编号：*PB23000102*

作品类别：☑软件设计　□硬件制作　□工程实践

2025 年第 X 届全国密码技术竞赛作品设计报告

题目：单表代换密码分析与辅助破译工具

王泓锔 PB23000102

2025 年 6 月 1 日

中国密码学会

基本信息表

编号：PB23000102

作品题目：单表代换密码分析与辅助破译工具

作品类别：☑软件设计　□硬件制作　□工程实践

作品内容摘要：

本作品是一款基于 Python 和 Tkinter 开发的图形用户界面（GUI）工具，专注于单表代换密码的分析与辅助破译。它集成了密码的加密、解密基础功能，并提供了强大的辅助破译模块。手动破译方面，工具能进行字母频率分析、提供替换建议，并允许用户灵活地导入、设置和微调密钥。自动破译方面，采用了结合 N-gram 统计（单字母至四字母）和带词长加权的词典匹配的适应度函数，并利用模拟退火算法进行多轮次智能搜索，同时支持用户预设部分已知密钥以指导搜索方向。工具旨在为密码学学习者和爱好者提供一个直观、易用的单表代换密码分析平台，加深对经典密码体制及其分析方法的理解。

作品特色：

1）**用户友好的图形界面：**基于 Tkinter 构建，提供直观的加密、解密、手动分析和自动破译操作界面。

2）**综合的辅助分析功能：**手动模式下提供字母频率统计、与标准英文频率对比的替换建议，支持密钥的逐字设定和完整导入。

3）**智能的自动破译引擎：**采用模拟退火算法，结合了从单字母到四字母的 N-gram 统计模型和按单词长度加权的词典匹配作为适应度函数，有效搜索密钥空间。

4）**用户引导的自动搜索：**允许用户在自动破译前手动锁定部分确定的密钥映射，显著提高复杂情况下破译的效率和准确性。

5）**多轮次与全程最优追踪：**自动破译支持多轮次迭代运行，并持续追踪和显示当前任务的全程最优解，方便用户比较和获取最佳结果。

6）**过程可视化与日志记录：**自动破译过程中，高亮显示用户锁定的映射，并通过日志记录关键的寻优步骤。

关键词：

单表代换密码；密码分析；模拟退火算法；N-gram 分析；适应度函数；图形用户界面；辅助破译工具

目录

1 作品概述 **1**

1.1 引言 1

1.2 研究背景与意义 1

1.3 国内外研究现状 1

2 设计实现与方案 **3**

2.1 核心模块设计 3

2.2 主要功能实现方案 3

2.2.1 加密与解密 3

2.2.2 手动辅助破译 3

2.2.3 自动辅助破译 4

3 系统测试与结果 **5**

3.1 测试环境 5

3.2 功能测试与结果分析 5

3.2.1 加密与解密功能 5

3.2.2 手动辅助破译功能 5

3.2.3 自动辅助破译功能 6

4 应用前景与总结 **7**

4.1 应用前景 7

4.2 总结 7

5 代码链接 **7**

1 作品概述

1.1 引言

单表代换密码作为一种历史悠久的古典密码体制，其原理是将明文字母表中的每一个字母替换为密文字母表中的另一个唯一对应的字母。尽管其加密方式相对简单，但在密码学的发展史上占有重要地位，并且是学习和理解替换密码和频率分析等基本密码分析方法的经典案例。例如，著名的凯撒密码就是单表代换密码的一个特例 [1]。本作品“单表代换密码分析与辅助破译工具”旨在设计并实现一个用户友好的软件，通过集成多种分析手段，辅助用户完成对此类密码的加密、解密及核心的破译任务，从而降低学习和实践密码分析的门槛，提升分析效率。

1.2 研究背景与意义

在信息安全日益重要的今天，密码学的研究和应用已渗透到各个领域。虽然现代密码体制（如 AES、RSA 等）的安全性远超古典密码，但对古典密码的研究和分析仍然具有重要的学术价值和教育意义。它们不仅是现代密码学思想的基石，也是培养密码分析思维、理解统计攻击等基本密码分析技巧的有效途径。

