

广度优先破解二叉树加密算法

任广永

(安庆师范学院 计算机与信息学院, 安徽 安庆 246003)

摘要:通过介绍二叉树加密算法,分析二叉树加密算法是利用加密二叉树的树形对明文信息进行加密的处理加密过程和其存在的漏洞,提出了基于广度优先搜索算法的二叉树加密算法的破解方法,基于广度优先搜索算法的破解算法是由于在二叉树加密算法的密文传输过程中存在着可以被截获的公钥,对于公钥可以分析成一个可能的加密森林,在应用广度优先算法的搜索过程中生成解密链表,最后达到了破解的目的。

关键词:二叉树; 先根序搜索序列; 后根序搜索序列; 广度优先

中图分类号: TP301.6

文献标识码: A

文章编号: 1673-629X(2008)02-0156-03

BFS Crack Binary Tree's Encryption Algorithm

REN Guang-yong

(College of Computer and Information, Anqing Teachers College, Anqing 246003, China)

Abstract: Introduces binary tree encryption algorithm, and analyzes binary tree encryption algorithm that uses binary tree's shape. Then found BUGs, and hold up BFS crack binary tree's encryption algorithm. BFS crack binary tree's encryption algorithm based on binary tree's encryption algorithm's public key, which can be intercepted when it is transporting. So a encryption's forest can be found out by the public key. And an encryption-link-list can be gained. So BFS crack binary tree's encryption algorithm is successful.

Key words: binary tree; preorder sequence search; postorder sequence search; BFS

0 引言

二叉树^[1]是一种非常重要的非线性的数据结构,是以分支关系定义的层次结构。在哈夫曼树编码已经被广泛应用的今天,现在把二叉树与密码学联系到了一起,进行加密编码,已经被很多人接受。二叉树加密算法是基于在仅仅知道二叉树的先根序遍历序列和后根序遍历序列的基础上无法唯一确定一个树的理论之上的加密算法。但是这并不是无懈可击的,在已知先根序遍历序列和后序遍历的序列后,结合广度优先搜索,能够对二叉树加密算法进行破解。

1 二叉树加密算法

1.1 二叉树加密算法简述

树是 $n(n \geq 0)$ 个结点的有限集。在任意一棵非空树中:有且仅有一个特定的称为根的结点;当 $n > 1$

时,其余结点可分为 $m(m > 0)$ 个互不相交的有限集 T_1, T_2, \dots, T_m , 其中每一个集合本身又是一棵树,并且称为根的子树^[1]。

二叉树也是递归定义的,二叉树是另一种树型结构,它的特点是每个结点至多只有二棵子树(即二叉树中不存在度数大于2的结点),并且,二叉树的子树有左右之分,其次序不能任意颠倒。逻辑上二叉树有五种基本形态:

- (1)空二叉树;
- (2)只有一个根结点的二叉树;
- (3)右子树为空的二叉树;
- (4)左子树为空的二叉树;

(5)完全二叉树(Complete Binary Tree):若一棵二叉树至多只有最下面的两层上结点的度数可以小于2,并且最下一层上的结点都集中在该层最左边的若干位置上,则此二叉树称为完全二叉树。

尽管二叉树与树有许多相似之处,但二叉树不是树的特殊情形^[2]。

基于二叉树的加密算法就是利用二叉树来对二进制信息进行编码。加密二叉树的任意一个结点用唯一的序号来表示,加密二叉树是一棵树形不确定的任意

收稿日期:2007-08-07

基金项目:国家自然科学基金资助项目(60773128);安徽省教育厅自然科学基金资助项目(2006KJ081B);安徽安庆科技重点公关资助项目(200705)

作者简介:任广永(1957-),男,安徽固镇人,高级实验师,硕士,研究方向为计算机应用与软件开发。

二叉树。并且约定二叉树的左分支表示字符“0”,右分支表示字符“1”。

1.2 二叉树加密算法的加密解密过程

1)基于二叉树的加密算法的加密过程如下:

首先,用于加密的二叉树由信息的管理者开始选取,利用二叉树的树型对信息加密,将明文信息转化为密文信息。加密之后,信息的管理者将二叉树的先根序遍历序列和后根序遍历序列作为公钥随密文传出。

2)基于二叉树的加密算法的解密过程如下:

