
互联网食品违法违规行为监测报警

技术规范

Technical specification for monitoring and alarming
of food violations on the Internet

前 言

本文件参照 GB/T 1.1-2020《标准化工作导则 第 1 部分：标准化文件的结构和起草规则》的规定起草。

本文件起草单位：国家食品安全风险评估中心、香港中文大学（深圳）、浙江清华长三角研究院、国家药监局信息中心、食品安全与营养（贵州）信息科技有限公司、中国人民大学、北京林业大学、江南大学、天津科技大学。

本文件主要起草人：王华丽、陈山泉、张朝正、王秀娟、蔡强、陈亚龙、陆颖、强文佳、陶光灿、林丹、生吉萍、高笑歌、宿文凡、王建新、马东、王建华。

互联网食品违法违规行为监测报警技术规范

1 范围

本技术规范是依据项目中课题三的研究成果产出：互联网食品销售违法违规行为数据挖掘算法及预处理技术方案和互联网食品销售违法违规行为检测报警系统使用说明书进行制定相关内容。

本技术规范主要梳理了目前主要互联网食品安全违法违规行为，研究了互联网食品安全违法违规监测方法，通过对互联网食品销售违法违规行为多源异构大数据挖掘和融合平台进行规范操作，实现对互联网食品销售违法违规行为的风险识别与评价、预警、追溯和智慧监管。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注明日期的引用文件，仅所注日期的版本适用于本文件。凡是不注日期的应用文件，其最新版本（包括所有修改单）适用于本文件。

《网络食品安全违法行为查处办法》

《网络食品经营监督管理办法》

《网络餐饮服务食品安全监督管理办法》

3 术语和定义

下列术语和定义只适用于本规范。

3.1 互联网食品交易平台

在互联网食品经营活动中为双方或多方提供网页空间、虚拟经营场所、交易规则、交易撮合、电子订单等服务，供交易双方或多方开展交易活动的信息网络系统。

3.2 互联网食品交易平台提供者

获批工商营业执照并提供食品经营第三方交易平台服务的公民、法人或其他组织。

3.3 互联网食品生产经营者

通过第三方交易平台或自建网站进行交易的食品生产经营者。

3.4 Scrap 框架

Scrapy 框架 是用 Python 语言实现的一个为了抓取网页数据、提取结构性数据而编写的应用框架，该框架是封装的，包含 request（异步调度和处理）、下载器（多线程的 Downloader）、解析器（selector）和 twisted（异步处理）等。

3.5 BERT 模型

BERT 是“Bidirectional Encoder Representations from Transformers”的首字母缩写，整体是一个自编码语言模型（Autoencoder LM），并且其设计了两个任务来预训练该模型，即用了 Masked LM

和 Next Sentence Prediction 两种方法分别捕捉词语和句子级别的表达。BERT 模型的主要输入是文本中各个字/词(或者称为 token)的原始词向量；输出是文本中各个字/词融合了全文语义信息后的向量表示。

3.6 AHP 层次分析法

一种层次权重决策分析方法，将决策问题按总目标、各层子目标、评价准则直至具体的备投方案的顺序分解为不同的层次结构，然后用求解判断矩阵特征向量的办法，求得每一层次对各元素对上一层次某元素的优先权重。

3.7 LFTD 算法

是指基于用于边缘检测的轻量化快速人脸检测器（LFFD）和一种基于目标检测方法的文本检测模型（CTPN）而生成的一种轻量化快速文本检测器（LFTD）

3.8 CTC loss 函数

是指一个损失函数，计算一个输入序列 x ，模型预测序列中所有能映射到标签的输出序列概率总和的负对数，主要用在没有事先对齐的序列化数据训练上，比如语音识别，ocr 识别等，主要的优点是可以对没有对齐的数据进行自动对齐。

4 互联网食品生产经营违法违规行为

4.1 篡改生产经营证照行为

（1）篡改生产经营者名称（店名）

互联网食品生产经营者利用 Photoshop 等软件篡改名称（店名）的行为，使得互联网食品生产经营者的店名与其生产经营许可证、营业执照上的名称不一致。

（2）篡改生产场所和经营场所

互联网食品生产经营者可能由于搬迁等原因，导致实际地址与生产经营许可证、营业执照中地址不符。

（3）篡改生产经营许可证项目

互联网食品生产者超过许可的类别范围销售食品、互联网食品经营者超过许可的经营项目范围从事食品经营的。例如证照上只有热食类食品制售，但实际上出售冷食如果汁和奶茶等食品。

