

基于互联网信报控制协议的灾情信息 获取技术研究与应用*

李兆隆, 吴艳梅, 李 敏, 李永强

(云南省地震局, 云南 昆明 650224)

摘要: 通过建立云南省 IP 地址定位与状态变化数据基础数据库, 使用互联网信报控制协议, 研发基于互联网信报控制协议的灾情信息收集系统, 实现对云南省 IP 地址主机设备震前、震后的在线状态比对, 短时间内生成震区灾情判断数据, 为震害评估及应急决策提供依据。经过 2017 年云南 2 次地震检验, 该系统收集数据及分析结果满足地震应急相关工作需求。

关键词: 地震应急; 灾情获取; IP 地址

中图分类号: P315.941

文献标识码: A

文章编号: 1000-0666(2017)02-0311-06

0 引言

云南省地震局地震应急技术系统开发、应用部署于“十五”期间, 从开始使用至 2017 年 3 月, 经历了多次的数据、模型本地化与系统的技术优化、软件升级(曹彦波等, 2015)。现已发展成包含地震应急评估系统、应急信息发布服务、综合数据查询服务 3 个主要功能模块为主体的、共计 40 余个服务接口的综合应急技术系统。虽然有多年的优化与升级来保证应急资料、评估信息的全自动化产出、推送, 但是, 随着互联网技术的高速发展与网络基础设施的不断完善, 各类先进的技术迅速得到普及与应用, 互联网大潮下的信息产出与需求, 早已超出了现有地震应急技术系统的接受与服务能力: (1) 初评估阶段的地震影响场是由静态模型计算后得到, 缺乏必要的实时数据支撑, 且其形状规则不能准确地反应地震破坏情况与影响范围; (2) 地震灾害发生后, 缺乏便捷、准确的地震现场数据获取手段, 无法在后方指挥部最需要灾情信息的时间段内获取到有价值的灾情信息, 极大地制约着后方指挥部的应急决策能力(程陈, 史文博, 2013)。因此, 建立一套主动、便捷、高效、低成本的、能与现有软件系统融合的灾情获取系统成为地震应急技术系统

发展的必要(王喜双等, 2014)。

针对以上地震应急技术系统的问题, 根据云南省地震局现有技术系统的运行、管理特点, 本文借助互联网大数据思维, 利用因特网信报控制协议, 编写基于互联网信报控制协议的灾情信息收集软件系统, 对特定地域内的各类互联网联网设备(TCP/IP)的在线状态进行采集与分析, 积累长期的背景数据, 找到其在线状态、传输时间与地震影响状况间的对应关系, 在震后有针对性的对震区数据进行快速、加密收集, 同时与常态数据进行对比分析, 从而在震后极短时间内完成灾情数据的收集及相关灾情的判别。

1 云南省 IP 地址定位与状态变化数据基础数据库

收集、录入云南省的 IP 地址基础数据(表 1),

表 1 基础数据采集来源

Tab. 1 Collection source of basic data

名称	资料来源建议
IP 地址列表	百度地图开放平台
IP 地址定位数据	RTBAsia Open Data Exchange
IP 地址扫描数据	后台软件
IP 地址常态背景数据	后台软件

* 收稿日期: 2017-03-27.

基金项目: 中国地震局震灾应急救援司专项课题“地震应急公共服务平台研发”和“基于 IP 的灾区灾情信息与分布技术研究及模型”共同资助。

包括 IP 地址列表、IP 地址定位、IP 地址行政区划归属、IP 地址分时段在线状态（比例）、IP 地址分时段平均延迟、IP 地址常态扫描数据记录等数据，建立云南省 IP 地址定位与状态变化基础数据库，数据库详细内容见表 2。

通过前期对数据需求的调研及监测部门监测

要素常态值运算方法的研究，设定了一系列的数据项目采集存储要求，同时对各类数据的存储方法、存储类型加以规定（张方浩等，2016）。主要包括以下几个方面：（1）IP 地址列表及定位数据；（2）IP 地址扫描数据结果数据；（3）IP 地址常态背景数据；（4）IP 地址软件计算用状态变化数据。

