatic in the evaluation of intrusion detection systems. In general, several recent studies exhibit that KDD data sets cannot entirely identify recent low impression attacks, also contain ing unbalanced normal and abnormal records. UNSW-NB15 is a data set that collected by ACCS over IXIA Perfect Storm tool to form hybrid real recent normal and imitated abnormal traffic records of the network. The UNSW-NB15 data set comprises 49 features including float, binary, integer, and nominal records. The Bro, Argus-IDS tools and twelve algorithms in C# are used to expect these features. The main information of the data sets is depicted in Table 2.

Hybridized feature selection technique which incorporates GOA and SA algorithms have implemented under MATLAB R2016a environment. All the observed experiments in this study have conducted on a computer with a Pentium Core i7 CPU with 16 GB RAM running Windows 8 operating system. In this paper, the detectors have trained and tested utilizing the NSL-KDD and UNSW-NB15 data sets to evaluate the performance of our approach. Take SVM as a classifier for choosing the best parameters that can affect the performance of our approach unwillingly, so we have employed the LibSVM tool[55] . For the proposed method GOSA, a number of iterations are set as 100, meanwhile, we perceived that after 100 iterations, there is no progress in the performance of classifier for the proposed technique and population size is set as 50. To reduce unit variances of features and to avert features having large value ranges in comparison to those with small value ranges, we convert the data in the normalized form in a range [0, 1]. To train and test the proposed intrusion detection technique, a stratified 10-fold validation technique is employed.

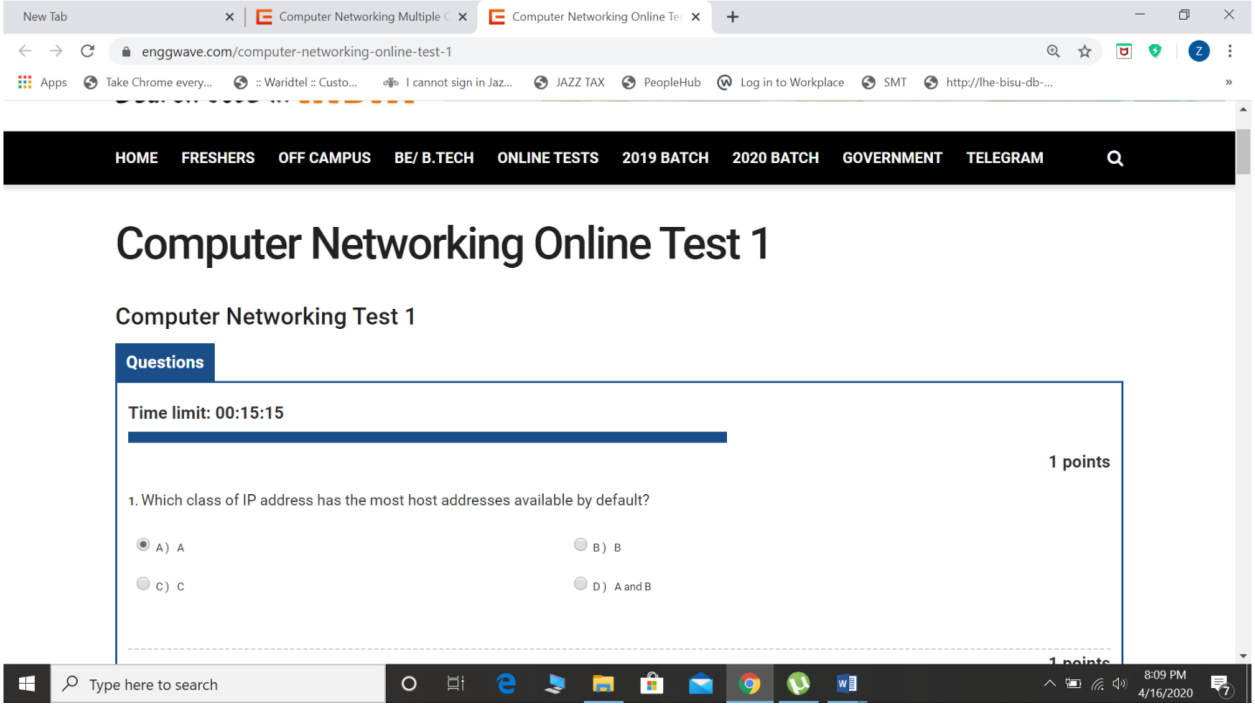
In this segment, the experimental outcomes of the proposed GOSA technique is investigated to conclude the efficacy of its in the intrusion detection system. In this context, acquired outcomes are also compared to the result of another methods. To estimate the efficacy of the proposed approach, we have utilized the numerous tests as a part of our investigations on NSL-KDD and UNSW-NB15 data sets (can see Table 1 and Table 2). From this data set, we have generated small training data set along with testing dataset for each attack class utilizing stratified 10-fold validation. This dataset has employed for the experimental purpose. In the available literature, a lot of researchers performed several experiments on the NSL-KDD and UNSW-NB15 data sets and implemented various experimental tools and techniques.

