基于 SSL 和国密算法的安全传输系统设计

代 乾 坤
(公安部第三研究所 上海 201204)

摘 要 针对某政务系统数据安全传输的需求,设计基于 SSL 及国密算法的数据安全传输系统。通过对 SSL、SM3、SM4、防中间人攻击、防重放攻击的研究,设计一套通过 SSL 通道交换预分配密钥加密的随机数,完成系统间双向握手,建立会话过程的系统实现机制。采用国密算法 SM4,既保证数据的安全,又因为采用对称加密算法,保证了数据加解密传输效率。通过时间戳有效防止应用数据重放攻击,通过消息鉴别码有效防止应用数据中间人攻击,增强了系统的安全性,达到了数据安全传输的设计目的。

关键词 SSL HTTPS SM3 SM4 重放攻击 中间人攻击

中图分类号 TP391 **文献标志码** A DOI:10.3969/j.issn.1000-386x.2023.02.051

DESIGN OF SECURE TRANSMISSION SYSTEM BASED ON SSL AND NATIONAL SECRET ALGORITHMS

Dai Qiankun
(*The Third Research Institute of the Ministry of Public Security, Shanghai 201204, China*)

Abstract Aimed at the requirement of data security transmission in a government system, a data security transmission system based on SSL and national secret algorithms is designed. Through the research of SSL, SM3, SM4, man-in-the-middle attack and replay attack, a set of random numbers of pre-distributed key encryption exchange was designed through the SSL channel, which completed the two-way handshake between systems, and established the system implementation mechanism of session process. The national secret algorithm SM4 ensured the security of data, and the symmetric encryption algorithm ensured the efficiency of data encryption and decryption transmission. The time stamp could effectively prevent the application data replay attack, the message authentication code could effectively prevent the application data man-in-the-middle attack, which enhanced the security of the system, and achieved the design purpose of data security transmission.

Keywords SSL HTTPS SM3 SM4 Replay attack Man-in-the-middle attack

0 引 言

没有网络安全就没有国家安全,就没有经济社会稳定运行,广大人民群众利益也难以得到保障。事实上,重大的网络安全事件对世界范围的网络安全威胁和风险日益突出。如 Facebook 泄露 5 000 万用户数据,一度造成该公司市值下跌 1 000 亿美元。2017 年,WannaCry 勒索病毒使得至少 150 个国家遭受了攻击,受害的电脑超过了 30 万台,造成损失达 80 亿美元。时至今日,勒索病毒一直威胁着网络信息安全。

目前的网络应用大部分基于 TCP/IP 协议设计开发,TCP/IP 协议最初设计并非为安全通信设计,所以在用户认证授权、数据安全传输等许多安全机制方面存在大量的安全漏洞,恶意第三方可以通过侦听破译、截获、篡改等非法手段,对业务应用系统进行破坏[1]。因此政务系统的数据敏感性、网络的虚拟性和开放性决定了包括 Web 应用系统在内的安全需要强有力的身份认证和数据安全传输机制来保证。本文设计的安全传输系统作为某部数据备案系统的一部分,录入的

收稿日期:2020-07-03。2017 年国家重点研发计划项目(2017YFB0802301,2017YFB0802302);上海优秀技术带头人基金支持项目(16XD1420200)。**代乾坤**,助理研究员,主研领域:网络信息安全。

数据较为敏感,因此数据的安全传输尤为重要。为了能够让备案数据在系统间进行安全传输,设计了本系统。

1　技术路线

1.1　备案系统网络拓扑结构

备案系统采用“门店-市-省-部”四级部署模式,备案客户端部署在各企业门店,根据各省具体情况,可选择省级部署或者“市-省”两级部署,各门店通过备案客户端采集备案数据,完成数据备案,备案客户端采集到的数据备案到市级数据平台,各地市级数据平台通过专用安全传输系统,上报备案数据到省级数据平台,省级数据平台在接收到各地市上报的数据后,放入消息队列,完成解密、解析并向部级数据平台报送,系统部署架构见图 1。本文针对“市-省”“省-部”之间的数据安全交换机制进行设计,并最终设计实现架构中的安全传输系统。

