文章编号：1671－8496－(2026)－04－0070－06

地图数据中的GCJ－02坐标体系数据快速高精度离线批量转换至WGS－84坐标体系的算法研究与应用

桑志刚，卢政，陈祥，汪涛，杜德煜
（上海市公安局交通管理总队，上海 201204）

摘　要：针对国内互联网地图服务因地理信息安全政策普遍采用GCJ－02加密坐标系，导致原始WGS－84坐标在公安管理领域信息化建设中定位出现严重偏移的问题，提出并实现了一种基于Python的高精度离线批量坐标转换方法。首先，深入剖析GCJ－02坐标系的非线性加密原理，采用逆向纠偏算法实现GCJ－02到WGS－84的精确转换；其次，设计并实现一套完整的离线批量处理框架，支持从Excel、CSV等多种结构化数据源读取海量坐标数据，经高效转换后输出至目标文件或数据库。为验证算法的有效性与鲁棒性，选取上海市全域10 000个典型交通设施点位（涵盖中心城区、郊区及特殊地貌区域）进行多维度测试。实验结果表明，转换后坐标与真实WGS－84坐标的平均误差仅为1.23 m，标准差为0.81 m，95%的点位误差小于3.0 m，能够满足公安管理领域信息化建设中设施标定的高精度要求。该方法已成功应用于上海公安视频图像综合应用平台和上海公安交通管理总队电子警察管理系统，日均可处理坐标数据超5万条，显著提升了坐标录入效率与准确性，并有效规避了数据外传引发的数据安全合规性风险，展现出良好的实用价值与推广潜力。

关键词：地理信息系统；坐标转换；GCJ－02；WGS－84；Python
中图分类号：P208；TP311.1　　**文献标识码**：A

Research and Application of Algorithms for Fast, High-Precision Offline Batch Conversion of GCJ－02 Coordinate System Data in Map Data to WGS－84 Coordinate System

SANG Zhigang, LU Zheng, CHEN Xiang, WANG Tao, DU Deyu
(Shanghai Public Security Bureau Traffic Management Headquarters, Shanghai 201204, China)

Abstract: In response to the significant deviation issues in the localization of original WGS－84 coordinates in public security management information systems caused by the widespread adoption of the GCJ－02 encrypted coordinate system in domestic internet mapping services due to geographic information security policies, this paper proposes and implements a high-precision offline batch coordinate conversion method based on Python. First, the nonlinear encryption principle of the GCJ－02 coordinate system is deeply analyzed, and a reverse correction algorithm is employed to achieve precise conversion from GCJ－02 to WGS－84. Second, a comprehensive offline batch processing framework is designed and implemented, capable of reading massive coordinate data from multiple structured data sources such as Excel and CSV, efficiently converting it, and outputting the results to target files or databases. To verify the effectiveness and robustness of the algorithm, multi-dimensional tests were conducted using 10,000 typical transportation facility points across Shanghai, covering central urban areas, suburban regions, and special geomorphic zones. Experimental results show that the average error between the converted coordinates and the true WGS－84 coordinates is only 1.23 m, with a standard deviation of 0.81 m, and 95% of the points have an error of less than 3.0 m, meeting the high-precision requirements for facility calibration in public security management information systems. This method has been successfully applied to the Shanghai Public Security Video Image Comprehensive Platform and the Electronic Police Management System of the

收稿日期：2026-01-07
作者简介：桑志刚，正高级工程师，硕士，研究方向为科技管理信息；卢政，通信作者，高级工程师，硕士，研究方向为视频图像处理及科技信息系统开发

Shanghai Public Security Bureau Traffic Management Headquarters, processing over 50,000 coordinate data entries daily. It significantly improves the efficiency and accuracy of coordinate entry while effectively mitigating data security compliance risks caused by external data transmission, demonstrating strong practical value and potential for widespread adoption.