目前，密码学教育和爱好者进行古典密码分析时，常需要手动进行大量的统计和尝试，效率较低且易出错。一个功能完善、操作便捷的辅助工具能够：

- **提高学习效率**：帮助初学者直观理解单表代换密码的原理、频率分析方法和启发式搜索策略。
- **增强实践能力**：为密码学爱好者和 CTF 竞赛参与者提供一个实用的分析平台，用于解决相关的古典密码题目。
- **促进算法理解**：通过观察自动破译的过程和参数调整，用户可以更深入地理解模拟退火等优化算法在密码分析中的应用。

因此，开发这样一款单表代换密码分析与辅助破译工具，对于密码学知识的普及、分析技能的培养以及相关研究的辅助均具有积极意义。

1.3 国内外研究现状

对单表代换密码的破译方法研究已相当成熟。频率分析是最基础且有效的方法之一，通过统计密文中各字母、字母对（bigrams）、三字母组（trigrams）乃至更长 N-gram 的出现频率，与目标语言（如英文）的标准统计数据进行对比，可以获得关于密钥映射的线索。

国内外已有不少文本介绍了这些经典的分析技术 [1]，也有一些爱好者和研究者开发了相关的辅助软件或脚本。这些工具的功能通常包括自动计算字母和 N-gram 频率、基于频率进行初步的密钥猜测，以及提供用户交互界面进行密钥的修改和部分解密的查看。

在自动破译方面，除了简单的频率匹配外，更高级的工具会采用启发式搜索算法，如爬山算法、模拟退火算法、遗传算法等，通过定义一个适应度函数（fitness function）来评估解密文本的“自然语言相似度”，并在密钥空间中搜索使适应度函数最优的密钥。这些适应度函数通常综合考虑 N-gram 统计和词典匹配。

本作品的设计目标是在现有研究的基础上，整合多种分析手段，特别是引入了带词长加权的词典匹配和支持用户部分锁定的模拟退火多轮迭代机制，力求在用户友好性和破译能力之间取得良好平衡。相较于

一些仅提供单一分析方法或命令行界面的工具，本作品通过 GUI 集成了手动和自动破译流程，并强调了用户引导在自动搜索中的作用，以及多轮迭代和全程最优解追踪的策略，旨在提供更全面的辅助。

2 设计实现与方案

本作品“单表代换密码分析与辅助破译工具”采用 Python 语言作为主要的开发语言，利用其内置的 Tkinter 库构建图形用户界面（GUI），确保了工具的跨平台性和易用性。整体设计遵循模块化思想，主要包括以下几个核心模块：

2.1 核心模块设计

- **密码逻辑模块 (`cipher_logic.py`)**：负责单表代换密码的底层加密和解密算法实现，以及密钥的合法性校验。
- **英文统计模块 (`english_stats.py`)**：存储预先收集的英文统计数据，如单个字母的标准频率、常见的二元组和三元组列表等。
- **手动分析辅助模块 (`analysis_helpers.py`)**：提供支持手动破译的功能，包括计算输入密文的字母频率、根据频率生成替换建议、应用部分已知密钥进行实时解密预览等。
- **适应度评估模块 (`fitness.py`)**：核心功能是计算一段给定文本的“适应度分数”，该分数衡量了文本接近自然英文的程度。它综合考虑了单字母、双字母、三字母和四字母（N-grams）的对数概率得分，以及词典匹配得分（支持按单词长度加权）。
- **自动求解器模块 (`auto_solver.py`)**：实现了模拟退火（Simulated Annealing）这一启发式搜索算法，用于自动搜索最佳密钥，并支持用户预设的部分锁定映射。
- **图形用户界面模块 (`main_gui.py`)**：作为用户与工具交互的接口，集成了上述所有功能，提供清晰的选项卡式布局和交互控件。