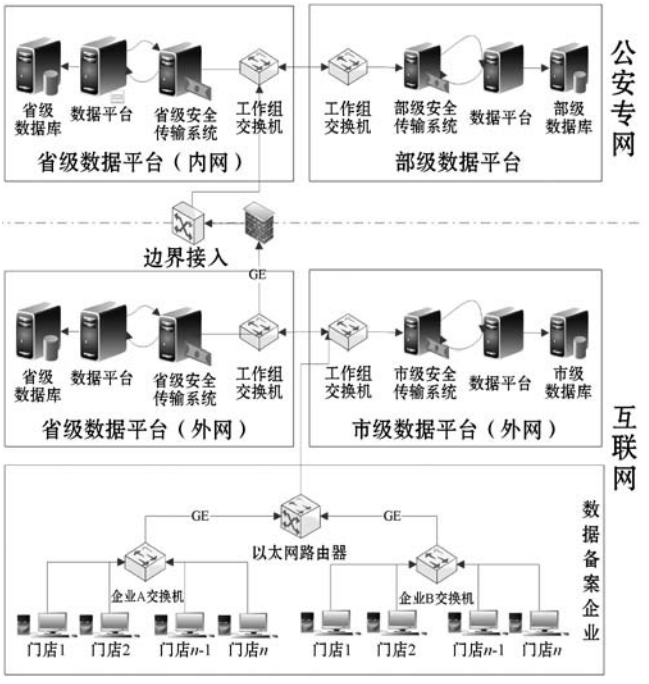

图 1　数据交换网络拓扑图

1.2　系统实现框架

安全传输系统采用多层架构设计,使用 TCP/IP 交互协议。根据各级单位业务量可采用均衡负载部署,均衡负载使用 Nginx 做业务分发,底层多节点服务集群。为了有效提高数据传输效率,接收到的待传输数据采用 ActiveMQ 消息中间件临时存储数据。MyBatis 作为数据库 ORM 框架,Durid 作为数据库连接池,Redis 缓存数据库资源。底层数据存储在磁盘阵列中,数据库采用 MySQL 主从模式部署,KeepAlived 保证主从切换机制。系统架构见图 2。

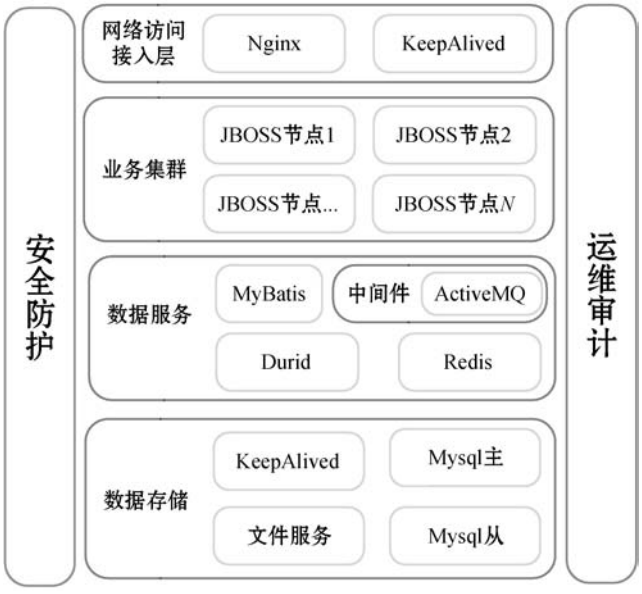

图 2　系统架构

1.3　安全传输系统工作流程

各级数据平台配备本级安全传输系统,平台需要上报或接收的数据,通过安全传输系统完成。安全传输系统在接收到平台传送的数据后,按照给定的目的地址完成安全传送,到达指定目的地址后,由目的地址数据平台完成数据解密、解析、存储、转发工作。例如,图 3 有数据平台 A、数据平台 B,分别部署了安全传输系统 A、安全传输系统 B。数据平台 A 与数据平台 B 的数据交互流程如下:

数据平台 A 上报数据到数据平台 B:

1）数据平台 A 把需要上报的数据通过图 3 中流程 1,明文交给安全传输系统 A,并指定要传送的目的地址数据平台 B,安全传输系统 A 把要传输数据放入消息中间件。