To observe the performance of the multi-class classification problem, we utilize the ROC (Receiver Oper ating Characteristics) curve. It is one of the most significant assessment metrics for testing the performance of any classifier. In regard to a binary classifier, the ROC curve plots the true positive rate versus the false positive rate, across different settings of the classifier process. In Figure 6, the curve's x-axis displays the false positive rate, while the y-axis designates the true positive rate and illustrates that the area found under curve

with the value of (1) in case of the proposed method, point out that it is a suitable classifier to determine attack in both kinds of intrusion dataset compared to other wrapper technique such as GA,GOA,PSO,DE,SA. Figure 7 provides the stratified 10-fold validation outcomes of proposed GOSA,GOA, GA, PSO, DE, SA, and PCA using SVM concerning DR. A comparison of the detection rates for proposed (GOSA), GA-SVM, GOA-SVM, DE-SVM, SA-SVM, and PCA-SVM is shown in Figure 7. The figure illustrates that the detec ion rate of proposed hybrid technique is far higher compared to other wrapper and filter techniques with SVM on both data set.

In order to maximize classification performance and minimize the computational time of the intrusion detection system, in the available literature, several hybrid evolutionary algorithms have been introduced. Different techniques have the inimitable understanding to tackle cyber-security problems, so integration of more than one algorithm that can reduce the shortcomings of one another is the best way to increase the performance of IDS system. Although hybrid evolutionary algorithm leads to higher performance but it has few drawbacks such as high computational time and poor diversity. To overcome the existing issues, we have introduced a new hybrid evolutionary algorithm for feature selection, called GOSA which integrates the characteristics of GOA and SA algorithms. The proposed method not only improves the classification performance but also reduces the computational time. To improve the search capability and robustness during the evolution process, we have applied the SA mechanism after the completion of the GOA phase so as to obtain best solution. In addition, GOSA technique is also used to select suitable SVM parameters, which avoid over-fitting concern of SVM. Two intrusion data sets such as UNSW-NB15 and NSL-KDD have used for experimental work. It is conceded in two scenarios: firstly, the classifier is trained with all attributes and then classifier is trained with feature subset acquired from proposed (GOSA) technique. The rational results attained from the experimental study demonstrated that GOSA performed better than the prevailing techniques concerning different performance measures such as detection rate, false alarm rate, accuracy, execution testing time, and execution training time. The proposed technique has provided a high detection rate of 99.86% accuracy of 99.89% and a low false alarm rate of 0.009 in NSL-KDD and high detection rate of 98.85%, the accuracy of 98.96% and a low false alarm rate of 0.084 in UNSW-NB15. This quality constructs the model computationally more efficient.

