

基于互联网的重要基础地理变化信息快速更新技术研究

省级基础地理信息数据（DLG）的五年两轮传统更新模式，无法保证数据的现势性，直接影响其使用价值和使用范围，为满足江苏国民经济建设高质量发展对地理信息动态化更新服务日益提高的需求，江苏省基础地理信息中心开展了基于互联网的重要基础地理信息变化快速更新技术研究工作。由传统基础测绘更新模式转变为主动更新机制，形成了一套基于互联网信息挖掘的省级重要基础地理信息快速更新技术方案，并通过生产实践验证了变化发现信息的精确性。

一、建设内容

1) 目标网站及变化信息词汇收集归纳：根据重要地理信息类型，收集可能包含相关地理信息变化信息的网页，形成目标网站汇总，归纳总结重要基础地理信息要素特征相关变化词汇。

2) 形成了一套变化信息快速采集方案：根据不同重要地理要素自身特征及信息分布情况，设计实现对应的互联网的地理空间变化监测与更新方法流程，形成一套基于互联网的重要基础地理信息要素变化信息快速发现路线。

3) 研究了多源多类型变化信息快速提取：互联网数据来源多样（文本信息、POI 数据、影像等），变化类型包含属性变化及点线面多种图形变化。需要依据不同情形快速高效处理，对于采集获取的数据分别利用信息提取、矢量变化检测、影像智能变化检测等信息挖掘技术提取变化信息。

4) 成果数据统一入库管理：对变化数据进行汇集，按规范建立

Geodatabase 空间数据库，利用基础测绘库管系统进行变化相关信息入库，保证每项更新唯一性，最后实现统一管理分发。

二、应用推广

2018 年至今，项目研究的基于互联网的重要基础地理变化信息快速发现技术，改变了传统基础测绘变化发现作业模式，探索了一种变化信息主动发现模式，支撑了基础地理变化信息有侧重分要素采集、季度更新入库。至 2019 年上半年，更新全省重要基础地理变化信息共 541 处。

项目研究成果应用于天地图·江苏、全省基础性地理国情监测和江苏省特色小镇地理国情监测等多个重点项目，缩短了重要基础地理信息变化发现周期，提高了更新效率。”

一、立项背景

江苏省是国家重大发展战略的叠加区域，江苏省政府明确提出要在十三五期间加快实施“1+3”重点功能区战略，基本形成以城市群为主体、大中小城市和小城镇协调发展的城镇格局。加快现代基础设施建设。着力构建现代化综合交通运输体系，积极推进高铁、航空、水运、过江通道、公路等重大交通基础设施建设。在江苏省新一轮省域均衡大发展的新形势下，传统的省级基础地理信息数据（DLG）的五年两轮的更新模式，数据的现势性显然跟不上变化速度。

重要基础地理信息及时更新的重要前提是变化信息快速发现，随着互联网技术的发展，地理信息与日俱增，网络信息呈现爆炸式增长，越来越多表示地物发生变化的信息正以互联网为载体展现出来，这些信息时效性好、描述信息丰富、信息获取方便。其中政府官方网站、

省市级官方新闻媒体、互联网地图服务网站内容可信度、地理信息描述准确度、属性完整性、更新频率可基本得到保障，而且重要地理信息要素由于其重要社会经济意义，是互联网网站报道和采集的重点和焦点。因此，可以基于互联网信息挖掘技术实现重要基础地理信息变化快速发现，据此开展重要基础地理信息快速更新工作，作为常规更新工作的一项重要补充。

本项目主要解决以下问题：

1) 以互联网信息挖掘为主、辅以矢栅变化检测等技术快速发现江苏省境内的水系、交通、居民地及设施、自然经济文化区域、境界与政区和地名等重要基础地理要素变化信息。

2) 匹配现有测绘成果数据，实现要素变化主体的时间、空间大致范围、地名和属性等信息的抽取，通过各种技术手段其进行快速采集。

3) 对发现和采集的变化信息进行整理分析，提取增量数据，数据入库，进行统一管理与任务分发。

二、科技含量

1) 总体思路

项目构建快速全省各类重要基础地理信息变化完整技术流程为目标。对照重要地理信息类型，从权威性强与时效性充分保证的政府网站、官方媒体新闻以及互联网地图三类网站筛选相关网站形成采集目标网站库，整理形成包括重要基础地理要素特征词汇、要素变化特征词汇等在内的词汇库；组合特征词汇及时空参数制定各类要素

