# 基于深度学习的隐蔽性有害信息特征研究

张安康，刘加兵
（国家计算机网络应急技术处理协调中心湖北分中心，湖北 武汉 430074）

**摘要：**针对互联网上的有害信息不断通过各种方式将核心内容隐蔽，从而逃避识别和检索的问题，文章以研究隐蔽性有害信息的特征为目的，以多家互联网公司的原始数据为基础，基于深度学习的经典 BERT 算法模型设计了一种方案，分三个阶段实现海量隐蔽性有害信息的智能文本分类，得到隐蔽性有害信息的三类特征的定量表征，为研究隐蔽性有害信息提供参考。

**关键词：**隐蔽性有害信息；文本识别；深度学习；特征分类
**中图分类号：**TP391　　**文献标识码：**A　　**文章编号：**2096-9759( 2023 )06-0146-04

## 1 引言

随着国内互联网的蓬勃发展，互联网上的诈骗、反动、钓鱼等各类有害信息也层出不穷[1-3]，严重危害了国家网络安全和社会公共利益。由于互联网信息量大、形式丰富，因此人工智能技术被广泛应用于数据挖掘，通过提取样本数据，利用深度学习模型实现有害信息的智能识别[4-6]。

然而，随着智能识别技术的不断发展，有害信息自身也在不断进化，通过各种方式将核心内容隐藏起来，从而逃避识别和检索，本文称之为隐蔽性有害信息。为研究隐蔽性有害信息特点，我们收集了本地多家互联网公司约 5 千万条原始数据，抽取其中部分有害信息进行初步分析发现，隐蔽性有害信息普遍不再直接展示完整网站、QQ 号码等内容，而是通过正常文字加特殊字符形式隐藏关键有害信息，增加识别难度。

本文为研究隐蔽性有害信息的特点，以前期收集到的本地多家互联网公司的原始数据为基础，基于深度学习的经典 BERT 算法模型，设计了一种方案，实现海量隐蔽性有害信息的智能文本分类，得到隐蔽性有害信息多维特征的定量表征，进而为智能识别隐蔽性有害信息提供一种参考方法。

## 2 基本模型简介

2018 年，谷歌旗下研究人员首先提出了 BERT 模型[7]。BERT 是一个预训练表征模型，采用 maskedlanguage model（MLM）技术来生成深层次的文本特征表示，而非像传统的 BiLSTM 模型采用浅层拼接双向语言模型的方法进行预训练，并引入了遮蔽式语言模型来随机遮蔽文本序列中的部分词语，BERT 模型是基于 Transformer 模型建立起来的，因此理解 BERT 模型首先需要了解 Transformer 模型。

Transformer 模型在 2017 年由 Shazeer 等人首先提出[8]，该模型采用多头自注意力计算方法，完全使用注意力机制来提取文本特征，与传统机器学习的 CNN 或 RNN 的计算方式完全不同。Transformer 模型的多头自注意力机制不依赖于前一时刻的计算，比传统的 CNN 或 RNN 模型的算法更优。Transformer 模型关注文本序列中的每一个单词，在解决长距离依赖问题上，由于最大路径长度为 1，因此在捕获长距离依赖关系问题上处理的更好[9]，使得该模型在自然语言处理领域效果显著。