（4）篡改生产经营许可证的有效期

商家在生产经营许可证的有效期失效时间段内，经过修改篡改证照中的有效期保持生产经营活动。

（5）其他

商家在互联网食品交易平台上使用假证从事食品生产经营活动。假冒其他生产经营者名义经营

食品。

4.2 生产加工阶段违法违规行

- (1) 使用变质食品原料；
- (2) 食品加工程序不当；
- (3) 食品中混有异物、或存在非法添加的行为；
- (4) 生产经营环境不卫生、不规范。
- (5) 虚假标注生产企业名称和食品生产许可证号。
- (6) 将非食品标注为食品。
- (7) 将化工产品标注为食品级。
- (8) 其他违反食品安全法律法规的行为。

4.3 贮存、运输配送阶段违法违规行

- (1) 使用不合格（不具备食品级别）包装材料，或使用不卫生的包装材料，包装方法不当；
- (2) 不具备对贮存、运输有特殊要求食品的保质条件（例如保鲜、保温、冷藏或者冷冻）。

4.4 销售阶段违法违规行

- (1) 网上刊载的食品名称、成分或者配料表、产地、保质期、贮存条件，生产者名称、地址等信息与食品标签或者标识不一致；
- (2) 非保健食品信息明示或者暗示具有保健功能；
- (3) 婴幼儿配方乳粉产品信息明示或者暗示具有益智、增加抵抗力、提高免疫力、保护肠道等功能或者保健作用；
- (4) 对贮存、运输、食用有特殊要求的食品，没有标识和说明；
- (5) 销售过期食品；
- (6) 互联网上销售食品质量与实体店销售不一致；
- (7) 不处理或延迟处理投诉举报事件
- (8) 不提供销售凭证
- (9) 提供虚假信息（网络上刊登虚假信息）
- (10) 虚假刷单及虚假评论
- (11) 销售特殊医学用途配方食品中特定全营养配方食品

4.5 其他违法违规行

未及时采取有效措施排查、消除安全隐患，未落实食品安全责任等。

5 违法违规行爲监测技术

5.1 检测证照真伪

从饿了么、美团等互联网交易平台获取互联网商家的营业执照，利用图像篡改特征监测互联网交易平台中执照图像是否被篡改过。

图像篡改形式主要包括以下形式：

- (1) 图像拼接：往图像中添加来源于其他图像的物体。
- (2) 图像复制粘贴：将图像中的某个物体复制粘贴一次或多次并添加到原图像中。
- (3) 图像局部去除：移除图像本身的部分内容。

采用 Resnet50 提取图像特征，对残差进行融合，减少梯度弥散问题。使用 CBAM 注意力机制放大特征图谱中的空间和通道维度中的重要特征。通过上采样对深层特征和浅层特征进行融合。采用多任务训练方法，输出图像块的篡改可能性，融合后的特征图则输出不同尺度图像块的篡改区域。采用 Focal loss 作为损失函数来进行训练。构建证照篡改检测系统。

5.2 提取证照文字与识别生产经营范围

从饿了么、美团等互联网交易平台爬虫互联网商家的营业执照，利用图像文字提取功能监测证照文字是否被改动。

被篡改的文本具有以下某些特点：

- (1) 大多数以长矩形形式存在，即长宽比一般较大或较小。这与普通的目标监测中的物体不一样（长宽比较接近 1）。
- (2) 普通物体（比如猫）存在明显的闭合边缘轮廓，正常文本没有边缘轮廓。
- (3) 文本中包含多个文字，并且文字之间有间隔。

在 LFTD 的基础上设计文本识别模块，使用 CTC loss 求损失函数，实现文本监测和文字识别的端对端算法图像文字识别模型。

5.3 风险评估消费者评论文本数据技术

对消费者的评论文本数据进行风险评估程序如下：从互联网交易平台（例如美团、饿了么）上采集消费者评论文本数据、对文书数据预处理、标注文本数据、对文本数据进行风险评估。