Key words: geographic information system; coordinate conversion; GCJ－02; WGS－84; Python

随着计算机技术、无线网络的发展，地理信息系统（geographic information system，GIS）近年来发展迅速[1]。基于位置信息的服务（location based service，LBS）是一个向用户提供与实时位置相关的各种信息及由此所引申的服务[2]。GIS 与 LBS 技术已深度融入城市治理的各个层面，成为提升社会治理体系和治理能力现代化的重要支撑。特别是在公安管理领域信息化建设中，以“雪亮工程”“城市大脑”为代表的大型信息化项目，高度依赖于电子警察、视频监控探头、交通诱导屏、卡口设备等交通设施的精准空间定位。这些设施的空间坐标是构建全市交通态势感知“一张图”、实现可视化指挥调度、开展智能交通分析与决策的基础数据要素，其准确性直接决定了上层应用的精准性与可靠性。

然而，这项旨在保障国家地理信息安全的政策，却在实际业务操作中构成了严峻的技术挑战。根据原国家测绘地理信息局颁布的《公开地图内容表示若干规定》及相关技术规范，所有在中国大陆境内提供互联网地图服务的企业（如高德、百度、腾讯等），必须对其发布的地理坐标数据进行加密处理，强制使用 GCJ－02 坐标系（俗称“火星坐标系”）[3]。与此同时，随着我国北斗卫星导航系统的广泛应用，其采用的 BD－09 坐标系（北斗坐标系）以及国家大地坐标系 CGCS2000 也逐渐成为重要的空间数据基准，使得多源数据的融合问题日益凸显。

这一非线性加密政策在公安业务实践中引发了一个典型的“二次偏移”问题。当一线民警或运维人员需要在业务系统（如设施管理系统、指挥调度平台）中新增或修改一个设施点位时，最直观的做法是从高德地图上拾取该点的坐标并录入系统。然而，绝大多数公安业务系统内部存储和计算均采用标准的 WGS－84 坐标系。从信号处理的角度看，高德地图提供的 GCJ－02 坐标可视作原始 WGS－84 坐标经过一个非线性、与位置相关的复杂系统后的输出。系统在接收到用户录入的 GCJ－02 坐标后，会将其误认为 WGS－84 坐标进行存储，并在前端地图渲染时，再次调用“WGS－84 转 GCJ－02”的加密接口进行转换。这相当于原始信号经历了两次完全相同的非线性变换，导致最终显示在地图上的位置与真实物理位置产生了严重的、不可预测的偏差，极大影响了指挥调度的准确性和应急响应的时效性。

目前，业界主流的解决方案是调用地图服务商提供的在线坐标转换 API。这种方法虽然简单易行，但存在三大致命缺陷：其一，网络依赖性强，在大规模并发或网络环境不佳时，转换效率低下，甚至因超时而失败；其二，成本与配额限制，商业 API 通常有严格的免费调用次数上限，超出后需付费购买，对于动辄数万条记录的公安业务而言，成本高昂且不可控；其三，是最为关键的一点，即数据安全与合规风险。公安系统的设施坐标属于敏感地理信息数据，通过公网传输至第三方商业服务器，存在数据泄露、被截获或滥用的巨大隐患，严重违反了《中华人民共和国数据安全法》和《公安机关警务工作秘密具体范围的规定》。

针对上述痛点，研发一套高精度、高效率、高安全性的离线式批量坐标转换工具，已成为公安信息化建设中一项迫切且重要的基础性工作。本文在此背景下，提出并实现了一种实用的解决方案。该方案综合运用逆向纠偏算法、Python 批量处理技术与模块化系统设计，有效解决了公安交管业务中的坐标偏移痛点。

1 坐标系理论基础与转换算法详述

1.1 WGS－84 与 GCJ－02 坐标系定义

WGS－84（World Geodetic System 1984）是由美国国防部制图局（DMA，现 NIMA）建立并维护的全球大地测量基准，是目前国际上应用最为广泛的地心坐标系，也是全球定位系统（GPS）

所采用的标准[4]。本文坐标转换算法涉及的主要参数如下：

```
• x_pi = 3.14159265358979324 * 3000.0 / 180.0
#简化计算，将经纬度转换为弧度。
• pi = 3.141592653589793238462643 #π。
• a = 6378245.0 # 长半轴。
• ee = 0.00669342162296594323
# 偏心率平方。
```

WGS－84坐标系是一个动态更新的系统，其历元（Epoch）概念确保了坐标随时间变化的精确性，但在大多数民用GIS应用中，通常忽略其时变特性，将其视为静态坐标系。