在接收到密文后,首先根据加密树的先根序遍历序列和后根序遍历序列,在已存的现有加密树中搜寻信息加密用的加密树,在密文接收者那里存有能够由先根序遍历和后序遍历能够唯一确认的二叉加密树。基于二叉树的解密算法是加密算法的逆过程。密文中的每一个字符对应二叉树的一个结点,在已知树形的情况下,从根结点出发到密文表示的相应结点的路径上,分支字符组成的字符串作为该密文字符所对应的明文段。将所有密文字符所对应的明文组合在一起,即是明文信息。加密算法的密钥就是二叉树的树形。从二叉树的根结点出发,检查根结点或子树的根结点:如果该结点不是需要查找的密文字符,则搜索该结点的左子树,如果左子树中存在密文字符,则字符串增加一位0;以根结点的左孩子为根结点继续搜索,如果左子树中不存在,那么密文字符位于右子树当中,字符串增加一位1;以根结点的右孩子为根结点继续搜索,直到搜索到密文字符,此时就完成了对密文字符的解密工作^[3]。

1.3 二叉树加密算法加密树的选取

首先构造一棵完全二叉树,然后随机对二叉树多次进行移位和剪枝两种处理,由此形成加密二叉树。对结点1、2进行移位处理是指:交换二叉树中分别以1和2两个结点为根的子树的位置。对结点3点进行剪枝处理是指:将以3为根的子树任意插入到一个孩子数小于2的结点之下,作为左孩子或右孩子。

加密二叉树的构造是任意的、随机的,最主要的就是在一棵二叉树与另一棵二叉树能够存在相同的先根序搜索序列和后根序搜索序列时,只能选择其中的一棵二叉树,但是有些二叉树的结构是不能被使用的,例如只存在左分支或者右分支的二叉树,这种二叉树在编码时,可以编码的信息有限,编码包括许多连续的1,特征明显,容易破解。完全二叉树也是不被允许的,因为编码所用的二叉树是在完全二叉树的基础之上变化而来的。

1.4 二叉树加密算法举例

将需要加密的明文信息转化为二进制字符串,则

加密过程是分解明文中的字符串,从根结点出发,按字符“0”或“1”确定查找该结点的左孩子或右孩子,直至叶结点或者只有一个孩子的树结点,便得到了该子串相对应的结点序号。将结点序号按照明文的顺序排列在一起,即得到基于二叉树加密算法的密文。

信息管理者应用的加密树如图1所示。

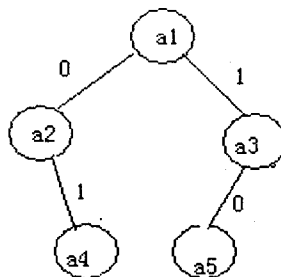


图1 加密树

假设需要加密的明文为:001101。

通过加密树的加密,知道a1是NULL,a2是0,a3是1,a4是01,a5是10。所以以上的明文经过加密之后的密文为:a2a4a3a5。加密的公钥是:先根序遍历序列a1a2a4a2a5和后根序遍历序列a4a2a3a5a1。

在信息接收者接收到公钥和密文时,根据公钥,即密树存储库中搜索对应的加密树,然后,对密文的解密之后得到对应的明文为:001101。这样最后再把二进制转化为自然语言,就得到了信息管理者传输的自然语言明文。

2 广度优先破解算法

根据二叉树理论,具有 n 个结点的二叉树具有 $C_{2n}^n/(n+1)$ 种排列方式^[4]。当二叉树具有128个结点时,共有超过 2.22×10^{291} 种不同的排列^[2],这样也就有 2.22×10^{291} 种二叉树存在,但当知道了树的前序遍历序列和后根序遍历序列后,可能的二叉树的数量也减少到了可以计算的数量,这就给在知道先根序遍历序列和后根序遍历序列后破解密文成为了可能。

2.1 广度优先算法

广度优先搜索算法又称为宽度优先搜索算法,即BFS(Breadth First Search),常常与深度优先搜索并列提及。这是一种相当常用的图算法,其特点是:每次搜索指定点,并将其所有未访问过的近邻加入搜索队列,循环搜索过程直到队列为空^[5]。

把图论中的这个算法应用到二叉树中,对二叉树进行遍历搜索,就是以树的根节点开始,逐个将各个遍历到的结点的兄弟节点加入搜索树,访问到的节点出队,如此反复直到队列为空为止。二叉树的广度遍历就是在图的广度优先算法扩展而来。广度优先是一种

控制结点扩展的策略,这种策略优先扩展深度小的结点,把问题的状态向横向发展。

2.2 破解算法

首先,在获取到公钥后,根据先根序遍历序列和后根序遍历序列构造出所有可能应用的加密树(假设有 N 种),形成一个可能加密森林。建立一个节点 $root$,建立一个基于这个可能加密森林的 N 叉可能加密树。

