COMP5423 Internet Infrastructure and Protocols

Survey on Text Classification

Written Report

**Group Members & Contribution**

|  |  |
| --- | --- |
| Student ID | Name |
| 19074889G | WONG CHI HANG |

Table of Contents

[Task Definition 3](#_Toc70432360)

[Single Label Classification 3](#_Toc70432361)

[Multi-Label Classification 3](#_Toc70432362)

[Binary Classification 3](#_Toc70432363)

[Challenges 4](#_Toc70432364)

[Ambiguity 4](#_Toc70432365)

[Entity identification 4](#_Toc70432366)

[Ambiguity 4](#_Toc70432367)

[Typical methods 4](#_Toc70432368)

[TF-IDF 4](#_Toc70432369)

[FTPS on Windows 5](#_Toc70432370)

[State-Of-The-Arts 5](#_Toc70432371)

[Transformer 5](#_Toc70432372)

[Bert 5](#_Toc70432373)

[XLNet 5](#_Toc70432374)

[Application scenarios 6](#_Toc70432375)

[Email Filtering 6](#_Toc70432376)

[Conclusion 6](#_Toc70432377)

[Reference 7](#_Toc70432378)

Task Definition

FTP is commonly used network protocol for file transfer over the network. However, it was not designed to be a secure protocol. Many security issues have arisen and listed in RFC 2577.

To overcome the security issue, Secure FTP (SFTP) and FTP-SSL (FTPS) are often used as an alternative. However, so many confused them with each other. Although both of them are serving same propose for file transfer, they are entirely different protocol.

In this project, we are going to compare these two quite similar protocols to find out what are the differences and then we try to implement both protocols on Linux and Windows server. Finally, we try to apply one of it as example to solve a real-world stimulated case.

Here is some brief information about FTPS and SFTP.

Single Label Classification

FTPS (FTP- SSL) is an extension of File Transfer Protocol (FTP) that adds support for the Secure Sockets Layer (SSL) and Transport Layer Security (TLS). However, the former (SSL) has been prohibited due to many serious vulnerabilities discovered.

Multi-Label Classification

FTPS (FTP- SSL) is an extension of File Transfer Protocol (FTP) that adds support for the Secure Sockets Layer (SSL) and Transport Layer Security (TLS). However, the former (SSL) has been prohibited due to many serious vulnerabilities discovered.

Binary Classification

FTPS (FTP- SSL) is an extension of File Transfer Protocol (FTP) that adds support for the Secure Sockets Layer (SSL) and Transport Layer Security (TLS). However, the former (SSL) has been prohibited due to many serious vulnerabilities discovered.

Challenges

For the authentication, SFTP supports SSH Key authenticate, while FTPS supports X.509 certificate.

FTPS supports anonymous login, however, SFTP must be an account for login.

FTPS supports server-to-server copy but SFTP is not.

Ambiguity

FTPS (FTP- SSL) is an extension of File Transfer Protocol (FTP) that adds support for the Secure Sockets Layer (SSL) and Transport Layer Security (TLS). However, the former (SSL) has been prohibited due to many serious vulnerabilities discovered.

Entity identification

FTPS (FTP- SSL) is an extension of File Transfer Protocol (FTP) that adds support for the Secure Sockets Layer (SSL) and Transport Layer Security (TLS). However, the former (SSL) has been prohibited due to many serious vulnerabilities discovered.

Ambiguity

FTPS (FTP- SSL) is an extension of File Transfer Protocol (FTP) that adds support for the Secure Sockets Layer (SSL) and Transport Layer Security (TLS). However, the former (SSL) has been prohibited due to many serious vulnerabilities discovered.

Typical methods

Bag of Words

Folder Structure

All connection can be successfully created, expect the connection to FTPS on Linux server.

The below error was prompted.

TF-IDF

To see the what are the difference in server-side implement, FTPS and SFTP server were implemented in Windows and Linux environment. The servers were set in Virtual Machine by using oracle virtual box.

The VM configuration show as below:

Word-Embedding

To see the what are the difference in server-side implement, FTPS and SFTP server were implemented in Windows and Linux environment. The servers were set in Virtual Machine by using oracle virtual box.

The error may due to incorrect configuration and bug in old software version. There could be a work-around solution, may try to find out later.

State-Of-The-Arts

Transformer

In client-side implementation, Java is used as the programming language for the development.

Two java class are developed. One for FTPS, and the other for SFTP.

Bert

For the FTPS implementation, Apache Commons Net™ library is used to access ftps server.

Below is the execution result that upload and download completed successfully.

XLNet

For the FTPS implementation, Apache Commons Net™ library is used to access ftps server

GLoVe

For the FTPS implementation, Apache Commons Net™ library is used to access ftps server

Application scenarios

From the study and implementation have done earlier, it seems that using SFTP in Windows environment is the most easy and efficient way for file transfer. To prove that, I have tried to implement a solution for a case scenario by using spring integration in Java.

Email Filtering

As the scheduler application was running, data files were generated at sftp download folder and ready for client to download.

News Filtering

As the scheduler application was running, data files were generated at sftp download folder and ready for client to download.

SftpDemo Module

The module is used to download the data file from server. To stimulate data processing, it appends a string to the file then upload to file server.

Also, uploaded to the sftp server.

Conclusion

Below files are submitted via blackboard.

1. FTPS\_and\_SFTP\_Comparison\_and\_Implementaion\_Group34\_19074889G.ppt
2. FTPS\_and\_SFTP\_Comparison\_and\_Implementaion\_Group34\_19074889G docx
3. FTPS\_and\_SFTP\_Comparison\_and\_Implementaion\_Group34\_19074889G\_workspace.zip

Reference

Computing for geeks (2019). *Install and Configure FTP Server on Windows Server 2019*. [online] Available at: https://computingforgeeks.com/install-and-configure-ftp-server-on-windows-server-2019/ [Accessed 13 May. 2020].

Microsoft. (2019). *New-SelfSignedCertificate.* [online] Available at: https://docs.microsoft.com/en-us/powershell/module/pkiclient/new-selfsignedcertificate?view=win10-ps [Accessed 13 May. 2020].

Computing for geeks. (2019). *Install and Configure OpenSSH Server on Windows Server 2019.* [online] Available at: https://computingforgeeks.com/install-and-configure-openssh-server-on-windows-server [Accessed 13 May. 2020].

Microsoft. (2019). *OpenSSH Server Configuration for Windows 10 1809 and Server 2019.* [online] Available at: https://docs.microsoft.com/en-us/windows-server/administration/openssh/openssh\_server\_configuration [Accessed 13 May. 2020].

server-world.info. (2019). *OpenSSH : SFTP only + Chroot.* [online] Available at: https://www.server-world.info/en/note?os=Ubuntu\_18.04&p=ssh&f=5 [Accessed 13 May. 2020].

How to forge. (2019). *How to install and configure VSFTPD.* [online] Available at: https://www.howtoforge.com/tutorial/how-to-install-and-configure-vsftpd/ [Accessed 13 May. 2020].