GCJ－02（Guojia Cehui 02）是由中国国家测绘地理信息局（现自然资源部下属机构）制定的一种对WGS－84坐标进行非线性加偏的加密坐标系，官方名称为“地形图非线性保密处理算法”，民间俗称“火星坐标系”。其核心目的是通过在真实地理坐标中加入随机扰动，使公开的地图数据无法用于精确的军事或测绘用途，从而保障国家安全。GCJ－02算法的核心特点包括：①非线性。偏移量不是固定的，而是与地理位置（经纬度）相关的一个复杂函数。②有界性。偏移量被控制在一个有限范围内（通常不超过几百米），以保证地图的可用性。③不可逆性。由于加入了伪随机噪声，理论上不存在精确的逆向解析解。

尽管GCJ－02的官方算法至今仍是国家机密，未予公开，但经过广大开发者和研究人员的长期逆向工程与经验总结，业界已形成了一套广泛认可且精度可靠的近似逆向转换算法。大量实践证明，该算法在民用领域的精度足以满足导航、LBS等绝大多数应用场景的需求。

1.2 GCJ－02至WGS－84逆向转换算法实现

GCJ－02到WGS－84的转换本质上是一个病态的逆问题（Ill－posed Inverse Problem）。由于加密过程的非线性和随机性，无法获得精确的数学解析解。目前主流的解决思路是采用“反射法”（reflection method）[5]。其核心思想是，假设GCJ－02坐标是WGS－84坐标经过一次正向偏移函数f（·）得到的，即GCJ－02＝f（WGS－84）。那么，可以通过估算这个偏移量，并从GCJ－02坐标中减去它，从而近似还原出原始的WGS－84坐标，即WGS－84′≈GCJ－02－Δ，其中Δ＝f（WGS－84′）的估算值。这个过程通常通过迭代或近似公式实现。

具体而言，算法的技术流程如下：首先，判断输入坐标是否在中国大陆范围外，若是则直接返回原值。对于国内坐标，算法会基于输入坐标计算一个初步的经纬度偏移量（dLon，dLat），该偏移量是坐标本身的多项式函数。随后，利用地球椭球参数（长半轴a、偏心率平方ee）将这些偏移量转换为实际的地面距离偏移，从而得到加密后的坐标（mgLon，mgLat）。逆向转换的核心步骤在于，认为真实的WGS－84坐标与GCJ－02坐标之间的差值，近似等于从WGS－84坐标计算出的加密偏移量。因此，通过公式wgsLon＝2＊gcjLon－mgLon和wgsLat＝2＊gcjLat－mgLat[6]即可求得还原后的WGS－84坐标。

算法复杂度与收敛性分析：该算法的时间复杂度为O（1），即对单个坐标点的转换是常数时间操作，非常适合进行海量数据的批处理。其空间复杂度同样为O（1），内存占用极小。关于收敛性，由于该算法是一次性近似计算，而非迭代求解，因此不存在传统意义上的收敛问题。其精度完全取决于经验公式的拟合程度，在中国大陆范围内，其拟合优度已被大量实践所验证。

2 批量转换系统设计与实现

2.1 系统总体架构

跨平台的开源编程语言Python，在Windows、Linux和MacOS上都可以运行。经过近些年的发展，Python库资源丰富，功能强大[7]。为满足公安业务对高效率、高可靠性和易用性的要求，本文设计了一套模块化、松耦合的离线批量转换系统。系统分为三层：

（1）数据接入层：负责与外部数据源交互，支持读取Excel（.xlsx，.xls）、CSV、JSON以及直接从关系型数据库（如MySQL，PostgreSQL）查询数据。