MCQES



Computer Networking Online Test 1

Computer Networking Test 1

Questions

Time limit: 00:15:15

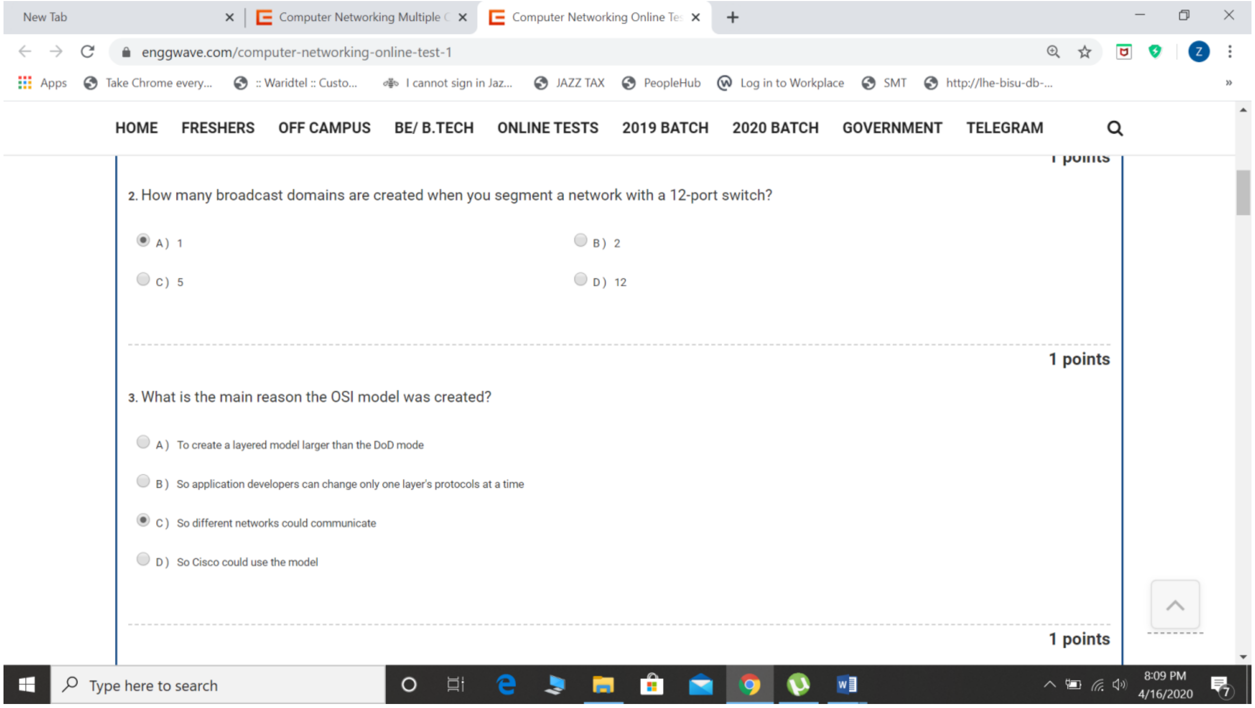
1 points

1. Which class of IP address has the most host addresses available by default?

● A) A ○ B) B

○ C) C ○ D) A and B

1 points



1 points

2. How many broadcast domains are created when you segment a network with a 12-port switch?

● A) 1 ○ B) 2

○ C) 5 ○ D) 12

1 points

3. What is the main reason the OSI model was created?

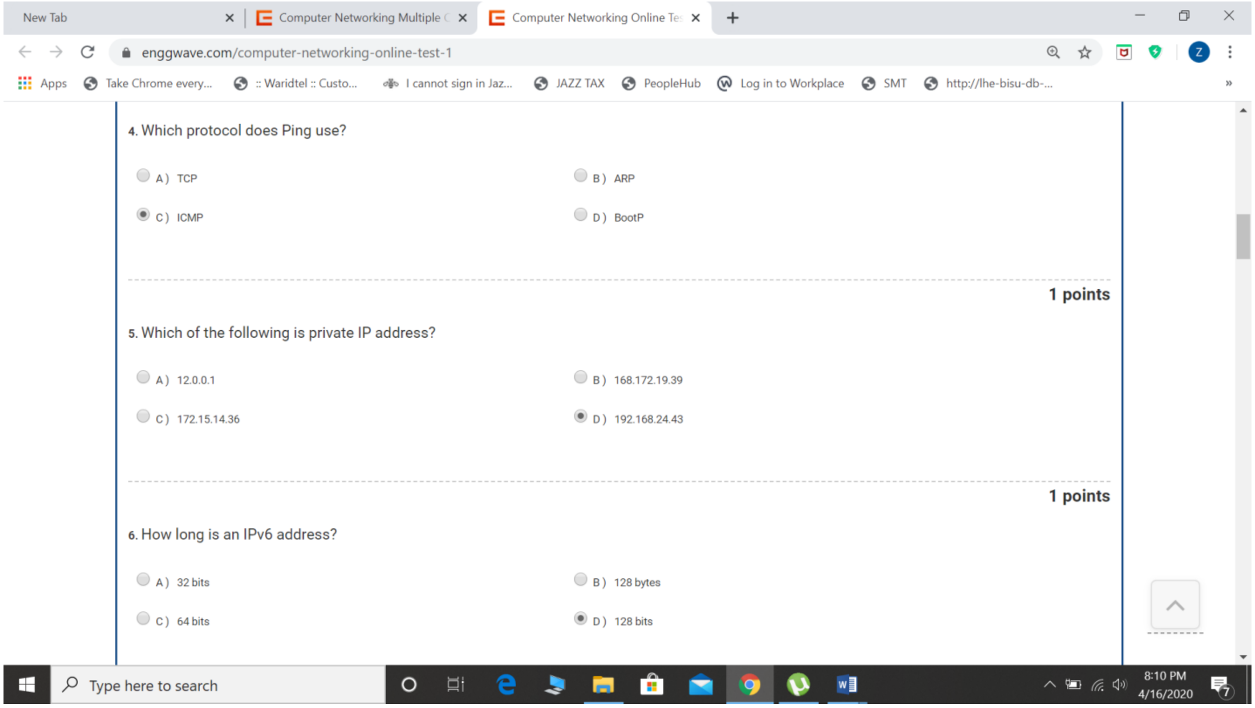
○ A) To create a layered model larger than the DoD mode