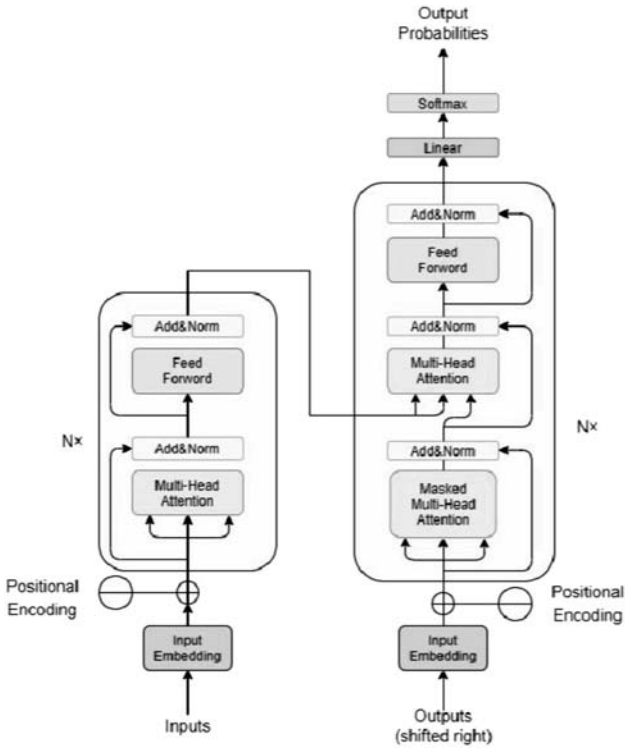


图 1 Transformer 模型架构示意图

Transformer 模型架构如图 1 所示，Transformer 模型由编码器和解码器两个部分组成。编码器由 n 个相同的编码层组成，并且由两个分层构成每个编码层。同样的，解码器也是由 n 个相同的解码层组成，但不同的是，解码层中有三个分层。解码层和编码层结构的不同在于解码层中多加入了一个子层，该子层是添加了 Masked 的多头自注意力层，Masked 操作是将文本输入序列中的某一个或多个词进行掩盖，这样在进行多头自注意力计算时掩盖后的词将不起作用。这层注意力层可以确保未来的信息不会影响预测第 i 个位置。

Transformer 模型除了编码层和解码层，还增加了数据预处理模块。Transformer 模型将文本输入的序列进行词嵌入，但 Transformer 模型在特征表示过程中仅使用了注意力机制，没有关注序列时间结构，所以不会考虑序列的顺序特征。为解决这一问题，增加了文本序列预处理模块，在该阶段阶段对文本进行位置嵌入。

**收稿日期：**2023-02-17
**作者简介：**张安康（1989-），男，湖北宜昌人，硕士研究生，中级工程师，主要研究方向为信息安全、通信信息系统；刘加兵（1993-），男，湖北黄石人，硕士研究生，初级工程师，主要研究方向为网络安全、通信信息系统。

文本序列输入中各个词向量的绝对位置和词向量间的相对位置需要在特征提取前加入序列中，因此Transformer模型使用不同频率的正弦和余弦函数，在特征提取前在文本序列中添加位置信息。主要规则为，将余弦编码使用在奇数位置，正弦编码使用在偶数位置，位置信息计算方法如公式(1)和(2)所示：

$$PE_{(pos,2i)} = \sin\left(pos/10000^{2i/d_{model}}\right) \quad (1)$$

$$PE_{(pos,2i+1)} = \cos\left(pos/10000^{2i/d_{model}}\right) \quad (2)$$

其中，pos表示单词的位置，i代表词嵌入中的第几维，表示单词的维度，不同的正弦曲线对应位置编码的每个维度。

Transformer模型是在没有添加其他结构的情况下，第一个完全用注意力机制搭建的模型[10]，采用多头自注意力机制对文本序列进行特征表示。Transformer模型和传统CNN模型和RNN模型采用完全不同的处理思路，在多个自然语言处理任务中明确缩短训练速度。

BERT模型在Transformer模型上更进一步优化，因此成为目前机器学习领域广泛应用的算法模型。BERT模型主要分为输入层、预训练层和输出层3个阶段。在输入层，由Token Embeddings、Segment Embeddings和Position Embeddings三个部分加权构成，以获得更多的文本信息。在预训练层，采用MLM对双向的Transformers进行预训练，以生成双向的深层文本特征表示。在输出层，对模型内部的参数等进行微调，最后得到的文本特征结果。

如图2所示，Transformer中的编码层组成了BERT模型中最基础模块BERT Layer（图中虚线小方框内部分），多个BERT Layer模块叠加组成BERT Encoder(图中虚线大方框内部分)，BERT Encoder最终构成完整的BERT模型。