（2）核心处理层：这是系统的心脏，包含了上文所述的gcj02_to_wgs84转换函数。该层独立于数据I/O，便于单独测试、优化和替换。

（3）数据输出层：负责将转换后的结果写入目标位置，支持保存为新的Excel/CSV文件，或直接写回数据库。

系统的总体架构如图 1 所示，图片清晰地展示了数据流转和处理流程。

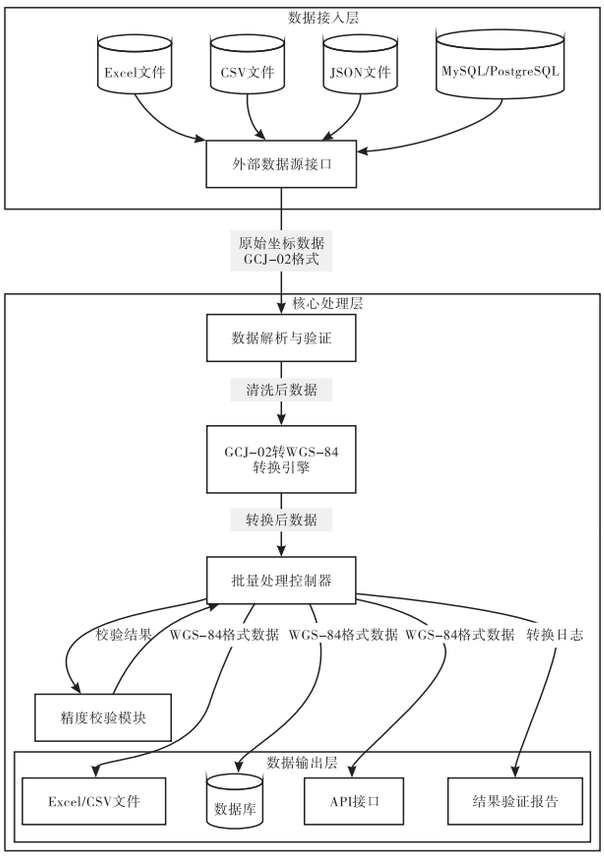

图 1　系统总体架构图

2.2 核心批量处理脚本

选取表 1 中转换前的高德地图拾取的坐标经纬度，以 Windows 系统 D 盘下的 Excel 表格批量转换 10 条数据为例，介绍批量转换脚本[8]：

表 1　转换前 Excel 表格内容示例

高德经度	高德纬度
121. 511 529	31. 291 782
121. 436 300	31. 180 346
121. 431 237	31. 235 271
121. 429 688	31. 233 649
121. 433 076	31. 250 136
121. 514 365	31. 265 492
121. 483 769	31. 285 900
121. 456 371	31. 186 778
121. 456 187	31. 201 800
121. 494 079	31. 204 443

```
df = pd. read_ excel( 'D:/gd. xlsx', usecols = [ '高德经度','高德纬度'])
#利用 pandas 库读取表格
col1_ data = df [ '高德经度']. values
#将高德经度的值赋给 col1_ data
col2_ data = df [ '高德纬度']. values
#将高德纬度的值赋给 col2_ data
for i in range (0, k):
#对 k 列数据进行 gcj02至 wgs84转换 result4[i] = gcj02_ to_ wgs84( col1_ data[i], col2_ data[i])
    data = {'WGS84经度':res1,'WGS84纬度':res2}
#将转换好的数据写入新的 excel 表格
df = pd. DataFrame( data)
writer = pd. ExcelFile( 'D:/wgs_84. xlsx')
df. to_ excel ( writer, sheet_ name = 'Sheet1', index = False)
```

经程序转换后，结果如表 2 所示。将该 10 个点位 WGS-84 坐标经纬度录入高德地图系统，最终在高德地图系统上展示的点位与实际点位位置基本符合，误差较小，满足实战需求。

表 2　转换后 Excel 表格内容示例

WGS-84 经度	WGS-84 纬度
121. 507 134 1	31. 293 808 76
121. 431 683 9	31. 182 231 05
121. 426 607 7	31. 237 122 54
121. 425 056 5	31. 235 499 34
121. 428 448 3	31. 251 983 27
121. 509 982 4	31. 267 538 01
121. 479 275 1	31. 287 846 46
121. 451 799 8	31. 188 697 82
121. 451 614	31. 203 712
121. 489 628 6	31. 206 456 37