○ B) So application developers can change only one layer's protocols at a time

● C) So different networks could communicate

○ D) So Cisco could use the model

1 points



4. Which protocol does Ping use?

○ A) TCP ○ B) ARP

● C) ICMP ○ D) BootP

1 points

5. Which of the following is private IP address?

○ A) 12.0.0.1 ○ B) 168.172.19.39

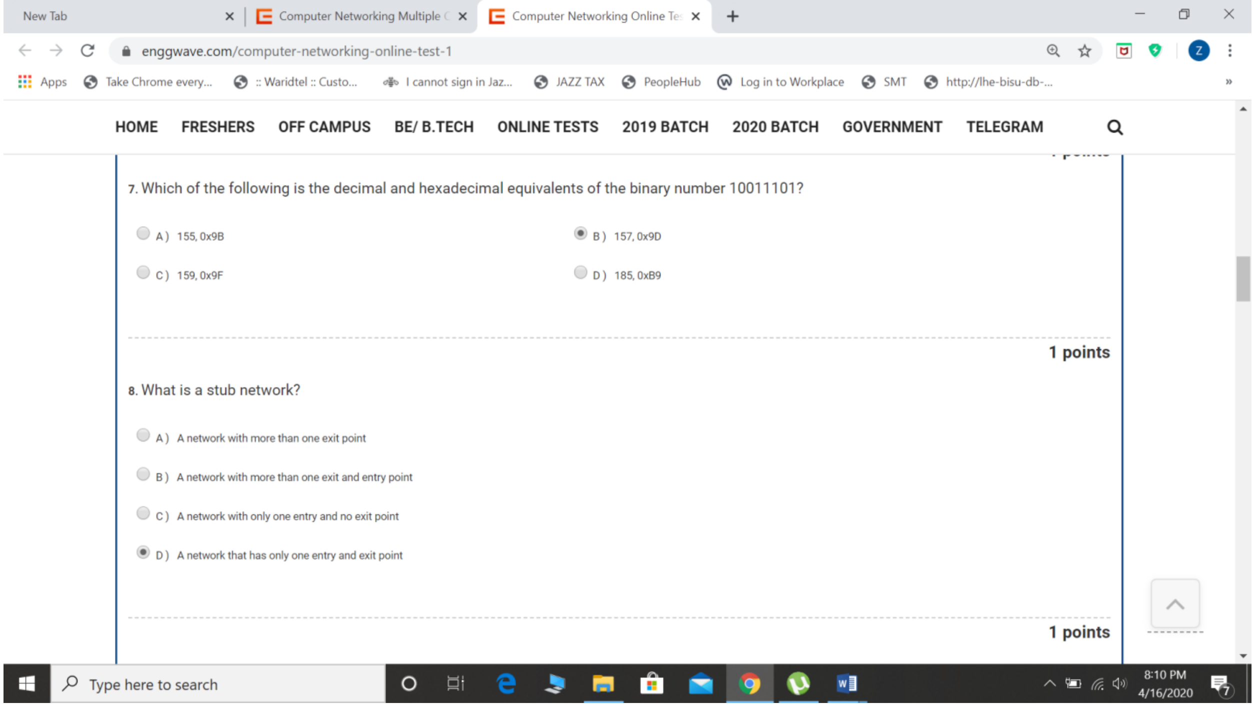
○ C) 172.15.14.36 ● D) 192.168.24.43