2）安全传输系统 A 在接收到数据平台 A 的传送指令后,根据目的地址,与安全传输系统 B 建立双向握手。

3）握手完成后,通过生成的对称密钥,完成数据加密。

4）加密后的数据通过图 3 中流程 5,完成到安全传输系统 B 的发送。

5）安全传输系统 B 在接收密文后,放入消息中间件,交给消息驱动 Bean 进行数据解密、解析。

6）解密后的明文数据放入消息中间件,消息中间件的消息驱动 Bean 通过图 3 中的流程 3 完成到数据平台 B 的传输。

7）数据平台 B 在接收到数据后,完成解析、入库,至此数据安全上报完成。

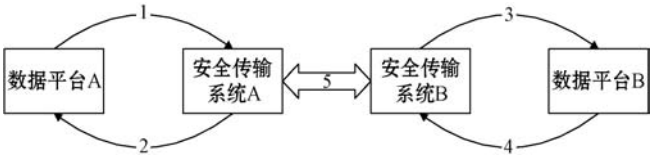

图3 安全传输系统工作流程示意图

数据平台 B 下发数据到数据平台 A：

1）数据平台 B 把需要下发的数据通过图 3 中流程 4，明文交给安全传输系统 B，并指定要传送的目的地址数据平台 A，安全传输系统 B 把要传输数据放入消息中间件。

2）安全传输系统 B 在接收到数据平台 B 的传送指令后，根据目的地址，与安全传输系统 A 建立双向握手。

3）握手完成后，通过生成的对称密钥，完成数据加密。

4）加密后的数据通过图 3 中流程 5，完成到安全传输系统 A 的发送。

5）安全传输系统 A 在接收密文后，放入消息中间件，交给消息驱动 Bean 进行数据解密、解析。

6）解密后的明文数据放入消息中间件，消息中间件的消息驱动 Bean 通过图 3 中的流程 2 完成到数据平台 A 的传输。

7）数据平台 A 在接收到数据后，完成解析、入库，至此数据安全下发完成。

2 应用安全关键技术

2.1 防中间人攻击设计

中间人攻击又称“MITM 攻击”（Man-in-the-Middle Attack），是一种间接的入侵式攻击，通过如 ARP 欺骗的方式收到客户端原应发给服务器的连接请求，再通过对服务器证书篡改，获得客户端的信任，同时采取一系列技术手段最终得到被加密的通信数据。显然，这是在通信双方毫不知情的情况下秘密进行的[2-3]。

在实施中间人攻击时，攻击者拦截客户端发送给服务器的请求，然后伪装成客户端与服务器进行通信，将服务器返回给客户端的内容发送给客户端，伪装成服务器与客户端进行通信。通过这个中间位置，便可以获取到客户端与服务器之间通信的所有内容。

防止中间人劫持的有效措施是能够校验客户端与服务器端的真实性。安全传输系统为了能够在两个安全传输系统间完成身份的真实性验证，增加了预分配密钥作为间接凭证。在双向握手过程中，随机数通过密文方式传输，传输时增加消息鉴别码，对随机数预分配密钥加密后进行 SM3 杂凑值计算得到消息鉴别码，因为随机数密文传输，保证了消息鉴别码的不可伪造，从而保证了中间人获取通信明文的可能，达到防止中间人攻击的目的。

2.2 防重放攻击设计

重放攻击（Replay Attacks）也称为回放攻击，即攻击者把以前窃听到消息或消息片段原封不动地重新发送给接收方达到对主体进行欺骗的攻击行为，其主要用于破坏认证正确性。重放攻击是攻击行为中危害较为严重的一种[4-5]。假如网络账户中的资金转出操作，一条消息表示用户发送了一个转账请求，攻击者在窃听到消息体后，可以通过多次发送这条消息而偷窃账户余额，从而使客户账户遭受损失。

安全传输系统中为了防止重放攻击进行了设计，增加了数据新鲜性检查机制。数据传输发起方在双向握手及发送数据过程中，增加了时间戳请求参数，且时间戳以密文方式传输，只有取得预分配密钥才能对密钥进行解析。接收方在接收到数据后，进行解密，并判断数据是否为新鲜发送的数据。