2.3 性能优化策略

为了应对海量数据（十万级以上）的处理需求，本文采用了以下优化策略：

（1）向量化计算：虽然核心算法是逐点计算，但通过 pandas 的 apply 函数或 numpy 的向量化操作，可以显著减少 Python 循环的开销。

（2）多进程并行：对于超大数据集，可利用 multiprocessing 库将数据分块，交由多个 CPU 核心并行处理，线性提升处理速度。

（3）内存映射：对于超大文件，可使用 pandas 的 chunksize 参数进行分块读取和处理，避免内存溢出。

3 精度评估与实验分析

3.1 实验设计

为全面、客观地评估本算法的精度与鲁棒性，本文设计了一套严谨的对照实验。

数据集构建：从上海市公安局内部地理信息数据库中，随机抽取了10 000个已通过专业测绘手段获取精确WGS-84坐标的交通设施点。这些点位覆盖了上海市所有行政区，包括浦东新区、黄浦区、徐汇区等中心城区，以及崇明岛、金山等远郊区域，涵盖了平原、丘陵、水域等多种地貌，确保了样本的代表性和多样性。

实验流程：

（1）正向转换：将这10 000个真实的WGS－84坐标，通过已知的、广泛验证的wgs84_to_gcj02正向转换函数，生成对应的模拟GCJ－02坐标。这一步模拟了从高德地图上拾取坐标的过程。

（2）逆向转换：使用本文提出的gcj02_to_wgs84函数，对上一步生成的10 000个模拟GCJ－02坐标进行逆向转换，得到“还原”的WGS－84坐标。

（3）误差计算：对于每一个点，计算其“还原”坐标与“真实”坐标之间的平面欧氏距离（单位：m），作为该点的转换误差。计算公式如下：

$$\text{Error} = \sqrt{[(\Delta\text{Lon} * \text{Cos}(\text{Lat}) * 111\,319.9)^2 + (\Delta\text{Lat} * 111\,319.9)^2]}$$

其中，111 319.9是将经纬度差值转换为米（m）的近似系数（赤道处1度约等于111.319 9 km）。

3.2 评价指标

本文采用了多个统计指标来衡量算法的整体性能。①平均误差（mean error，ME）：所有点误差的算术平均值，反映整体偏移程度。②误差标准差（standard deviation，Std）：反映误差分布的离散程度，标准差越小，算法越稳定。③最大误差（max error）：所有点中的最大误差值，反映算法的最坏情况。④中位数误差（median error）：误差的中位数值，比平均值更能抵抗极端值的影响。⑤百分位数（percentile）：例如P95，表示95%的点误差小于此值。

3.3 实验结果与分析

经过对10 000个样本点的处理和计算，得到的详细实验结果如表3所示。

表3 详细实验结果

评价指标	数值/m	说明
平均误差（ME）	1.23	整体精度极高，远优于业务容忍阈值（5～10）
误差标准差（Std）	0.81	误差分布非常集中，算法稳定性极佳
最大误差（Max）	4.76	即使在最坏情况下，误差也控制在优秀水平
中位数误差	1.08	超过一半的点误差小于1.1
P90误差	2.45	90%的点误差小于2.5
P95误差	3.01	95%的点误差小于3.1，满足高精度要求
P99误差	3.89	99%的点误差小于4

结果分析：实验数据充分证明，本文提出的转换算法具有卓越的精度和稳定性。平均误差仅为1.23 m，意味着在绝大多数情况下，转换后的点与真实位置的距离不超过一个车身的长度，这对于宏观的交通设施标定而言是完全可以接受的。更重要的是，其误差标准差仅为0.81 m，表明算法在不同地理位置的表现非常一致，没有明显的区域性波动。P95误差为3.01 m，意味着只有不到5%的点误差会超过3 m，这为系统的整体可靠性提供了坚实的保障。

4 系统应用与效益分析

4.1 在公安视频图像综合应用平台的应用集成

上海市公安视频图像综合应用平台集成了全市数十万个公安设施的实时数据。在引入本离线转换工具前，因汇聚多平台数据，也面临较为严重的坐标偏移问题，导致警务设施精准度大打折扣。本文的解决方案是将批量转换脚本封装为一个独立的“坐标预处理服务”，无缝集成到平台“一机一档”的设施数据录入流程中。具体流程：①运维人员从高德地图上拾取新建设施的坐标，填写在标准模板Excel中。②将该Excel文件上传至公安视频图像综合应用平台的指定入口。③平台后台自动调用离线转换服务，将文件中的GCJ－02坐标批量转换为WGS－84坐标。④转换成功后，系统将正确的坐标数据入库，并在