5.3.1 数据获取

从互联网交易平台（例如美团、饿了么）上采集消费者评论文本数据，选取 java 语言，采用 Scrap 框架采集文本数据，从手机移动端、PC 客户端、PCH5 页面上采集销售者评论文本。

5.3.2 文本预处理

由于文本数据是直接从互联网交易平台上采集的，采集的文本数据存在重复值、异常值或缺失

值，我们在使用前需要进行文本数据清洗。清洗数据如下：

- (1) 删除评论时间为空值或缺失的数据；
- (2) 删除评论内容为空值、或只有标点符号、或只有表情的数据。

采用文本处理模型对文本数据进行预处理：

- (1) 将输入的文本切分为单个字。例如，将“感觉不新鲜”切分为：“感”“觉”“不”“新”“鲜”；
- (2) 在文本首尾添加特殊标记。在序列头部增加[CLS]标记，尾部增加[SEP]标记。例如，[CLS]“感”“觉”“不”“新”“鲜”[SEP]；
- (3) 对序列进行截断或扩充操作，使序列长度为预定义长度，如 128。扩充操作具体是在序列末尾添加[PAD]标记，直至序列长度达到预定义长度。例如，[CLS]“感”“觉”“不”“新”“鲜”[SEP][PAD]；
- (4) 进行注意力标记，扩充标记对应的位置为 0，其余为 1。例如，11111110；
- (5) 根据词典，将序列中的标记转换为相对应的索引值。例如 [101,1996,3899,2003,10140,2002,102,0]。

5.3.3 文本数据标注

从文本角度制定出一套互联网食品销售风险评估指标体系，该指标体系有 3 层。目标层为互联网食品销售风险评估，准则层分为生产加工阶段，配送阶段和销售阶段。在生产加工阶段，商家为了追求利益使用变质的原材料，并且由于市场监管的疏松，食品加工的卫生状况也存在较大问题。在配送阶段，主要存在外卖包装卫生状况与包装规范问题。在销售阶段，商家也会出现售卖以次充好的商品的情况，对于客户的反馈与举报，采取置之不理或搪塞的态度，更加恶劣的是采取刷单等措施给顾客造成虚假的商家信息。由以上现象得出了方案层的若干指标。

最后通过 AHP 层次分析法得到每个指标的权重，请专家指导和人工标注的形式标注数据。

5.3.4 文本风险评估

对消费者的评论文本进行风险评估，采用 BERT 模型用标注的评论训练模型。首先将消费者的评论随机划分为训练集和验证集，将训练集评论输入 Bert 模型，Bert 模型输出向量后再连接一个多层神经网络，输出评论的分类概率。针对一个评论预测出结果，将该结果与真实的标注结果进行计算，算出损失函数，反向传播。每训练一轮后，在验证集上验证模型表现，持续训练直至验证集表现不再提升。至此完成训练过程，保持模型参数。

模型训练完成后，推理时，将得到的消费者评论进行异常值清洗等预处理后，然后放入模型中，即可算得每个预定义的维度的概率。对每个维度的概率加权求和，即得到每家店铺的风险值。

5.4 数据挖掘与融合平台

基于互联网食品销售违法违规行为多源异构大数据的“多源异构大数据识别”、“数据融合计算”与“知识深度演化与贝叶斯推理”三个科学与技术，通过多源异构大数据结构和非结构数据统一表示、数据融合与特征提取、质量综合评价、实体识别与关联分析等融合技术，通过大数据聚类、复杂网络理论、深度学习、贝叶斯推理、知识图谱与演化学习等技术理论，形成整套互联网食品销售违法违规多源异构大数据挖掘与融合平台。用于互联网食品销售违法违规行为的风险识别与评价、预警、追溯和智慧监管（具体操作见附件）。

