**华侨大学**

**本科毕业论文**



|  |  |
| --- | --- |
| **题目：** | **基于Visio矢量图自生成的信息** |
|  | **隐藏技术研究** |

**院（系）** 　 计算机科学与技术学院

**专 业**网络工程

**届 别** 2014届

**学 号** 1425132007

**姓 名** 董浩

**指导老师**  田晖 教授

**华侨大学教务处印制**

2018 **年**5 **月**

# 摘 要

信息隐藏是将隐秘信息隐藏到普通信息中进行传输且不让对手发觉。目前的信息隐藏方法大部分是基于有载体的，攻击者可通过隐写分析检测到载体中存在隐秘信息，甚至可通过分析载体特征集来获取其中的隐秘信息。因此，近年来无载体信息隐藏受到了越来越多研究者的关注。相较于基于载体修改的信息隐藏方法，无载体信息隐藏使得攻击者不可能用原始载体进行对照，从而令传统的检测方法难以奏效，从而具有更好的安全性。

本文以Visio矢量图为对象，提出了两种基于矢量图的无载体信息隐藏方法，其目标是在保证结构合理、感官不违和的前提下根据隐秘信息自动生成Visio矢量图，实现高效安全的隐蔽通信。

**关键词：**无载体信息隐藏；自由树；自生成算法；Visio；矢量图

# ABSTRACT

Information hiding is to conceal secret information into ordinary information for transmission and not let opponents find out. Current information hiding methods are mostly based on the carrier, and the attacker can detect the hidden information in the carrier by steganalysis, even can get the secret information through analyzing the characteristic of the carrier set. Therefore, in recent years, coverless information hiding has attracted more and more attention from researchers. Compared to the information hiding method based on carrier modification, the coverless information hiding makes it impossible for an attacker to use the original carrier for comparison, thus making traditional detection methods ineffective and thus having better security.

This thesis takes Visio vector-graph as an object, and proposes two methods of coverless information hiding based on vector-graph. Their goals are to automatically generate Visio vector-graph based on secret information under the premise of ensuring the structure is reasonable and the senses are not against, to achieve efficient and secure covert communications. Specifically, the main research content of this paper is as follows:

Firstly, a large number of Visio vector-graph samples were analyzed and summarized, and a Visio vector-graph general model based on a tree structure was established. This model can depict most of the Visio vector-graph, and has the advantages of clear hierarchy, clear structure, and easy extension; Further, a free-tree-based Visio vector-graph self-generating algorithm is proposed. The principle is to use the tree structure model to guide the generation of Visio vector-graph, and to control the determination and selection of shapes according to heuristic rules, and ultimately generate the basic Visio vector-graph.

Secondly, a coverless information hiding method based on Visio vector-graph node self-generated is proposed. The principle is to use the tree structure model to guide the generation of Visio vector-graph, and to control the determination and selection of the shapes according to the secret information, and ultimately generate the basic Visio vector-graph. The performance of the proposed method was evaluated by means of security analysis, undetectability analysis, hidden capacity statistics, and the like.

Finally, a coverless information hiding method based on Visio vector-graph shape attributes self-generated is proposed. The principle is to use the tree structure model to guide the generation of Visio vector-graph, to control the determination and selection of graphic elements according to heuristic rules, and to control the settings of the shape attributes according to the secret information, ultimately generate the basic Visio vector-graph. The performance of the proposed method was evaluated by means of security analysis, undetectability analysis, hidden capacity statistics, and the like.

**keywords：**Coverless Information Hiding; Free Tree; Self-generating Algorithm; Visio; Vector-graph

目 录

[摘 要 I](#_Toc515303701)

[ABSTRACT II](#_Toc515303702)

[第一章 绪 论 1](#_Toc515303703)

[1.1 研究的目的和意义 1](#_Toc515303704)

[1.2 国内外研究现状 2](#_Toc515303705)

[1.3 主要研究内容 3](#_Toc515303706)

[1.4 本文篇章结构 4](#_Toc515303707)

[第二章 基于自由树的Visio矢量图自生成算法 5](#_Toc515303708)

[2.1 Visio对象模型概述 5](#_Toc515303709)

[2.2 可行性分析 7](#_Toc515303710)