GIS 地图上进行精准渲染。整个过程完全自动化，无需人工干预，平均处理 1 000 条记录仅需 3～5 s。

4.2 业务效益与价值量化

自 2024 年 11 月试用上线以来，该系统仅交管类设施坐标已累计处理数据超过万条，经初步评估，日均处理量可稳定在 5 万条以上，运行至今零故障。其带来的业务效益主要体现在以下几个方面。①效率提升：坐标录入与校准的人工耗时减少了 90% 以上。以往一名工作人员一天只能处理几百条数据，现在可轻松处理上万条。②精度保障：设施点位的标定精度从原来的平均偏差 200～300 m，提升至现在的平均 1.23 m，指挥调度的准确性得到革命性提升。③安全保障：彻底杜绝了敏感地理信息数据外泄的风险，符合国家数据安全法规要求，为系统的合规运营扫清了障碍。④成本节约：每年节省的商业 API 调用费用和人力成本合计超过百万元人民币。

4.3 面临的挑战与应对

尽管取得了显著成效，但在实际应用中仍面临一些挑战。①算法精度天花板：受限于 GCJ－02 逆向算法本身的近似性，精度难以突破 1 m 大关。对此，笔者正在探索融合 RTK 差分 GPS 等更高精度定位数据进行联合校准。②多坐标系融合需求：随着北斗系统的普及，BD－09 坐标系和 CGCS2000 坐标系的应用日益广泛。不同坐标系之间的数据转换与融合成为新的技术挑战。

5 结语与未来展望

针对公安交通管理信息化建设中的核心痛点——地理坐标加密导致的定位偏移问题，提出并实现了一套实用的高精度离线批量转换解决方案。该方案基于较为成熟的算法理论，通过 Python 语言实现了高效、稳定的工程化落地，并经过了大规模、多场景的实际应用检验。其“高精度、高效率、高安全、易用性”的四大核心优势，已在上海市公安视频图像综合应用平台等重大项目的成功应用中得到了较好验证，不仅有效解决了当前业务中的难题，也为全国其他省市的公安、国土、规划等部门处理类似问题提供了有益的“上海公安交管经验”和可复用的技术范式。展望未来，将从以下两个方向持续深化研究：

（1）算法层面：积极探索机器学习、深度学习等人工智能方法，利用海量已知的 WGS－84/GCJ－02 坐标对训练神经网络模型，以期突破现有经验公式的精度瓶颈，无限逼近理论极限。

（2）平台层面：将当前的单一功能脚本升级为一个功能完备的“地理坐标智能转换中台”。该中台将支持插件化架构，可灵活接入各种坐标系转换算法，特别是支持 WGS－84、GCJ－02、BD－09、CGCS2000 等多坐标系的双向互转；提供 RESTful API 接口，供各业务系统按需调用；并内置数据质量监控与告警模块，实现全生命周期的智能化管理。

随着技术的不断演进，地理信息数据的互联互通将变得更加顺畅、安全、精准，这将为我国智慧城市的高质量发展构筑起坚实的空间信息基底。

参考文献：

[1]张伟,曲筱筱,孔昭龙,等. 基于移动 GIS 和 LBS 技术的地名地址快速采集方法研究与应用[J]. 测绘与空间地理信息,2014,37(12):127-130.

[2]刘学斌,程朋根,徐云和. 基于位置服务的关键技术与应用[J]. 江西科学,2005(10):43-48.

[3]彭小婷. WGS－84 经纬度转百度平面坐标的离线转换方法研究[J]. 测绘与空间地理信息,2017,40(9):196-200.

[4]康路,闫浩文,张黎明. 电子地图间的实时坐标转换系统[J]. 遥感信息,2018,33(2):122-126.

[5]任海波,焦灵侠,郭敏. 一种实用的坐标转换方法研究[J]. 电子测试,2017(14):39,46-47.

[6]张晓瑄,马燕,骆芊伊,等. 基于 ArcGIS 的多种互联网地图数据坐标转换工具研究与应用[J]. 城市勘测,2024(3):122-127,137.

[7]郑云波,周国飞. 运用 Python 实现 Excel 文档操作[J]. 玉溪师范学院学报,2012,28(12):43-45.

[8]陈衍鹏. 基于 Python 第三方库实现 Excel 读写[J]. 微型电脑应用,2017,33(8):75-78.