附件：互联网食品销售违法违规行为监测报警系统

互联网食品销售违法违规行为监测报警系统分为外部平台和内部平台两个部分。

外部平台面向大众，展示了网页的基本概况和安全数据。内部平台供专业人员使用，在外部平台数据的基础上进行了进一步的处理和分析，使数据更加直观。

1 外部平台

在外部平台中，网站主要包含了4个模块：首页、监测地图、食品资讯、关于我们。

1.1 登录

选择适配浏览器，输入网址 <http://www.fdmonitor.com/f>，进入系统首页，如图1所示。

1.2 首页



图1 系统外部平台首页

在首页中，输入店铺名，并点击【搜索】，系统根据关键词跳转到监测地图页，并查询特定店铺信息。如图2、图3所示：



图2 搜索



图 3 系统外部平台首页功能示意

1.3 监测地图

监测地图是查询店铺数据信息的入口。

点击页面右上角菜单栏【监测地图】，页面会默认显示商家的基本信息，其中，左边栏查看店家的名称，地址，风险值，右边的地图查看店家信息的地理位置信息。如图 4 所示。

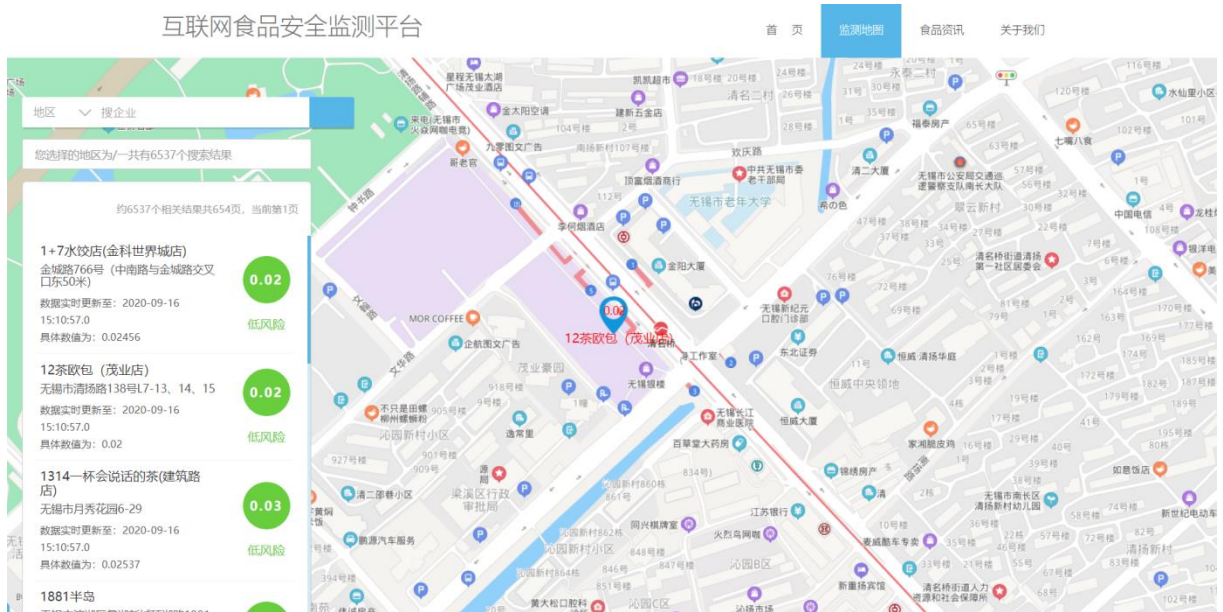


图 4 监测地图界面

监测地图同样支持检索功能，其中，地区选项卡选择具体省市，搜索框中搜索店家名字的关键词。“省”为必选项。

在选择完后，点击“蓝色”按钮，即可进行检索。如图 5 所示

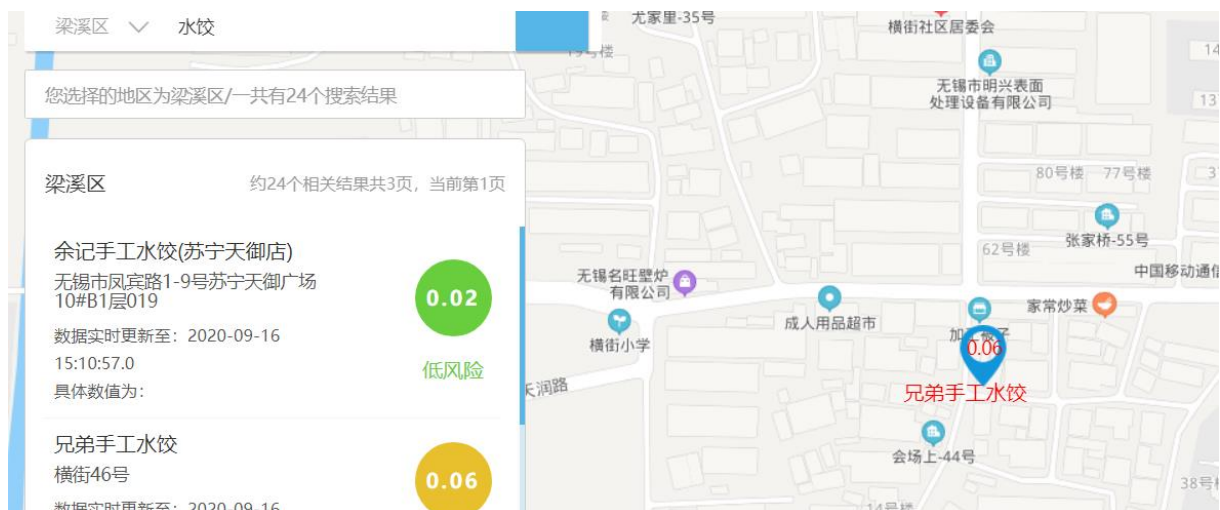


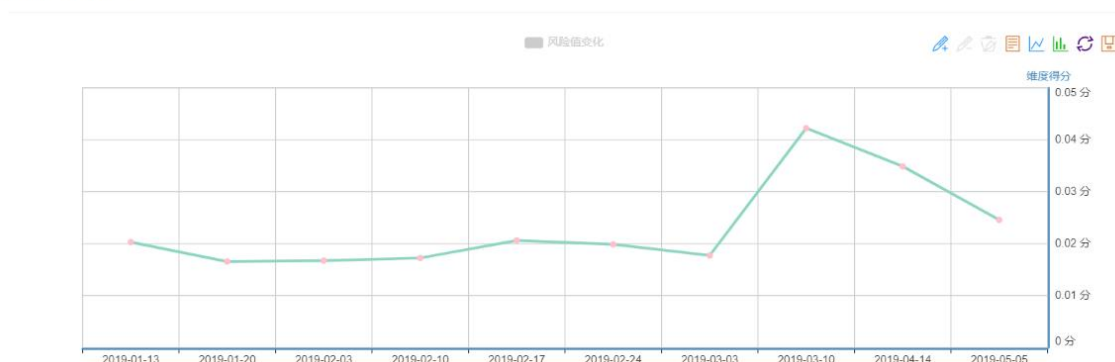
图 5 监测地图界面选项卡功能

如要进一步确认店家的数据信息，先在左侧点击店铺名，然后在右侧地图上的点击坐标点，即可进入对应店家的商家详情界面。