1 points

6. How long is an IPv6 address?

○ A) 32 bits ○ B) 128 bytes

○ C) 64 bits ● D) 128 bits



7. Which of the following is the decimal and hexadecimal equivalents of the binary number 10011101?

○ A) 155, 0x9B ● B) 157, 0x9D

○ C) 159, 0x9F ○ D) 185, 0xB9

1 points

8. What is a stub network?

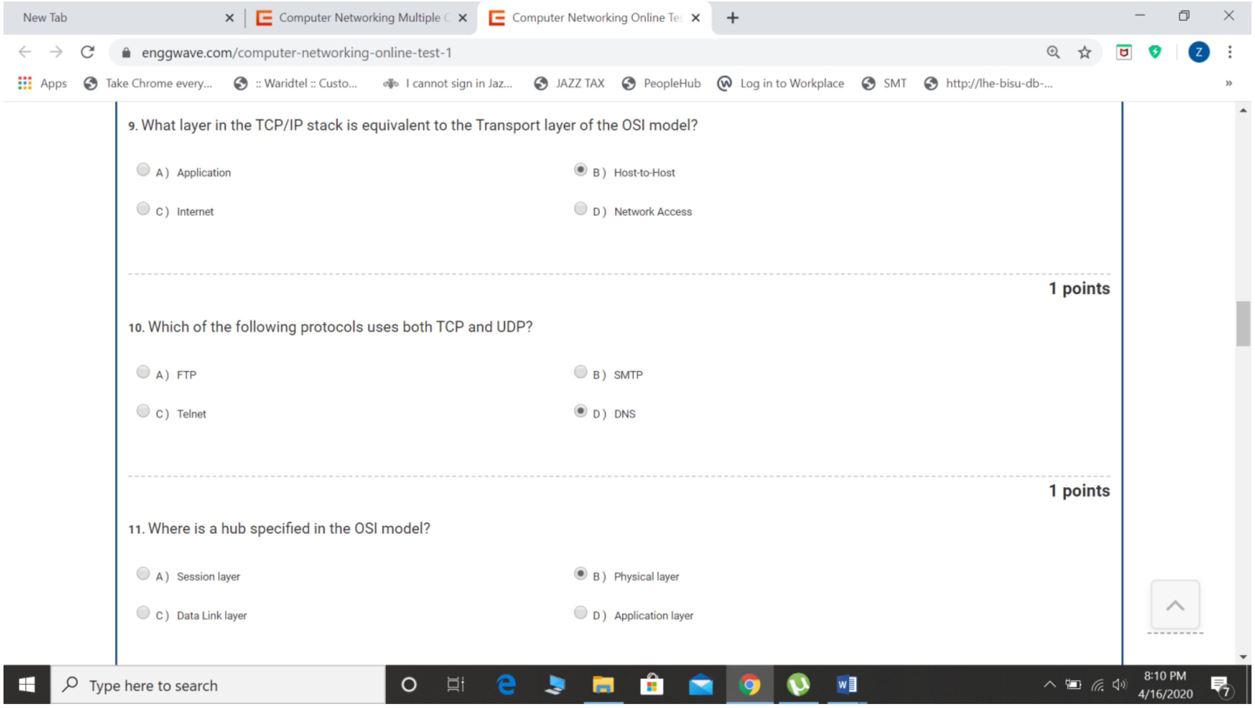
○ A) A network with more than one exit point