[2.3 本章小结 8](#_Toc515303711)

[第三章 基于矢量图节点自生成的无载体信息隐藏方法 9](#_Toc515303712)

[3.1 基于矢量图节点自生成的无载体信息隐藏模型 9](#_Toc515303713)

[3.2 隐藏算法 9](#_Toc515303714)

[3.3 提取算法 12](#_Toc515303715)

[3.4 性能分析 14](#_Toc515303716)

[3.4.1 不可感知性 14](#_Toc515303717)

[3.4.2 不可检测性 15](#_Toc515303718)

[3.4.3 安全性 15](#_Toc515303719)

[3.4.4 隐藏容量 15](#_Toc515303720)

[3.5 本章小结 16](#_Toc515303721)

[第四章 基于矢量图图元属性自生成的无载体信息隐藏方法 17](#_Toc515303722)

[4.1 基于矢量图图元属性自生成的无载体信息隐藏模型 17](#_Toc515303723)

[4.1.1 图元特征集 17](#_Toc515303724)

[4.2 隐藏算法 21](#_Toc515303725)

[4.3 提取算法 23](#_Toc515303726)

[4.4 性能分析 27](#_Toc515303727)

[4.4.1 不可感知性 27](#_Toc515303728)

[4.4.2 不可检测性 28](#_Toc515303729)

[4.4.3 安全性 28](#_Toc515303730)

[4.4.4 隐藏容量 29](#_Toc515303731)

[4.5 本章小结 29](#_Toc515303732)

[第五章 结论与展望 30](#_Toc515303733)

[5.1 总结 30](#_Toc515303734)

[5.2 进一步工作展望 30](#_Toc515303735)

[参 考 文 献 32](#_Toc515303736)

[致 谢 35](#_Toc515303737)

# 第一章 绪 论

## 研究的目的和意义

在当今的信息化社会，信息技术给人们的生产和生活方式均带来了翻天覆地的变化。生产车间里繁忙的工人逐渐被智能化的机械替代，生产力也由以体力劳动为主转变为以脑力劳动为主。人们可以通过网上购物浏览全国各地的商品，相隔万里的朋友通过微信、QQ这些交流平台就能实现实时沟通，而且在线支付、送货上门这些便捷功能更是几乎让我们做到了“足不出户”。信息技术使得人们的工作和生活更加的便利，同时也给人类带来了新的威胁和挑战。如何保障信息安全就是其中的一个典型问题。人们在购物时使用在线支付是一件很便捷的事情，一旦支付信息泄露，别人盗取你账户上的钱款也很“便捷”。这只是其中一个方面，还有垃圾短信，骚扰电话，冒名办卡、犯案等等，所以保障信息安全是一件非常重要的事情。

最初的手段是使用加密术，将信息经过一系列的加密算法最后转化为无法识别的乱码，使之在缺少必要信息时不可读。但加密术在保证了安全性的同时，却忽略了隐蔽性这个问题，甚至反而容易成为第三方攻击的焦点。世界上最早的密码情报就是从波斯帝国回雅典送信的雅典间谍身上的腰带情报，但却因其杂乱的字符引起了斯巴达的注意，虽然通过各种排列组合都无法获得正确信息，但最后还是被破解了，并获得了波斯军队要袭击斯巴达的情报。由此可见，仅仅对信息进行加密是不够的，再复杂的加密方式都有被破解的可能，只是时间的长短问题。

隐写技术是最近十几年来广受关注的一种信息安全技术。根据信息论的观点，信息和数据冗余构成了数据，所以数据本身是存在冗余的（数据特性冗余）。人类的眼睛对于色度变化、高频信号这些相对不敏感，其中一些数据对于人眼是多余的，即视觉冗余[1]。隐写技术利用这些冗余，在不影响载体使用价值和实际含义的前提下，将隐秘信息隐藏在多媒体文件信息中，使得隐秘信息可以在不引起其他人注意的情况下实现传输。与传统的通过使得信息内容在缺少必要信息时“不可读”的加密技术相比，隐写技术使得隐秘信息变得“不可见”，从而进一步提高了隐秘通信的安全性。