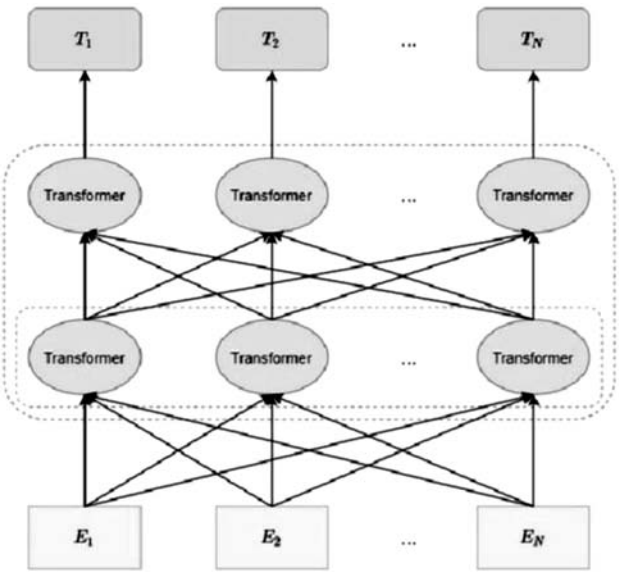


图2 BERT模型架构示意图

BERT采用了双向Transformer结构来确保获取所有的文本信息，每个单向的Transformer的每一个token只会关注到当前往左的token，通过双向的Transformer实现每一个token关注到左右两边的所有token，保障单词两侧的上下文信息都会被捕获。

BERT分为预训练和微调两个步骤。BERT通过在海量文本上进行预训练，为单词学习到一个特征表示，之后在下游特定的NLP任务中进行微调，来满足某一种具体的应用任务的需求。跨不同任务的统一体系结构是BERT的一个显著特点，预训练和最终下游任务的差异很小，将预训练模型使用到下游任务时只需要对输入输出层做一些修改就能使用，十分方便。

## 3 方案设计

本文设计一种方案，分为准备阶段、训练阶段、应用阶段3个阶段，利用BERT模型完成隐蔽性有害信息的特征分析，如图3所示。

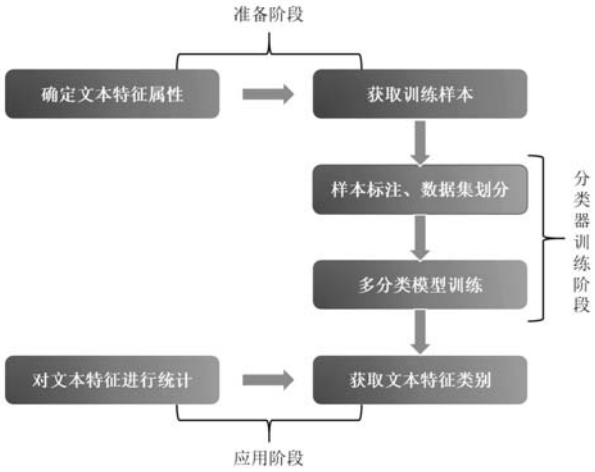


图3 隐蔽性有害信息方案设计示意图

### 3.1 准备阶段

在准备阶段，首先针对一定量的隐蔽性有害信息文本进行初步特征分析，分析的维度包括语言特征、文字特征、逻辑特征等。其次，根据初步分析的结果，制定每个特征维度的细化归纳分类，注意要保证各特征维度下的类别互斥。如语言特征可归纳为：仅含有特殊字符、不含特殊字符、正常文字和特殊字符同时包含三种完备的情况。最后，针对一定量的隐蔽性有害信息文本按照归纳的特征进行标注。如语言特征中可将仅含有特殊字符的文本标注为0、不含特殊字符的文本标注为1、同时包含特殊字符和正常文字的文本标注为2。