然后根据密文,在应用广度优先搜索算法的基础上,对上述 N 叉可能加密树进行广度优先算法查找出所有的对应的密文与二进制明文之间的对应关系,同时把所有论域中所有的加密符号对应的二进制明文全部按加密树森林的对应顺序存储起来,形成一个密文与明文二进制相对应的破解库。

最后,把二进制明文转换成自然语言的明文,存储到相应的密文与明文的破解库中,替代原来的二进制明文,按照密文的先后顺序,找到查找密文与明文破解库中的对应节点,形成一个按照密文顺序排列的破解链表,再按照自然语言的语义,就能够得到对应密文的明文了。

2.3 举例

由第一节给出的加密例子给出对应的破解过程。

根据先根序遍历序列和后根序遍历序列,也就是公钥: $a_1a_2a_4a_2a_5$ 和 $a_4a_2a_3a_5a_1$,得出有可能的加密树如图 2 所示。

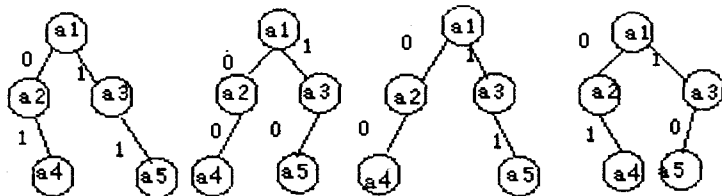


图 2 可能的加密森林

根据以上的可能的加密树,由广度优先的破解方法,要生成可能加密森林的一个 N 叉树,因此,根据图 2 可能的加密树得出加密森林,如图 3 所示。

在可能的加密 N 叉树上应用广度优先搜索算法,

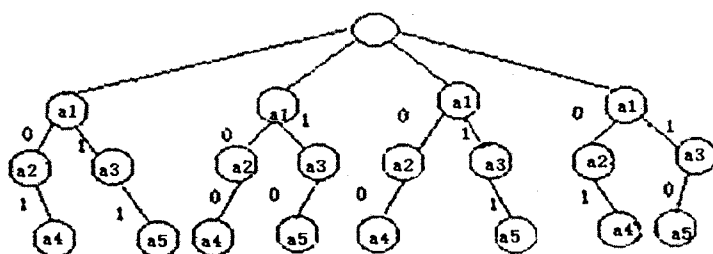


图 3 可能的加密 N 叉树

把遍历到的节点按对应顺序写到解密链表中,如图 4 所示。

p	a 1	a 2	a 3	a 4	a 5
	NULL	0	1	0 1	1 1
	NULL	0	1	0 0	1 0
	NULL	0	1	0 0	1 1
	NULL	0	1	0 1	1 0

图 4 解密链表

然后根据广度优先搜索算法,将这个链表中的二进制转换成自然语言,最后按照自然语义得到相应的明文。这样就达到了破解的目的。

3 总结

通过分析了二叉树加密算法的加密与解密及其传输过程,总结出了基于图的广度优先搜索算法的破解算法,主要针对在基于二叉树的先根序搜索序列和后根序搜索序列能够得到一个加密树森林,通过加密森林得到一个 N 叉加密树,再对 N 叉树进行广度优先搜索,逐步建立解密链表,达到破解的目的。

参考文献:

- [1] 严蔚敏,吴伟明.数据结构[M].北京:清华大学出版社,2003.
- [2] 臧武军.数据加密技术[J].教育信息化,2005,12:28-30.
- [3] 陈伟,付宇洁,秦科.基于二叉树的加密算法[J].实验科学与技术,2006(2):81-85.
- [4] 张乐星.一种新的网络数据加密方案[J].计算机时代,2004(3):10-11.
- [5] 唐明华.用改进的广度优先搜索算法计算点的出行范围[J].茂名学院学报,2006,16(3):35-38.

(上接第 142 页)

参考文献:

- [1] Abraham, Ajith. Business Intelligence from Web Usage Mining[J]. Journal of Information & Knowledge Management, 2003,2(4):375-390.
- [2] 费爱国,王新辉.一种基于 Web 日志文件的信息挖掘方法

[J]. 计算机应用,2004,24(6):57-59.

- [3] 谢芳,王波.基于关联规则个性化推荐的改进算法[J].计算机应用,2006,26:149-151.
- [4] 胡可云,陆玉昌,石纯.基于概念格的分类和关联规则的集成挖掘方法[J].软件学报,2000,11(11):1478-1484.
- [5] 石晶,龚震宇,袁抗萍.基于 Web 使用挖掘的个性化服务系统[J].电子科技大学学报,2002,31(4):399-403.