3 数据安全传输设计

3.1 基于 SSL 及 SM4 的双向握手

TCP/IP 协议组是目前使用最广泛的网络互联协议，作为 Internet 使用的标准协议集，是黑客实施网络攻击的重点目标。SSL 安全传输协议能够认证客户端与服务器，在客户端服务器间建立加密通道；加密数据以防止数据中途被窃取；维护数据的完整性，确保数据在传输过程中不被篡改。SSL 协议可分为三层：SSL 记录协议（SSL Record Protocol），它建立在可靠的传输协议（如 TCP）之上，为高层协议提供数据封装、压缩、加密等基本功能的支持。SSL 握手协议（SSL Handshake Protocol），它建立在 SSL 记录协议之上，用于在实际的数据传输开始前，通信双方进行身份认证、协商加密算法、交换加密密钥。SSL 警报协议，客户机和服务器发现错误时，向对方发送一个警报消息，如果是致命错误，则算法会立即关闭 SSL 连接，删除中间数据[6]。SSL 保证 TCP/IP 的通信应用程序间的隐私保护与完整性，Internet 的超文本传输协议 HTTP 可以使用 SSL 来实现安全的通信，即 HTTPS。

超文本传输安全协议（Hypertext Transfer Protocol Secure，HTTPS）是以安全为目标的 HTTP 通道，即 HTTP 下加入 SSL 层，HTTPS 的安全基础是 SSL。数据安全

传输系统采用 HTTPS 协议,建立会话后,采用 SM4 国密算法进行数据加解密。

根据 SSL 协议的优点,本文设计基于 RSA 服务器证书通过 HTTPS 请求建立 SSL 通道,通过 SSL 通道交换系统随机数密文,利用系统预分配密钥解密密文得到随机数,由双方随机数生成 SM4 加密密钥。至此基于 SSL 通道的双向会话连接过程结束,系统间建立了采用国密算法的数据加密机制。

安全传输系统间通过 SSL 交换随机数,随机数通过系统预分配密钥加密传输。以安全传输系统 A(简称系统 A)与安全传输系统 B(简称系统 B)的双向握手过程为例,获取 A 系统发送的随机数密文 A_{RM},用预分配密钥 SM4 解密后得到 A 系统随机数 A_R,同时 B 系统生成随机数 B_R,A_R 与 B_R 异或后得到加密密文 SK_B,同时返回 B 系统生成的随机数密文 B_{RM}。A 系统收到 B 系统响应后,获取 B 系统响应的随机数密文 B_{RM},用预分配密钥 SM4 解密后得到 B 系统随机数 B_R,A_R 与 B_R 异或后得到加密密文 SK_A,SK_A 与 SK_B 一样,至此双向握手建立完成,A 系统和 B 系统之间的安全传输通道建立完成。A 系统与 B 系统的双向握手过程(详见图 4)如下:

1) 系统 A 生成随机数 A_R = Random( ),并对随机数 A_R 用预分配密钥加密得到 A_{RM}。

$$A_{RM} = \text{SM4ENC}(KEY1, A_R)$$

2) 计算请求 MAC 值 A_{Hmac},获取时间戳 A_T。

$$A_{Hmac} = \text{SM3}(A_R)$$

$$A_T = \text{new NOW}( )$$

3) HTTPS 发送请求到系统 B。

4) 系统 B 完成接收,解析得到 A_{RM}、A_{Hmac}、A_T。

5) 预分配密钥解密 A_{RM},得到系统 A 的生成的随机数 A_{RL}。

$$A_{RL} = \text{SM4DEC}(KEY1, A_{RM})$$

6) 计算 A_R 的 MAC 值 A_{HmacL},判断 A_{Hmac} 与 A_{HmacL} 是否一致,不一致返回完整性校验失败,握手失败。

$$A_{HmacL} = \text{SM3}(A_{RL})$$

7) 获取服务器时间 N_T,计算时间戳 A_T 与服务器时间的差值是否大于 60 s,大于 60 s 返回请求超时。

$$N_T = \text{new NOW}( )$$

$$A_T - N_T > 60 \text{ s}$$

8) 获取系统 B 随机数 B_R = Random( ),分别计算 B_R 密文 B_{RM}、B_R 的 MAC 值 B_{Hmac}。

$$B_{RM} = \text{SM4ENC}(KEY1, B_R)$$

$$B_{Hmac} = \text{SM3}(B_R)$$

9) 计算会话密钥 SK_B 并放入内存。

$$SK_B = \text{SM4ENC}(KEY2, A_{RL} \text{ XOR } B_R)$$

10) 响应并返回 B_{RM}、B_{Hmac}。

11) 系统 A 获取响应参数 B_{RM}、B_{Hmac},解密 B_{RM} 得到 B_{RL}。

$$B_{RL} = \text{SM4DEC}(KEY1, B_{RM})$$

12) 计算 B_R 的 MAC 值 B_{HmacL}。

$$B_{HmacL} = \text{SM3}(B_{RL})$$

13) 计算 B_{RL} 的 MAC 值 B_{HmacL},判断 B_{Hmac} 与 B_{HmacL} 是否一致,不一致返回完整性校验失败,握手失败。

$$B_{HmacL} = \text{SM3}(B_{RL})$$

14) 计算会话密钥 SK_A 放入内存,至此双向握手结束。

$$SK_A = \text{SM4ENC}(KEY2, A_R \text{ XOR } B_{RL})$$

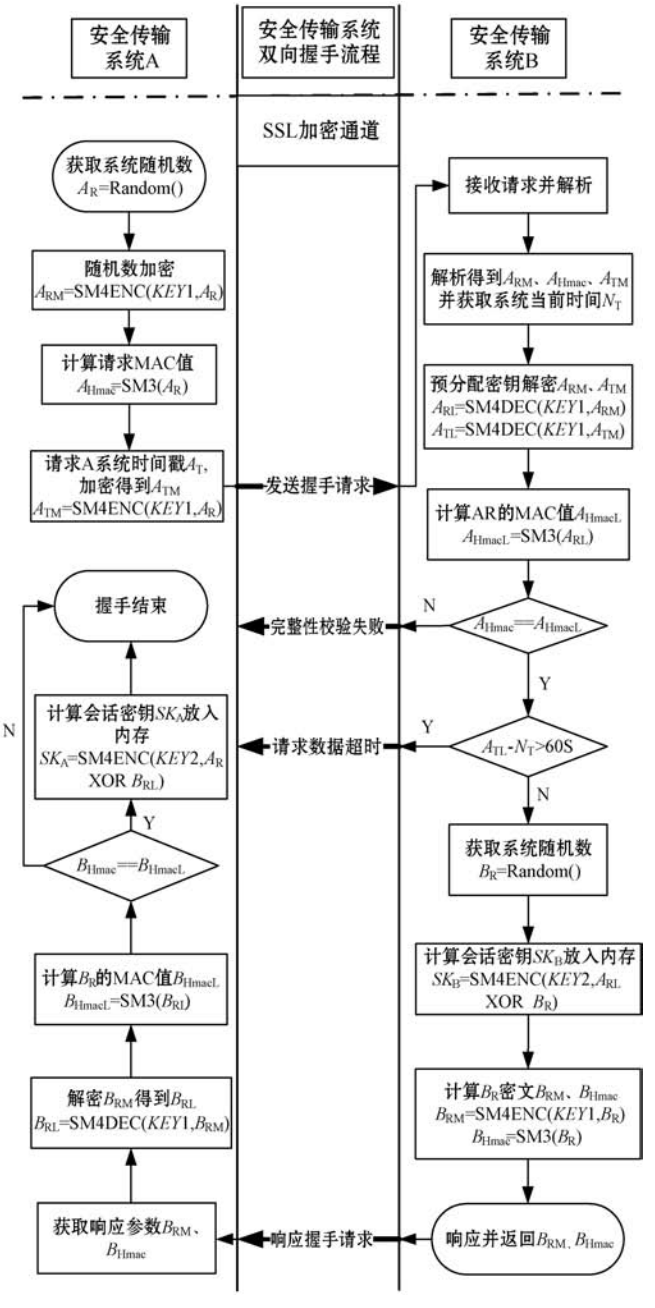

图 4　安全传输系统双向握手流程

3.2 SM4 加密算法保护数据安全

SM4 是我国商用分组密码算法。在商用密码体系中,SM4 是一种对称加密算法,其算法公开,分组长度与密钥长度均为 128 bit,加密算法与密钥扩展算法都采用 32 轮非线性迭代结构,S 盒为固定的 8 bit 输入 8 bit 输出[7-8]。