表 2 IP 地址定位与状态变化基础数据库表设计（截至 2017-03-19）

Tab. 2 Design of IP address location and state change database (As to 2017-03-19)

数据类别	数据表描述	英文表名称	记录数
数据存储表（属性数据）	IP 地址段列表	IP_list	5 933
	IP 地址定位信息表	IP_location_info	3 862 325
	IP 地址分时段常态数据表	Ip_normal_state_with_time	3 862 325
	IP 地址扫描记录表	Ip_scan_history	758 679 057
	IP 地址事件扫描记录表	IP_scan_history_with_event	32 930 157
基础地理图（表）（空间数据）	IP 地址定位信息表	IP_location_info	3 862 325
系统运行参数表（属性数据）	IP 地址状态变化表	Ip_change_state	3 862 325
	IP 地址扫描状态控制表	Ip_scan_state	1
	紧急事件表	Event_list	2

截至目前，该基础数据库共收集 IP 地址 3 862 325 个，可定位 IP 地址 1 820 793 个，全省分布相对均匀，可定位 IP 地址主要集中于地市、县区等行政中心，乡村分布较少。以乡镇为单元统计 IP 地址分布情况，个别乡镇目前尚无定位 IP 地址分布，少量乡镇可定位 IP 地址分布较少，大部分乡镇的 IP 地址数在一千个至数千个之间，部分临近城区乡镇数量可达数万个（图 1）。数据库

数据完全符合 IP 地址后台扫描软件系统数据需求与成果产出的质量要求。

2 基于互联网信报控制协议的灾情信息收集系统的研发与应用

以云南省 IP 地址及云南省基础地理数据为基础，建立云南省 IP 地址定位与状态变化数据基础数据库，使用互联网信报控制协议完成全省 IP 地址的常态与异常数据收集，并对数据进行分析整理，根据相关模型产出震区 IP 地址在线状态变化图等结果资料。软件系统业务流程如图 2 所示。

2.1 系统功能

按照云南省地震局现有应急指挥技术系统特点，基于 IP 的灾区灾情信息收集系统全部部署于阿里云服务器，共分 IP 地址后台扫描端、主控分析端两大部分，相关软硬件环境需求见表 3。扫描端以云计算节点方式长期独立运行并生产数据，与主控分析端通过 MSMQ（微软消息队列）方式进行数据传输与作业交接；所有的数据读取、分发、存储工作由主控分析端完成，并负责响应由软件系统定期发起或由用户随机发起的各类数据核算、查看、分析作业（钱文静，邓仲华，2009）。