○ B) A network with more than one exit and entry point

○ C) A network with only one entry and no exit point

● D) A network that has only one entry and exit point

1 points



9. What layer in the TCP/IP stack is equivalent to the Transport layer of the OSI model?

○ A) Application ● B) Host-to-Host

○ C) Internet ○ D) Network Access

1 points

10. Which of the following protocols uses both TCP and UDP?

○ A) FTP ○ B) SMTP

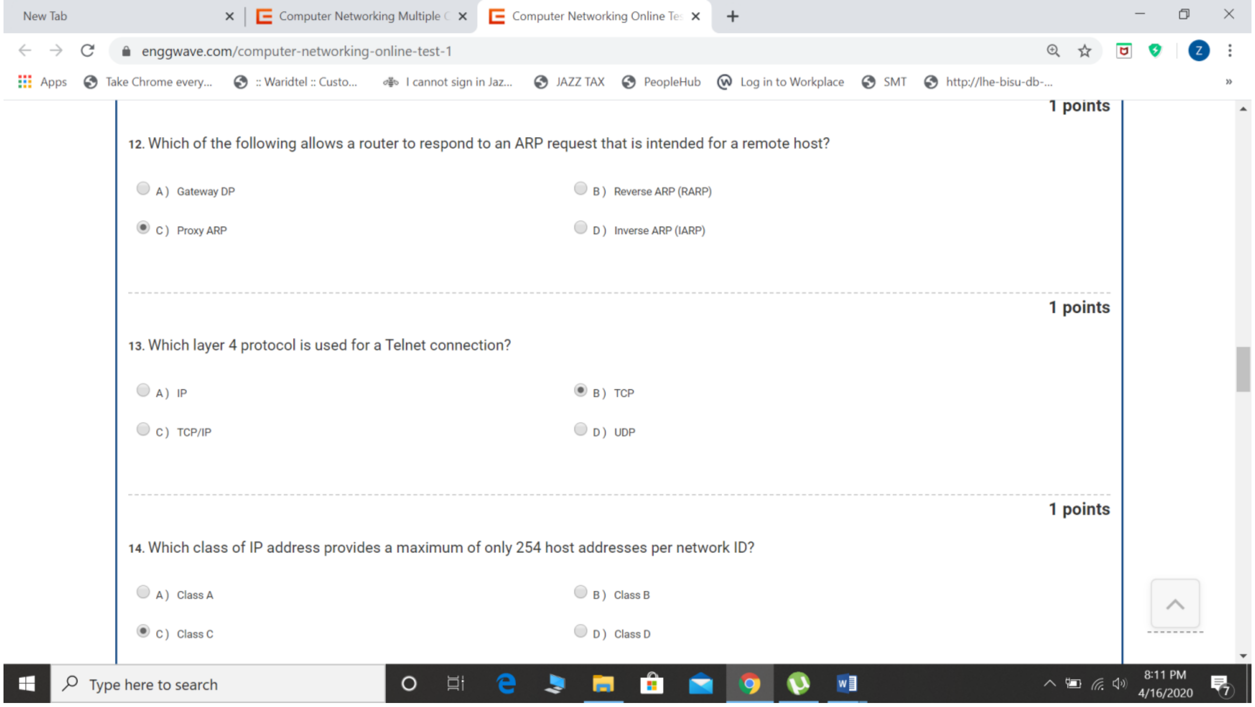
○ C) Telnet ● D) DNS

1 points

11. Where is a hub specified in the OSI model?

○ A) Session layer ● B) Physical layer

○ C) Data Link layer ○ D) Application layer



1 points

12. Which of the following allows a router to respond to an ARP request that is intended for a remote host?

○ A) Gateway DP ○ B) Reverse ARP (RARP)

● C) Proxy ARP ○ D) Inverse ARP (IARP)