对一定量的数据观察后，本文将隐蔽性有害信息的文本特征按照语言特征、文字特征和逻辑特征3个维度进行划分。

语言特征表示隐蔽性有害信息是否为正常中文语言，具体可细分为仅包含特殊字符、不含特殊字符和特殊字符与正常文字同时包含三个完备情况。

文字特征表示隐蔽性有害信息种单个文字的种类，具体可细分为仅包含存一种特殊文字字体和多种字体混杂两种情况。

逻辑特征表示隐蔽性有害信息在内容方面的特点，具体可细分为：含有手机、QQ、微信等联系方式且不含有网址；含有手机、QQ、微信等联系方式且同时含有网址；仅含有手机、QQ、微信等联系方式且不含网址；不含手机、QQ、微信等联系方式但含有网址；既不含有网址也不含有手机、QQ、微信等联系方式共四类互斥情况。

### 3.2 训练阶段

在分类器训练阶段，针对每一个特征，如语言特征，将标注好的文本进行训练，形成该特征的分类器。同样以语言特征为例，以标签比例为参考进行数据集的等比例划分。划分好数据集后构建分类器模型进行训练。训练流程中监测训练集和验证集的recall、precision和F1值，以保证模型健壮可用。

在分类器的选择上，需要考虑隐蔽性有害信息的数据特点，选择较为合适的分类器。通过模型特点和数据特点双向考虑，最终选择BERT作为模型。理由如下：

隐蔽性有害信息中包含大量特殊字符，而 BERT 词表中包含有大量字符，当遇见不在词表中的字符是也会自动识别为特殊 token。使用 BERT 首先省去了人工收集词表的过程。其次，BERT 的词向量具有较好的表征能力。相较于训练时获取的词向量，在分类任务上有着更好的表现。最后，BERT 具有一定的语义理解能力，即对于是否含有特殊字符等问题天然具有更好的甄别能力，对特殊字符也天然的更为敏感。

利用 BERT 实现文本分类的方法如图 4，文本输入到 BERT 中经过 Embedding 层进行高位映射。BERT 提取特征后将输出特征的第一个向量作为输出，下游接入多分类模型获取各类别概率，为每一个文本特征形成一个分类器。

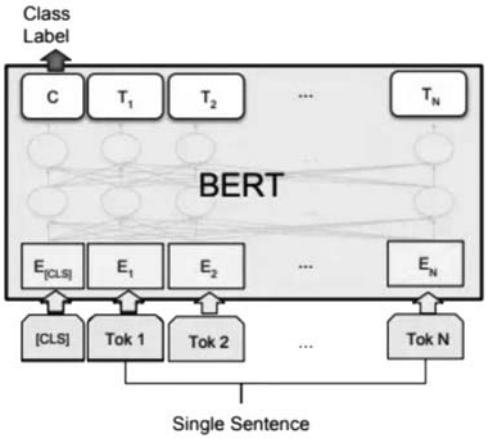


图 4 BERT 的文本分类训练示意图

### 3.3 应用阶段

在准备阶段的特征分析中，我们确定了隐蔽性有害信息的文本特征属性的 3 个维度，并分析罗列出各维度下的特征子类型。在训练阶段，经过标注和训练获取了对于特定文本特征子类型的分类器。

在应用阶段，遍历所有文本，以各文本特征的分类器对文本进行判别，利用分类器对海量隐蔽性有害信息进行特征统计，最后将统计特征汇总作为文本特征分布的参考依据。

流程示意图如图 5 所示。海量文本分别通过三个文本特征分类器，获取各文本对应隐蔽性有害信息特征属性的子类别。在特征分布统计中，对各子类别数量进行统计，以获得各子类别的分布情况。

