Open VPN Application in COVID-19 Pandemic

Zhengchun Zhou* and Tongcheng Huang
College of Information Engineering, Shaoyang University, Shaoyang 422000, China

*Corresponding author: 20156@hnsyu.edu.cn

Abstract. With the use of COVID-19 information, the way in which COVID-19 prevention and control agencies expand their internal networks is addressed. This article introduces the concept of open VPN technology and puts forward a design scheme using open VPN gateway access, which solves the problem of direct reporting of users accessing the system from the external network. At the same time, the paper uses big data analysis based on the ArcGIS Dashboard design to implement a visual and interactive real-time monitoring system of the global epidemic situation, which can timely and effectively obtain the development of the 2019 coronavirus disease epidemic. The system emphasizes fast and convenient online dynamic display of the global distribution of the epidemic, real-time statistical analysis of the spatial distribution of the epidemic, dynamic display of the severity of the epidemic in various countries and regions and the development trend of the global epidemic, and promotes epidemic monitoring to a certain extent Work development.

Keywords: COVID-19, open VPN, epidemic prevention and control, real-time monitoring.

1. Introduction
The epidemic caused by COVID-19 has disrupted people's daily life and physical health. With the spread of the epidemic, various epidemic problems and social contradictions have surged, and after the Spring Festival, a large number of people have returned from their hometowns to work in the city. This has once again formed a peak in population flow. The task of clearing stocks and controlling increments is particularly important. The data volume of the epidemic-related information is relatively large, or the classification set in a short time is not enough accurate. With the development of epidemic prevention and control work, hot issues in different periods are also changing in real time, such as where the masses gather during the epidemic period, which communities where tenants cannot enter the communities, which merchants are driving up prices, and which companies are experiencing problems with resumption of work and production? If such a large amount of epidemic-related hotspot information is manually screened and classified, there are problems such as slow speed, easy to miss and make mistakes, insufficient classification, insufficient focus, and real-time analysis that cannot keep up with the needs of decision-making [1]. If the information is insecure in the transmission process, it will cause information leakage or tampering, which will reduce the quality of the report and the normal progress of the subsequent prevention and control work, especially the reliable information used for the prevention and treatment of infectious diseases. Great harm, even causing social panic,
and endangering social stability. In order to ensure the security of information transmission, the use of VPN technology is a more cost-effective solution.

2. OpenVPN protocol

VPN (Virtual Private Channel) is a tunnel for secure data transmission between enterprises and enterprises or between individuals and enterprises. It provides identity authentication, data encryption, integrity protection and access control security services. OpenVPN is an application layer VPN based on the OpenSSL library. It is widely deployed due to its simple and easy-to-use features. OpenVPN relies on the security of OpenSSL, and establishes a TCP/UDP (usually UDP by default) secure tunnel on the client and server through designated ports, and then encrypts it in the TLS tunnel the communication data and the tunnel are shown in Figure 1.

OpenVPN provides two completely different authentication modes: TLS mode and pre-shared static key (PSK) mode. Pre-shared static key mode uses pre-shared static key to authenticate and encrypt identity. TLS mode uses certificate-based SSL/TLS the protocol performs authentication, establishes a secure tunnel, exchanges symmetric session keys, and uses the session key to encrypt the data tunnel. The TLS mode can ensure the secure distribution and update of symmetric keys, and thus has stronger security in real applications [2]. Therefore, this article mainly studies the OpenVPN protocol based on the TLS authentication mode. The OpenVPN secure communication process is shown in Figure 2, which relies on the following sub-protocols.
Figure 2. OpenVPN communication process in TLS mode