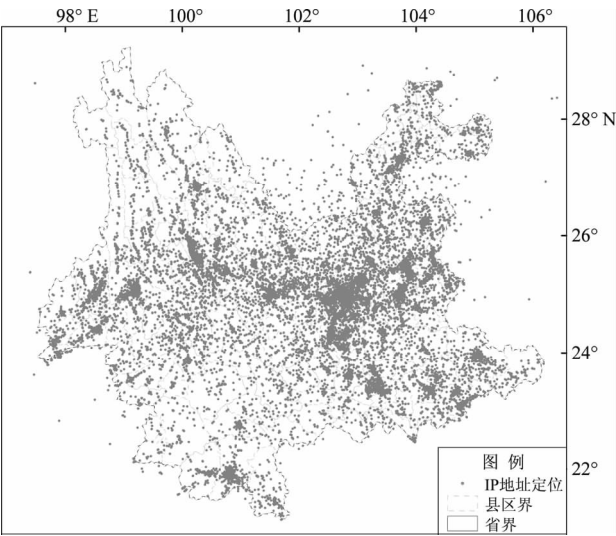


图 1 IP 地址地理位置分布图

Fig. 1 Geography map of IP addresses

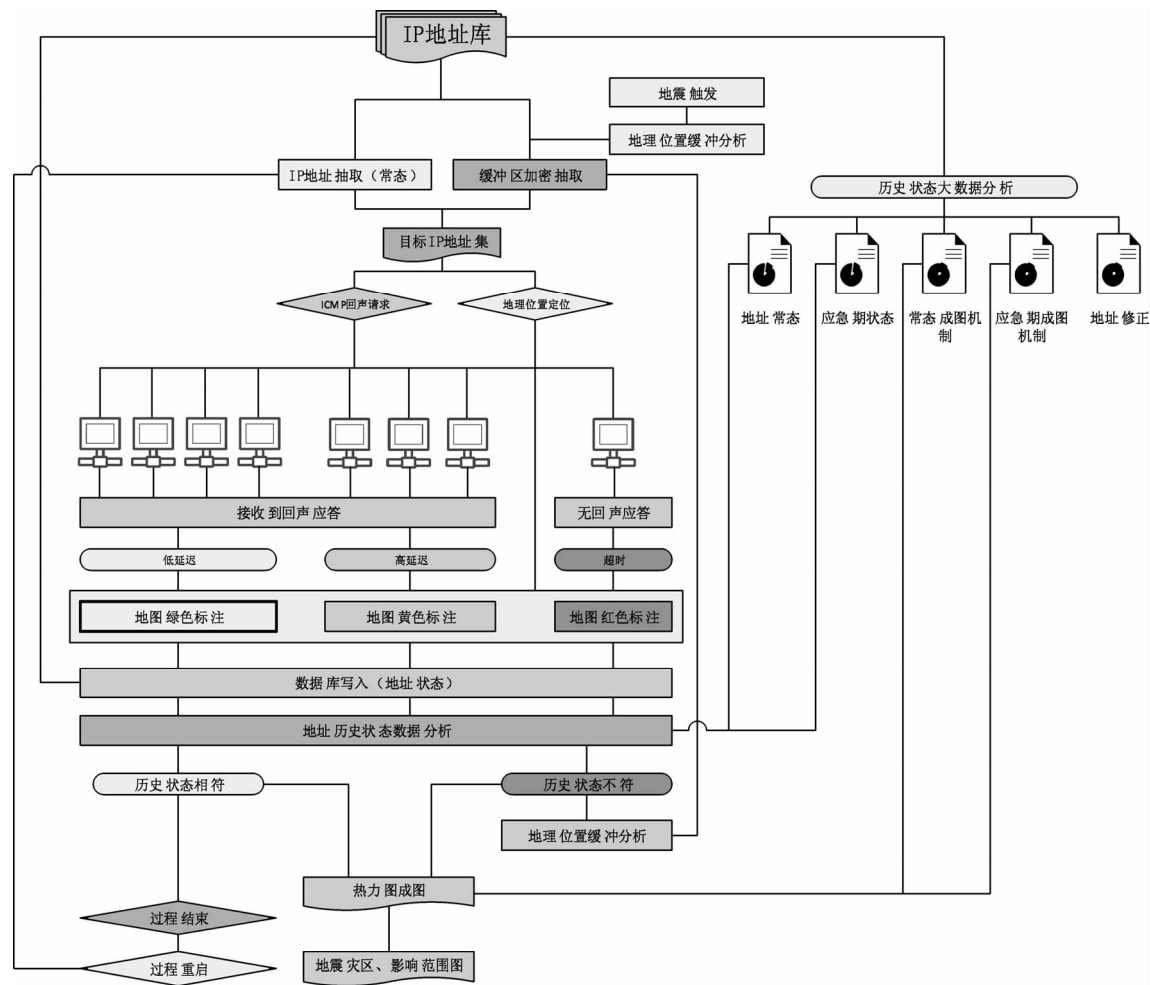


图 2 软件系统业务流程图

Fig. 2 Flow chart of the software system business

软件系统功能总体包含系统参数设置、IP 地址扫描、IP 地址常态值核算、数据产出（已完成方法研究，软件功能未实现）4 个方面。功能模块根据应用时间、服务对象不同，对上述 4 方面功能进行了进一步的分割或结合，以保证系统运行期各模块之间互联、互通便捷且不影响（陶彩霞等，2013），具体功能设计见表 4。