2.2 主要功能实现方案

2.2.1 加密与解密

用户输入明文（或密文）和 26 个字母的代换密钥。程序首先通过 `validate_key` 验证密钥的合法性。加密时，根据密钥建立明文字母到密文字母的映射表；解密时，则建立逆映射表。遍历输入文本，对每个字母进行替换，非字母字符保持不变。同时提供了随机生成合法密钥的功能，方便用户快速获取测试密钥。

2.2.2 手动辅助破译

手动模式下，用户首先输入密文并点击分析。程序会计算并显示密文中各字母的出现频率，并与标准英文频率对比，给出初步的“密文 $\rightarrow$ 明文”替换建议。用户可以根据这些建议和自己的判断，通过界面：

1. **导入完整密钥**：直接输入一个完整的 26 字母密钥串，程序会将其转换为内部的映射表。
2. **逐字设定/取消映射**：用户可以指定单个密文字母应该映射到哪个明文字母，或取消已有的映射。程序会进行冲突检查。

所有操作都会实时更新“当前密钥映射”状态区和“部分解密结果”预览区，后者将未确定的字母用下划线 (_) 显示。

2.2.3 自动辅助破译

自动破译采用模拟退火算法，其核心流程如下：

1. **初始化**：用户输入密文、可选的锁定映射（如 `X=e`）以及希望执行的模拟退火轮次。程序根据锁定映射生成一个初始密钥（未锁定部分随机填充）。
2. **迭代寻优（单轮 SA）**：
   - **邻域搜索**：在每一步迭代中，算法从当前密钥出发，通过 `modify_key_with_locks` 函数随机交换两个未被用户锁定的字母映射，从而产生一个候选密钥。
   - **适应度评估**：使用该候选密钥解密密文，并通过 `fitness.py` 中的 `calculate_fitness` 函数计算解密后文本的适应度分数。
   - **状态转移与接受准则**：若候选密钥的适应度高于当前密钥，则接受。若低于当前密钥，则根据当前“温度”和适应度差相关的概率（Metropolis 准则）接受它。
   - **降温**：“温度”参数随迭代次数增加而逐渐降低。
   - **记录本轮最优**：在单轮 SA 运行中，持续记录遇到的适应度最高的密钥。
3. **多轮迭代与全程最优**：GUI 主控模块会循环执行上述单轮 SA 过程达到用户设定的轮次。每轮 SA 结束后，其找到的最佳解会与一个“全程最优解”比较并更新。
4. **结果展示与日志**：界面上始终显示“全程最优解”的密钥、对应的解密文本（高亮锁定部分）和适应度分数。日志区会记录重要优化节点的信息。
5. **任务管理**：“清空当前任务结果和日志”按钮可以重置全程最优解和日志，但保留用户输入的锁定映射。

3 系统测试与结果

为验证本工具的各项功能，我们进行了系统性的测试。

3.1 测试环境

- 操作系统：Windows 11
- Python 版本：Python 3.9+
- 主要依赖库：Tkinter (Python 内置)
- N-gram 及词典数据：
  - Monograms, Bigrams, Trigrams, Quadgrams: 基于大型英文文本语料库统计的频率数据，格式为 `NGRAM COUNT`。
  - `common_words.txt`: 包含约 10,000 个常用英文单词的词典。

3.2 功能测试与结果分析

我们针对工具的三个主要功能模块设计了测试用例：

3.2.1 加密与解密功能

- **用例 1 (正确性)**: 输入已知明文"the quick brown fox jumps over the lazy dog" 和密钥"qwertyuiopasdfghjklzxcvbnm"。
- *预期与结果*: 加密功能正确，生成密文"zit jxoea wkgvf ygb pxdhl gctk zit sqmn rgu"。使用相同密钥解密此密文，准确还原原始明文。密钥输入对大小写不敏感。随机密钥生成功能正常。
- **用例 2 (无效密钥)**: 输入无效密钥如"abc"。
- *预期与结果*: 点击加密或解密时，弹出错误提示“无效密钥!”。功能符合预期。