但隐写术也不是万能的，第三方仍然可以通过隐写分析等方法检测到载体中存在隐秘信息，甚至可通过分析载体特征集来获取其中的隐秘信息。于是本文提出的基于Visio矢量图自生成的信息隐藏技术，通过制定的生成规则自生成Visio矢量图，并在矢量图自生成的过程中嵌入隐秘信息。由于没有初始载体，可避免现阶段的隐写分析等方法的检测。

Office办公系列软件广受人们欢迎，Visio作为其中一款专业的办公制图软件，因其强大的表现能力和丰富的内容而被越来越多的人们使用。所以用Visio矢量图来进行信息隐藏是一个不错的选择。

## 国内外研究现状

从已有文献来看，隐写术在近年来取得了迅猛的发展[2]。隐写对象已从早期的图像载体发展到今天的几乎所有的数字媒介[3]，包括音频[4]、视频[5]、流媒体[6]、网络协议[7]、文本[8]等。基于图像的隐写主要利用的是视觉冗余，人眼对于图像中一些色度的变化是不敏感的，所以图像载体的冗余信息很多，隐藏容量大，而且隐藏简单，是目前研究成果最多、应用最为广泛的隐写对象。基于音频的隐写研究成果相对较少，因为人耳的敏感性要比人眼要高，冗余相对较少；其隐写思路与图像隐写类似，通过修改语音数据的冗余信息来实现，隐写算法可分为回声隐写[9]、相位隐写[10]等。基于视频的隐写研究也逐渐受到人们关注，其利用人眼对于运动图像的相对不敏感性的视觉冗余，而且视频是帧组成的，其隐藏空间巨大，可实现在保证视频文件感官质量的前提下完成大文件的隐藏[11]。基于文本的隐写主要可分为三类：基于格式的文本隐写，将信息隐藏在行间距、字体、字体颜色等格式数据中[12]；基于语义的文本隐写，根据隐秘信息替换或自动生成语句[13]；基于扩展空间的文本隐写，此类隐写方法主要是将文本转换为其他格式的载体对象再根据其载体冗余来进行隐写，如图像、视频、流媒体等。基于网络的隐写是最新的一种信息隐藏技术，可以说是非常年轻，但其发展却十分迅猛，研究成果颇多，有一股后浪之势；主要利用OSI模型设计时存在的冗余和缺陷，以各种通信协议为载体来进行信息隐藏；因现在网络协议的广泛使用和通信协议的多种多样，其研究内容丰富、潜力巨大，如FTP协议[14]、CSMA/CD协议[15]、TCP/IP协议[16]等。

隐写技术也从空域中简单的最低有效位替换隐写发展到今天的各种变换域隐写。空间域的算法优点是隐藏容量大，嵌入提取过程简单但是鲁棒性差，最低有效位隐藏（Least Significant Bit, LSB）是其中最经典的，但由于其抗隐写分析能力不好，易被卡方检测法检测，故又有许多改进算法：文献[17]提出了一种更为简单有效的像素最优化调整算法（Optimal Pixel Adjustment Process, OPAP）；针对LSB算法的抗隐写分析能力不好的缺陷；文献[18]提出了LSB匹配隐写术（LSB Matching），通过对载体嵌入后的分布特征进行改进，增强了其不可检测性，抗隐写分析能力大大增强；Mielikainen提出的基于像素对匹配(Pixel Pair Matching, PPM)的LSB 匹配算法[19]，具有更好的隐写性能，缺点是隐藏容量太小，像素对中仅仅只能嵌入2bit秘信。相对而言变换域算法的优点是抗攻击能力强，即鲁棒性好，但是隐藏容量小，算法复杂，可分为量化索引调制隐写[20]、扩频隐写[21]、统计隐写[22]等。

随着无载体信息隐藏技术的提出，隐写术又迈出了新的一步。由于其无需通过对载体进行修改来实现隐秘信息的隐藏，且可抵抗各种隐写分析手段，越来越受到人们的关注，并有很多成果[23-26]。