在商家详情界面，查看店家的月销售量，风险值变化等情况，并且可以对店家进行评价。如图 6 所示



风险变化趋势



1+7水饺店(金科世界城店)的评价

说说商家在服务过程中的表现? 您的评价将提升他们的服务质量! 最多输入300字

图 6 商家详情界面

其中，风险变化趋势通过选择图标右上角的功能键，以不同的形式显示。如图 7 所示。

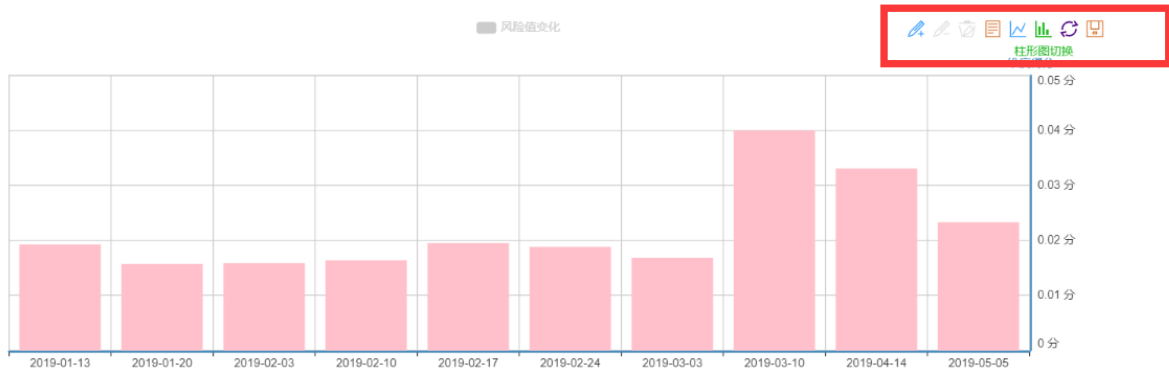


图 7 风险变化趋势功能

1.4 食品资讯

食品资讯界面用于展示食品行业的相关新闻、法律法规、饮食贴士。

点击系统页面右上角菜单【食品资讯】按钮，进入食品资讯界面，如图 8 所示



图 8 食品资讯界面

在点击对应报道的【查看详情】按钮，可以在同页面查看文章的详情，如图 9 所示。

- 相关新闻
- 相关新闻
- 法律法规
- 饮食小贴士

相关新闻

无锡市质检院全力保障企业疫情防护用品免费检测

发布者: 系统管理员 点击数: 120 发布时间: 2020-05-10 17:31:56 更新时间: 2020-05-10 17:50:19

无锡市质检院积极响应无锡市政府和无锡市市场监管局关于疫情防控的相关要求，充分发挥技术支撑和保障职能，对本市企业生产的疫情防护用品实施免费检测。全院干部职工以强烈的责任意识和大局意识，齐心协力、忘我工作，展开了一场争分夺秒的“质量守护战”，为群众的健康安全贡献一份力量。

2月11日，无锡市质检院接到无锡联华日用科技有限公司委托信息，需要对该企业生产的84消毒液、75%酒精等2类产品进行检测。企业负责人表示，为满足疫情防控所需，企业正在24小时轮班加急生产，希望质检院在质量把控方面提供支撑。无锡市质检院第一时间响应，主动联系厂家对送检样品执行标准、检验项目、检验报告时效要求等进行沟通确认。下午4点多，公司样品送达质检院。为做好双方人员防护，质检院在一楼大厅设置了“送样专用通道”和临时样品接收处。“全副武装”的质检技术人员早已在岗位待命，迅速与企业人员办理了样品交接手续。