3.2.2 手动辅助破译功能

- **用例 3 (频率分析)**: 输入一段较长的英文单表代换密文（例如小说《老人与海》的第一段）。
- *预期与结果*: “统计与建议”区准确显示各密文字母频率，并按降序给出与英文标准频率对应的替换建议，参考价值较高。
- **用例 4 (密钥设定与导入)**: 手动设定高频字母映射，或导入已知正确密钥。
- *预期与结果*: “当前密钥映射”区正确显示设定，“部分解密结果”区实时更新。导入功能正确应用完整密钥。冲突处理机制能有效提示。

3.2.3 自动辅助破译功能

- **用例 5 (完全自动破译)**: 使用 500 字符，由随机密钥加密的英文文本。设置执行 30 轮。
- *预期与结果*: 程序通常能在数轮内找到接近完美的解，全程最优适应度分数显著提升（例如，从约-20 提升至-6 左右），解密文本可读性高，但可能存在 1 到 2 对映射错误，不过手动调整映射很容易。日志区清晰记录了“发现本轮更优”和“发现新的全程最优”的时刻。
- **用例 6 (带锁定映射破译)**: 对同一密文，预先锁定 3 个正确的映射（例如，`X=e`, `Q=t`, `J=a`）。
- *预期与结果*: 收敛到高质量解的速度明显加快，有时仅需更少轮次即可达到或超过无锁定情况下的最优结果。解密文本中，锁定的部分被正确高亮为蓝色粗体。
- **用例 7 (短文本破译)**: 使用约 80 字符的密文。
- *预期与结果*: 破译难度增加，对 N-gram 数据和词典的依赖性增强。多次运行后，仍有机会找到较好的解，但稳定性不如长文本。
- **用例 8 (清空任务功能)**: 运行后点击“清空当前任务结果和日志”。
- *预期与结果*: 全程最优解、日志等被清空，但“手动锁定密钥”框内容保留，方便在初次破译不够完善时手动设置参数二次破译。功能符合预期。

测试表明，工具的各项功能基本达到设计要求。自动破译的性能在有良好统计数据和足够迭代时表现出色，用户锁定映射能有效提升效率。

4 应用前景与总结

4.1 应用前景

本“单表代换密码分析与辅助破译工具”在以下方面具有实际的应用前景和价值：

1. **密码学科普与教育：**作为教学辅助工具，帮助学生理解古典密码原理、频率分析及启发式搜索算法。
2. **密码学爱好者与 CTF 竞赛：**为爱好者和竞赛参与者提供分析古典密码的实用平台。
3. **历史文献与简单加密场景分析：**辅助分析可能使用简单替换加密的历史通信或非关键信息。
4. **算法研究与扩展基础：**可作为研究更复杂古典密码或集成先进 NLP 技术的基础。

4.2 总结

本作品成功设计并实现了一个功能较为全面的单表代换密码分析与辅助破译工具。通过图形用户界面，集成了加密解密、手动频率分析、密钥交互式设定以及基于模拟退火（支持用户锁定和多轮迭代）的智能自动破译功能。自动破译模块结合 N-gram 统计和带词长加权的词典匹配进行适应度评估，有效指导密钥搜索。

测试结果验证了工具各项功能的正确性和有效性。手动分析为用户提供了有力的统计支持和交互手段；自动破译功能在数据和算力支持下，展现了较强的破译能力，用户引导的加入进一步增强了实用性。

本工具主要局限在于对统计数据质量的依赖，以及对非标准或过短密文的处理能力有限。模拟退火算法虽能有效跳出局部最优，但不保证总能找到全局最优解。

综上，本作品为单表代换密码分析提供了一个实用、易学的辅助平台，达到了设计目标，具有良好的应用价值。未来可从算法优化、支持更多密码类型、提升用户体验等方面进一步完善。

5 代码链接

Simple_Substitution_Cypher_tool，by: 王泓镔 PB23000102,GitHub

参考文献

[1] Stallings, W. (2017). *Cryptography and Network Security: Principles and Practice (8th Edition)*. Pearson.