综上所述，在信息隐藏技术方面的成果是十分丰富的，其发展也十分迅猛。基于有载体的信息隐藏技术基本已十分成熟，而无载体信息隐藏技术则潜力巨大，具有很大的研究价值。

## 主要研究内容

目前的信息隐藏方法大部分是基于有载体的，第三方可通过隐写分析等方法检测到载体中存在隐秘信息，甚至可通过分析载体特征集来获取其中的隐秘信息。本文提出了基于Visio矢量图自生成的信息隐藏技术。通过制定的生成规则自生成Visio矢量图，并且在矢量图自生成的过程中嵌入隐秘信息。由于无原始载体文件进行对照，可避免现阶段大部分的隐写分析等方法的检测。

本文提出的基于Visio矢量图自生成的信息隐藏技术，在保证结构合理、感官不违和的前提下自生成Visio矢量图，并在生成过程中实现信息隐藏。主要研究内容如下：

1. 对大量的Visio矢量图样本进行分析、总结，建立了基于树形结构的Visio矢量图一般模型，该模型可描绘绝大部分Visio矢量图，且具有层次分明、结构清晰和易于扩展等优点；进一步，提出了基于自由树的Visio矢量图自生成算法，其原理是利用树形结构模型指导Visio矢量图的生成，并根据启发式规则控制图元的确定和选择，最终生成基本的Visio矢量图。
2. 提出一种基于Visio矢量图节点自生成的无载体信息隐藏方法，其原理是利用树形结构模型指导Visio矢量图的生成，并根据隐秘信息控制图元的确定和选择，最终生成基本的Visio矢量图。通过安全性分析、不可检测性分析、隐藏容量统计等方式对提出的方法的性能进行了评测。

提出一种基于Visio矢量图图元属性自生成的无载体信息隐藏方法，其原理是利用树形结构模型指导Visio矢量图的生成，根据启发式规则控制图元的确定和选择，再根据隐秘信息控制图元属性的设置，最终生成基本的Visio矢量图。通过安全性分析、不可检测性分析、隐藏容量统计等方式对提出的方法的性能进行了评测。

## 本文篇章结构

第一章：介绍本文提出的基于Visio矢量图自生成的信息隐藏技术的研究目的和意义，再根据阅读的文献概括了国内外关于信息隐藏技术的研究现状，最后介绍一下该技术的主要研究内容。

第二章：通过对Visio对象模型的研究掌握了使用C#创建和生成Visio矢量图的方法，提出了基于自由树的Visio矢量图自生成算法，并对其进行了可行性分析。

第三章：分析节点自生成时图元集选择特点，提出了基于Visio矢量图节点自生成的无载体信息隐藏方法，介绍了该方法在隐藏与提取上的详细步骤，最后对该方法进行了性能分析。

第四章：分析节点图元的图元特征集，提出了基于Visio矢量图图元属性自生成的无载体信息隐藏方法，介绍了该方法在隐藏与提取上的详细步骤，最后对该方法进行了性能分析。

第五章：对本文进行总结并提出进一步工作的展望。

# 第三章 基于矢量图节点自生成的无载体信息隐藏方法[[1]](#footnote-1)

## 基于矢量图节点自生成的无载体信息隐藏模型

通过前文的分析可知，本文提出了一种基于自由树的Visio矢量图自生成算法，其中对于各节点图元的选择是随机的。而为达到信息隐藏的目的，可以用隐秘信息来控制各节点图元的选择，从而驱动生成一个包含隐秘信息的Visio矢量图，即基于Visio矢量图节点自生成的无载体信息隐藏方法。在生成过程中仅控制对图元的选择，未对图元作出任何修改，所以可避免隐写分析等手段的检测，保证了隐蔽性。

方法的具体过程如下。读取隐秘信息，将隐秘信息加密转化为二进制流数据。按照前序遍历顺序生成自由树，用加密后的隐秘信息控制每个节点图元的选择。最后形成带隐秘信息的Visio矢量图。算法模型如图3.1。



图3.2 基于Visio矢量图节点自生成的无载体信息隐藏方法模型

无载体信息隐藏的一个最大问题就是其不可感知性，用算法自生成的文件要保证感官不违和，且具有逻辑性。该方法中对于每个节点的图元集，并不是随便确定范围的。如该节点的父节点为集线器，那么该节点就不能是路由器，而应该从PC、打印机这些中选择。并且该方法采用的树形结构具有层次感，自上而下，从左往右，条理清晰，逻辑性强。所以该方法具有良好的不可感知性。