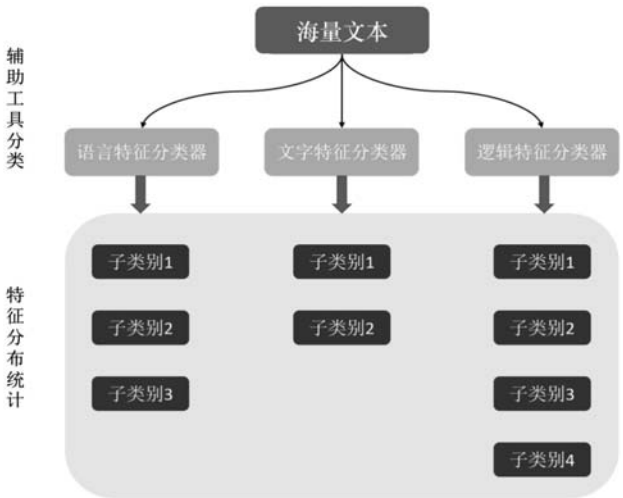


图 5 文本特征统计流程图

## 4 验证和结果

根据以上方案设计，我们使用 Ubuntu 18 平台和通用 BERT 模块实现本算法的文本特征分类器，对我们收集到省内多家互联网公司约 5 千万条有害信息原始数据进行分析，针对文本特征属性的 3 个维度分析统计结果如下。

（1）语言特征维度：通过分析输出结果发现，有害信息均含有部分特殊文字如火星文等，但并非全文都是特殊文字。

表 1 语言特征输出结果

| 子类别编号 | 子类别特征 | 占比 |
|---|---|---|
| 1 | 全部为特殊字符 | 3.1% |
| 2 | 不含特殊字符 | 2.4% |
| 3 | 特殊字符和正常文字同时包含 | 94.5% |

（2）文字特征：通过分析输出结果发现，有害信息中的特殊文字一般不局限于某一类，繁体字、异体字、形近字等都会出现。

表 2 文字特征输出结果

| 子类别编号 | 子类别特征 | 占比 |
|---|---|---|
| 1 | 只有一类字体或只有一个特殊文字 | 4.3% |
| 2 | 多种字体混杂 | 95.7% |

（3）逻辑特征：将语言特征中包含特殊字符的样本输入，通过分析输出结果发现，有害信息一般同时含有网址、手机号、微信号等，否则无法传递关键信息，仅含有特殊字符但不含以上任何信息的文本基本不包含有害内容。

表 3 逻辑特征输出结果

| 子类别编号 | 子类别特征 | 占比 |
|---|---|---|
| 1 | 含有手机、QQ、微信等联系方式且不含网址 | 31.4% |
| 2 | 含有手机、QQ、微信等联系方式且同时含有网址 | 51.6% |
| 3 | 不含手机、QQ、微信等联系方式但含有网址 | 1.8% |
| 4 | 不含任何以上信息（非有害信息） | 15.2% |

## 5 结语

本文针对目前有害信息识别工作的痛点隐蔽性有害信息进行研究，通过收集多家互联网公司原始数据，使用深度学习中的 BERT 模型设计文本分类方案，通过样本训练和特征统计，总结出隐蔽性有害信息的特征。主要结论为：有害信息基本为正常文字和特殊字符混合，特殊文字一般不局限于某一类，一般都含有网址、手机号、微信号等，主要目的为在隐蔽性和可读性之间寻求平衡。本文提出的方法，能够对隐蔽性有害信息的特征进行有效分析，进而为智能识别隐蔽性有害信息提供一种参考方法。

**参考文献：**

[1] 张慧嫦, 李力卡. 基于机器学习的通信信息诈骗识别模型[A]. 中国通信学会.2019 中国信息通信大会论文集(CICC 2019)[C]. 中国通信学会:人民邮电出版社电信科学编辑部, 2019:4.

# Web 服务器动态性能提升的策略研究

虞安骥[1]，陈旭瑶*[2]，胥志强[2]
（1.江西开放大学，江西 南昌 33001；2. 江西应用科技学院，江西 南昌 33000

**摘要：**由于经济全球一体化进程加剧，互联网+的全面铺开，互联网的利用率和普及率不断提高，使得网络用户和网络需求量不断增加，使得对 Web 服务器性能的要求越来越高。文章通过分析大规模 Web 服务器（IBM 奥运会网站）访问用户响应时间，演示了在不同请求流量强度下 Web 服务器性能的问题，并针对提升服务器性能提出了解决办法，寻找 Web 服务器性能提升的途径。

**关键词：**Web 服务器；动态性能提升；策略研究
**中图分类号：**TP309 TP302.2　　**文献标识码：**B　　**文章编号：**2096-9759( 2023 )06-0149-03

## Research on Strategies for Improving Dynamic Performance of Web Server

YU Anji[1], CHEN Xuyao[2], XU Zhiqiang[2]
(1. Jiangxi Open University, Nanchang, Jiangxi 33000, China;
2 Jiangxi Institute of Applied Science and Technology, Nanchang, Jiangxi 33000, China