2月12日上午7点多钟，检验室张诚、陈作王、吕妍凤、尤晓栋等同志便早早赶到实验室，开始了紧张的检测工作。直到中午12点多，几位同志顾不上吃午饭，还在盯着仪器设备和检测数据。他们说，“现在防护用品这么紧缺，尽早完成检测，便让更多急需的人尽快得到放心的防护！”

图 9 食品资讯界面文章详情

1.5 关于我们

关于我们界面主要用于展示行业资讯，同时，用户可以通过此部分提交留言，实现和平台的交互。



图 10 关于我们-行业资讯

公众留言

如果您在使用我们产品时遇到问题，或者对我们有任何意见建议，欢迎在此留言，我们将关注您的问题并尽快与您联系。

姓名

邮箱

电话

分类 请选择

留言

验证码 303Q

立即提交

图 11 关于我们-公众留言

点击系统菜单栏【关于我们】按钮，进入关于我们主页面，如图 10 所示。

点击左侧导航栏【行业资讯】，查看行业资讯。点击【公众留言】，填入个人信息，进入留言页面。如图 11 所示。

2 内部平台

内部平台可视化展示监测中采集数据量、监测地区情况、监测评论情况、监测商家及高风险商家情况。如图 12 所示。

页面最上方展示了监测中总采集数据量、当月采集数据量、监测地区数、监测评论数、监测商家数、高风险商家数。

中部为类似于外部平台的地图功能，更加直观的展示了监测地区的食品安全状况。地图可以拖动和缩放，更加便于管理。

页面的左右两侧用图标的形式展示数据，包括各地区采集数据量，各地区月销售量，风险值雷达图，以及用户评论的词云分布等情况。



图 12 系统内部平台