基于Visio矢量图节点自生成的无载体信息提取算法流程如图3.4所示。

## 性能分析

设计一个基于Visio矢量图节点自生成的无载体信息隐藏方法，在保证正确性同时还要考虑该方法在各方面的性能表现。本节将从不可感知性，不可检测性，安全性、隐藏容量四个方面对算法进行性能分析。

### 不可感知性

任意选择Visio矢量图图元进行绘制，不仅难以找到规律开制定相应的规则，而且绘制完成的Visio矢量图的感官和谐性无法得到保证。按照树形结构来生成Visio矢量图，避免了随机生成图元造成的杂乱无章的感觉，图3.5为通过基于Visio矢量图节点自生成的无载体信息隐藏方法生成的Visio矢量图 ,可见与一般人工绘制的Visio矢量图并无太大差别，不可感知性良好。

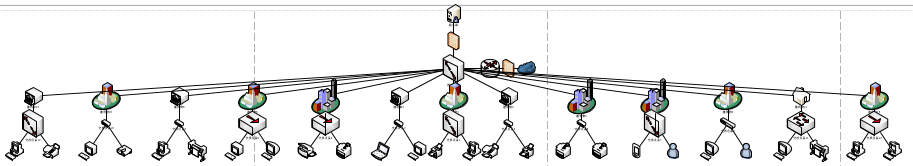


图3.5 根据基于Visio矢量图节点自生成的无载体信息隐藏方法生成的Visio矢量图（实验结果）

# 参 考 文 献

1. 中兴通讯学院. 对话多媒体通信[M]. 人民邮电出版社, 2010.
2. E. Zielińska, W. Mazurczyk, K. Szczypiorski, Trends in Steganography[J], Communications of the ACM, 2014, 57(3):86-95。
3. A. Cheddad, J. Condell, K. Curran, P. M. Kevitt. Digital Image Steganography: Survey and Analysis of Current Methods[J], Signal Processing, 2010, 90 (3): 727–752
4. F. Djebbar, B. Ayad, K. A. Meraim, H. Hamam. Comparative Study of Digital Audio Steganography Techniques[J], EURASIP Journal on Audio Speech & Music Processing, 2012, 2012(1):25.
5. M. M. Sadek, A. S. Khalifa, M. G. M. Mostafa. Video Steganography: A Comprehensive Review, Multimedia Tools and Applications[J], 2015, 74 (17): 7063–7094
6. W. Mazurczyk. VoIP Steganography and Its Detection - A survey[J], ACM Computing Surveys, 2013, 46, article no. 20
7. A. El-Atawy, Q. Duan, E. Al-Shaer. A Novel Class of Robust Covert Channels Using Out-of-Order Packets[J], IEEE Transactions on Dependable and Secure Computing, 2017, 14(2): 116-129
8. E. Satir, H. Isik. A Huffman Compression Based Text Steganography Method[J]. Multimedia Tools & Applications, 2014, 70(3):2085-2110.
9. Y. Xiang, I. Natgunanathan, D. Peng, et al. A Dual-Channel Time-Spread Echo Method for Audio Watermarking[J]. IEEE Transactions on Information Forensics & Security, 2012, 7(2):383-392.
10. M. NUTZINGER, J. WURZER. A Novel Phase Coding Technique for Steganography in Auditive Media[C]. Proceedings of IEEE International Conference on Availability, Reliability and Security, Vienna, Austria: IEEE, 2011:91-98
11. 梁惠, 郭立, 陈运必. 基于视觉感知模型的大容量视频隐写算法[J]. 中国科学院大学学报, 2010, 27(1):55-62.
12. 陈青, 牛月蕊, 周黎敏, 等. 一种基于字符间距编码的双重水印嵌入的文本水印方法: CN, CN103049682A[P]. 2013.
13. 徐迎晖, 杨榆, 钮心忻, 等. 基于语义的文本隐藏方法[J]. 计算机系统应用, 2006, 15(6):91-94.
14. 李彩秋, 周亚建, 魏更宇. FTP协议信息隐藏方法的研究与改进[J]. 计算机工程与应用, 2013, 49(21):111-113.
15. 姜绍君, 王颖. 基于CSMA/CA协议的隐藏终端问题的改进[J]. 计算机技术与发展, 2015(1):33-36.
16. 刘朝晖, 孙星明. 基于TCP/IP协议的信息隐藏新方法[J]. 计算技术与自动化, 2004, 23(3):109-112.
17. C. K. Chan, L. M. Cheng. Hiding Data in Images by Simple LSB Substitution[J], Pattern Recognition, 2004, 37(3):469-474.