1 points

13. Which layer 4 protocol is used for a Telnet connection?

○ A) IP ● B) TCP

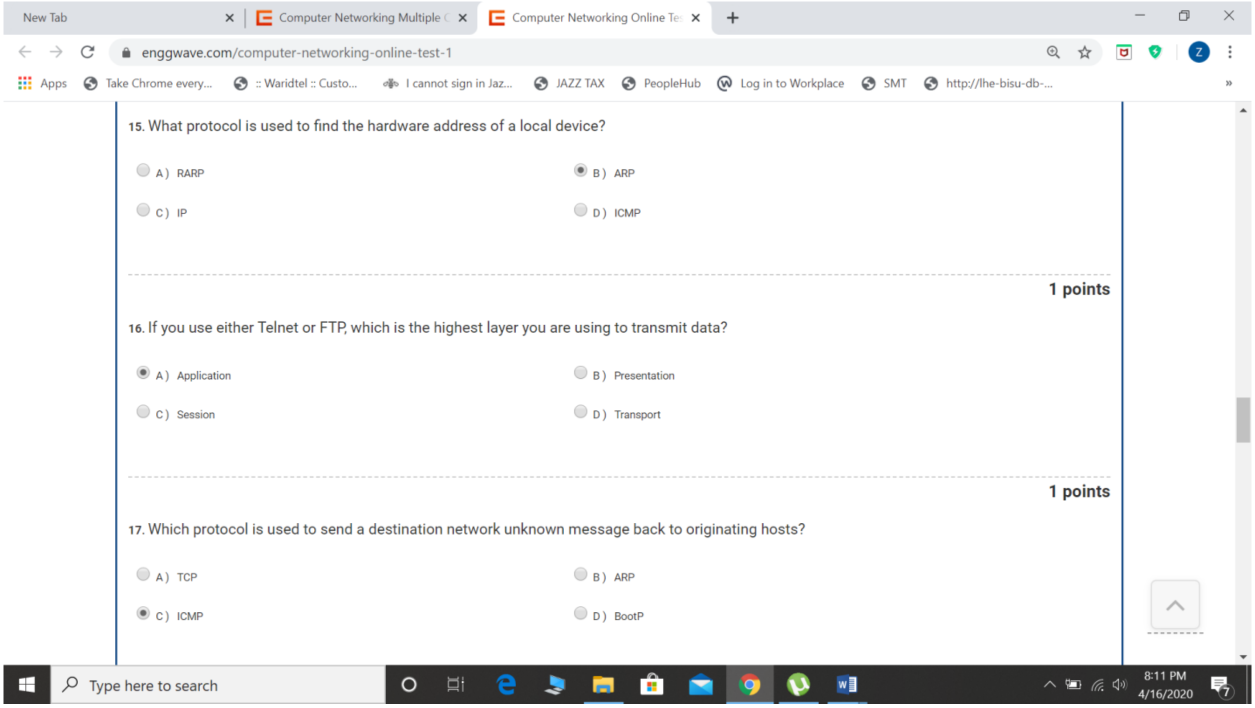
○ C) TCP/IP ○ D) UDP

1 points

14. Which class of IP address provides a maximum of only 254 host addresses per network ID?

○ A) Class A ○ B) Class B

● C) Class C ○ D) Class D



15. What protocol is used to find the hardware address of a local device?

○ A) RARP
● B) ARP
○ C) IP
○ D) ICMP

1 points

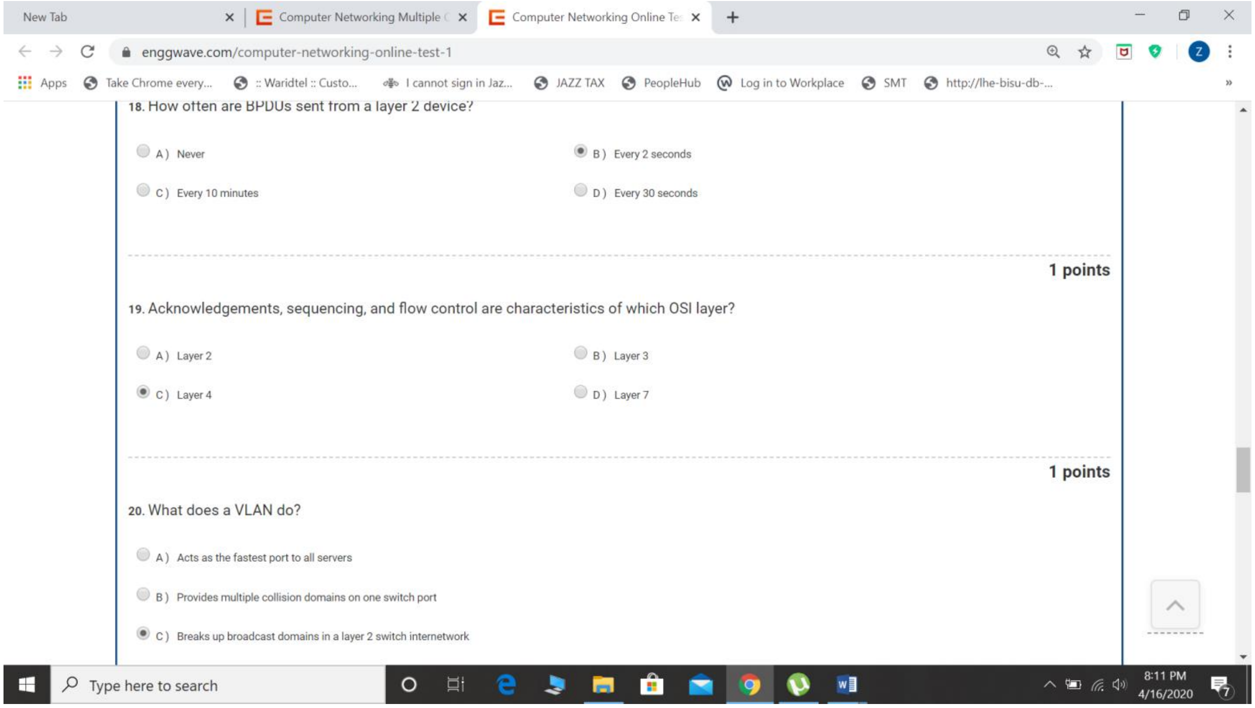
16. If you use either Telnet or FTP, which is the highest layer you are using to transmit data?

● A) Application
○ B) Presentation
○ C) Session
○ D) Transport

1 points

17. Which protocol is used to send a destination network unknown message back to originating hosts?

○ A) TCP
○ B) ARP
● C) ICMP
○ D) BootP



18. How often are BPDUs sent from a layer 2 device?

○ A) Never
● B) Every 2 seconds
○ C) Every 10 minutes
○ D) Every 30 seconds

1 points

19. Acknowledgements, sequencing, and flow control are characteristics of which OSI layer?

○ A) Layer 2
○ B) Layer 3
● C) Layer 4
○ D) Layer 7

1 points

20. What does a VLAN do?

○ A) Acts as the fastest port to all servers
○ B) Provides multiple collision domains on one switch port
● C) Breaks up broadcast domains in a layer 2 switch internetwork