2.2 运行状态

截至 2017 年 3 月，IP 地址后台扫描软件系统已完成设计部署，试运行 24 d，未发现重大设计失误或数据偏差，运行界面见图 3。试运行期间，设定 IP 地址扫描速率为 750 次/s，重访周期约为 30 ~ 45 min；期间共收集数据 7 亿余条，完成一次常态值核算，核算常态值 3.6 亿余条。

经校验，常态值的分布情况与预想值较为接近。图 4 为较典型的 IP 地址在线状态分布图，由连续 3 周的观测数据汇总得到，统计时为得到较

表 3 软件系统部署要求

Tab. 3 Deployment requirements of the software system		
应用类别	环境	环境要求
扫描端	硬件环境	CPU: Intel Xeon E5 - 2680V3 2.5GHz 2 核 内存: 4 096 MB (I/O 优化) 硬盘存储: 50 G 互联网带宽: 5 Mbps
	软件环境	Windows Server 2012 R2 Framework 4.5.1
数据库 主控端 分析端	硬件环境	CPU: Intel Xeon E5 - 2680V3 2.5GHz 4 核 内存: 16 384 MB (I/O 优化) 硬盘存储: 1 T
	软件环境	Windows Server 2012 R2 数据中心版 SQLServer 2014 Framework 4.5.1
开发环境	软件环境	Visual Studio 2015, C#

为典型的曲线，已将休息日数据剔除。图中 IP 地址为疑似某小型企业固定 IP 地址在线状态分布，其在线时间段与法定工作时段有明显的正相关关系，法定工作时间其在线率接近 100%，而其他时间则几乎为 0%，0 点的数据突跳为测试用异常数据。

由此类常态值可分析、分辨出 IP 地址对应的用户个体（群体），此类分析结果可直接作用于震后人员伤亡、迁移评估，对震后的灾情估算有较大价值，尤其对人员伤亡情况判断有重要意义。此类数据的分析、分辨模型仍需较长时间进行研究，以便获得更为准确的判别结果。

表 4 IP 地址扫描软件功能

Tab. 4 Function of IP address scanning software

系统功能模块	系统功能描述
系统参数设置	设置扫描端软件系统运行参数等内容
IP 地址随机抽取	按照一定规则对 IP 地址实现随机抽取，保证抽取结果分布均匀，覆盖全面
IP 地址快速扫描	对抽取地址集逐一快速扫描，并将扫描结果写入数据库
IP 地址常态值核算	对扫描结果数据集进行加工、整理、统计，完成 IP 地址常态数据的核算、更新、校验
IP 地址加密监测	对紧急事件发生区域 IP 地址集进行短间隔加密扫描
紧急事件影响范围估算	依据加密监测数据集，根据相关转化模型，对紧急事件影响范围进行估算（已完成方法研究，软件功能未实现）
数据输出	对上述数据进行图形化、图表化数据输出（已完成方法研究，软件功能未实现）