安全传输系统采用 HTTPS 进行随机数交换,交换后的随机数异或后采用 SM4 加密后得到会话密钥,系统采用 ECB 模式进行传输数据的 SM4 加密。

4 结 语

针对某政务系统数据上报、下发的需求,采用多层架构设计及 TCP/IP 交互协议,设计开发了基于 SSL 及国密算法的数据安全传输系统。通过研究 SSL、SM3、SM4、防中间人攻击、防重放攻击等关键技术,设计一套通过 SSL 通道交换预分配密钥加密的随机数,完成系统间双向握手,建立会话过程的系统实现机制。采用 ActiveMQ、MyBatis、Durid、Redis、业务集群等技术,有效保证了系统的稳定高效传输。系统采用国密算法 SM4,既保证数据的安全,又因为采用对称加密算法,保证了数据加解密传输效率。通过时间戳有效防止应用数据重放攻击,通过杂凑值有效防止应用数据中间人攻击,增强了系统的安全性,达到了数据安全传输的设计目的。

参 考 文 献

[1] 韦俊琳,段海新,万涛. HTTPS/TLS 协议设计和实现中的安全缺陷综述[J]. 信息安全学报,2018,3(2):1 - 15.

[2] 康荣保,张玲,兰昆. SSL 中间人攻击分析与防范[J]. 信息安全与通信保密,2010(3):85 - 87,90.

[3] 金敏捷,秦飞龙. 基于中间人攻击的 SSL 防范对策探究[J]. 船舶,2017,28(4):92 - 94.

[4] 陈宇琦. 一种基于时间戳的无线射频重放攻击抵御方案[J]. 现代计算机,2012(6):24 - 25,29.

[5] 肖斌斌,徐雨明. 基于双重验证的抗重放攻击方案[J]. 计算机工程,2017,43(5):115 - 120,128.

[6] 徐静,常朝稳. SSL 协议的安全性分析[J]. 微计算机信息,2006,22(9):19 - 21.

[7] 杨润东,李子臣. 基于国密算法的新型电子邮件加密系统研究与实现[J]. 信息安全研究,2018,4(11):1046 - 1051.

[8] 伍娟. 基于国密 SM4 和 SM2 的混合密码算法研究与实现[J]. 软件导刊,2013(8):127 - 130.

(上接第 325 页)

[5] 曹哲超,王轶骏,薛质. 基于页面标签和文本特征的暗网重要站点识别[J]. 通信技术,2019,52(12):3021 - 3026.

[6] MacQueen, J. Some methods for classification and analysis of multivariate observations[C]//Proceedings of Berkeley Symposium on Mathematical Statistics & Probability,1965.

[7] Kim Y. Convolutional neural networks for sentence classification[EB/OL]. [2023 - 01 - 01]. https://arxiv. org/abs/1408. 5882.

[8] Hinton G E. Distributed Representations[M]. Cambridge: MIT Press, 1986.

[9] Mikolov T, Chen K, Corrado G S, et al. Efficient estimation of word representations in vector space[EB/OL]. [2023 - 01 - 01]. https://arxiv. org/abs/1301. 3781? ref = hackernoon. com.

[10] Mikolov T, Sutskever I, Chen K, et al. Distributed representations of words and phrases and their compositionality [C]//Proceedings of the 26th International Conference on Neural Information Processing Systems, 2013.

[11] Joulin A, Grave E, Bojanowski P, et al. Bag of tricks for efficient text classification[C]//Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics,2017.

[12] Pennington J, Socher R, Manning C. Glove: Global vectors for word representation[C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP),2014.

[13] Peters M E, Neumann M, Iyyer M, et al. Deep contextualized word representations[EB/OL]. [2023 - 01 - 01]. https://arxiv. org/abs/1802. 05365.

[14] Devlin J, Chang M, Lee K, et al. BERT: Pre-training of deep bidirectional transformers for language understanding [EB/OL]. [2023 - 01 - 01]. https://arxiv. org/abs/1810. 04805v1.

[15] Dumais S T, Chen H. Hierarchical classification of web content[C]//International ACM SIGIR Conference on Research and Development in Information Retrieval,2000:256 - 263.

[16] Frank E, Bouckaert R R. Naive bayes for text classification with unbalanced classes[C]//European Conference on Principles of Data Mining and Knowledge Discovery, 2006: 503 - 510.

[17] Trstenjak B, Mikac S, Donko D. KNN with TF-IDF based Framework for Text Categorizatio [J]. Procedia Engineering, 2014,69: 1356 - 1364.

[18] Lin T, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//International Conference on Computer Vision, 2017: 2999 - 3007.