采集规则，利用互联网爬虫、POI 调用等技术采集数据，经过文本解析、语义分析、变化检测、地名地址匹配、道路定位等技术工序抽取变化要素名称、时间、空间范围等变化信息；对变化数据进行汇集，建立 Geodatabase 空间数据库，并进行数据去冗余工作，形成最后变化信息空间库，进行统一管理与任务分发；若其精度不符合要求，则通过内外业一体化技术、快速更新辅助系统、无人机移动测量系统、车载移动测量系统等手段对其进行快速采集。对采集的变化信息进行整理分析，提取增量数据，基于基础测绘数据库管理系统采用增量更新模式进行数据入库。

2) 技术方案与创新成果

项目提出了一种对重要基础地理信息变化快速发现的技术，该技术集成了网络发现、地理位置信息文本提取、地名地址匹配与矢量影像变化检测等多种方法，并结合现有的基础测绘、地理国情监测、POI 等数据，实现了对互联网蕴含的重要基础地理信息的快速采集、提取、校正与更新。

具体技术方案与创新成果如下：

项目根据基础地理类型特征差异及资料收集情况，以互联网信息挖掘为主，利用多技术手段获取变化信息。在变化信息快速发现后，区别于传统内外业采集模式，项目将无人机遥感测量、车载移动测量、内外业一体化生产平台及快速更新辅助系统等技术作为重要基础地理快速采集技术手段，体现新型测绘技术在生产实践中的应用。利用基于大数据云计算技术开发的新版省级基础测绘数据库管理系统，采

用增量更新模式进行数据入库，满足对不同时期数据的时空回溯、分析挖掘等需求。

3) 实施效果

2018 年至今，项目研究的基于互联网的重要基础地理信息变化快速发现技术，改变了传统基础测绘变化发现作业模式，探索了一种变化信息主动发现模式，支撑了基础地理变化信息有侧重分要素采集、季度更新入库。2018 年至 2019 年上半年重要要素快速更新共 541 处，其中高速国省道变化信息 22 处，铁路变化信息 2 处，重要交通附属设施变化信息 255 处，国省级自然经济文化区（旅游文化）变化信息 12 处，国省级自然经济文化区（经济）变化信息 29 处，县级及以上行政境界、乡镇级及以上行政地名调整变化信息 70 处，医院变化信息 21 处，学校变化信息 56 处，水系变化信息 71 处，重要水系附属设施变化信息 3 处。

项目研究成果应用于天地图·江苏、全省基础性地理国情监测和江苏省特色小镇地理国情监测等多个重点项目，缩短了重要基础地理信息变化发现周期，提高了重要基础地理信息的更新效率。

三、创新点

项目集成了网络发现、地理位置信息文本提取、地名地址匹配与矢量影像变化检测、大数据云计算等多种方技术法，并结合现有的基础测绘、地理国情监测、POI 等数据，实现了对互联网蕴含的重要基础地理信息的快速采集、提取、校正与增量更新。

1) 基于互联网的重要基础地理变化快速发现技术

本项目本着灵活实用的原则，针对政府网站、官方媒体新闻以及互联网地图定向挖掘，从技术体系层面研究了一种省级重要基础地理变化信息快速发现方法。

(1) 根据每类重要基础地理信息类型特点构建搜索规则，主要包括地理信息发生变化的地理要素名称关键词组、变化描述关键词组和被搜索的网站。对网站按对应地理要素进行归类整理。最终整合形成各类重要基础地理信息要素搜索规则。

(2) 根据重要地理要素和目标网站相关性，提取要素特征词汇和变化特征词汇组合，并结合时空参数形成搜索参数模版，结合爬虫、API 调用技术，对应开发每一类重要地理变化信息采集脚本，完成采集方案制定。

2) 集成多手段的重要基础地理变化信息快速采集更新技术

本项目实施的重要基础地理信息快速更新是新型基础测绘的一种探索，项目灵活应用多种生产平台及信息技术实现了目标信息的快速采集，实现了对传统基础测绘生产模式的发展和革新。

(1) 在变化信息快速发现后，区别于传统内外业分离的采集模式，项目将无人机遥感测量、车载移动测量、内外业一体化生产平台等多种技术作为重要基础地理快速采集技术手段，体现新型测绘技术在生产实践中的应用。

(2) 项目采用了变化信息快速更新辅助系统，该系统基于在线地图服务 API 和数据库技术的互联网地图资源实现在线查询定位、在线标绘、在线入库及坐标转换方法，可快速有效根据辅助方位信息定位