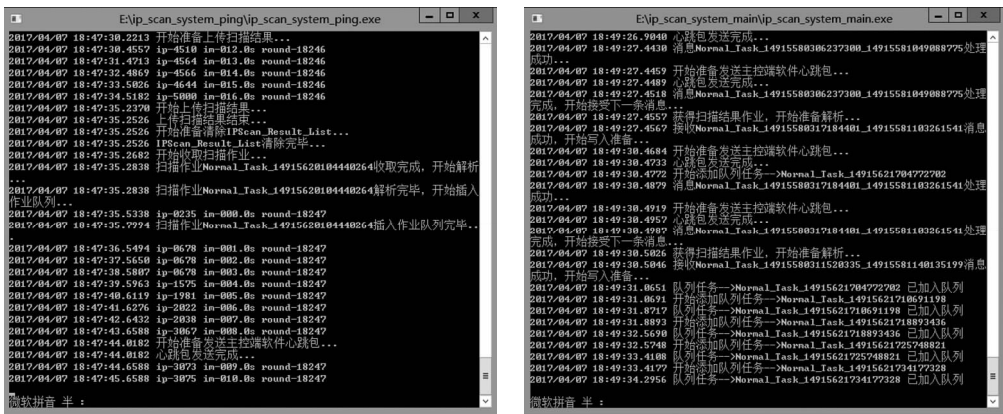


图 3 软件系统运行界面

Fig. 3 Running interface of the software system

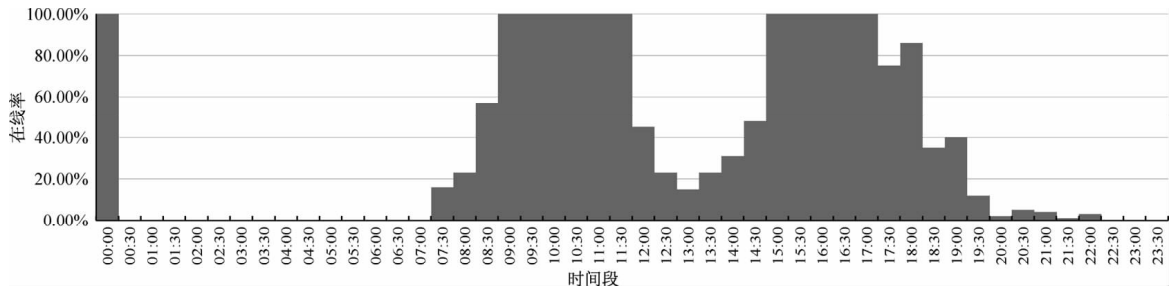


图 4 疑似小型企业固定 IP 地址在线状态变化图

Fig. 4 Online state change curve of the fixed IP address of suspected small business