(1) OpenVPN handshake protocol, similar to TCP's three-way handshake process, the handshake data packet includes: the handshake request packet P_CONTROL_HARD_RESET_CLIENT_V1/V2 initiated by the client (V1/V2 represents the subsequent two different key negotiation methods, corresponding to each PSK authentication mode and TLS authentication mode, OpenVPN 2.0 and above use P_CONTROL_HARD_RESET_CLIENT_V2 by default), the server-side request response packet P_CONTROL_HARD_RESET_SERVER_V1/V2 (the key negotiation method is consistent with the request), and the client confirms the server-side response P_ACK_V1;

(2) OpenVPN control protocol, the control protocol includes two stages of TLS handshake and key negotiation, which are encapsulated in the P_CONTROL_V1 packet. When the TLS handshake is completed, the TLS encrypted tunnel has been established, and the session key negotiation information will be encapsulated in the TLS record Secure transmission in the layer. When needed, you can use P_CONTROL_SOFT_RESET_V1 to request session key negotiation;

(3) OpenVPN records the protocol. After the establishment of a secure tunnel and key exchange, the two parties conduct encrypted data communication P_DATA_V1. There is no detailed official specification for OpenVPN. In 2016, the Tomas Novickis method gave the expected expectations in the TLS mode. The OpenVPN state machine is shown in Figure 3. Novickis tried to derive the real state machine implemented by OpenVPN based on Leanly, but it encountered the following difficulties in the research process: (1) Using the python library Scaly can quickly construct data packets, but only tries to use the packet payload captured by Wireshark before For multiplexing, each data packet field is not carefully designed according to the protocol principle, resulting in failure to pass certificate verification or HMAC verification; (2) P_CONTROL_V1 There are still some problems when the data packet load is too large and needs to be transmitted in blocks. Therefore, the author got OpenVPN the goal of realizing the state machine has not been achieved [3].
3. COVID-19 OpenVPN demand analysis and solution deployment

3.1. Demand analysis
With the informatization process of the health department and the timely upload of the epidemic network report, especially the infectious disease network direct report users need to report the epidemic report to the disease control department daily, which requires effective and rapid access to the infectious disease information system. If the direct network user directly accesses the infectious disease information system through the public network, it may be stolen or tampered with by illegal users, which will cause serious health information security incidents. In order to ensure the security of the epidemic information system, the system is only open to legitimate users of the CDC, so that all direct network reporting users such as hospitals cannot access the system, resulting in a serious lack of system functions [4].

3.2. Solution deployment
Deploy VPN SSL gateway equipment on the firewall without changing the existing network topology. Configure NAT on the firewall and use the set IP address to access the Internet. Since OPEN VPN is an application layer gateway, it is safe and reliable. The application layer strategy can effectively protect the internal application server. The fourth port does not touch the Internet server. As long as the firewall's SSL (TCP443) port is opened, this design can guarantee the internal security of the web server. Connect an OPEN VPN device in the firewall and assign the set legal IP address [5]. Then configure the IP address, internal network interface, mask, gateway, etc., and then complete other corresponding settings. The network topology is shown in Figure 4.
4. Epidemic reporting system based on real-time monitoring

4.1. Text processing

The Chinese text after word segmentation is split from a string of string sequences into separate string lists, which are essentially unstructured character data, and input training model training in the later stage does not use computational operations. Therefore, the classifier needs to encode the finished text in a certain format. This process is called text encoding or text representation [6]. The selection of text representation model directly affects the effect of text feature extraction. In general, the model selects the appropriate text representation method according to the method of text feature selection. In order to solve the problem of high dimensionality and high sparsity caused by discrete representation, a distributed representation model is usually used to represent text. In the CBOW model, the input text is a word vector \( \{x_1, x_2, ..., x_n\} \) that has undergone one-hot encoding. By weighting and averaging the input vector, the output of each node is obtained as:

\[
u_j = \frac{1}{C} W (\Sigma_i x_i)
\]

Among them: \( v_{w_j} \) is the j column of the output matrix \( W \), then there is an output weight \( y_{c,i} \).

\[
y_{c,i} = p(w_{r,i} | w_i, \Lambda, w_c) = \frac{\exp(u_j)}{\sum_{j=1}^{v} \exp(u_j)}
\]