变化信息涉及数据与区域。

(3) 项目更新采用增量更新入库的方式进行数据入库。一方面，增量更新入库方式仅需对增量数据包进行检查和处理，能够大大提高数据质检和入库效率；另一方面，采用增量更新入库，能够突破传统基础测绘采用版本建库模式无法实现数据可回溯的制约，建立可动态维护的基础测绘时空数据库，满足对不同时期基础测绘数据的时空回溯、分析挖掘等需求。

3) 一种高效率的自动矢量变化检测及伪变化信息剔除技术

项目开展了基于地理实体的多源矢量空间数据匹配技术研究，从多个维度和层次实现对矢量地理数据的整合与融合。

(1) 利用叠置分析和几何约束等关键技术，获取地理实体在空间和属性上的双特性，解决点、线、面同源矢量空间匹配。研究了基于多特征组合的点实体匹配、基于 Fréchet 距离的线实体匹配、基于拓扑及空间相似性的面实体匹配技术，解决异源矢量空间数据。

(2) 矢量空间数据变化检测应用中，一般很难得到真实无误的变化信息，检测结果受很多因素影响，如人为误操作、数据转换和变化检测参数设置等引起的伪变化信息难以避免，利用空间分析技术研究伪变化信息的自动剔除方法。

(3) 同时以基础测绘实际生产需求为依托，从提高变化数据更新作业效率角度出发，研究分析拼音拼写规范化细则，并制定基础测绘地名数据库中汉语地名拼音规范化自动化方案。

四、保密方面

无。

五、国际比较

1) 科技查新

2019年7月，项目委托了江苏省科技查新咨询中心（国家一级科技查新咨询单位）进行了科技查新（查新报告编号：201932B2508729），结论如下：项目基于重要基础地理要素建立词汇库及采集规则，利用互联网爬虫及POI调用技术采集数据，构建文本解析、语义分析，利用矢量变化检测、面向对象影像智能分析、地名地址匹配、道路定位的要素抽取工序快速发现重要要素，在已检文献中未见具体述及；项目采用了变化信息快速更新辅助系统，该系统基于在线地图服务API和数据库技术的互联网地图资源实现在线查询定位、在线标绘、在线入库及坐标转换方法，在已检文献中未见具体述及。

2) 科技成果评价

2020年3月26日，江苏省测绘地理信息学会组织专家召开本项目科技成果评价会，成果主要科技成果如下：

1) 提出了一种基于互联网资源挖掘的重要基础地理要素快速更新技术。该技术集成了网络发现、地理位置信息文本提取、地名地址匹配和矢量影像变化检测等多种方法，能实现重要基础地理要素变化信息的快速检测。

2) 提出了一种伪变化信息的剔除检测方法。综合考虑数据的大小、类型、填写规则等因素，从多个维度和层次实现对矢量地理数据

的整合，能有效剔除伪变化信息，提高数据更新的工作效率。。

以东南大学胡伍生教授为主任的评价委员会对该项目进行了评价，一致认为项目的技术路线正确，方法先进，较好地完成了省级基础地理信息快速更新工作，成果总体达到国内领先水平。” ”

项目成果已成功应用于“天地图·江苏”、全省基础性地理国情监测和江苏省特色小镇地理国情监测等多个重点项目。

1) 服务省级基础测绘地理工作

2018 年至今，项目研究的基于互联网的重要基础地理信息变化快速发现技术，改变了传统基础测绘变化发现作业模式，探索了一种变化信息主动发现模式，支撑了基础地理变化信息有侧重分要素采集、季度更新入库。以国省道及铁路设施为例，可快速获取交通部门网站的道路更新信息，及时采集数据信息，完成数据动态更新。

2018 年至 2019 年上半年重要要素快速更新共 541 处。变化发现成果在经快速更新后，第一时间与智慧江苏城市时空云平台、天地图江苏、江苏省快速制图服务系统、航道普查等其他重要地理信息项目平台进行信息共享互用，避免重复采集，节约资金成本。《江苏省“十三五”基础测绘 1:10000DLG 第二轮更新》生产项目明确将项目成果列为省级 DLG 数据更新的重要补充数据。

在生产技术层面上，项目进一步补充完善了江苏省级基础地理信息数据动态更新技术体系，探索了一种重要基础地理信息动态更新模式，充分利用各种地理信息成果资料及互联网地理信息资源，整合了互联网地理信息挖掘、电子平板一体化调绘、无人机航空摄影、车载