**Abstract:** Due to the intensified process of economic global integration. The full spread of internet plus, Because of the network users and network demand continue to increase, so making the performance of Web servers increasingly demanding. This paper demonstrates the performance problems of the Web server under different request traffic intensities by analyzing the response time of the mass access Web server (IBM Olympic website), and proposes a solution to improve the performance of the server, looking for ways to improve the performance of the Web server.

**Keywords:** Web server; Dynamic performance improvement; Strategy research

## 1 研究问题

互联网技术的极速发展彻底改变了人们的生产生活方式以及学习方式。首先，以阿里巴巴为代表的电商平台给人们带来了足不出户的购物体验。阿里创建的电商平台能够同时支撑亿万用户在线访问，面对如此巨大的流量冲击，服务器很容易出现因过载而宕机的现象（房俊华，2017）；其次，百度搜索引擎用户达到了 7.66 亿人，每天响应的搜索请求超过 60 亿次；再者，网络教育公开课代表平台：慕课平台，全称为“大规模开放网络课程”，中文翻译为“慕课”人们依托慕课平台可以更加方便地学习，所以用户数量快速增长，一门慕课课程动辄上万人，甚至最多可以达到 16 万人，像是慕课如此大的规模资源网站每天都会产生海量的用户请求，这对服务器来说是个巨大的挑战。

就目前为止，影响 Web 服务器请求响应速度的主要因素有两个：（1）从客户机到服务器网络链路端到端的传输延迟；（2）Web 服务系统（包括 Web 服务器、数据库服务器、应用服务器等）的处理性能。虽然，近年来的网络技术发展突飞猛进，甚至于 10G 以太网等纷纷涌现，但就目前而言，对于宽带 IP 主干网 Web 服务器中依然存在由于 Web 服务器端的网络带宽小、用户访问量巨大以及其他的一些原因（如跨网络运营商、跨越广域网等）而导致对网站的访问速度缓慢的状况，对于

**收稿日期：**2023-03-24
**基金项目：**2020 年江西省教育厅科学技术研究重点项目；项目名称：MOOC 远程教育中 Web 服务器集群负载均衡算法研究（项目编号：GJJ209917）。
**作者简介：**虞安骥（1985-），男，江西南昌人，博士，副教授，研究方向：教育学。
**通信作者：**陈旭瑶（1994-），女，江西南昌人，硕士，全国信息化工程师，研究方向：财务大数据。

[2] 范亮，陈倩.人工智能在网络安全领域的最新发展[J].中国信息安全,2017(12):104-107.
[3] 张志勇,荆军昌,李斐,等.人工智能视角下的在线社交网络虚假信息检测、传播与控制研究综述[J].计算机学报,2021,44(11):2261-2282.
[4] 龚文全.人工智能在有害信息识别服务的应用和发展趋势[J].电信网技术,2018(2):10-14.
[5] 卢刚.面向网络社区的敏感信息语义计算方法研究[D].北京:北京邮电大学,2018.
[6] 吴珊,李跃新.智能设备网络虚假信息行为识别与控制技术研究[J].计算机测量与控制,2019,27(4):88-91,133.
[7] Devlin J, Chang M-W, Lee K, et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding [J]. North American Chapter of the Association for Computational Linguistics,2018:179-195.
[8] Vaswani A, Shazeer N, Parmar N, et al. Attention is All you Need[J].Advances in neural information processing systems, 2017,30:231-242.
[9] Tay Y,Dehghani M,Bahri D,et al.Efficient Transformers:A Survey[J].Learning,2020:1-39.
[10] Letarte G,Paradis F,Giguère P,et al.Importance of self-attention for sentiment analysis[C].In Proceedings of the 2018 EMNLP Workshop BlackboxNLP:Analyzing and Interpreting Neural Networks for NLP,2018:267-275.