18. T. Sharp. An Implementation of Key-Based Digital Signal Steganography[C]// Proceedings of International Workshop on Information Hiding, Springer-Verlag, 2001:13-26.
19. J. Mielikainen, LSB matching revisited[J], IEEE Signal Processing Letters, 2006, 13(5):285-287.
20. S. Li, Y. Jia, C.-C. Jay Kuo. Steganalysis of QIM Steganography in Low-Bit-Rate Speech Signals[J], IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2017, 25(5): 1011 – 1022
21. M. Li, M. K. Kulhandjian, D. A. Pados, S. N. Batalama, M. J. Medley. Extracting Spread-Spectrum Hidden Data from Digital Media[J], IEEE Transactions on Information Forensics and Security, 2013, 8(7): 1201 – 1210
22. V. Sedighi, R. Cogranne, J. Fridrich. Content-Adaptive Steganography by Minimizing Statistical Detectability[J], IEEE Transactions on Information Forensics and Security, 2016, 11(2): 221 – 234
23. X. Chen, S. Chen, Y. Wu, Coverless Information Hiding Method Based on the Chinese Character Encoding[J], Journal of Internet Technology, 2017，8(2):91-98.
24. C. Yuan, Z. Xia, X. Sun, Coverless Image Steganography Based on SIFT and BOF[J], Journal of Internet Technology, 2017，18(2):209-216.
25. Z. Zhou, Y. Mu, C. N. Yang, N. Zhao, Coverless Multi-keywords Information Hiding Method Based on Text[J], International Journal of Security and Its Applications, 2016，10(9):309-320.
26. J. Zhang, H. Huang, L. Wang, H. Lin, D. Gao, Coverless Text Information Hiding Method Using the Frequent Words Hash[J], International Journal of Network Security, 2017,19(6): PP.1016-1023.
27. 熊雄. 浅析利用程序处理Visio文档[OL]. Available： http://www. paper.edu.cn/html/releasepaper/2009/06/445/, 2018-01-30
28. A. Castiglione, A. D. Santis, C. Soriente. Taking Advantages of A Disadvantage: Digital Forensics and Steganography Using Document Metadata[J]. Journal of Systems & Software, 2007, 80(5):750-764.

# 致 谢

吋光飞逝，转眼间四年紧张而又充实的大学生活即将画上句号。在这四年的学习生活中，我得到了许多来自老师、同学和朋友的关怀和帮忙。在学位论文即将完成之际，我要向所有在大学期间给予我支持、帮忙和鼓励的人献上我最诚挚的谢意。

首先，我要感谢我的指导老师田晖老师对我的教导。从毕设课题的选择，开题报告，算法的改进，到论文的撰写，田晖老师都给了我悉心的指导和热情的帮忙，让我能从一开始对无载体信息隐藏领域的0基础到一步步深入对该领域的了解，进而提出自己的研究思路，实现算法并最后完成论文。田晖老师对工作的认真负责、对学术的钻研精神和严谨的学风，都是值得我终生学习的。

不积跬步何以至千里，本设计能够顺利的完成，也归功于各位任课老师的认真负责，使我能够很好的掌握和运用专业知识，并在设计中得以体现。在此向华侨大学，计算机学院的全体老师表示由衷的谢意。感谢他们四年来的辛勤栽培

最后，感谢我的家人对我的关爱和鼓励，以及所有陪我一路走来的同学和朋友，正是由于他们的支持和照顾，我才能安心学习，并顺利完成我的学业。

董浩

2018 年 5 月 20 日

1. 1 本章的研究成果已申请专利：田晖,董浩,卢璥,陈永红,王田,蔡奕侨. 一种基Visio拓扑结构图自生成的的隐写方法，中国发明专利. [↑](#footnote-ref-1)