移动测量等手段，实现了省级重要基础地理信息的快速采集更新。

2) 为自然资源调查监测提供数据和技术支持

项目积极对接全省地理国情监测需求，变化信息发现成果直接作为 2018 年度地理国情参考资料。此外，针对年度专题性需求，积极通过网络爬虫和定向搜索获取 A 级以上景区、开发区保税区名录、高校名录、化工园区等资料，及时共享下发至作业单位，与地理国情监测项目工作组形成良好的互动协作更新机制。

在专项监测方面，项目互联网地理信息挖掘技术应用到江苏省特色小镇国情监测项目教育、餐饮、商业设施等专题数据的获取，无人机快速采集、电子平板一体化调绘等技术直接应用到特色小镇专题数据生产实践中。

项目组承接了江苏省发改委委托的省级“特色小镇”创建对象年度考核实地核查任务，充分利用无人机、内外业一体化作业平台，在多源数据整合基础上，首创了一种实时化、智能化、网络化的特色小镇现场考核模式，是将快速更新项目快速、精准、机动经验模式落到实处的一次有益探索。

此外，江苏省 InSAR 地面沉降监测二期监测项目中，专题应用研究与分析涉及到地铁沿线沉降监测，面临现有基础测绘数据缺乏部分详细的地铁空间信息问题，利用快速更新辅助系统，快速地从互联网在线地图采集了所需地铁信息，保障了专项监测工作的顺利进行。

3) 辅助基础测绘数据日常工作

本项目设计了一种矢量数据自动变化检测方法，该自动变化检测

方法已经应用到全省基础性地理国情监测和基础测绘生产中。实践证明，该方法提高耗时比约 17，有效的提高了矢量空间数据变化检测的效率，该方法通过图形的几何约束结合填写规则获取的要素变化信息有着比较高的准确率，其中面状无缝地表覆盖数据几乎达 100%，面状、线状和点状地理国情要素也达到 97%以上。同时该方法不仅可以快速自动检测变化信息，而且能有效剔除伪变化信息。设置伪变化信息自动剔除容差，可在一定程度上提高了变化信息识别质量，减少了数据库更新的工作。

来自东南大学、南京大学、南京师范大学、江苏省测绘产品质量监督检验站、江苏省测绘地理信息学会等单位的 5 位专家，对项目技术方案及丰富应用成果表示充分肯定，一致认为，该成果总体达到国内领先水平。” 项目提出了一种对重要基础地理信息变化快速发现的技术，实现了对互联网蕴含的重要基础地理信息的快速采集、提取、校正与更新。申请发明专利 2 项、软件著作权 1 项，快速更新辅助系统 1 套、发表论文 5 篇。项目研究成果应用于天地图·江苏、全省基础性地理国情监测和江苏省特色小镇地理国情监测等多个重点项目，缩短了重要基础地理信息变化发现周期，提高了更新效率。

” 基于互联网的重要基础地理变化信息快速更新项目提出了一种对重要基础地理信息变化快速更新的技术，该技术集成了网络发现、地理位置信息文本提取、地名地址匹配与矢量影像变化检测等多种方法，并结合现有的基础测绘、地理国情监测、POI 等数据，实现了对互联网蕴含的重要基础地理信息的快速采集、提取、校正与更新。

项目针对政府网站、官方媒体新闻以及互联网地图三类网站定向挖掘，从技术层面研究了各类省级重要基础地理变化信息快速发现方法；灵活应用多种生产平台及信息技术实现了目标信息的快速采集，实现了对传统基础测绘生产模式的发展和革新；采用实时采集和定期增量更新入库相结合的方式，实现了全要素统一更新向分要素侧重更新转变，形成了每年度不少于四次的重要基础地理信息更新成果数据库。

在科技成果转化应用方面，项目申请发明专利 3 项、软件著作权 2 项，开发快速更新辅助系统 1 套、发表论文 5 篇。项目研究成果应用于天地图·江苏、全省基础性地理国情监测和江苏省特色小镇地理国情监测等多个重点项目，缩短了基础测绘重点要素变化发现周期，探索了一种按需更新的新型基础测绘生产更新模式，为基础测绘和自然资源调查监测各类生产研究项目提供了有效数据和技术支撑，产生了显著的社会和经济效益。

推荐其参加 2020 地理信息科技进步奖一等奖评选。”