

工程成果	基于特征优选融合的隐写分析与系统
申请人	赵险峰
团队成员	黄炜、曹纭、夏冰冰、张纪宇、安宁钰、周楠、张弘、邓果、孙磊
申请人自述	<p>(请简述工程成果的目的和意义, 解决了什么问题, 有何贡献或影响, 在何处应用, 应用效果等。总字数不超过 1000 字, 可附页)</p> <p>隐写 (Steganography) 通过在媒体中隐藏机密信息, 试图掩盖保密通信或者存储的事实, 而隐写分析 (Steganalysis) 的目的是判断媒体中隐写信息的存在, 其方法主要基于判断媒体统计特征在隐写后的变化。据媒体多次报道, 隐写不但已经被情报人员和犯罪组织利用, 也被恶意程序用于隐藏代码和隐蔽通信, 因此, 国家有关部门急需建立有效的隐写分析技术手段。</p> <p>当前, 隐写分析研究人员已经提出了很多隐写特征, 可发掘的空间逐渐缩小, 因此, 通过特征优选和融合选取和组合特征与寻找新特征同样重要。但是, 当前发现、优选和融合隐写特征尚不存在流程上比较确定的方法, 主要依靠对隐写算法的人工分析, 依靠经验确定相应的特征组合用于分类识别, 这往往使得特征不能得到优化使用, 特征之间存在较大冗余, 影响了识别效果和效率; 以上情况也使得隐写分析的设计时间较长, 对新出现的隐写往往不能在短期内形成识别方法。</p> <p>为部分解决上述问题, 2012 年, 我们在隐写分析方面主要研发基于特征优选融合的隐写分析与系统, 主要贡献如下:</p> <p>1、特征及其关系评价指标。通过多方面估计特征和特征向量在隐写前后的变化, 得到了隐写特征敏感性评价指标, 主要包括互信息 (Mutual Information, MI) 指标、最大平均偏差 (Maximum Mean Discrepancy) 指标以及通过回归分析得到的综合指标; 根据指标的变化关系, 也得到了特征之间的关联性质。</p> <p>2、根据以上指标, 从统计特征集中发现了一批有效隐写特征, 优选了现有的特征。新发现的特征包括 JPEG 因子直方图特征函数</p>

特征、GIF 隐写 Walsh 谱特征、MP3 隐写块间量化不均匀特征、视频隐写运动向量校准特征、量化隐写差值特征等，优选的特征包括联合校准特征、Markov 特征、SPAM 特征等。

3、通过以上得到的特征间关联性以及互信息估计、主成分分析 (Principal Component Analysis, PCA) 等方法融合特征和去冗余。得到了一批新的融合特征，包括 SKLOIS JPEG PCA 融合特征、SKLOIS JPEG MI 融合特征、SKLOIS 位图 PCA 融合特征等。

本成果的形式包括专利、软件、论文等 (参证证明材料)，其中，面向增加网络防护设备隐写分析功能的中间件于 11 月在“2012 年深圳高技术交易会”展出，受到大量关注，一些企事业单位表示了合作意向。在核心的性能指标——分析准确率方面，以上融合特征隐写分析针对当前普通配置下的高隐蔽性隐写识别准确率超过 80%，相比特征融合前的结果普遍提高 4% 以上；尤其是，以上隐写分析的设计流程有较强的确定性，已经用软件实现自动和半自动的辅助设计，针对新出现的隐写，可以在特定的特征范围内自动确定敏感特征，一般能在 24 小时内给出新的隐写分析方法和软件模块。成果目前已经或者正在得到如下应用：

1、安全部系统应用。本成果的重要部分——“SS-03 特征优选融合隐写分析系统”于 2012 年 8 月交付安全部某部门使用，2012 年 9 月通过集中测试，目前使用效果良好，分析准确率“在国内处于领先地位” (参见“应用证明”)。本团队已与该部门签订了新的合作合同 (30 万)，另一项在分布式高性能计算环境下深入运用本成果的合作也基本确定。

2、保密系统应用。2012 年，本成果主要部分进入国家保密系统某重大建设工程，子系统预算金额 1226 万元，用于网络环境下的隐写检查；在主机环境下的隐写检查方面，本团队也于 2012 年 11 月得到了保密机构的定向项目支持 (40 万)。