2.3 预期成果及地震应用

在地震灾害（紧急事件）发生后的一定时间段内，使用上述方法及软件系统，对震区 IP

地址逐一进行短周期重访操作，将这些目标主机因电力中断、线缆破坏、人为关机等原因而引起的通信中断（曹刻等，2008），因民众仓

促避难无暇上网而引发的网络通畅,因灾民集中使用网络而引发的网络阻塞等网络状态进行

收集,并逐一根据位置信息在地图上进行分时段标绘,即得到图5。

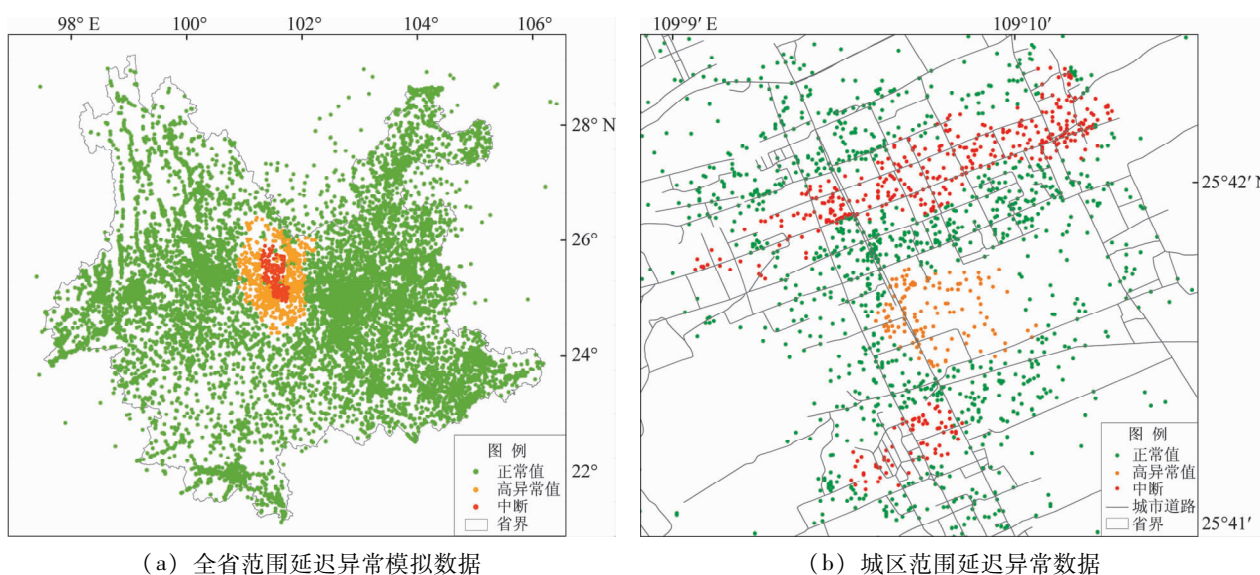


图5 震区IP地址状态分布图(模拟数据)

Fig. 5 Distribution map of IP address state in seismic area (simulation data)

从此类图件中可快速准确地判读出地震灾害极重灾区(高烈度区)、通信中断、阻塞、生命线工程破坏等内容(帅向华等,2014)。如图5a为模拟全省IP地址震后状态分布图,其中红色区域为通信中断地区,由此可认定该区的地面线路基本中断,造成原因可能为地面通信线路故障;黄色区域为高延迟区域,此区域可能为震感强烈区域,导致该区居民集中使用网络资源进行报灾、发送微博微信等引发区域网络阻塞。图5b为震后城市IP地址状态分布图,可明显判读出一呈条带状分布的通信中断区域,此区域的通信中断极有可能是由为该区域提供电力的输电线路损坏而引起。

2017年2~3月,软件系统覆盖区共发生2次

4.0级以上地震,分别为2月8日云南鲁甸4.9级、3月12日云南鲁甸4.5级地震。这2次地震发生后,软件系统均根据EQIM数据自动创建并开始进行紧急事件扫描,2月8日地震事件持续扫描为5h,扫描半径50km;3月12日地震事件持续扫描时间为6h,扫描半径60km。

这2次地震事件均未收集到有效的网络状态突变数据。由图6可知,常在线设备的延迟数据并未因地震事件发生较大波动,震后延迟波动范围与常态值基本保持一致,并未发生预计中的突跳,其原因在于两次地震震区均为地震常发区,且震级偏小,当地居民对此类不会造成较大损失或生命线工程毁坏的地震并不会采取较为激烈的应急动作。

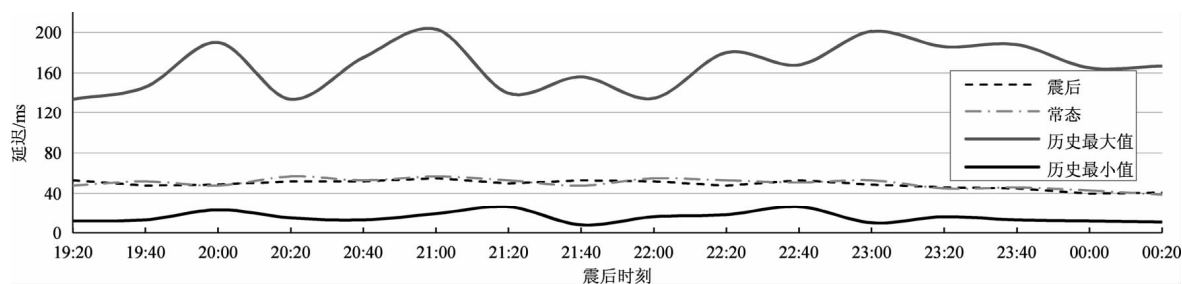


图6 2017年3月12日鲁甸4.5级地震震区某常在线IP地址延迟数据对比图

Fig. 6 Comparison of the frequent online IP address delay data in the Ludian $M_s4.5$ seismic area on Mar. 12, 2017