Ship-Gram uses back propagation algorithm and stochastic gradient descent algorithm to update the weights. First, define the conditional probability of input words to output words:
\[ E = -\log (w_j | w_i) = -v^T h - \sum_{j=1}^{V} \exp (v^T h) \]  

(3)

4.2. Overall system design

Based on the above two key technologies of text classification and named entity recognition, the author’s team has developed an intelligent epidemic prevention and control platform based on business needs, which mainly realizes the display of epidemic prevention and control maps, intelligent analysis of epidemic data, intelligent research and judgment of epidemic data, and Epidemic data spatial analysis and other functions [7]. The intelligent epidemic prevention and control platform is shown in Figure 5.

![Intelligent epidemic prevention and control platform](image)

Figure 5. Intelligent epidemic prevention and control platform

5. Conclusions

In summary, the ease of use and security level of OPEN VPN is higher than that of SLL VPN. With the increase in the amount of epidemic data worldwide, a visual and interactive global real-time monitoring system for the epidemic is designed based on OPEN VPN. Through the real-time epidemic monitoring system, the location and number of confirmed COVID-19 cases in all affected countries are displayed, the number of deaths and the recovery status, real-time data is fed back through the geographic area, which visually shows the spatial distribution and dynamic development of the epidemic. The system uses map big data to predict the population density, monitor the movement of the population related to the epidemic, report the situation of the epidemic inspection with one click, generate the epidemic data report in real time, and publish it to the society in a timely manner as needed, to study the situation of the epidemic to enhance the people’s awareness of prevention and control Detect outbreaks or epidemic clusters in high-incidence areas to provide powerful early warning functions, provide certain technical support for the establishment of a unified, efficient, fast, and accurate epidemic monitoring system, directly promote the development of epidemic monitoring, and ensure the sustainability of the national economy Development and social stability.

Acknowledgments

This dissertation is supported by Scientific Research Project of Hunan Provincial Department of Education (Research on virtual private network based on openvpn technology. Grant No:20C1686).
References

[1] Williamson, E. J., Walker, A. J., Bhaskaran, K., Bacon, S., Bates, C., Morton, C. E., & Cockburn, J. Factors associated with COVID-19-related death using OpenSAFELY. Nature, 584(7821) (2020) 430-436.

[2] Zhou, B., Wu, Q., Zhao, X., Zhang, W., Wu, W., & Guo, Z. Construction of 5G all-wireless network and information system for cabin hospitals. Journal of the American Medical Informatics Association, 27(6) (2020) 934-938.

[3] Palchaudhuri, S., Gabre, J., Prenner, S., & Solga, S. COVID-2019.2 Reboot: Returning a GI Fellowship to Pre-pandemic Practices. Digestive Diseases and Sciences, 65(9) (2020) 2461-2465.

[4] Wendelboe, A. M., Amanda Miller, J. D., Drevets, D., Salinas, L., Miller, E. J., Jackson, D., & Public Health Working Group. Tabletop exercise to prepare institutions of higher education for an outbreak of COVID-19. Journal of Emergency Management, 18(2) (2020) 1-20.

[5] Gajendran, N. (Ed.). Blockchain-Based secure framework for e-learning during COVID-19. Indian journal of science and technology, 13(12) (2020) 1328-1341.

[6] Shumba, K. G., Nyamaruze, P., Nyambura, V. P., & Meyer-weitz, A. Politicising the COVID-19 Pandemic in Zimbabwe: Implications for Public Health and Governance. African Journal of Governance & Development, 9(1) (2020) 270-286.

[7] Li, X., Guo, L., Sun, L., Yue, P. W., & Xu, H. Teleultrasound for the COVID-19 Pandemic: A Statement from China. ADVANCED ULTRASOUND IN DIAGNOSIS AND THERAPY, 4(2) (2020) 50-56.