鉴于上述 2 次地震案例, 该软件系统在应对云南省内较小地震的过程中尚存在一定不足, 其原因主要为: (1) 软件系统设计的应对目标地震为震级较大、具有较大破坏力的地震, 设计采用的数据收集方式、精度不能很好地应用于较小地震; (2) 软件系统数据分析方式、方法不能较为准确地体现数据的微小波动。后期将对上述两个问题进行详细研究并提出解决方案与更优的数据分析方法。

2.4 其他说明事项

本软件系统使用的主机在线状态探测方法是互联网点对点联通状态测试的最常用方法, 其回报内容有限且不涉及隐私, 非有意进行特定特殊参数设置并进行大量并发操作不具有危害性。软件系统试运行阶段所有运行参数 (探测及重访周期等) 设置均在合理、合法范围内, 公网出入流量较小且稳定, 不具备有害行为特征, 不会被认定为有害软件或黑客软件 (冯登国等, 2014)。

3 结论

本文所介绍的基于 IP 的灾区灾情信息收集系统是云南省地震局在地震应急大数据应用方面进行的重要探索, 主要以快速、准确、主动获取震区灾情为目的, 是基于成熟的互联网技术与低廉的硬件服务实现的, 是具有较强地域通用性 (由

数据层控制适用地域范围) 的灾区一手灾情数据获取软件系统。经一般小型地震检验, 系统功能可以满足地震应急中对震区灾情判断的需求。后续需进一步对 IP 地址进行提纯与高精度定位数据更新, 并加强异常数据筛选与热力图向地震影响范围转化机制等方面的研究, 以保证该系统可以更好地服务于地震应急工作。

参考文献:

- 曹刻, 王锋, 李永强, 等. 2008. 云南宁洱 6.4 级地震灾区范围的快速判断[J]. 灾害学, 23(2): 76-79.
- 曹彦波, 李兆隆, 李永强, 等. 2015. 云南地震应急快速评估模型本地化集成研究[J]. 地震研究, 38(1): 148-154.
- 程陈, 史文博. 2013. 大数据挖掘分析在地震科研中的应用[J]. 信息系统工程, (12): 27-28.
- 冯登国, 张敏, 李昊. 2014. 大数据安全与隐私保护[J]. 计算机学报, 37(1): 246-258.
- 钱文静, 邓仲华. 2009. 云计算与信息资源共享管理[J]. 图书与情报, (4): 47-52.
- 帅向华, 胡素平, 郑向向. 2014. 基于互联网信息快速估计汶川地震有感范围[J]. 地震地质, 36(4): 1094-1105.
- 陶彩霞, 谢晓军, 陈康, 等. 2013. 基于云计算的移动互联网大数据用户行为分析引擎设计[J]. 电信科学, 29(3): 27-31.
- 王喜双, 赵邦六, 董世泰, 等. 2014. 油气工业地震勘探大数据面临的挑战及对策[J]. 中国石油勘探, 19(4): 43-47.
- 张方浩, 和仕芳, 吕佳丽, 等. 2016. 基于互联网的地震灾情信息分类编码与初步应用研究[J]. 地震研究, 39(4): 664-672.

Study of Disaster – information Acquisition and its Application Based on the Internet Control Message Protocol

LI Zhaolong, WU Yanmei, LI Min, LI Yongqiang

(Earthquake Administration of Yunnan Province, Kunming 650224, Yunnan, China)

Abstract

In this paper, we firstly built a foundation database of the IP address information of on-line equipments and their dynamic state, then taking the advantage of Internet Message Control Protocol (IMCP), we developed a system for disaster – information acquisition, which was capable of comparing the state of the on-line equipments before and after an earthquake, and outputting the estimated situations in the earthquake – stricken area, to help the emergency personnel to assess the losses and make decisions. The system was applied into the emergency work for the magnitude 4.9 earthquake occurred on 8th, Feb. and magnitude 4.5 earthquake on 12th, Mar., 2017, which was proved to be practicable.

Keywords: earthquake emergency; disaster acquisition; IP address